

R. Bruce Martin
David B. Burr
Neil A. Sharkey
David P. Fyhrie

Skeletal Tissue Mechanics

Second Edition



Springer

Skeletal Tissue Mechanics

Springer Science+Business Media LLC New York

R. Bruce Martin David B. Burr Neil A. Sharkey
David P. Fyhrie

Skeletal Tissue Mechanics

Second Edition

With 194 Illustrations



Springer

R. Bruce Martin, retired
Orthopaedic Surgery
UC Davis Medical Center
Olga, WA, USA

Neil A. Sharkey
Kinesiology, Orthopedics
and Rehabilitation
Pennsylvania State University
University Park, PA, USA

David B. Burr
Department of Anatomy
and Cell Biology
Indiana University School of Medicine
Indianapolis, IN, USA

David P. Fyhrie
Orthopaedic Surgery, Biomedical
Engineering
UC Davis Medical Center
Sacramento, CA, USA

ISBN 978-1-4939-3001-2 ISBN 978-1-4939-3002-9 (eBook)
DOI 10.1007/978-1-4939-3002-9

Library of Congress Control Number: 2015942705

Springer New York Heidelberg Dordrecht London
© Springer Science+Business Media New York 2015

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made.

Printed on acid-free paper

Springer Science+Business Media LLC New York is part of Springer Science+Business Media
(www.springer.com)

Dedication

This book is dedicated to R. Bruce Martin, one of the smartest and most thoughtful scientists any of us has known. It is unusual for a book to be dedicated to its first author. However, in this case, it is entirely appropriate. This book was Bruce's brainchild, based on many years of lecturing and thinking about the skeleton and how it works both from an engineering perspective and from a biological one. Many of the ideas expressed in this second edition of *Skeletal Tissue Mechanics*, as in the first, were original with Bruce, and they still retain a freshness and originality of thought that continue to make them valuable insights into the workings of our musculoskeletal system. He thought of them; he developed them; he applied them. He, more than most, understood both the mechanical and the biological aspects of the skeleton; he could integrate them effectively, and he could explain them even to the novice; and with humility and objectivity, he rigorously critiqued them. We know, because he did this for us. Bruce's skill in integration and his interest in education make this book not just useful for teaching. It provides a foundation for further experimental and theoretical research in the skeletal tissues.

Bruce was not able to participate in the revision for this edition, although he wanted to. Several years before we began this revision, he tried to push us to bring the sadly out-of-date text into compliance with modern theoretical constructs, and informed by the many new experimental discoveries in areas such as mechanical signaling, the importance of bone architecture, bone quality markers, and bone fatigue, among others. The reader will easily see, by comparison to the first edition, that the science surrounding skeletal biology and its mechanical implications has by no means stood still. More sophisticated experiments and greater biological depth to computational modeling have produced insights that were unknown just a few years ago. Significant progress has been made in a few short years. We hope that, in Bruce's spirit, we have captured some of these advances in this second edition of *Skeletal Tissue Mechanics*.

Preface

The first edition of *Skeletal Tissue Mechanics*, published in 1998, grew out of Bruce Martin's lecture notes for a course by the same name, which was part of the biomedical engineering curriculum at the University of California at Davis. Given this history, it is not surprising that the book is intended to be used as a textbook, not a reference book. But it also is intended to comprehensively review the biology of all the skeletal tissues and their mechanical properties and biomechanical functions. The integration of anatomy and physiology with structural and material behavior was one of the book's principal objectives, and this remains the goal of the second edition as well. The interdisciplinary nature of the book is strengthened by the complementary blend of perspectives brought by the authors, who have backgrounds in Functional Anatomy, Anthropology, Physiology, Biomedical and Mechanical Engineering, and Orthopedics.

The first edition was primarily written for graduate and advanced undergraduate students in biomedical engineering. It fit a niche between textbooks focused on the biomechanics of locomotion and sport, and those aimed more specifically at orthopedic biomechanics, which often do not concentrate on the mechanical behaviors of skeletal tissues or the biologic processes that underlie them. In practice, the book was used more broadly by residents in orthopedic surgery and students at various levels who were studying anthropology, kinesiology, veterinary medicine, and skeletal physiology.

An explosion of new information—both experimental and computational—has improved our understanding of the influences on, and adaptability of, skeletal tissues. Our approaches to understanding both the dynamic and the static properties of skeletal tissues have become more sophisticated as well. Because of this, the first edition of *Skeletal Tissue Mechanics* became woefully dated, and its continued use as a text mandated that we update it and expand upon certain topics.

Besides updating and refreshing the content, and correcting some errors present in the first edition, the second edition of this book differs from the first in a number of important ways. The Biology of all skeletal tissues (bone, cartilage, ligament, and tendon) is now incorporated into a single chapter

and described using a hierarchical approach, focusing on the mechanics of the different biological levels of organization in each. The Mechanical aspects of each tissue are moved to subsequent chapters, with mechanics related to bone now at the end of the book. Having used the book ourselves, we have attempted to reorganize the chapters into a sequence that makes more sense from a didactic standpoint.

We have added new Historical Notes, set off from the main text in boxes, which we hope will increase student interest but also broaden their perspective. Additional boxes with Technical Notes will help students understand how to collect, analyze, and use information related to the textual material. We have updated the supplementary readings at the end of each chapter. We revised the exercise set we developed for students that we hope will help them to consolidate and use the information found in the chapters. Few of these exercises are simply practice at solving problems demonstrated in the book; rather, most are designed to stimulate the student to synthesize concepts based on material discussed in the text and to develop analytical problem-solving skills that will serve them in the future. Finally, figures have been updated and redrawn in some cases to improve content or make them more easily understandable, and color figures are included where we felt it would enhance learning.

We appreciate both our students and our colleagues who have contributed directly and indirectly to this book. Feedback and critical thought from students were particularly helpful in identifying gaps, or confusing presentations. Discussions with colleagues are always enlightening, and informed some of the changes. Beyond that, the coauthors themselves engaged in many entertaining discussions, some philosophical and others scientific, which expanded the boundaries of our own thinking. Finally, we would like to thank Drew Brown for his help in drawing (and redrawing, often several times) some of the new figures.

Olga, WA
Indianapolis, IN
University Park, PA
Sacramento, CA

R. Bruce Martin
David B. Burr
Neil A. Sharkey
David P. Fyhrie

Contents

Preface	vii
Chapter 1. Functional Musculoskeletal Anatomy	1
1.1 Introduction	1
1.2 An Illustrative Experimental Model	2
1.3 Static Analysis of Forces in Joints	4
1.3.1 Forces in the Elbow Joint	4
1.3.2 Forces in the Hip Joint	9
1.3.3 Clinical Significance of High Joint Forces	13
1.4 Hip Forces in Human Ancestors	13
1.5 Dynamic Joint Loading	17
1.6 Indeterminate Joint Problems	19
1.7 Equine Fetlock Forces	23
1.8 Summary and Further Reading	26
1.9 Exercises	27
References	32
Chapter 2. Skeletal Biology	35
2.1 Introduction to Bone	35
2.2 The Shapes of Bones	40
2.3 Types of Bone and Tissue	41
2.3.1 Cancellous Versus Compact Bones	43
2.3.2 Lamellar Versus Woven Bone	44
2.3.3 Primary and Secondary Bone	47
2.4 The Collagen-Mineral Composition of Bone	51
2.4.1 Quantitative Analyses of Bone Composition	56
2.4.2 Three Dimensional Imaging of Bone Using High Resolution Computed Tomography (microCT)	57
2.5 Bone Cells	60
2.6 Introduction to Cartilage	65
2.6.1 Mechanical Significance of Cartilage	70
2.6.2 Organization of Articular Cartilage	71

2.6.3 The Role of Cartilage in Growth	75
2.7 Longitudinal Growth of Bones	76
2.7.1 Development of Metaphyseal Trabeculae	79
2.7.2 Growth of the Physis	79
2.7.3 Closure of the Physes	80
2.8 Introduction to Ligament and Tendon	81
2.8.1 Structure and Composition	83
2.8.2 Injury and Repair	86
2.9 Summary for Further Reading	87
2.10 Exercises	87
References	89
 Chapter 3. Growth, Modeling and Remodeling of Bone	 95
3.1 Introduction	95
3.1.1 Growth	96
3.1.2 Modeling Versus Remodeling	96
3.1.3 The A-R-F Sequence and BMUs	99
3.1.4 The Six Phases of an Osteon's Lifetime	100
3.1.5 Remodeling Cycle Duration	102
3.1.6 Bone Structural Units Are Produced by Remodeling	104
3.1.7 Remodeling on the Different Bone Envelopes	105
3.2 Fracture Healing	108
3.2.1 Basic Concepts	109
3.2.2 Important Tissues in Fracture Healing	109
3.2.3 Three Biological Phases	111
3.2.4 Four Biomechanical Stages	114
3.2.5 Stability of Fixation	116
3.2.6 Primary Union	118
3.2.7 Posttraumatic Osteoporosis	118
3.2.8 Electrical Phenomena and Fracture Healing	119
3.3 Analysis of Bone Remodeling	119
3.4 Mathematical Notation for Bone Remodeling	121
3.5 Histomorphometric Measurement of Remodeling	122
3.5.1 Assumptions in Histomorphometry	123
3.5.2 Histomorphometric Measurements for Cortical Bone	125
3.5.3 Calculating BMU-Level Results	126
3.5.4 Histomorphometric Measurements for Cancellous Bone	131
3.5.5 True vs. Histologic BMU Activation Frequency in Cortical Bone	134
3.6 Remodeling Details	138
3.6.1 Activation Is Prime	138
3.6.2 Resorption: Out with the Old	139

3.6.3 Refilling: In with the New	142
3.6.4 Computation of FP	148
3.6.5 The Equal Lamellar Work Areas Principle	148
3.6.6 Why Does the Apposition Rate Slow Down?	150
3.7 Long-Term Effects of Osteonal Remodeling	154
3.7.1 Implications for the Aging Skeleton	154
3.7.2 Osteonal Overlapping	155
3.7.3 The Random Remodeling Assumption	155
3.7.4 Mathematical Theories	156
3.7.5 Another Theoretical Approach to Bone Remodeling	162
3.8 Summary and Further Reading	164
3.9 Exercises	165
References	168

Chapter 4. Mechanical Properties of Ligament and Tendon 175

4.1 Quasi-static Tensile Properties	175
4.2 Viscoelastic Properties	178
4.3 Mathematical Modeling	181
4.4 Measuring Structural and Material Properties	184
4.4.1 Storage and Testing Environments	184
4.4.2 Measuring Cross-Sectional Area	186
4.4.3 Gripping Slippery Viscoelastic Tissues	187
4.4.4 Determining Resting Length	189
4.4.5 Measuring Specimen Elongation	190
4.5 Functions at Junctions	193
4.5.1 The Myotendinous Junction	193
4.5.2 Ligament and Tendon Insertions to Bone	194
4.6 Functional Adaptation and Specialization	198
4.6.1 Determinants of Tendon Architecture	200
4.6.2 Sites of Tendon Compression	201
4.6.3 Flexors and Extensors	204
4.7 Age and Mechanical Behavior	206
4.8 Exercise and Immobilization	209
4.9 Pathology and Healing	210
4.9.1 Ligament	210
4.9.2 Tendon	211
4.9.3 Surgical Repair	213
4.9.4 Anterior Cruciate Ligament	213
4.9.5 Supraspinatus Tendon (Rotator Cuff)	215
4.9.6 Flexor Tendons of the Hand	216
4.10 Summary and Further Reading	217
4.11 Exercises	218
References	220

Chapter 5. Synovial Joint Mechanics	227
5.1 Introduction	227
5.1.1 Functions of a Synovial Joint	228
5.1.2 Joint Diseases	229
5.2 Mechanical Properties of Cartilage	230
5.2.1 Initial Points	230
5.2.2 Structure of Articular Cartilage	233
5.2.3 Permeability	234
5.2.4 Indentation Testing	235
5.2.5 Tensile Tests	237
5.2.6 Poroelastic Theory of Articular Cartilage	239
5.2.7 Other Models for Cartilage Mechanics	244
5.2.8 Modeling Tissue Charge	244
5.2.9 Inhomogeneity of Modulus	245
5.2.10 Effect of Strain Rate on Tensile Tests	246
5.2.11 Effect of Strain Rate on Compression Tests	247
5.3 Lubrication of Joints	248
5.3.1 Friction	248
5.3.2 Wear	250
5.3.3 Types of Lubrication	250
5.3.4 Synovial Joint Lubrication	260
5.3.5 Diversity of Joint Architecture	263
5.4 Summary and Further Reading	265
5.5 Exercises	266
References	270
 Chapter 6. Mechanical Adaptability of the Skeleton	 275
6.1 Introduction	275
6.2 The Historical Context	276
6.3 Mechanical Adaptability and Evolutionary Adaptability	280
6.3.1 Somatic Versus Evolutionary Adaptation	280
6.3.2 Somatic Versus Evolutionary Effects	282
6.4 What Is the Design Goal: What Does Bone Adapt to?	284
6.4.1 Optimization, Mechanical Adaptability, and Structure-Sensitive or Insensitive Properties	285
6.5 How Does Bone Adapt: What Are the Signals and How Are They Transduced?	286
6.5.1 Measurement of Strain in Living Animals	287
6.5.2 Principal Stresses, Strains and Trabecular Orientations	288
6.5.3 Compressive Strains Are Larger than Tensile Strains	290
6.5.4 Canine Disuse Experiments	295

6.5.5	The Mechanostat and the Equilibrium Strain Range	295
6.5.6	Mechanical Compensation and the Roles of Bone Modeling and Remodeling	297
6.5.7	Optimization and the Mathematical Basis of Wolff's Law	302
6.5.8	Some Final Thoughts on Optimization, Adaptation and the Scientific Method	303
6.5.9	How Frequently Should Dynamic Loads Be Applied: The Role of Strain Rate	306
6.5.10	Effects of High Frequency Loading: Vibration	308
6.5.11	When Theory Meets Reality	309
6.5.12	Is Strain a Direct Stimulus to Bone Cells? Do Strain Gradients and Distributions Regulate Regional Bone Cell Activity?	310
6.5.13	Fluid Flow	313
6.5.14	Summary	315
6.6	Some Additional Theories	316
6.6.1	Pauwels' Stress Magnitude Theory: Control of Modeling Drifts	317
6.6.2	Adaptive Elasticity Theory: Control of Density or Modeling Drifts	319
6.6.3	What Controls Osteonal Tunneling Directions?	320
6.6.4	Adaptive Finite Element Models: Control of Density	323
6.6.5	Self-Trabeculating Models: Control of Density and Trabecular Alignment	327
6.6.6	Related Engineering Models	331
6.6.7	Synthesis	332
6.6.8	Experimental Consistency	332
6.6.9	Relationship of Mechanically Adaptive Responses to Other Control Factors	334
6.6.10	Mechanical Adaptability and Damage Repair	335
6.7	Mechanical Adaptability in Cartilage	336
6.7.1	The Carter–Wong Chondral Calcification Theory	336
6.7.2	Frost's Chondral Modeling Theory	338
6.8	Summary and Further Reading	339
6.9	Exercises	340
	References	343

Chapter 7. Mechanical Properties of Bone	355
7.1 Introduction	355
7.2 Fundamentals of Solid Mechanics	356
7.2.1 Strength and Stiffness of a Structure	356
7.2.2 Stress and Strain	358
7.2.3 Principal Directions	360
7.2.4 Strength and Modulus of a Material	360
7.2.5 Generalized Hooke's Law: Anisotropy	362
7.3 Determinants of the Strength of a Whole Bone	365
7.3.1 Mechanics	365
7.3.2 Examples	367
7.3.3 Mechanical Failure of Whole Bones	371
7.4 Material Properties of Cortical Bone	374
7.4.1 Properties of Individual Secondary Osteons	374
7.4.2 Combined Loading Effects	377
7.5 Summary	378
7.5.1 Effects of Osteons on Cortical Bone Mechanical Properties	380
7.5.2 Anisotropy of Cortical Bone Mechanical Properties	385
7.5.3 Anisotropy Effects Determined by Load-Deformation Testing	385
7.5.4 Anisotropy Effects Determined by Ultrasonic Measurements	387
7.5.5 Determinants of Osteonal Bone Mechanical Properties	389
7.6 Material Properties of Cancellous Bone	399
7.6.1 Stress-Strain Curves for Cancellous Bone	399
7.6.2 Strength in Tension Versus Compression	400
7.6.3 Modeling Cancellous Bone as a Cellular System of Plates or Struts	407
7.6.4 Invariance of Yield Strain Hypothesis	408
7.7 Predicting Material Properties: Bone as a Composite Material	409
7.8 Summary and Additional Reading	410
7.9 Exercises	411
References	414
 Chapter 8. Fatigue and Fracture Resistance of Bone	 423
8.1 Introduction	423
8.2 Basic Fracture Mechanics	425
8.2.1 Linear Elastic Fracture Mechanics	425
8.2.2 Kinds of Cracks	427
8.2.3 Modes of Cracking	428
8.2.4 Maximum Tolerable Flaw Size	431
8.2.5 It Takes Energy to Propagate a Crack	431

8.2.6 Beyond the Linear Theory: Real Cracks Have Ears	432
8.2.7 Beyond the Linear Theory: Real Cracks Are Complicated	434
8.2.8 Crack Growth and Fatigue	437
8.3 Fatigue Behavior of Bone	438
8.3.1 The S-N Curve	438
8.3.2 Fatigue Damage in Bone	439
8.4 Creep Behavior of Bone and Its Relationship to Fatigue	442
8.5 Fatigue Behavior of Fiber-Reinforced Composite Laminates	445
8.5.1 The Birth and Growth of Cracks	445
8.5.2 Material Strength and Fiber Diameter	446
8.5.3 The Road to Failure	447
8.6 Osteonal Bone as a Fibrous Lamellar Composite Material	448
8.6.1 Osteonal Bone's Road to Failure	448
8.6.2 Bone Toughness	451
8.6.3 The Conflict between Strength and Toughness in Bone	452
8.6.4 Controlled Crack Propagation Studies	452
8.6.5 Mathematical Analysis of Osteonal Pullout	453
8.6.6 Comparing Theory with Data	454
8.6.7 Crack Initiation	456
8.6.8 Crack Stopping	457
8.6.9 Strain Rate	457
8.6.10 Effect of Remodeling	458
8.7 Modeling Fatigue Damage Effects in Osteonal Bone	459
8.7.1 Why Would Cracks Be Stopped in Tension But Not in Compression?	461
8.7.2 If Cracks Are Self-Limiting in Tensile Fatigue But Not in Compressive Fatigue, Why Is Fatigue Life Longer in Compression Than in Tension?	462
8.8 The Role of Fatigue in Activating Bone Remodeling	463
8.8.1 Random or Directed Repair	463
8.8.2 The Physiological Basis for Targeted Remodeling	464
8.8.3 Microdamage and Bone Fragility in the Elderly	466
8.9 Modeling Stress Fractures	468
8.10 Summary and Additional Reading	474
8.11 Exercises	475
References	477
Index	483

1

Functional Musculoskeletal Anatomy

...mechanical science is of all the noblest and most useful, seeing that by means of this all animate bodies which have movement perform all their actions...

Leonardo Da Vinci (1452–1519)

Abstract This chapter describes how engineering principles can be used to estimate joint forces. Principles of static and dynamic analysis are reviewed, with examples of static analysis applied to the hip and elbow joints and to the analysis of joint forces in human ancestors. Applications to indeterminate problems of joint mechanics are presented and utilized to analyze equine fetlock joints.

Keywords Joints • Forces • Statics • Dynamics • Kinematic • Kinetics

1.1 Introduction

The skeleton is first and foremost a mechanical organ. Its primary functions are to transmit forces from one part of the body to another and to protect certain other organs (e.g., the brain) from mechanical forces that could damage them. Therefore, the principal biologic role of skeletal tissues is to bear loads with limited amounts of deformation. To appreciate the mechanical attributes that these tissues must have to perform this role, it is necessary to learn something about the forces which whole bones normally carry. In most cases, these forces result from loads being passed from the part of the body in contact with a more or less rigid environmental surface (e.g., the heel on the ground when walking) through one or more bones to the applied or supported load (e.g., the torso). In addition to the forces transmitted in bone-to-bone contact, muscle and ligament forces act on the bones, and these forces (especially the muscle forces) are large and important.

Most muscle, ligament, and bone-to-bone forces act in or near the body's major diarthrodial joints. The purpose of this chapter is to explain how

conventional engineering analysis may be used to estimate joint forces, and to provide some practical experience in solving such problems.

1.2 An Illustrative Experimental Model

A major advantage of conventional static and dynamic analyses is that the forces and moments imposed on bones can be estimated with reasonable precision from non-invasive external measurements. Actual physical measurement of internal loading is more challenging, for both experimental and ethical reasons. Although physical measurements of skeletal loading have been made in live subjects (we will explore an example later in this chapter) these types of studies are relatively rare and always limited in scope. One way around the ethical dilemma posed by invasive measurement is to apply life-like loading environments to cadaver specimens donated for scientific research, thereby enabling direct, detailed measurement of internal kinetics and its physical consequences within the tissues themselves. Orthopaedic research has a long and rich history using this approach and much has been learned about normal skeletal loading and the biomechanical consequences of pathology, surgical reconstruction, and joint replacement. Before tackling classic static analyses, we will examine some data sets taken from a series of cadaver experiments focused on the lower extremity, an exercise that will lend further appreciation of the large forces routinely imposed on our skeletons.

The data presented in Fig. 1.1 were gathered using an apparatus that reproduces the kinematics and kinetics of normal gait in cadaver lower extremities (Sharkey and Hamel 1998). One set of computer controlled actuations reproduces the motions of the proximal tibia as recorded in a standard gait lab while a second set simultaneously reproduces the actions of six major extrinsic muscle groups of the foot and ankle. These muscles reside in the shank and include the triceps surae (gastrocnemius and soleus), tibialis posterior, tibialis anterior, peronei (peroneus longus and brevis) flexor hallucis longus and flexor digitorum longus. The vertical axis of the plot is force, given in units of body weight (BW) for easy comparison, and the horizontal axis represents increments of time when the foot is in contact with the ground during normal walking. The vertical force imposed by the ground is given by the heavy dark double-peaked plot and reflects the classic force-time curve for gait whereby the first peak is a component of the braking force occurring shortly after heel-strike and the second peak a component of the push-off force to propel the body forward. The shape and magnitude of this plot confirms life-like loading equivalent to that routinely measured in standard gait labs.

The remaining plots in the graph represent the internal tissue forces present during normal walking. Peak tension in the Achilles tendon (a function of triceps surae contraction needed to propel the body forward)

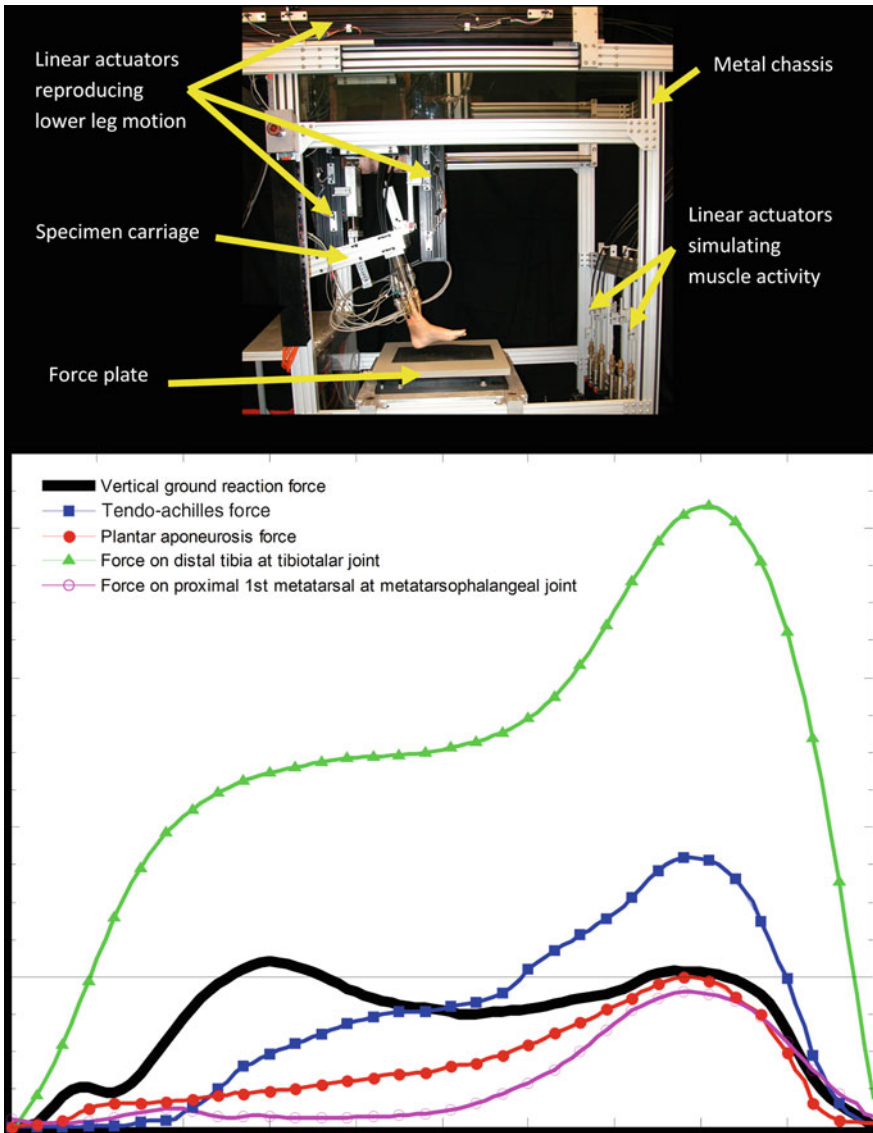


FIGURE 1.1. *Top*: Image of an experimental apparatus enabling detailed measurements of biomechanical phenomena in skeletal tissues under normal or pathologic conditions. *Bottom*: Plots of the forces carried in the select skeletal tissues during normal walking.

is approximately $1.8 \times \text{BW}$ while the peak tensile force carried in the plantar fascia, a tough band of ligamentous tissue that spans the sole of the foot is roughly equivalent to body weight (Erdemir et al. 2004). Perhaps most striking are the plots of joint reaction force measured in the tibiotalar

(ankle) and first metatarsophalangeal (MTP; at the base of the great toe) joints. The routine load transmitted through the ankle is over $4.0 \times \text{BW}$ (Sharkey and Hamel 1998) while that carried through a very small digital joint in the first ray of the foot is almost equal to BW (Kirane et al. 2008). This is a remarkable feat (no pun intended). With each step we take day in and day out, both our hard and soft skeletal tissues are exposed to tensile and compressive forces considerably greater than body weight and this is only for walking; the forces are much larger during more strenuous activities like running and jumping. The mechanical explanation for these high forces is examined in detail within the next sections of this chapter, while future chapters examine the mechanisms enabling such an incredibly strong and durable system.

1.3 Static Analysis of Forces in Joints

1.3.1 *Forces in the Elbow Joint*

To extend this subject, let us consider the human elbow joint. We choose this joint because it works more or less as a simple hinge, and because it is familiar to us, conforming to the image that we acquired in grade school of the skeleton as a system of levers. Although we use this simplicity in solving for the forces, we should spend a few moments examining the anatomy to appreciate the simplifying assumptions that we are making.

A lateral view of the bones of the elbow is shown in Fig. 1.2a. The olecranon of the ulna wraps around the distal end of the humerus to form the major part of the elbow's hinge joint. The proximal joint surface of the radius articulates with the distal joint surface of the humerus and also with a cartilage-covered notch (the radial notch) on the lateral aspect of the ulna; this means there are three articulations in the elbow joint. These articulations allow for flexion and extension of this joint as well as pronation and supination¹ of the forearm. The radius and ulna are bound together by the interosseous membrane along the central part of their length and by several ligaments at their proximal and distal ends (not shown). We assume that these structures cause the two forearm bones to act as one structural unit. We also assume that the wrist joint is stabilized by its musculature so that the hand and forearm flex about the elbow as a simple hinge joint.

When the elbow is flexed with the palm upward, two muscles are primarily responsible: the biceps and the brachialis (Fig. 1.2b, c). The biceps has two heads (i.e., points of origin). The short head originates on the coracoid process of the scapula, while the long head runs through the shoulder joint

¹ For the definition of these and other anatomical terms, see a human gross anatomy text, e.g., Grant's (Agur 1991) or Gray's (Williams 1995).

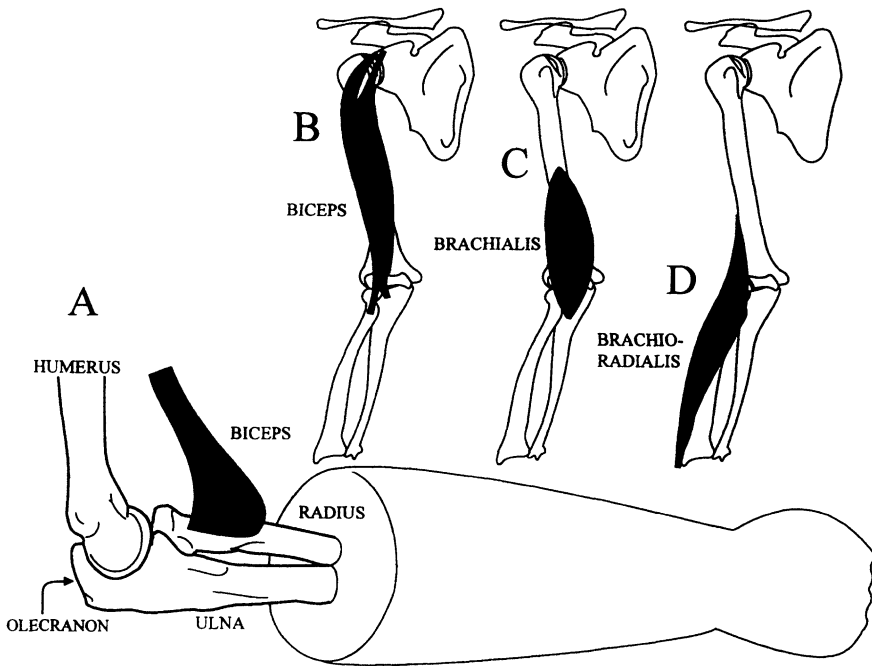


FIGURE 1.2. (a–d) Anatomical drawings of the elbow joint. (a) Overall view of the flexed joint and the biceps inserting primarily on the radius. (b, c, d) Approximate locations of the biceps, brachialis, and brachioradialis muscles.

to attach to the superior lip of the glenoid fossa (the scapular part of the shoulder joint). The biceps divides distally into a tendon that inserts on the proximal radius and an aponeurosis (tough band of connective tissue) which blends with other muscles in the proximal forearm (Fig. 1.2b). The brachialis originates along the anterior surface of the central humerus and inserts on the proximal ulna (Fig. 1.2c). When the forearm is flexed rapidly, or a large force must be applied during flexion, the brachioradialis also acts between the distal humerus and the distal radius (Fig. 1.2d). Other muscles also cross the elbow and participate in flexion. However, to simplify matters, and because the present problem is for a static situation, we assume that the biceps and brachialis muscles act as one in holding the elbow in a flexed position against the action of a weight held in the hand. We want to find the force required in the “biceps-brachialis muscle” to support the flexed forearm and the total force exerted on the distal end of the humerus by the radius and ulna. These forces are expressed as multiples of the weight in the hand.

To solve this problem and others like it, three steps are necessary:

1. Draw a free-body diagram.
2. Write the equations of static equilibrium.
3. Solve these equations simultaneously to obtain the unknown forces.

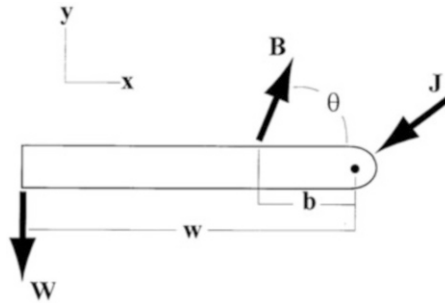


FIGURE 1.3. Free-body diagram for the elbow force calculation. **J** is the joint reaction force, **W** is the weight in the hand, and **B** is the biceps force.

We simplify the analysis by assuming that it is two dimensional and that the forces all act in a sagittal plane containing the humerus and the forearm. The free-body diagram is constructed by isolating the structure being analyzed such that the internal forces to be determined are exposed and replacing all the forces acting on the structure by vectors. The free-body diagram for the elbow problem is shown in Fig. 1.3. The forearm is made a free body by “amputating” through the elbow joint, whose internal forces we wish to reveal. The forearm lies in the x - y plane. The weight in the hand, the muscle force, and the joint reaction force are represented by the vectors **W**, **B**, and **J**, respectively.² The joint force is assumed to pass through a fixed center of rotation for the joint, shown by a dot. The insertion point of the muscle is b meters (m) along the forearm from the joint center, and the center of gravity of the weight in the hand is w meters from this point. The equations for static equilibrium in two dimensions are

$$\begin{aligned}\sum M &= 0 \\ \sum F_X &= 0 \\ \sum F_Y &= 0\end{aligned}\tag{1.1}$$

where M stands for moments about an arbitrary point, and F_X and F_Y are force components in the x - and y -directions. Because, in two dimensions, there are three such equations, one may solve for three unknown force *components*. Alternatively, as usually happens, one may solve for one force *vector* and the *magnitude* of another force of known direction.

For the elbow problem, taking moments about the joint center, the equilibrium equations are

²In this chapter, vector quantities are represented in bold type and scalars (such as vector magnitudes) in ordinary type.

$$\sum M = wW - bB \sin \theta = 0 \quad (1.2)$$

$$\sum F_X = B \cos \theta - J_X = 0 \quad (1.3)$$

$$\sum F_Y = B \sin \theta - W - J_Y = 0 \quad (1.4)$$

(Here, any moment of the x -component of $\mathbf{B}$ about the joint center is ignored.) Solving these equations yields

$$B = wW / (b \sin \theta) \quad (1.5)$$

$$J_X = B \cos \theta \quad (1.6)$$

$$J_Y = B \sin \theta - W \quad (1.7)$$

If $\theta = 75^\circ$, $w = 0.35$ m, and $b = 0.04$ m, then $B = 9.1$ W, $J_X = 2.3$ W, and $J_Y = 7.8$ W. The magnitude of the joint reaction force is $J = (J_X^2 + J_Y^2)^{1/2} = 8.1$ W, and its orientation is $\arctan(J_Y/J_X) = 74^\circ$ with respect to the x -axis. Thus, the muscle and joint reaction forces are eight to nine times greater than the weight held in the hand. This result is typical of virtually all the joints in the body in that the skeleton works at a mechanical disadvantage in terms of force, and as a consequence the forces acting on bones are high relative to the forces applied by (or to) the environment.

There is, of course, a good reason for this. Muscles can only contract about 30 % of their length. In the case of the biceps, for example, the overall length is about 25 cm, so the maximum contraction is 7–8 cm. The lever action of the forearm magnifies this distance by the ratio w/b in Fig. 1.3, enabling the hand to move much further (and also much faster). Of equal importance, the muscles insert proximally on the radius and ulna and do not create a “web” of flesh across the front of the elbow. Therefore, magnification of motion enables larger movements to be accomplished with a compact structure. The price that is paid for this compactness and magnification of motion is high forces in the muscles, across the joint surfaces, and within the bones.

Box 1.1 Giovanni Borelli on the Movement of Animals

Giovanni Alfonso Borelli (1608–1679) was the greatest of early biomechanicians. He held the chair in mathematics at Pisa in Italy, where he was a close friend of Malpighi, who was the professor of theoretical medicine. Together, they did much to persuade seventeenth-century physicians of the importance of physics in understanding medicine and physiology. Borelli’s treatise, *De Motu Animalium* (*On the Movement of Animals*) (Borelli 1989), has been translated into English and is a marvelous testament to his genius and ability to explain musculoskeletal mechanics clearly and simply. In the preface to his translation, Maquet points out that later biomechanicians unwittingly duplicated several of Borelli’s important discoveries. While some of his results

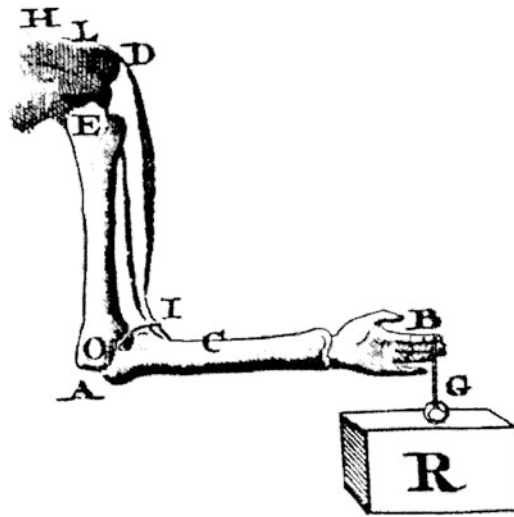


FIGURE 1.4. Diagram of elbow force problem from Borelli's *On the Movement of Animals* (Borelli 1989). With kind permission of Springer Science + Business Media.

contain errors, it must be remembered that Newton's laws were not published until 1687. Thus, although Borelli understood very well the principle of balancing moments about a fulcrum, he could not have been expected to fully understand static equilibrium as we do, based on Newton's laws. One of Borelli's greatest achievements was the discovery that animals' joints work at a mechanical disadvantage in terms of force. He discusses this in the following excerpt (Fig. 1.4).

Proposition VIII

"It is commonly thought that Nature raises considerable weights by using the machines of the muscles with a weak moving force.

The magnitude of the vital force of the muscles . . . sustains, raises, and moves not only an arm or a leg, but the whole animal machine, enabling it even to dance. Besides the mass of the animal, heavy enough by itself, this force carries, pulls and pushes considerable weights.

Aristotle . . . did not recognize the muscles but imagined spirits which pull and push the limbs. [He] remarked how difficult it would be for the huge mass of an elephant to be moved . . . by tenuous spirit or wind. He met the difficulty by saying that Nature moves the joints and limbs of the animal by using very small force . . . and said that the operation is carried out by way of a lever. Therefore, it is not surprising that huge weights can be moved . . . by a small force. Lucretius used the same example . . . Galen also says that a tendon is like a lever. He thinks that, consequently, a small force of the animal faculty can pull and move heavy weights.

This general opinion seems to be so likely that, to my knowledge, not surprisingly, it has been questioned by nobody. Who indeed would be stupid enough to look for a machine [in the body] to move a very light weight with a great force, i.e., not to save forces but rather to spend forces? And if this is rightly considered as stupid, how is it possible that wise Nature, everywhere looking for economy,

simplicity and facility, builds with great efficiency in animals machines to move, not heavy weights with a small force, but on the contrary, light weights with almost boundless force? ...I shall demonstrate that multiple and different machines actually are used in the motions of animals but that light weights are carried by large and strong force rather than heavy weights being supported by small force.”

1.3.2 Forces in the Hip Joint

The hip joint is the articulation between the femur and the acetabulum of the pelvis. The hip is one of the most important joints in the body from a medical perspective. Especially in aging individuals, overall health is promoted by the exercise that goes with walking. The ability to walk depends on having a healthy, painfree hip. Two kinds of diseases, both very common among the elderly, preferentially affect the hip. *Osteoarthritis* is the primary reason that about 200,000 total hip replacement procedures are performed each year in the United States, and *osteoporosis* in the femoral neck results in several hundred thousand hip fractures each year, virtually all of which require surgical treatment with a metal fixation device or hip replacement. To better understand the mechanics of the hip and the demands on the implants used to correct its problems, it is important to know the forces across the hip during walking and their determining factors.

When one is walking, the lower extremities and other parts of the body are moving (i.e., accelerating), so the conditions of static equilibrium are not satisfied. However, most people do not usually walk very vigorously, so the accelerations are small relative to the forces produced by muscle pulls and gravity. Therefore, the problem is usually solved as though the person were simply standing on one leg, assuming that this approximates the conditions during the “single leg stance phase” of gait, that is, when all the weight is being carried by one leg. In addition to this assumption, we assume that the problem is two dimensional, in the frontal plane, and that only one muscle is acting. If you try to stand on one leg, the force pulling your center of gravity downward tends to rotate your torso toward the medial side of the leg you are standing on. The muscles that resist this movement are the same ones which you use to abduct your thigh when lying down (i.e., move your lower limb away from the body’s center line). Again, several muscles act to do this, but we can lump them all together and simply call them the “abductor muscles.”

Again, the first step in calculating the forces in the joint is to draw a free-body diagram, “amputating” through the joint in question to “reveal” the force vectors acting there. In this case we delete the lower extremity and study the remaining portion of the body. Figure 1.5 shows this situation. It is important to understand that the subject was standing on the leg that has been “amputated,” and that the other leg, of which only a portion is shown, was not touching the ground. The abductor muscle force is represented by the force vector P . The joint reaction force acting on the middle of the

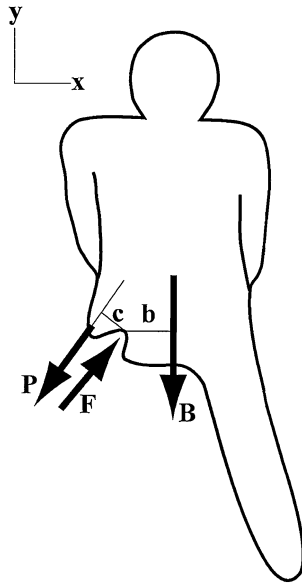


FIGURE 1.5. Free-body diagram for the calculation of the hip joint force while walking. The “amputated” leg is the one supporting the body. The leg in the free-body diagram is not in contact with the ground. **B** is the weight of the body (minus the amputated leg), **P** is the abductor muscle force, and **F** is the joint reaction force.

acetabulum in the pelvis is **F**. The weight of the body, represented by the vector **B**, is actually the entire body weight, W , minus the weight of the leg supporting the body. Because each lower extremity weighs about $1/6 W$, we let $B = 5/6 W$. This force acts downward slightly to the right of the centerline of the body (Fig. 1.5). Taking moments about the center of the acetabulum, we have

$$\sum M = cP - bB = 0 \quad (1.8)$$

$$P = (b/c)B = (b/c)(5/6)W \quad (1.9)$$

The lengths of the moment arms b and c have been estimated from antero-posterior pelvic radiographs. It was found that the b/c ratio ranges between 2 and 3.5. Following the lead of Frankel and Burstein (1973), we choose a conservative value of 2.4 and obtain the convenient result that $P = 2W$. That is, the force required in the abductor muscles to balance the body on the head of the weight-bearing femur during the stance phase of gait is twofold body weight.

To solve the rest of the problem, we write the equations for force equilibrium, assuming that the x -direction is horizontal and the y -direction is vertical:

$$\sum F_X = F_X - P_X = P_X - 2W \sin \theta = 0 \quad (1.10)$$

$$\sum F_Y = F_Y - P_Y - B = P_y - 2W \cos \theta - 5W/6 = 0 \quad (1.11)$$

where $\theta \approx 30^\circ$ is the angle that the abductor muscle line of action makes with the y-axis. Because $\sin 30^\circ = 0.5$, the components of the joint reaction force F are

$$F_X = W \text{ and } F_Y = (2 \cos 30^\circ + 5/6)W = 2.57W \quad (1.12)$$

The *horizontal* force of the femur on the pelvis is equal to body weight, and the vertical force is 2.5 times as much. The total joint reaction force is $F = (F_X^2 + F_Y^2)^{1/2} = 2.75W$. This force acts at an angle to the horizontal $\theta = \arctan(F_Y/F_X) = 68.7^\circ$.

We have seen that the magnitude of the forces in the hip joint depends critically on the ratio of the body weight moment arm to the abductor muscle moment arm. Thus, anything that increases the former or decreases the latter increases the abductor muscle force required for gait and consequently the force on the head of the femur as well. People with short femoral necks have higher hip forces, other things being equal. More significantly, people with a wide pelvis also have larger hip forces. This tendency means that women have larger hip forces than men because their pelvises must accommodate a birth canal (Burr et al. 1977). This fact may be one reason that women have more hip fractures and hip replacements because of arthritis than men do. It is also conceivable that this places women at a biomechanical disadvantage with respect to some athletic activities, although studies do not always show gender differences in the biomechanics of running, particularly endurance running (Atwater 1990).

Box 1.2 Measurement of Hip Joint Forces In Vivo

To verify the estimates of hip joint forces made using free-body calculations, several groups of investigators have implanted devices that allowed measurement of hip joint forces into hip replacement patients (Rydell 1965, 1966; English and Kilvington 1979; Davy et al. 1988; Kotzar et al. 1991; Bergmann et al. 1988, 1993, 2001; Taylor et al. 1997). When humans are used as experimental subjects, ethical concerns demand that the subject's health not be endangered in any way. This ethical mandate has extensive technical consequences for measurement of hip joint forces. To begin with, a healthy human subject cannot be used; experimentation must be restricted to a subject with a diseased or injured hip who will be undergoing replacement of the proximal femur with a prosthesis by way of treatment for this condition. This condition provides the opportunity for using an instrumented prosthesis (Fig. 1.6) that can measure joint forces instead of a standard prosthesis. However, the instrumented prosthesis must be as strong and as durable as a standard one, which imposes important constraints on the instrumentation design. Moreover, wires cannot be run from the prosthesis to the surface of the skin, where they could pose a risk of infection

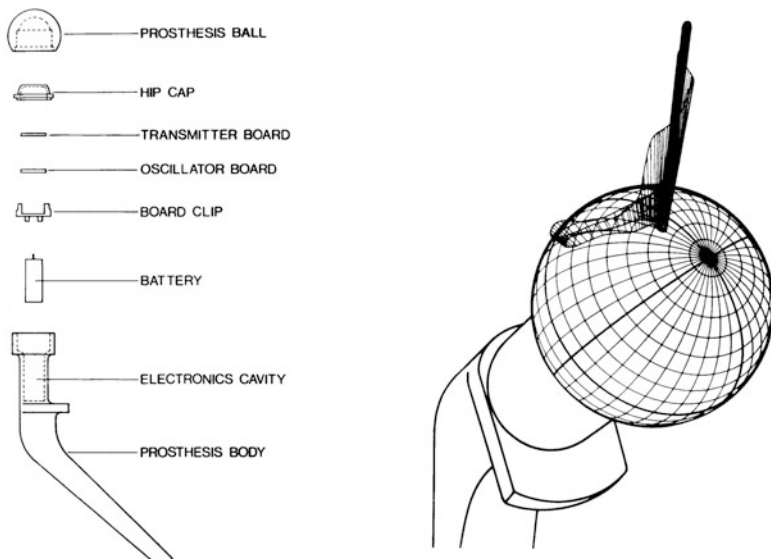


FIGURE 1.6. *Left:* Exploded view of telemetry electronics inside an instrumented hip prosthesis (Davy et al. 1988). *Right:* Three-dimensional plot of hip joint force vector during a gait cycle (with crutches). Orientations and lengths of lines emanating from the surface of the prosthesis head indicate force direction and magnitude, respectively (Davy et al. 1988).

or other problems. Consequently, the instrument must be fully contained within the prosthesis, with a working space of a few cubic centimeters, and it must be capable of transmitting data to a receiver outside the body while the patient walks.

The solution to this technical problem has been to use strain gauges (3–12 in number) to measure strains on interior surfaces of the prosthesis, and microelectronic circuitry to process the strain signal, which is broadcast as an FM signal to a receiver held against the skin. The device must be fully calibrated in the laboratory before insertion into the subject, so that the measured strains can be converted to force components on the ball or head of the prosthesis. The power source in the prosthesis used by (Davy et al. 1988) Davy et al. was a battery located in the stem of the prosthesis and activated by a magnetic switch turned on by a magnet held outside the patient's leg (Fig. 1.6). Because batteries carry a slight risk of releasing toxic components, and because they have a finite life, other investigators (Bergmann et al. 1993) used an inductive power source. A magnetic coil wrapped around the subject's leg induced current in a receiver coil within the prosthesis.

One might question whether a patient who has a degenerated joint, has recently had surgery on it, and is in a laboratory environment, is capable of walking “normally”—but these are the best data available. Findings from all studies are consistent with one another and with the theoretical estimates. During the single leg stance phase

of gait, hip joint forces between 2.5 and 3.3 times body weight were measured. The highest force recorded was in an individual who was attempting static single-leg balancing; during recovery from a momentary loss of balance, a force of 5.5 body weight was measured in the hip (Davy et al. 1988).

1.3.3 Clinical Significance of High Joint Forces

Because diarthrodial joints work at a mechanical force disadvantage so that limbs can move far and rapidly with short muscle contractions, high stresses are produced in the tissues of the bones and joints. Normally, these tissues carry their loads without causing pain, but various diseases and injuries can damage the tissues so that the deformations associated with loading are painful. To some degree, the pain is proportional to the amount of force carried by the tissues; in other words,

$$\text{Pain} = \text{Force} \times \text{Disease}$$

There are no nerves in cartilage, and the source of joint pain is poorly understood, but experience shows that reducing joint forces often alleviates pain. Often the physician is not able to do much about the disease, so the first consideration in controlling pain may be to reduce the forces in the joint. For example, the patient can lose weight, or walk with a cane, and thereby reduce the forces transmitted across the joint.

1.4 Hip Forces in Human Ancestors

Physical anthropologists have analyzed hip joint forces in skeletons of various hominids. Of particular interest is *Australopithecus* because of collateral evidence (footprints) that these individuals may have walked very much like modern humans. Figure 1.7 is a depiction of the famous *Australopithecus* skeleton known as Lucy. The differences between the anatomy of the femur and pelvis of *Australopithecus* and *Homo sapiens* include factors that seem to be very pertinent to hip joint force. For example, the neck of the *Australopithecus* femur; shaded in Fig. 1.8, overlaying the outline of a modern femur; was proportionately longer than ours, but the bone was smaller overall. The pelvis was also smaller overall, but the ilium (shown shaded in Fig. 1.9, on one side of a modern human pelvis of similar size) was more outwardly flared, moving the line of action of the abductor muscles away from the hip joint. These factors affect the moment arms of both the abductors and the body weight vector. The analysis of Lovejoy and co-workers (Lovejoy et al. 1973) indicates that the hip joint force in *Australopithecus* was about 2.5-fold body weight, a value somewhat less than that of modern humans.

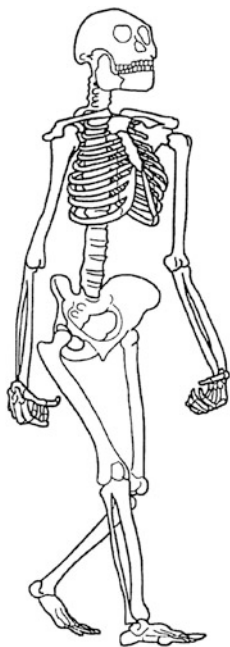


FIGURE 1.7. Artist's conception of the skeleton of *Australopithecus afarensis* ("Lucy") walking (McHenry 1994).



FIGURE 1.8. Lucy's proximal femur (shaded) compared to that of *Homo sapiens* (Lovejoy et al. 1973).

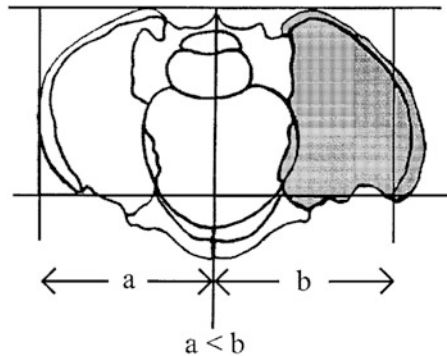


FIGURE 1.9. Lucy's ilium (shaded, on right) compared to that of *Homo sapiens* (on left). The left and right vertical lines would contain the modern pelvis; notice that Lucy's ilium projects well beyond the line on the right (Lovejoy et al. 1973).

However, when they performed a similar analysis on the skeletons of Native Americans excavated at archeological sites (and thus more comparable in their condition to fragmented *Australopithecus* skeletons), they obtained a value of 2.5 for these as well. Therefore, Lovejoy et al. concluded that *Australopithecus* and modern humans experienced similar hip forces (Lovejoy et al. 1973). They then considered the *pressure* on the head of the femur in these two species. Considering the degree to which the head of the *Australopithecus* femur was smaller than ours, and the estimated difference in body weights, it was concluded that the pressure on the *Australopithecus* cartilage was about half that on our femoral heads. However, these estimates are quite tenuous because of the scarcity and fragmented condition of the *Australopithecus* skeletons. Later in their paper, Lovejoy et al. conceded that other observations suggest that the *Australopithecus* hip force may have been substantially less than that of *Homo sapiens*. Two of the exercises at the end of this chapter allow you to explore this application of skeletal biomechanics.

Box 1.3 Reproducing Joint Forces in the Laboratory

As introduced previously, cadaver specimens are frequently used to model normal, pathologic, or reconstructed joints. For these laboratory simulations, researchers often rely on static calculations to approximate the loading conditions experienced by the joint in life. For example, a typical experiment examining the stability of a new hip prosthesis might use cadaveric hip joints implanted with the new component and then mounted into a testing machine for loading. But in what direction should the load be applied? What should the load magnitude be? Static analysis can be used to answer these questions.

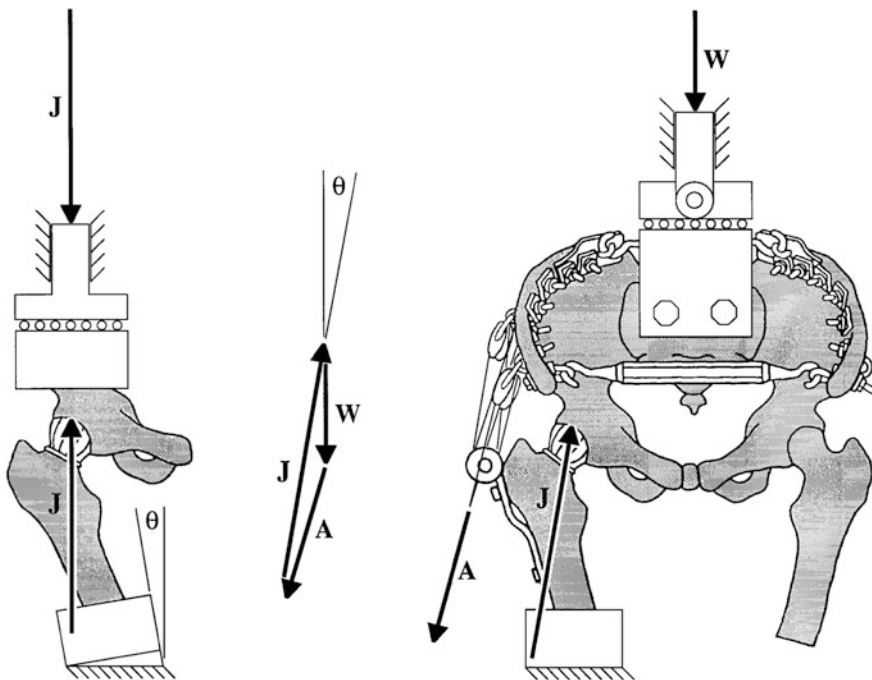


FIGURE 1.10. The *left* and *right* diagrams show two different experimental setups for simulating hip joint loading using cadaver bones. The vector diagram in the *middle* is for the case at *right* (Bay et al. 1997).

Once the investigator has determined the direction and magnitude of the joint reaction force for a given condition, it is simply a matter of aligning the specimen such that the vector of the joint force coincides with the axis of the load delivery system (Fig. 1.10, left). Many important biomechanical studies have been executed in just this fashion. However, like static analysis itself, this type of simulation incorporates several assumptions and the results should be interpreted with caution.

Let us revisit the hip example. A more physiologically correct model might use the whole pelvis and incorporate abductor muscle forces to reproduce the *in vivo* loading environment (Fig. 1.10, middle and right). In this case the proper joint reaction vector is produced by simulating abductor contraction using a cable system attached at the muscle origin and insertion sites. Shortening the cable produces a moment about the hip that is opposed at the sacrum. In theory, when the vertically directed force at the sacrum achieves upper-body weight, the joint reaction force at the hip will be the same as for the extracted hip joint model and the same as that produced during single-leg stance in life.

In a comparison of these two laboratory methods (Bay et al. 1997), whole intact pelvises were first loaded under simulated muscle action and the contact pressure distribution on the articular surfaces was recorded. The hips were then extracted from the pelvis and reloaded. The direction and magnitude of the joint reaction force was identical for both cases, but the distribution of contact pressure within the

joint was quite different. Why? Remember, static analysis assumes rigid bodies but musculoskeletal tissues do not always behave as rigid bodies. In the intact model the bones of the pelvis were free to flex and deform under load, while in the simpler joint model the pelvis became more rigid because of its altered geometry and the constraints placed upon it.

1.5 Dynamic Joint Loading

The overriding take-home message in this chapter is that the forces carried in bones and joints are magnified by the small lever arms of our muscles, producing internal joint forces many times the external forces imposed by the environment. It may seem odd to engineer a system to increase loads on joints, but from an evolutionary standpoint, our limbs and joints evolved for range of motion and speed of movement. The natural trade-off for this is short lever arms and increased forces on the joints (which was not such a problem, until we began to live so long, causing small amounts of damage that accumulate over time to progress to significant age-related joint deterioration, or osteoarthritis, in some people). So far, and because it is difficult to do otherwise, all of the examples used in this chapter have been for static loading, e.g., holding a ball in your hand or standing on one leg. Under normal conditions, however, we don't hold still for very long. (Think of a 3 year old child!) The *dynamics* of the body are quite different from a static analysis and are, generally, beyond the scope of this book. But to introduce the topic we will solve a simple dynamic problem to demonstrate how static force calculations *underestimate*, often by many times, the joint loads present during routine activities of daily living.

The model problem is to catch a dropped ball in an outstretched hand bent 90° at the elbow as in Fig. 1.11 and Borelli's drawing Fig. 1.4. The model includes only the forearm and assumes that the arm (i.e., the upper arm) remains static during the impact. To further simplify, there are several other assumptions that we can make:

- 1) The damping in the system during the initial impact is negligible, so the dynamic behavior from impact to maximum deflection is entirely elastic.
- 2) The hand catches the ball "perfectly" so that the momentum of the ball is completely transferred into momentum of the hand and ball together.
- 3) The initial force in the biceps required to hold the hand horizontal is ignored.
- 4) The mass of the forearm and hand (about 3 % of body weight) are concentrated at the hand, with the remainder of the forearm acting as a massless lever.
- 5) And (so that a single answer can be found) the hand moves downwards only about 2 cm before the motion is stopped ($\gamma_{\max} \approx [0.02/0.35]$ ($180/\pi = 3.27^\circ$)).

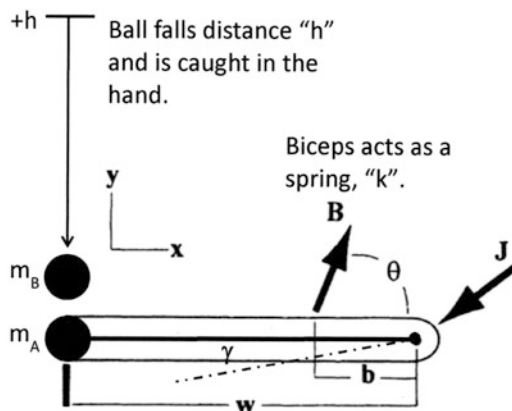


FIGURE 1.11. The ball, m_B , is dropped through a distance (h), and caught in an (invisible) hand. To simplify the problem, assume that all of the mass of the arm (m_A) is concentrated in the hand and the rest of the arm is massless. The arm rotates (dynamically) through an angle $\gamma(t)$ around the (fixed) elbow and the biceps force (B) creates a moment around the elbow to keep the forearm flexed.

To calculate the momentum of the ball ($M_B = m_B v_B$) before being caught, the velocity (of the ball for the fall of distance “ h ”) is calculated using simple mechanics and the acceleration of gravity (g) as:

$$\begin{aligned} h &= \frac{1}{2} g t_{\max}^2 \\ v_B &= g t_{\max} \\ M_B &= m_B \sqrt{2gh} \end{aligned} \quad (1.13)$$

Conservation of momentum when the hand catches the ball means that $M_B = (m_B + m_A) v_b$, where m_A is the mass of the arm and v_b is the initial speed of the hand. The initial speed of the hand, therefore, is:

$$v_b = \frac{m_B}{m_A + m_B} \sqrt{2gh} \quad (1.14)$$

The increase in length of the biceps (at 70°) is calculated using trigonometry, relationships from the previous static problem (Fig. 1.3) and the observation that the maximum angle of the arm motion is $\gamma \approx y/w$, where “ y ” is the distance the hand moves and w is the length of the forearm.

At the biceps insertion, the hand moves downward a distance $b\gamma$ which translates immediately into a change of length (dL) of the biceps of: $dL = \gamma b \sin(\theta)$. The biceps force (B) is assumed to be caused by a spring, so the force increases as the length of the biceps increases ($B = kdL$). The biceps force component in the “ y ” direction creates the moment holding up the arm:

$$M_B = bB_y = bkdL \sin(\theta) \approx kb^2 \gamma \sin^2(\theta) \quad (1.15)$$

This moment can be transferred into a force at the hand by zeroing moments at the elbow using: $F_H w = M_B$, so that the force at the hand that slows the falling of the combined ball-and-hand is:

$$F_H = k \left(\frac{b}{w} \sin(\theta) \right)^2 y \quad (1.16)$$

where we used the approximation for the angle, $\gamma = y/w$ (This moment transfer is valid because the mass of the arm is concentrated in the hand and the remainder is massless.).

After all of these calculations, we have what is needed to construct a differential equation: “Force = mass $\times$ acceleration” which, along with the initial velocity (above) and the initiation displacement of the hand (zero) can be integrated numerically (we used the Runge-Kutta algorithm). Collected together, the equations to integrate are:

$$\begin{aligned} \frac{d^2 y}{dt^2} &= \frac{F_H}{m_A + m_B} \approx k \left(\frac{b}{w} \sin(\theta) \right)^2 \frac{y}{m_A + m_B} \\ \frac{dy(0)}{dt} &= -\frac{m_B}{m_A + m_B} \sqrt{2gh} \\ y(0) &= 0 \end{aligned} \quad (1.17)$$

To obtain a specific integration for the derived equations, assume: $h = 2$ m, $g = 9.81$ m/s², $m_B = 2.23$ kg, $m_A = 2.1$ kg, $w = 0.35$ m, $b = 0.04$ m and $k = 850,000$ N/m. With these parameters (some similar to those for the static problem of the arm holding a weight) the *maximum angular deflection* of the arm is $\gamma = 3.46^\circ$ and the ratio of the maximum force in the biceps to the weight of the ball is $B/(m_B g) = 102$ (Fig. 1.12). In the static problem the ratio of the biceps force to the ball weight was only 9.1, so *the maximum dynamic biceps force for this example is an order of magnitude larger* than the for the equivalent static case (Fig. 1.12).

The important messages for this section are that even simplified dynamic models are significantly more complex than a static model and that the dynamic forces in muscles (and joints) can be many times greater than the static estimates—which are themselves many times larger than the forces applied to the hand or foot. Mechanical support for these very high loads (and the resulting high stresses) is provided by bone, cartilage, tendon and ligament, skeletal tissues that we will examine in detail throughout the remainder of this book.

1.6 Indeterminate Joint Problems

In addition to primarily examining the static case, the problems we have considered so far were limited to two-dimensional analyses in which it was possible to solve for three unknowns. To reduce the number of unknowns to

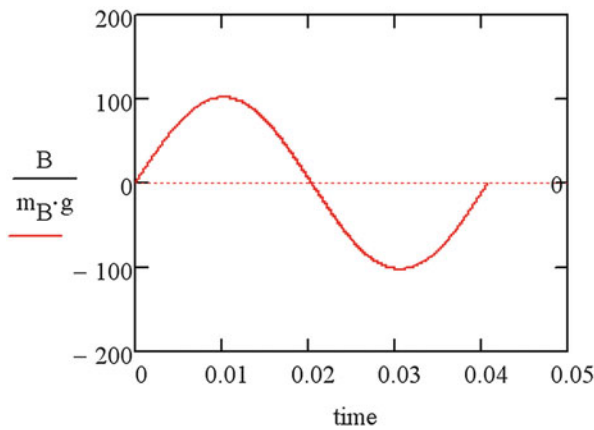


FIGURE 1.12. The dynamic biceps force ratio as a function of time in seconds for the parameters assumed in the text. The ball is stopped by the hand in 0.01 s and then, because of the assumption of zero damping, the arm oscillates at a frequency of 25 cycles per second. This simulation was run for a full cycle of the oscillation. Do you think that this approximate solution for catching a ball is correct after the first peak force? What about when the biceps force becomes negative?.

TABLE 1.1. Elbow muscle data (Winter 1990)

Muscle	Moment arm (cm)	PCA (cm ²)
Biceps	a = 4.6	α = 4.6
Brachialis	b = 3.4	β = 7.0
Brachioradialis	c = 7.5	γ = 1.5

the number of equations, we had to combine the function of some muscles while completely ignoring others. We could have obtained more equations by extending our analysis to three dimensions, but then we would have had to consider additional muscles as well. All real joints present static equilibrium problems that are mathematically indeterminate because there are always more unknowns than equilibrium equations. Thus, to solve for joint forces more realistically, it is necessary to find more equations to use in the solution.

To see an example of how this might be done, we return again to the static elbow problem. Previously, we noted that the primary flexors of the forearm are the brachialis and the biceps, and we combined their actions into a single force vector. Now, we separate these muscle forces. In addition, we include the third muscle, the brachioradialis (see Fig. 1.2). Table 1.1 shows the moment arms of the three muscles with respect to the center of rotation of the elbow when it is flexed at 90° (Winter 1990). It also shows the *physiologic cross-sectional area* (PCA) of each muscle. It is approximately true that the maximal force that a muscle can exert is proportional to its physiologic

cross sectional area. Another way to say this is that muscles have an upper limit to the stress that they can generate within themselves. Measured maximum stresses vary from 20 to 100 N/cm², depending on conditions (isometric or dynamic, parallel or pennate fibers, etc.). We will use 40 N/cm² as a rule of thumb for the maximal stress (Morris 1948; Maughan et al. 1983).

With all three muscles in the problem, there are five unknowns (three muscle force magnitudes and two components of joint force) but still only three equilibrium equations. We need two other relationships between the variables to obtain a solution. One way to get such additional equations would be to assume that all three muscles work together to equalize their individual stresses. That is, each muscle force is proportional to the muscle's physiologic cross-sectional area. Then the three muscle forces are

$$\begin{aligned} A &= k\alpha \text{ (biceps)} \\ B &= k\beta \text{ (brachialis)} \\ C &= k\gamma \text{ (brachioradialis)} \end{aligned} \quad (1.18)$$

where α , β , and γ are the respective cross-sectional areas and $k = 40 \text{ N/cm}^2$ is the constant muscle stress. This may be rewritten as

$$A/\alpha = B/\beta = C/\gamma \quad (1.19)$$

which constitutes the two additional equations that were needed. Using these to substitute for B and C in the moment equilibrium equation,

$$\sum M = aA + bB + cC - wW = 0 \quad (1.20)$$

one has

$$\begin{aligned} A &= [aw/(a\alpha + b\beta + c\gamma)]W = 2.86W \\ B &= (\beta/\alpha)A = 4.36W \\ C &= (\gamma/\alpha)A = 0.93W \end{aligned} \quad (1.21)$$

This solution indicates that the biceps, the muscle with the intermediate cross-sectional area and moment arm, also exerts the intermediate force. The brachioradialis, which has the largest moment arm and the smallest cross-sectional area, exerts the least force, and vice versa for the brachialis. Whether this has anything to do with the actual distribution of forces in the muscles is debatable! Many other equally plausible, but not necessarily so mathematically trivial, “additional conditions” may be imagined. For example, the force distribution may shift around to give different muscle fibers a rest, or to keep the weight from tipping sideways in the hand, or both. A number of investigators have speculated that the body resolves the indeterminacy of such problems by acting to optimize or minimize some biologically important variable, such as the total energy required by all the muscles acting. Collins (1995) provides additional discussion of this subject.

Box 1.4 Surgical Reconstruction of the Hip Joint

Total hip arthroplasty, which involves replacement of the ball and socket of the hip joint with manufactured components (Fig. 1.13), is one of the most successful surgical operations developed in this century. The procedure has relieved joint pain and increased the quality of life for millions of people. Before the development of successful hip replacement surgery by Sir John Charnley (1973), patients with debilitating degenerative arthritis of the hip were forced to either suffer with the condition or undergo hip joint fusion (called *arthrodesis*). As described in Sect. 1.2, the loads imposed on bones and joints are dramatic as a result of the mechanical disadvantage under which they function. Forces at the hip can easily exceed fivefold body weight. In addition, the hip joints of an average individual experience about a million load cycles (steps) each year. Thus total hip components should be both strong and durable.

A total hip replacement consists of two parts: a femoral component, the ball, and an acetabular component, the socket. Femoral components are made from high-strength alloys such as cobalt-chrome or titanium, and consist of a highly polished head atop an intramedullary stem. The stem is inserted into the canal of the femur and is usually fixed with an acrylic plastic [polymethyl methacrylate (PMMA)] called *bone cement*. This material does not actually cement the implant to the bone,

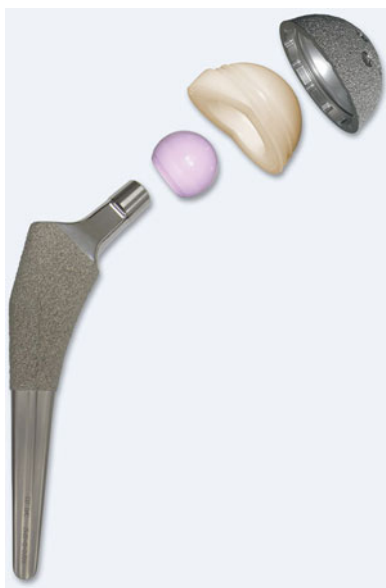


FIGURE 1.13. A modern hip joint replacement prosthesis. The spherical ceramic ball articulates with an ultrahigh molecular weight polyethylene socket fitted into a metal cup. The long metal stem is placed inside the medullary canal of the femur. The proximal stem and socket backing have porous surfaces designed for bone ingrowth in lieu of using polymethyl methacrylate (PMMA) for fixation to the bone.

however; it simply fills space so that the fit does not have to be exact, “grouting” the implant in place. More recent designs utilize a porous stem surface designed for fixation by bone ingrowth. The acetabular component is fabricated from ultrahigh molecular weight polyethylene, with or without a metal backing. A design with a metal backing is illustrated. Other combinations of materials for the bearing surface are sometimes used, including metal-on-metal, ceramic-on-ceramic, and ceramic-on-polyethylene. The PMMA and metal-on-polyethylene bearing surfaces innovations were important factors enabling Charnley to develop a system that featured good material and structural strength, good resistance to fatigue damage, and low friction at the bearing surface. Careful surgical technique remains important, but with current methods most patients, being elderly and relatively sedentary, can expect satisfactory performance for 15 years or more.

Despite the operation’s overall success, failures do occur, more frequently in young, active patients. The high forces imposed on the joint make maintenance of implant fixation difficult; component loosening and migration are not uncommon. High joint reaction forces also produce relatively high frictional forces at the metal–plastic interface. Over time, frictional wear debris, both metallic and polyethylene, elicits biologic responses that resorb bone and may further compromise implant stability. Improving the fixation and wear characteristics of total joint components is a major focus of orthopaedic research.

1.7 Equine Fetlock Forces

Horses, and in particular racehorses, place extraordinarily high loads on their limbs. This is especially true of the forelimbs, which are thought to carry about 60 % of the animal’s weight. Many racehorses are seriously injured each year by mechanical failures of the structures in the distal fore-limb. Because it is so difficult to repair these failures in such a way that the horse can stand and walk during healing, and because horses must be erect and mobile to survive, many animals die as a result of injuries that would not be life threatening in humans or other smaller animals.

Figure 1.14 shows a sketch of the anatomy of the distal forelimb of the horse. Remember, the horse does not have feet the way humans do; it walks on the tips of its “fingers.” The other bones of the foot (or hand) serve to lengthen the leg. P1 and P2 mark the first and second phalangeal bones in the figure; the third phalangeal bone is inside the hoof. The equine third metacarpal bone (MC3) has become extremely stout, and carries loads from the phalangeal complex up to the carpal bones and the radius, seen at the top of the diagram. (The second and fourth metacarpal bones are vestigial struts stuck to the sides of MC3; the other metacarpal bones no longer exist.) The joint between MC3 and P1 is called the *fetlock*; its biomechanical importance derives largely from the fact that the bones above it rise straight up to the torso, but the phalangeal bones below it angle sharply forward in a springy

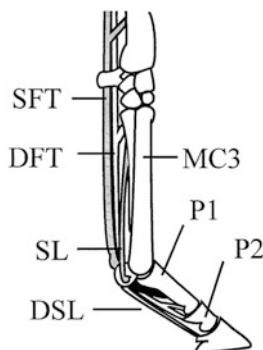


FIGURE 1.14. Equine distal forelimb anatomy, *sagittal view*. *SFT* superficial digital flexor tendon, *DFT* deep digital flexor tendon, *SL* suspensory ligament, *DSL* distal sesamoidian ligaments, *MC3* third metacarpal bone, *P1* proximal phalanx, *P2* middle phalanx.

cantilever arrangement. The joint between P1 and P2 is called the *pastern joint*.³

Behind the third metacarpus, which is also called the *cannon bone*, two tendons come down from the superficial and deep digital flexor muscles. The superficial flexor tendon (SFT) inserts on P2, and the deep flexor tendon (DFT) inserts on P3 inside the hoof. On their way to these attachment sites, these tendons pass over two small bones located just behind the distal end of the cannon bone. These bones are called the *proximal sesamoid bones*; their cartilaginous dorsal (anterior) surfaces articulate with the palmar (posterior) distal condyles of the cannon bone. The fetlock's sesamoid bones function very much like the patella in your knee, moving the line of action of the tendons away from the joint center so that they exert a larger moment about it. In addition to these tendons, there is a set of four ligaments that help stabilize the fetlock. The *suspensory ligament* (SL) runs from the proximal posterior surface of MC3 down to the sesamoid bones. From there, three *distal sesamoidian ligaments* connect to P1 and P2. These four suspensory ligaments, along with the sesamoid bones, are often called the *suspensory apparatus*. The cannon bone brings the weight of the horse down to the fetlock, which sits in the suspensory apparatus in much the same way as a person lies in a hammock. The flexor muscles and tendons help carry this load, with the sesamoid bones providing a bearing surface to protect the

³ P1 is also called the pastern bone. The name apparently comes from the fact that to pasture a horse without benefit of fences, ranchers would hobble it with a short rope connecting the P1 bones.

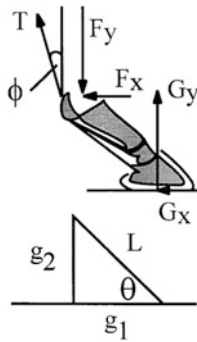


FIGURE 1.15. *Upper:* Free-body diagram for forces in the equine fetlock joint. G_x and G_y components of the ground reaction force, T tendon force, F_x and F_y components of the joint force. *Lower:* Diagram for the moment arms, g_1 and g_2 , of the ground reaction force. L distance from the ground contact to the joint center, θ angle of the phalanges with respect to the ground.

distal cannon bone, as well as increasing the moment arms of the tendons and ligaments, as previously mentioned.

To analyze the forces in the fetlock joint, we construct the free-body diagram shown in Fig. 1.15. The analysis is two dimensional in the sagittal plane. The “amputation” has been done in such a way that the free body includes the sesamoid bones and everything distal to the cannon bone. These structures are all assumed to be rigidly connected; the pastern joint is assumed to be fixed. The forces resulting from the flexor tendons and the suspensory ligament are all combined into a vector, T , which has components in the x - and y -directions. Considering the large weight of the horse, the weight of the free body is ignored. The remaining forces acting on the free body are the two components of the ground reaction force, G , and the two components of the joint reaction force, F . F is assumed to act at the joint’s center of rotation. Taking moments about this center, we have

$$\sum M = g_1 G_y - g_2 G_x - t T_y = 0 \quad (1.22)$$

where g_1 and g_2 are as defined in the lower portion of Fig. 1.15 and t is the moment arm of the vertical component of T . It is assumed that the x -component of T has a negligible moment arm with respect to the joint center (Fig. 1.15).

The force equilibrium equations are

$$\sum F_X = -T_x - G_x + F_x = 0 \quad (1.23)$$

$$\sum F_Y = T_y + G_y - F_y = 0 \quad (1.24)$$

These equations may be solved for various ground reaction forces and anatomical angles (ϕ and θ). First, let us assume that G and T are entirely

vertical (e.g., when the horse is standing). Then $\Phi = 0$, $T = (g_1/t)G = GL \cos \theta/t$ and $F = T + G$. If the horse weighs 500 kg (roughly 5000 N), and 60 % of that is divided equally between the two forelimbs, then $G = G_y = 1500$ N. For $L = 0.2$ m, $\theta = 40^\circ$, and $t = 0.05$ m:

$$T = (1500)(0.2)(0.766)/0.05 = 4596 \text{ N}$$

The force carried by the tendons and ligament behind the cannon bone is nearly equal to the entire weight of the animal! In fact, if the angle θ sags to 33° , T becomes 5000 N. The joint reaction force is, of course, also vertical, and is $F = T + G = 6096$ N, or 1.2 times body weight.

Now suppose that the ground reaction force is still entirely vertical, but the angle Φ is 20° . Then, from the moment equation,

$$T_y = (g_1/t)G = GL \cos \theta/t \quad (1.25)$$

so that

$$T = T_y / \cos \phi = (GL/t)(\cos \theta / \cos \phi) = 4891 \text{ N} \quad (1.26)$$

and from the force equations

$$F_y = T_y + G = [1 + (L/t) \cos \theta]G = 6096 \text{ N} \quad (1.27)$$

$$F_x = T_x \sin \phi = 1673 \text{ N} \quad (1.28)$$

The resultant joint force is 6321 N, and it acts at an angle of 15° to the vertical.

We return to this problem in the exercises that follow, and consider the effects of horizontal components of the ground reaction force on T and F . If this topic is of particular interest to you, you may want to read papers by Bartel et al. (1978) and Riemersma et al. (1988). A related work on the evolution of equine locomotion by Thomason (1991) is also of interest.

1.8 Summary and Further Reading

This chapter has considered the analysis of forces between one bone and another across joints. We have seen that such forces are always several times the external force being supported. This occurs because skeletal levers are arranged to magnify movement rather than force. Thus, muscle forces must be substantially greater than the forces being applied to the external environment. A person turning over in bed produces substantial hip joint forces, without supporting body weight, through the required muscle actions. Similarly, marine mammals such as dolphins and whales, whose buoyancy spares them the need to resist gravity, nevertheless have massive skeletons to transmit the muscle forces produced in swimming. The forces acting on bones can be estimated using the principles of static equilibrium. This method can

provide very useful information, but significant approximations must be made to solve such problems. In general, the joints of vertebrate animals are statically indeterminate; that is, the number of unknowns (muscles acting to move or stabilize the joint) is almost always greater than the number of equations to be solved.

Suggestions for further reading in this field begin with the classic textbook *Biomechanics of Human Motion* (Williams and Lissner 1977). Another well-known textbook, aimed at the medical community, is *Basic Biomechanics of the Musculoskeletal System* by Nordin and Frankel (1989). Steindler's *Kinesiology of the Human Body Under Normal and Pathological Conditions* (Steindler 1955) is excellent for its concentration on anatomy and pathomechanics. Early studies of the biomechanics of the human hip joint by Inman; McLeish and Charnley; and Rydell's paper on measurement of hip joint forces using an instrumented prosthesis, will give you a greater appreciation for the complexities of experimental work in this area. *Animal Mechanics* provides an approach to this topic well suited to zoologists. If you are historically inclined, by all means seek out Maquet's translation of Giovanni Borelli's wonderful seventeenth-century work, *De Motu Animalium* (*On the Movement of Animals*), sampled in Box 1.1 (Borelli 1989; Rydell 1966; Inman 1947; McLeish and Charnley 1970; Alexander 1968).

1.9 Exercises

- 1.1. Solve the second part of the hip force problem (the force equilibrium equations) by considering a free-body diagram of the weight-bearing lower extremity and using the result $P = 2W$. Remember to include the weight of the extremity itself. What is the disadvantage of using this free-body diagram to solve the moment equation?
- 1.2. Suppose that the distance b in women's pelvis is 10 % greater than men's because of their greater pelvic width. How does this change their abductor muscle force magnitude and the magnitude and direction of their hip joint reaction force, assuming the moment arm of the abductor muscles is independent of gender?
- 1.3. Consider the force required in the erector spinae muscles to stabilize the head of a student leaning over his book. Also of interest is the force between the fifth and sixth cervical vertebrae. Figure 1.16 is a free-body diagram of this situation: E is the erector spinae force, H is the weight of the head and neck acting downward at their combined center of gravity, and R is the vertebral reaction force acting at a center of pressure on the inferior end plate of C5. Determine the magnitude of E and the magnitude and direction of R . Assume that E is parallel to the spine and $\theta = 70^\circ$. Let the moment arms of E and H with respect to the point of action of R be 0.02 and 0.10 m, respectively. Express your results as multiples of H .

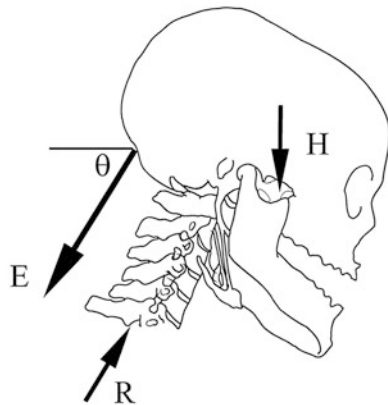


FIGURE 1.16. Diagram for Exercise 1.3, calculating the forces in the neck while studying. **H** is the weight of the head and neck acting at their center of gravity; **E** is the force of the erector spinae muscles and θ is its angle; and **R** is the vertebral reaction force.

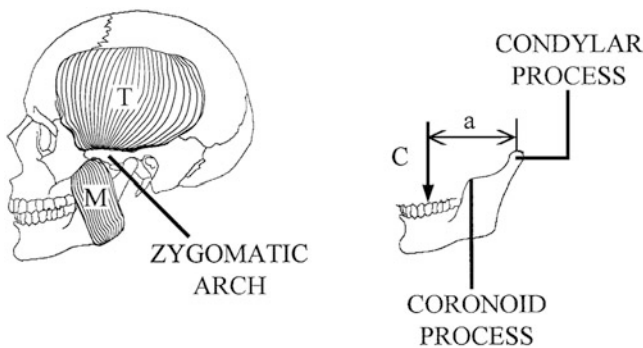


FIGURE 1.17. *Left*: Anatomical diagram showing locations of the temporalis (**T**) and masseter (**M**) muscles. The former passes through the zygomatic arch to pull the coronoid process up and to the right. *Right*: Detail of the mandible showing the assumed biting force, **C**, and its moment arm, **a**, with respect to the joint.

1.4. Consider the force produced in the temporomandibular joint when chewing. Figure 1.17 shows the location of the two major muscles active in a sagittal plane when chewing. The temporalis muscle arises from a broad area on the side of the skull, passes down through the space enclosed by the zygomatic arch, and inserts on the coronoid process of the mandible. In the picture of the skull, this point is hidden behind the other big chewing muscle, the masseter, which runs from the

anteroinferior edge of the zygomatic arch down to the “angle” of the jaw. Ignoring other masticatory muscles, draw a free-body diagram of the mandible, assuming that the problem is two dimensional and the same chewing forces act on each side of the jaw. Let $\mathbf{C}$ be a vertical chewing force acting on a molar with a 0.06-m moment arm about the condylar process in the temporomandibular joint (TMJ). Assume that the masseter force, $\mathbf{M}$, acts at an angle of 115° with respect to the x -direction, and that the temporalis muscle force, $\mathbf{T}$, acts at an angle of 80° . Their moment arms with respect to the TMJ center are $b = 0.04$ m and $d = 0.02$ m, respectively. After each of these forces is represented on the free-body diagram, write the equations for static equilibrium. Under what conditions could these equations be solved? How do the direction and magnitude of the joint reaction force, $\mathbf{R}$, depend on the relative magnitudes of $\mathbf{M}$ and $\mathbf{T}$? Try to obtain solutions to the equilibrium equations using at least two different kinds of additional conditions. You can read more about the mechanics of the human jaw in papers by Barbenel (1972) and Osborn (1996).

- 1.5. A male gymnast is performing the “iron cross” exercise on the rings, in which the arms are held straight out with a ring grasped in each hand, and the body hangs motionless and vertical between the rings (Fig. 1.18). Consult an anatomy book to determine the principal adductor muscles for the arm in this position, and estimate their angle of pull and insertion point. Draw a free-body diagram and calculate the approximate magnitude of the adductor muscle force and the magnitude and direction of the reaction force in the glenohumeral (shoulder) joint during the maneuver. State the principal assumptions necessary for your calculation.
- 1.6. Suppose, in Exercise 1.5, that the ropes holding the rings do not rise vertically from the gymnast’s hands, but pull outward or inward at an angle θ . How does this angle affect the forces in the shoulder joint?
- 1.7. In a classic paper, McLeish and Charnley (1970) described in detail the considerations necessary to estimate the forces in the hip during the stance phase of gait. As noted in Box 1.4, Sir John Charnley also developed the total hip replacement procedure in its modern form (Collins 1995). Read this remarkable paper.
- 1.8. For the equine fetlock joint, work out the magnitude and direction of the joint force when a horizontal component of the ground reaction force (i.e., G_x in Fig. 1.1) is present. Let $G_x = \pm 0.20 G_y$, and assume, as before, that $L = 0.2$ m, $\Phi = 0$, $\Theta = 40^\circ$, and $t = 0.05$ m.
- 1.9. To compare *Homo sapiens* and *Australopithecus* hip joint mechanics using similar skeletal materials, Lovejoy and co-workers (Lovejoy et al. 1973) studied the skeletons of Native Americans recovered in the

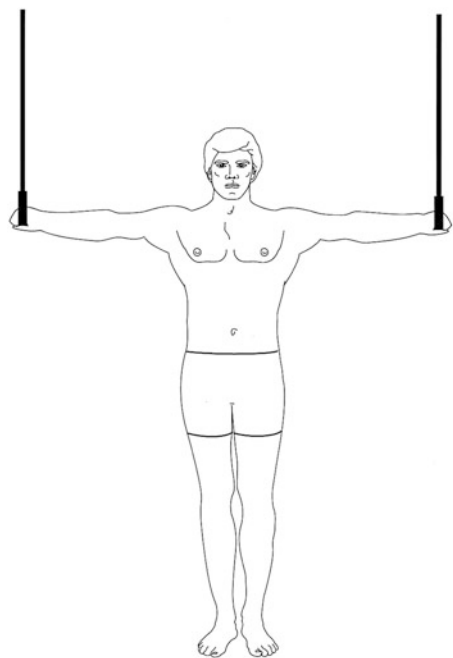


FIGURE 1.18. Sketch for Exercise 1.5 shows a gymnast performing the iron cross maneuver.

TABLE 1.2. Archeological skeletal data (Lovejoy et al. 1973)

Skeletons	Interfemoral head distance (mm)	Abductor moment arm (mm)	Abductor angle (°)
Australopithecus, STS 14	132	36	15
<i>Homo sapiens</i> , Amerindians, male ($n = 8$)	158.0 ± 10.3	50.6 ± 1.7	12.9 ± 2.2
<i>Homo sapiens</i> , Amerindians, female ($n = 12$)	165.8 ± 9.0	45.4 ± 3.8	13.3 ± 2.8

field rather than those from laboratory cadavers representing modern Americans. They estimated interfemoral head distance and abductor moment arm distance from the archeological remains of 20 *Homo sapiens* and 1 *Australopithecus*; these results are shown in Table 1.2. Assume that the body weight moment arm (b in Figs. 1.5 and 1.19) is slightly greater than half the interfemoral head distance (D_{IF}) shown in the second column of Table 1.2, i.e.,

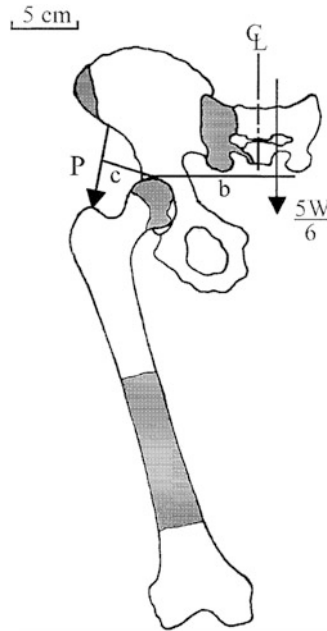


FIGURE 1.19. Diagram for *Australopithecus* hip force problem [Exercise 1.9]. P is the abductor muscle force and c is its moment arm with respect to the joint center; b is the moment arm of the body weight minus the support leg (i.e., $5W/6$). Shaded regions indicate reconstructed portions of the bones (Lovejoy et al. 1973).

$$b = k D_{IF} \quad (1.29)$$

where k is in the neighborhood of 0.5. Assume also that the abductor moment arm (c in Figs. 1.5 and 1.19) and angle are as shown in the third and fourth columns. Plot a graph of the force on the head of the femur as a function of k for the three kinds of skeletons shown. What can you deduce about the hip forces in Native American females and the *Australopithecus* individual compared to Native American males?

- 1.10. Another interesting factor in the development of human bipedalism is its relationship to cranial capacity. The diameter of the birth canal is primarily determined by the size of the infant's head at birth. As human evolution proceeded and cranial capacity increased, the birth canal had to become larger. Presumably, this led to widening of the pelvis and increased the moment arm of the body weight (b in Figs. 1.5 and 1.19) with respect to the hip joints of females. Assume the other parameters of the hip joint were

abductor muscle angle = 15°
 abductor muscle moment arm = $c = 36$ mm
 interfemoral head distance = $D_{IF} = 132$ mm
 body weight moment arm = $b = k D_{IF}$
 $k = 0.60$

and remained constant (refer to Exercise 1.9). Determine and graph the relationship between the hip abductor force $\mathbf{P}$ and the newborn cranial volume V , assuming that the head is a sphere and its newborn volume increases from 500 to 1000 cm^3 . Does encephalization (increasing brain size) appear to be an important factor in increasing hip joint forces during bipedal walking? For additional discussion of this issue, see Leutenegger (1972) and Ruff (1995).

- 1.11. Go to the library and read the paper by Zihlman and Hunter (1972) (Fig. 1.19). It is a more complex, three-dimensional analysis of hip mechanics in *Australopithecus*. Write a paragraph discussing its strengths and weaknesses.
- 1.12. Derive the analytical solution to the dynamic equations from Sect. 1.5:

$$\begin{aligned}
 \frac{d^2 y}{dt^2} &= \frac{F_H}{m_A + m_B} \approx k \left(\frac{b}{w} \sin(\theta) \right)^2 \frac{y}{m_A + m_B} \\
 \frac{dy(0)}{dt} &= -\frac{m_B}{m_A + m_B} \sqrt{2gh} \\
 y(0) &= 0
 \end{aligned}$$

Discuss the approximations used to develop the equations with particular emphasis on when and if they are valid. Give two examples of “catching a ball” where the equations derived in Sect. 1.5 are invalid.

- 1.13. Compare and contrast the mechanical stability of the shoulder joint and the hip joint. Which would you consider to be the most stable? Why? Discuss the role of bone shape, overlying muscles, associated tendons and ligaments in the maintenance of joint stability and force transmission.

References

- Agur AMR. Grant's atlas of anatomy. Baltimore: Williams & Wilkins; 1991.
 Alexander RM. Animal mechanics. Seattle: University of Washington Press; 1968.
 Atwater AE. Gender differences in distance running. In: Cavanaugh PR, editor. Biomechanics of distance running. Champaign: Human Kinetics Books; 1990.
 Barbenel JC. The biomechanics of the temporomandibular joint: a theoretical study. J Biomech. 1972;5:251–6.

- Bartel DL, Schryver HF, Lowe JE, Parker RA. Locomotion in the horse: a procedure for computing the internal forces of the digit. *Am J Vet Res.* 1978;39:1721–33.
- Bay BK, Hamel AJ, Olsen SA, Sharkey NA. Statically equivalent load and support conditions produce different hip joint contact pressures and periacetabular strains. *J Biomech.* 1997;30:193–6.
- Bergmann G, Graichen F, Siraky J, Jendrzynski H, Rohlmann A. Multichannel strain gauge telemetry for orthopaedic implants. *J Biomech.* 1988;21:169–76.
- Bergmann G, Graichen F, Rohlmann A. Hip joint loading during walking and running measured in two patients. *J Biomech.* 1993;26:969–90.
- Bergmann G, Deuretzbacher G, Heller M, Graichen F, Rohlmann A, Strauss J, Duda GN. Hip contact forces and gait patterns from routine activities. *J Biomech.* 2001;34(7):859–71.
- Borelli GA. On the movement of animals (1989 translation by P. Maquet). Berlin: Springer; 1989.
- Burr DB, Gerven DPV, Gustav BL. Sexual dimorphism and mechanics of the human hip: a multivariate assessment. *Am J Phys Anthropol.* 1977;47:273–8.
- Charnley J. Arthroplasty of the hip: a new operation (reprinted from *Lancet*, pp. 1129–1132, 1961). *Clin Orthop Relat Res.* 1973;95:4–8.
- Collins JJ. The redundant nature of locomotor optimization laws. *J Biomech.* 1995;28:251–67.
- Davy DT, Kotzar GM, Brown RH, Heiple KG, Goldberg VM, Heiple JG, Berrile J, Burstein AH. Telemetric force measurements across the hip after total arthroplasty. *J Bone Joint Surg.* 1988;70A:45–50.
- English TA, Kilvington TM. In vivo records of hip loads using a femoral implant with telemetric output. *J Biomed Eng.* 1979;1:111–5.
- Erdemir A, Piazza SJ, Hamel AJ, Fauth AR, Sharkey NA. Dynamic loading of the plantar aponeurosis in walking. *J Bone Joint Surg.* 2004;86A(3):546–52.
- Frankel VH, Burstein AH. Orthopaedic biomechanics. Philadelphia: Lea & Febiger; 1973.
- Inman VT. Functional aspects of the abductor muscles of the hip. *J Bone Joint Surg.* 1947;29A:607–19.
- Kirane YM, Michelson JD, Sharkey NA. Contribution of the flexor hallucis longus to loading of the first metatarsal and first metatarsophalangeal joint. *Foot Ankle Int.* 2008;29(4):367–77.
- Kotzar GM, Davy DT, Goldberg VM, Heiple KG, Berilla J, Heiple KG, Brown RH, Burstein AH. Telemeterized in vivo hip joint force data: a report on two patients after total hip surgery. *J Orthop Res.* 1991;9(5):621–33.
- Leutenegger W. Newborn size and pelvic dimensions of *Australopithecus*. *Nature.* 1972;240:568–9.
- Lovejoy CO, Heiple KG, Burstein AH. The gait of *australopithecus*. *Am J Phys Anthropol.* 1973;38:757–80.
- Maughan RJ, Watson JS, Weir J. Strength and cross-sectional area of human skeletal muscle. *J Physiol.* 1983;338:37–49.
- McHenry H. Tempo and mode in human evolution. *Proc Natl Acad Sci.* 1994;91:6780–6.
- McLeish RD, Charnley J. Abduction forces in the one-legged stance. *J Biomech.* 1970;3:191–209.
- Morris CB. The measurement of the strength of muscle relative to the cross section. *Res Q.* 1948;19:295–303.

- Nordin M, Frankel VH. Basic biomechanics of the musculoskeletal system. Philadelphia: Lea & Febiger; 1989.
- Osborn JW. Features of human jaw design which maximize the bite force. *J Biomech.* 1996;29:589–95.
- Riemersma DJ, Schamhardt HC, Hartman W, Lammertink JL. Kinetics and kinematics of the equine hind limb: in vivo tendon loads and force plate measurements in ponies. *Am J Vet Res.* 1988;49:1344–52.
- Ruff CB. Biomechanics of the hip and birth in early Homo. *Am J Phys Anthropol.* 1995;98:527–74.
- Rydell N. Forces in the hip joint. Part II. Intravital measurements. In: Kenedi RM, editor. Biomechanics and related bioengineering topics. Oxford: Pergamon Press; 1965. p. 351–7.
- Rydell N. Forces acting in the femoral head prosthesis. *Acta Orthop Scand Suppl.* 1966;88:1–46.
- Sharkey NA, Hamel AJ. A dynamic cadaver model of the stance phase of gait: performance characteristics and kinetic validation. *Clin Biomech.* 1998;13:420–33.
- Steindler A. Kinesiology of the human body under normal and pathological conditions. Springfield, IL: Charles C. Thomas; 1955.
- Taylor SJ, Perry JS, Meswania JM, Donaldson N, Walker PS, Cannon SR. Telemetry of forces from proximal femoral replacements and relevance to fixation. *J Biomech.* 1997;30(3):225–34.
- Thomason JJ. Functional interpretation of locomotory adaptations during equid evolution. In: Rayner JMV, Wootton RJ, editors. Biomechanics in evolution. Cambridge: Cambridge University Press; 1991. p. 213–27.
- Williams PL. Gray's anatomy. New York: Churchill Livingstone; 1995.
- Williams M, Lissner HR. Biomechanics of human motion. Philadelphia: Saunders; 1977.
- Winter DA. Biomechanics and motor control of human movement. New York: Wiley; 1990.
- Zihlman AL, Hunter WS. A biomechanical interpretation of the pelvis of Australopithecus. *Folia Primatol.* 1972;18:1–19.

2

Skeletal Biology

The human skeleton is a mechanically optimized biological system whose composition and organization reflect the functional demands made on it.

Thomas Einhorn (1992)

Nothing in biology makes sense except in the light of evolution.

Theodosius Dobzhansky (1973)

Abstract Chapter 1 has shown us that the skeleton must withstand very high forces; because our muscles can only contract a small percentage of their length, we must amplify movements using levers that “spend rather than save forces,” as Borelli put it. In this chapter we begin to develop an understanding of how the biology of bone, cartilage and the other connective tissues provides tissues that can support these large forces day in and day out for a lifetime.

Keywords Bone • Cartilage • Tendon • Ligament • Structure • Composition • Growth

2.1 Introduction to Bone

As more is learned about the mechanics of the skeleton, it becomes increasingly clear that mechanical demands heavily influence what happens biologically. Most of what we know about the basic cell- and tissue-level mechanisms that enable bone and its adjoining connective tissues to adjust to mechanical forces has been learned since the 1960s. Before then, skeletal research was focused on bone chemistry. Calcium is a metal and phosphorus a non-metal that, along with oxygen and hydrogen, are primary constituents of bone mineral. The serum concentrations of calcium and phosphorus are of great significance in physiology and medicine, and there was intense interest in the middle of the twentieth century in how bone contributed to the homeostatic regulation of these elements. Physiologists and endocrinologists realized that bone mineral served as a calcium and phosphorus reservoir for the rest of the body. Bone biology was seen largely in terms of its role in

transferring mineral to and from the skeleton. Moreover, it was recognized that bone's mechanical properties—its strength and stiffness—were to a significant degree determined by the bone's mineral content. Early work by John D. Currey of the University of York and others demonstrated the strong correlations between bone density and strength (Currey 1969a, 1969b) and, since bone mineral content could be measured in living animals using x-ray methods, it appeared that the main factors that cause bone tissue to be strong or weak were well understood. The roles of collagen alone and of the mineral-collagen composite were investigated to a much lesser degree, partially because collagen content and conformation could not be non-invasively measured in living animals.

Exploring the micro-anatomy of bone and the cellular constituents responsible for maintaining its strength was also challenging. Preparation of undecalcified bone sections entailed the tedious process of cutting thick sections of cortical bone and grinding them by hand using grinding powder or sandpaper to a thickness of 100–150 μm , where accurate analysis could take place. Although this was an employment possibility for orthopaedic residents and graduate students, it severely curtailed the number of specimens that could be made. With the advent of plastic embedding in the 1970s it became possible to cut thin sections using a microtome or diamond saw, greatly improving and accelerating the technical preparation of microscope slides. This fueled research by Ascenzi, Bonucci and others to study the microscopic structure of bone, and relate this to its mechanical properties (Ascenzi and Bonucci 1964, 1967, 1968, 1972).

Bone is much more, however, than its mineralized components. It is a hierarchically-organized material that is constructed as a fiber-reinforced composite material (Fig. 2.1).

Fiber-reinforced composites are often highly anisotropic (having different mechanical properties in different orientations) and can have low transverse strength because of the directionality of the fibers. Such a design can be used to increase strength in the predominant load direction without severely compromising properties in the perpendicular directions. For example, in a long bone like the femur the strong direction is along the femoral shaft and the weaker, but functionally adequate, directions are circumferential and radial. This is a good design for a femur, where most of the loading is along the long axis of the bone. For cancellous bone, the strong direction of the material is parallel to the surfaces of the trabeculae which, once again, makes sense as that is the predominant loading direction.

At the histological level, bone is either dense compact (cortical) bone, or more porous trabecular (cancellous) bone. The tissue that forms these microscopic architectures can be lamellar (sheet-like) or completely disorganized and nearly devoid of identifiable structure. Bone microstructure is often in layers, alternating between regions with highly oriented mineral (lamellae) and adjacent interlamellar regions. The regions with oriented crystals are

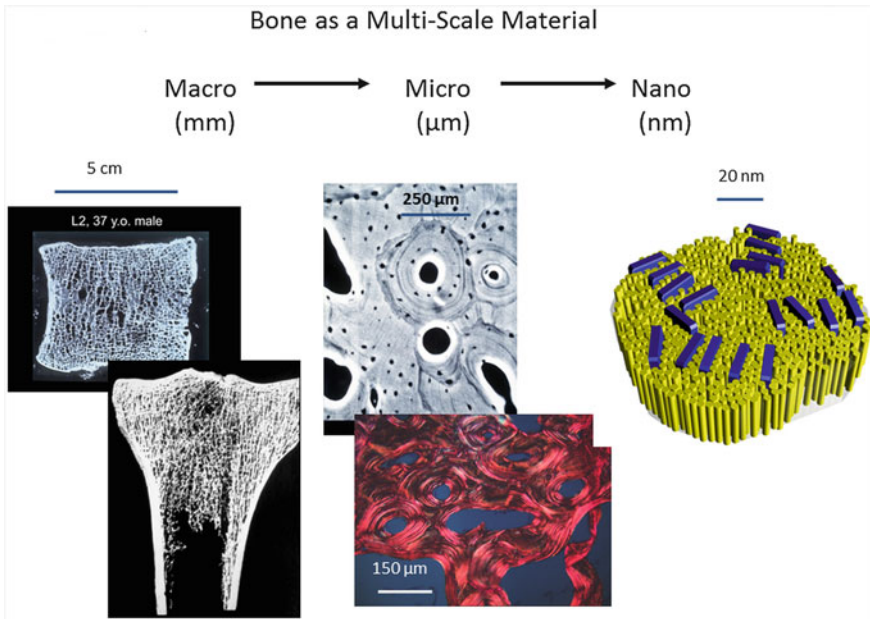


FIGURE 2.1. Bone is arranged hierarchically according to structure at size scales ranging from nanometers (collagen and mineral) to microns (trabecular plates) to milli- and centimeters (the bone as an organ). Structure at each level imparts a unique mechanical adaptation, and the combination across length scales forms a structure that is capable of lasting decades without failing.

20–50 % stiffer than the interlamellar regions, at least in longitudinal compression (Donnelly et al. 2006; Hengsberger et al. 2002; Xu et al. 2003).

At the microscopic level, secondary osteons (Haversian systems) about 200 μm in diameter can act as strong fibers embedded in a matrix composed of interstitial lamellae. Osteons are separated from the matrix by an interface called the cement line that has relatively low shear strength. The functional importance of cement lines lies in their ability to control fatigue and fracture processes, to absorb energy by stopping crack propagation, and to provide viscous damping in compact bone (Burr et al. 1988).

At the nanoscale, bone matrix is composed of cross-linked collagen fibers interspersed with mineral platelets, with noncollagenous proteins that control the assembly and size of these components. This organizational structure produces a composite material with mechanical properties superior to those of any of its constituents, including a remarkable ability to resist structural failure through the fatigue processes incurred by a lifetime of repetitive loading. As a first approximation bone's elastic properties are attributed to the mineral phase of the tissue and its inelastic properties to the organic (primarily collagen) matrix. This is clearly approximate and does not

consider either the interaction of the two phases or the roles that each plays in both the elastic and plastic deformation of bone tissue.

Box 2.1 What Is Hierarchy?

Bone and other connective tissues are made up from patterns of material that people categorize in a descending order of size as: anatomical, histological, molecular and atomic. All of the material patterns at all of the levels of size affect the mechanical properties. When authors write about bone mechanical properties they commonly say that ‘Bone is a hierarchical material.’ The notion that bone or other connective tissues are hierarchical is useful to help explain tissue mechanics since there is a large engineering literature about man-made hierarchical composites. In man-made hierarchical composites, the smaller design features are deliberately included to improve the functions of the larger parts. An example is the small ridges on the reinforcing steel bars used to strengthen the concrete of a bridge: The steel bars are smaller than the bridge and strengthen it and the ridges are smaller than the steel bars and strengthen the coupling between the steel and the concrete. As an inventor one could envision a roughening of the ridges to make them work better. In a hierarchy of size each smaller design feature improves the larger. A grave weakness of the hierarchy analogy for connective tissues is that the concept of design does not apply to a naturally evolved structure—we can’t know whether small features “help” the larger features to perform or if they are there for no reason at all. Evolution (selection) acts on the whole organism rather than on individual features of anatomy. As a result, heritable traits that are benign or even negative (when linked to an important positive trait) can become common in the population even if they don’t have a function. The human tendency to assign purpose and meaning to anatomical features helps in understanding some of the ways that connective tissues function, but our improved understanding does not mean we have identified a deliberate functional design plan (or hierarchy) inherent to the anatomy.

Recently developed techniques have provided the tools needed to study the finer fabric of bone structure. Quantitative image analysis systems have replaced traditional stereology and greatly improved the ability to discriminate small differences in trabecular architecture. Beginning in the 1980s, three-dimensional analysis using micro-computed tomography (μ CT) imaging made it possible to investigate full cancellous architecture using both advanced imaging theory and computational stress analyses. With improved spectroscopic imaging it is now possible to evaluate the structure and properties of bone at the molecular level using Fourier Transform Infrared Spectroscopy (FTIR) and Raman spectroscopy. Spectroscopic techniques are used to measure the mass of mineral and many of its characteristics (e.g. crystallinity, carbonate substitutions in the hydroxyapatite lattice) as well as the amount of collagen and its maturity through measurements of the relative content of younger divalent to more mature trivalent collagen cross-links.

Box 2.2 Fourier Transform Infrared Spectroscopy (FTIR) and Raman Spectroscopy

FTIR and Raman Spectroscopy detect molecular vibration to estimate some of the content and properties of the mineral and collagen of bone. FTIR measures infrared light absorption and reemission where Raman spectroscopy measures the energy change of photons that are scattered from the molecules (not absorbed and re-emitted). Both absorption and the Raman scattering depend on the molecular structure and give frequency dependent “peaks” in the raw data that are characteristic of tissue chemistry and structure. In both FTIR and Raman spectroscopy, it is necessary to decompose the signals using some form of Fourier Transform in order to quantify the measured signal.

There are several characteristic peaks associated with connective tissue composition. The Amide I band occurs from 1688 to 1623 cm^{-1} . Amide I results from vibration of a C=O peptide bond which, in connective tissues, occurs primarily in collagen. The Amide I band signals are, therefore, used to measure collagenous content and structure. For example, the ratio of mature (trivalent) to immature (divalent) collagen cross-links can be estimated from the ratio of peak intensities of the Amide I band at 1660 and 1690 cm^{-1} .

In bone or other calcified tissues the properties of the mineral can be evaluated in part by measuring asymmetric stretch of the P-O bond (ν_3/ν_1 , 1200–900 cm^{-1}), and by measuring in-plane bending (ν_4 PO4-3, 650–500 cm^{-1}). The ν_3/ν_1 PO4-3 phosphate peak can become saturated, indicating that a minimal amount of light is reaching the detector at those wavelengths, and some feel that the ν_4 PO4-3 band is a more accurate measure of mineral. Crystallinity can be determined as the ratio of the phosphate sub-bands at 1030 cm^{-1} and 1020 cm^{-1} . The intensities of the bands between wavelengths 905 and 825 cm^{-1} are determined by mineral crystallinity, structure and ionic substitutions within the apatite crystal. The ν_2 carbonate band has been found to be the best for identifying carbonate ions in hydroxyapatite because it is not obstructed by other bands, and provides a sharp peak. Overall mineralization of the tissue is represented as the ratio of phosphate to total protein (ν_4 PO4-3/Amide I), whereas carbonate substitutions are measured as the ratio of $\nu_2\text{CO}_3\text{-2}$ /Amide I. Both FTIR and Raman have become common and useful tools for measuring changes in bone composition in experimental animals in response to an intervention, treatment, or disease (Carden and Morris 2000; Ruppel et al. 2008; Gourion-Arsiquaud et al. 2009).

The mechanical properties of some of the individual components of bone tissue can now be measured directly. Nanoindentation (Rho et al. 1997) allows one to sample properties at specific locations and interfaces (cement lines, lamellae and the junctions between them) within the bone tissue. Energy dispersive X-ray diffraction techniques such as Small Angle X-ray Scattering (SAXS) and Wide Angle X-Ray Scattering (WAXS) (Gupta et al. 2006; Stock et al. 2011) can be used to measure mineral crystal thickness and spacing between the mineral plates, as well as lattice strains and the elastic modulus of the bone mineral itself. Atomic Force Microscopy (AFM) can now provide high resolution 3D images of surface features and image structure at the level of the collagen fibril (such as collagen spacing)

and the mineral platelets (Wallace et al. 2011). In a different mode, AFM can be used to make mechanical measurements of structure at a molecular level.

Evolution has had millions of years over which to find satisfactory solutions to the structural and materials engineering problems of bone. The ultrastructural, microstructural and even macroscopic organization of the constituents of bone interact in ways that reduce stress and allow the structure to sustain itself over a lifetime. Studies of these natural solutions to engineering problems provide insights into design criteria that can improve the way man-made materials function for particular applications (biomimetics).

2.2 The Shapes of Bones

While the shapes of our bones are quite variable, comparisons with those of other animals show that the architecture of a vertebrate skeleton is fairly conserved in nature and stereotypical across species. We usually can recognize a femur or a tibia, and we recognize that they have a characteristic positional relationship, regardless of what animal they came from. Yet within these broad architectural constraints, the skeletal morphology is quite variable, and we can imagine a rubber archetypical skeleton being stretched and molded to create such diverse skeletons as those of a flamingo, a bat, an elephant, a whale, or a human.

Not only is skeletal architecture conserved across species, but so too is the topology of individual bones within a species. Thus, a physical anthropologist can recognize an isolated bone, or even a fragment of a bone, as belonging to a human rather than one of the other primates. These constraints on architecture and topology give the impression that bones are determinate, static structures. On a scale of centimeters, this is approximately true. Yet the specific mechanical behavior of bones also relies on slight variations in the geometric distribution of bone, either within a cross section of compact bone, or by variations in trabecular architecture that correspond to the pattern of locomotion and the directions of the stresses and strains imposed by the animal and its environment. Differences in geometry (gross or macroscopic anatomy) can compensate for differences in the microanatomy (or nanoanatomy) that change the mechanical properties of the tissue. It can be readily appreciated that the strength of a whole bone depends on the material composing it, and its volume, cross-sectional area, three-dimensional shape, ratio of the smallest diameter to total length and other shape characteristics. These gross geometric properties of the bone are complemented by the material properties of the tissues themselves. Tissue material properties also are variable in space and time, responding to mechanical stimuli. The complex interrelationship between the physical properties of bone—those determined by the mineral-collagen composite—and the geometric properties, together determine the overall mechanical behavior of a whole bone as a structure.

2.3 Types of Bone and Tissue

Figure 2.2 is a sketch designed to show the general features found in most long bones, such as the humerus or metatarsal. This figure illustrates the relationships between many of the structural features of bones described in this chapter.

The nonmineralized spaces within a bone contain marrow, a tissue composed of blood vessels, nerves, and various types of cells. The chief function of the marrow is to generate the principal cells present in blood. The internal presence of marrow is a nearly universal feature of bones (the ossicles of the inner ear are an exception). The relationship of marrow and bone is both biologic and physical: they share common stem cells, and marrow never

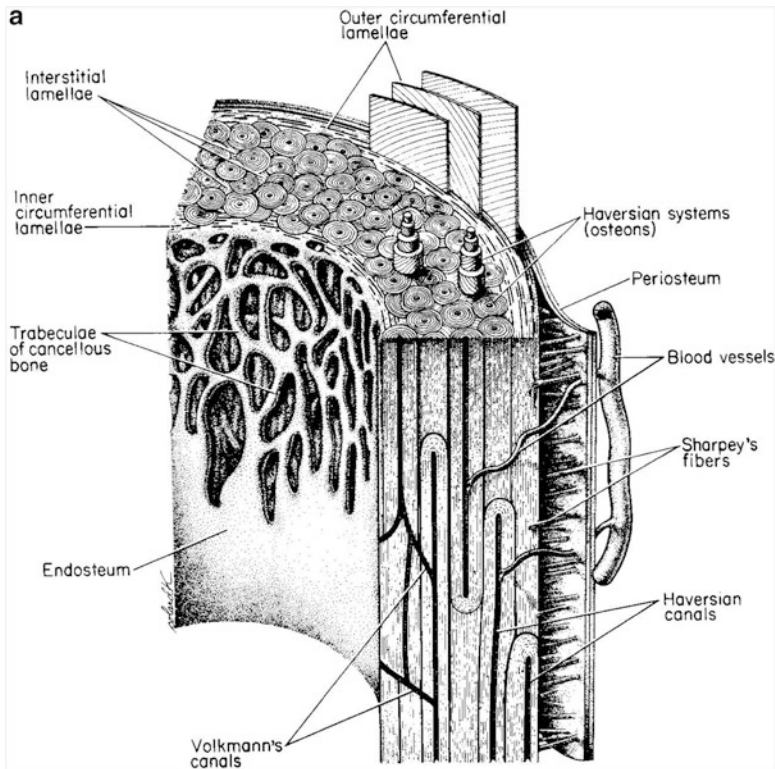


FIGURE 2.2. (a) Sketch of a hemi-section of bone showing important features three dimensionally of the cortical bone from a typical long bone. (b) Microradiograph (X-ray image of a thin cross section) showing compact bone. Haversian canals and resorption spaces are *black*; note variable mineralization of osteons. Outermost region (at top) contains well-mineralized primary bone. Field width, ~2 mm (Courtesy of Dr. Jenifer Jowsey).

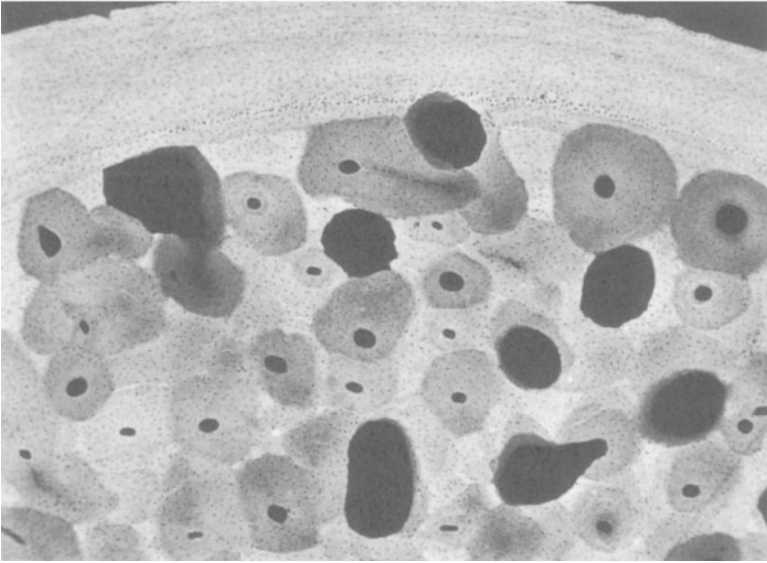
b

FIGURE 2.2. (continued)

exists outside of bone. Bone can be made to form outside the normal skeleton (e.g., in a muscle) by implanting osteogenic materials, but when this happens, a space containing marrow automatically forms in the nodule of extraskelatal bone. Moreover, marrow itself is a highly osteogenic material and can stimulate bone formation if placed in an extraskelatal location (Tavassoli and Yoffey 1983).

There are two types of marrow in bone, each with a different function. Bone is a blood forming organ, and red, or hematopoietic marrow, within the marrow cavity is abundant during rapid growth when there is need for an ever increasing volume of red and white blood cells. Red marrow begins to be replaced by yellow (fatty) marrow by the age of 13 as the growth rate slows down, with the red marrow in the diaphyses of long bones being completely replaced by yellow marrow by the middle of the 4th decade of life. The cells in yellow marrow (adipocytes) and bone forming cells (osteoblasts) arise from the same stem cell, creating what can appear to be an inverse relationship between the amount of fat and the amount of bone. In some pathologic conditions the inverse relationship is very strong. Although some red marrow is retained in the ends of the long bones, and in areas characterized by large amounts of cancellous bone (e.g. vertebral bodies, iliac crest and ribs), aging gradually converts more and more of this marrow to yellow or brown marrow.

2.3.1 Cancellous Versus Compact Bones

If you were to slice open a dozen bones of various sorts and wash away the marrow with a high-pressure stream of water, it would be clear that the remaining bone is of two distinct kinds, as determined by porosity (the volume fraction of soft tissues). While in principle the porosity of bone can vary continuously from 0 to 100 %, in fact most bone tissues are of either very low or very high porosity, with little bone of intermediate porosity. These two types of bone tissue are referred to as compact bone and cancellous bone, respectively (Fig. 2.3a).

Cancellous bone (also called trabecular or spongy bone) is porous bone found inside cuboidal bones (e.g. vertebrae, carpal bones), the flat bones (e.g. ilium, sternum), and the ends of long bones (Fig. 2.3b); its porosity is 75–95 %. The pores are interconnected and filled with marrow. The bone matrix is in the form of plates or rods called trabeculae, each about 100–150 μm thick. The arrangement of the trabeculae is variable. Sometimes they appear to be organized into orthogonal arrays; often, they are more randomly arranged. Cancellous bone tissue has a much lower density than cortical bone and allows the skeleton to build large but lightweight conical volumes of bone at the ends of long bones that carry the large loads of the joints with low stress at the wide end (the metaphysis) and higher stresses in the narrower shaft (the diaphysis). The density of the bone increases as the cone narrows, creating a means for the bone to funnel the high loads imposed at the joint to the stronger and more massive cortical bone comprising the middle portion of the bone (i.e. the diaphysis).

Compact bone is the dense bone found in shafts of long bones and forming a cortex or shell around vertebral bodies and other cuboidal and flat bones

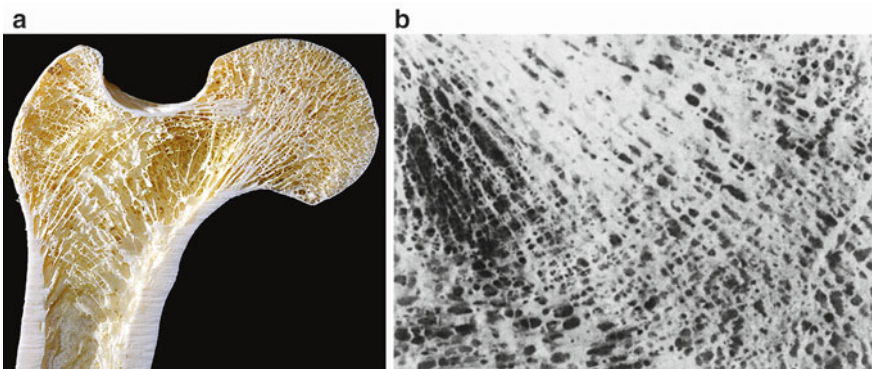


FIGURE 2.3. (a) A sagittal section of the proximal end of a human femur provides an appreciation for the organization of cancellous bone and its relationship to the dense cortical bone surrounding it. (b) Cancellous bone structure from the distal end of the femur at higher magnification shows more clearly the integration of trabecular rods and plates. Field width, ~ 1 cm.

(Fig. 2.3a, b). Hence, it is also called cortical bone. It is strong and provides both support and protection. Its porosity is 5–10 %, although this increases with age and with osteoporotic changes to the skeleton. Its pores consist of spaces categorized as follows.

Haversian canals are approximately aligned to the long axis of the bone, contain capillaries and nerve fibers, and are about 50 μm in diameter (about the diameter of a human hair). Haversian canals are named after an English physician, Clopton Havers (1691).

Volkman's canals are short, transverse canals connecting Haversian canals to each other and to the outside surfaces of the bone. These canals also contain blood vessels and nerve fibers. They are named after Richard von Volkman (1830–1889), a surgeon and early advocate of Lister's antiseptic surgical methods.

Resorption cavities are the temporary spaces created by osteoclasts in the initial stage of remodeling (see Chap. 3). Resorption cavities are about 200 μm in diameter.

It is important to remember that bone is a dynamic porous structure; its porosity may change as the result of a pathologic condition, with age, or in a normal adaptive response to a mechanical or physiologic stimulus. Cancellous bone may become more compact, or compact bone may become more porous. Such changes strongly affect bone's mechanical properties.

2.3.2 Lamellar Versus Woven Bone

Examining compact and cancellous bone at a still finer scale of resolution, it is evident that each may contain two major types of bone tissue.

Lamellar bone is slowly formed, highly organized bone consisting of parallel layers or lamellae comprising an anisotropic matrix of collagen fibers interdigitated with mineral plates. There may be a variety of patterns of lamellar arrangement at the level of the collagen fibril, and these likely each have important mechanical implications (Reisinger et al. 2011). Microscopically, however, the lamellae create optically bright circumferential bands of bone, each 3–7 μm thick, that give the appearance of tree rings, each separated by an optically darker layer about 1 μm thick that is mechanically less stiff than the thicker lamellae (Tai et al. 2007). The lamellae may be arranged around the endocortical or the periosteal circumference of the bone (primary or circumferential lamellae), or may be arranged concentrically around individual vascular channels (osteonal or concentric lamellae).

Two fundamentally different kinds of “plywood” architecture coexist within the general complexity of bone's lamellar structure (Giraud-Guille 1988). The first of these schemes corresponds to the classical view of lamellar structure: the collagen fibers are parallel in each lamella and change direction by 90° at the lamellar interface. That is, as a lamella is built up, layers of collagen are put down in one orientation. Then, when the next lamella is started, the orientation of collagen fibers suddenly changes so that they are

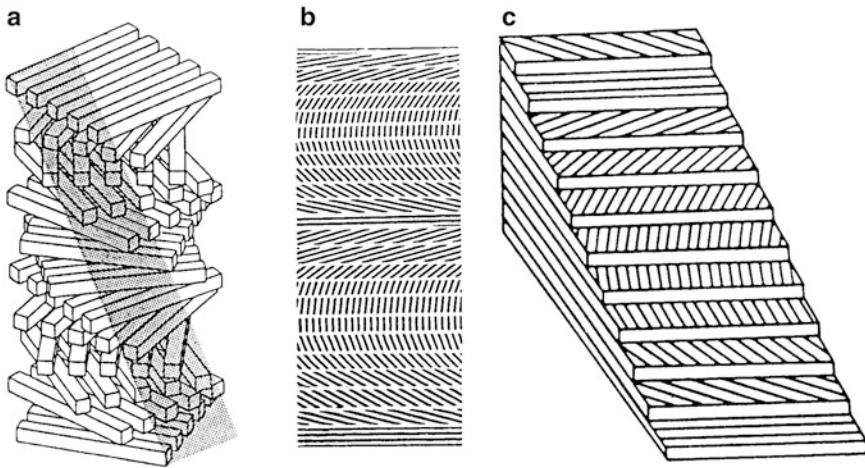


FIGURE 2.4. Diagram of helicoidal plywood structure. (a) Three-dimensional structure. (b) Visual effect of *arches* seen when an oblique section (as indicated by the *shaded plane* in (a)) is cut. (c) The *arches* seen in an oblique section are formed by the helicoidal lamellae (Neville 1984). With kind permission of Springer Science+Business Media.

laid down at right angles to the previous direction. In other regions of the bone, this orthogonal plywood-like structure is replaced by helicoidal plies (Fig. 2.4). In this scheme the collagen fibers continuously change their direction, so that in a sense there are no individual lamellae. However, the bone still shows a lamellar structure because as the orientation of the collagen fibers rotates through 180° cycles, the fiber orientation repeats itself, and layers appear when a histologic section is examined microscopically. A helicoidal arrangement of collagen fibers leads to a transversely isotropic structure at this level of organization as the elastic properties are similar in different directions (Reisinger et al. 2011). In her landmark paper, Giraud-Guille (1988) showed that both these architectures are present in human compact bone, but their relative distribution and interspecies variation have not been determined.

Both these forms of lamellar structure give rise to birefringence, the capacity of some fibrous structures to interact with polarized light. When a histologic section of bone is transilluminated with polarized light and viewed through a polarizing filter oriented perpendicular to the vibration plane of this incident light, the section appears dark (i.e., the observer sees a “dark field”) except where collagen fibers are parallel to the plane of the section. These collagen fibers rotate the light’s plane of polarization so it is no longer perpendicular to the viewing polarizing filter. Therefore, the light is not blocked and reaches the viewer’s eyes. Thus, in a bone section observed in a polarizing microscope, transversely oriented fibers are bright and longitudinally oriented fibers are dark. In the 1960s, Ascenzi and his co-workers

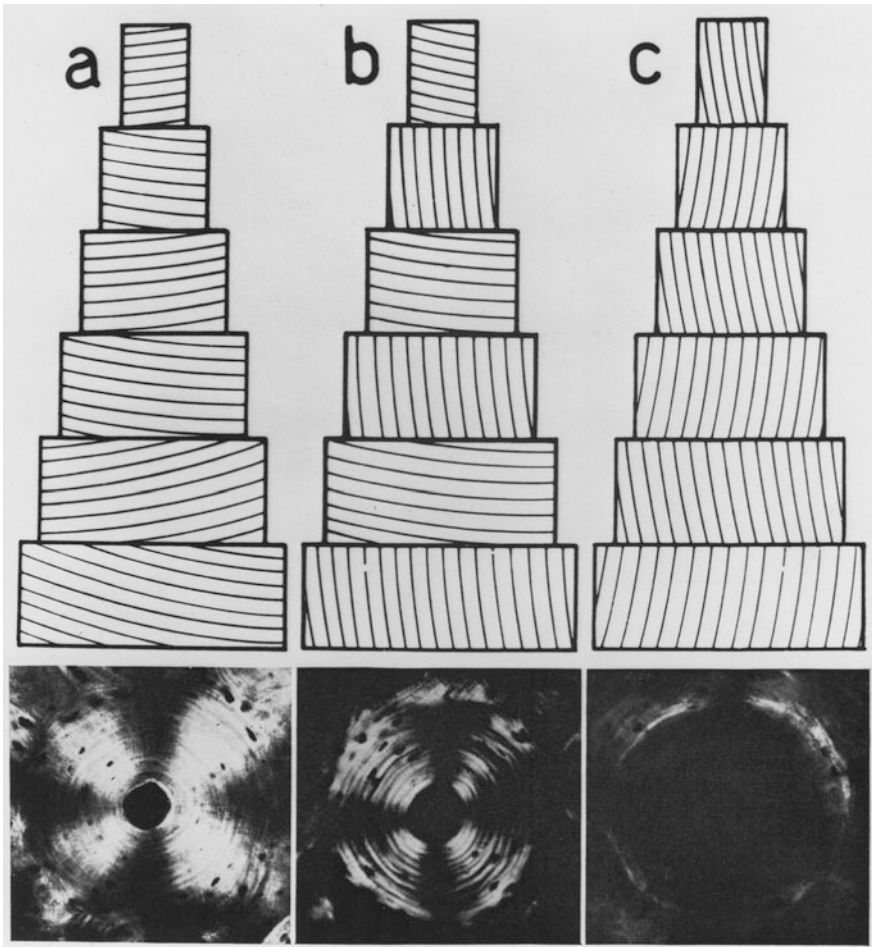


FIGURE 2.5. Three osteon types as defined by Ascenzi and Bonucci. Photomicrographs at *bottom* show appearance in plane-polarized light; diagrams *above* show hypothesized fiber arrangements in successive lamellae. (a) Type T or transverse (i.e., circumferentially wrapped) fiber orientation; (b) type A or alternating fiber orientations; (c) type L or longitudinal fiber orientation (Ascenzi and Bonucci 1968).

used this phenomenon to categorize osteons as having bright, alternating, or dark lamellae (corresponding to their collagen fibers exhibiting mostly circumferential, alternating, or longitudinal orientation) (Ascenzi and Bonucci 1967). These classifications are shown as parts a, b, and c, respectively, in Fig. 2.5. A dark “iron cross” pattern superimposed on the brighter osteons shows where the lamellae lie parallel to the polarization directions of the upper or lower polarizing filters.

Woven bone is a quickly formed, poorly organized tissue in which collagen fibers and mineral crystals are more or less randomly arranged. Woven

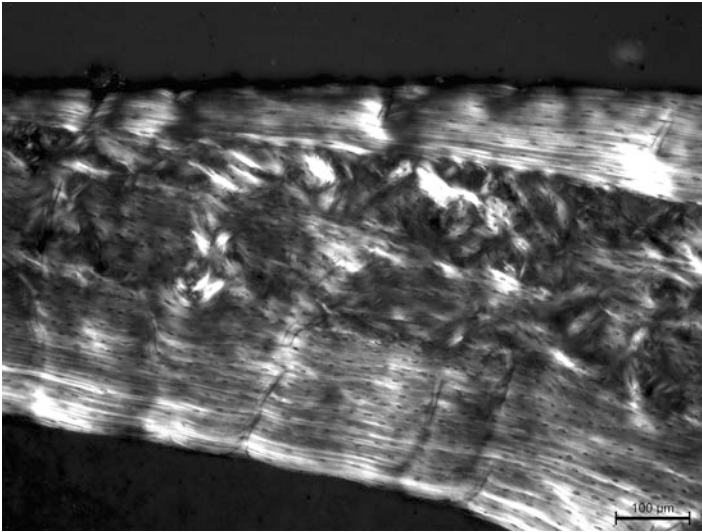


FIGURE 2.6. Polarized light image from a rat tibia showing disorganized fine-fibered bone in mid-cortex. This bone appears almost like woven bone, and is composed of very small collagen fibers, randomly arranged. Bone appearing more like primary lamellar bone is found along the periosteal and endocortical surfaces of the bone. Orig. mag. $\times 10$.

bone may become more highly mineralized than lamellar bone, which, mechanically speaking may help compensate for its lack of organization. Woven bone is common in a fracture callus and is formed early in the healing process to help stabilize the broken bone. In this case, woven bone can be considered calcified scar tissue that is rapidly formed for damage control and then remodeled and replaced over time with lamellar tissue reflective of the site's original organizational structure. The important generalization is that woven bone can be made more quickly than lamellar bone, but is weaker. Rats and other small animals also have fine-fibered bone in the cortices of their bones. In this bone the rather randomly arranged collagen fibers are smaller and closer together than in woven bone, so the tissue appears to be more organized (Fig. 2.6). It may be that the lamellar organization of bone can vary continuously; depending on factors such as how fast it is made.

2.3.3 *Primary and Secondary Bone*

Compact bone may be further characterized as primary or secondary bone.

Primary bone is tissue laid down *de novo* on an existing bone surface, such as the periosteal surface, during growth. It may be of two general types:

Circumferential lamellar bone, in which the lamellae are parallel to the bone surface (Figs. 2.7a–c). Blood vessels are incorporated into the lamellar

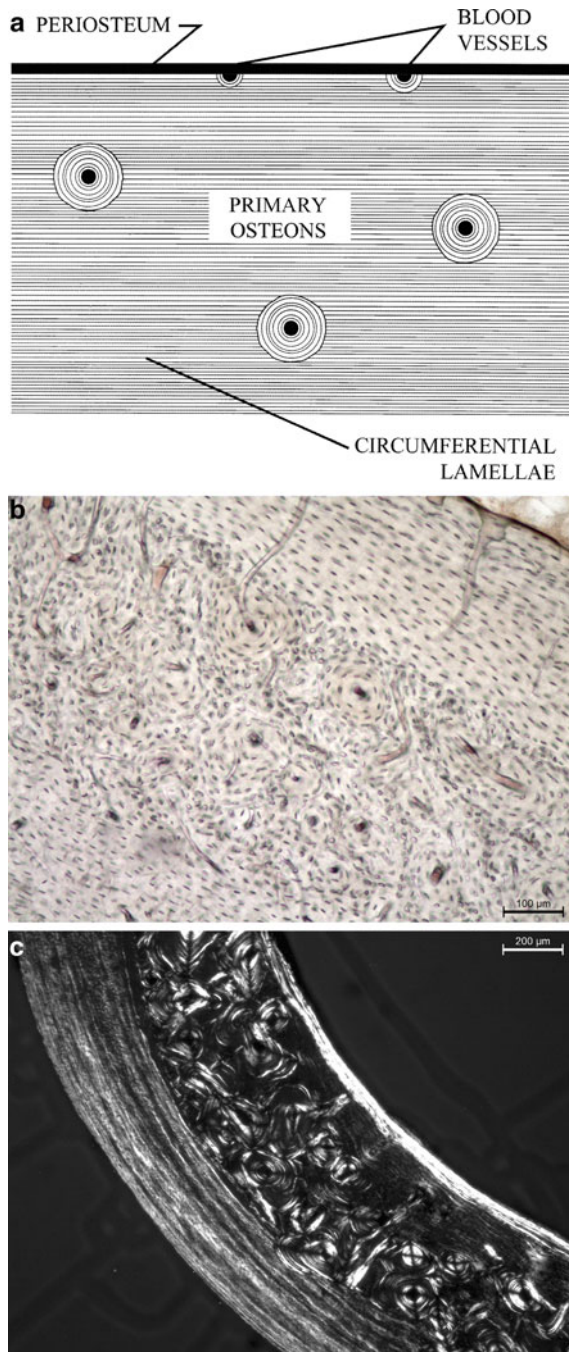


FIGURE 2.7. (a) Sketch of primary circumferential lamellar bone structure. Primary osteons form when blood vessels on the bone surface become incorporated into the new periosteal bone. They usually have several concentric lamellae, but their *cement line* is not scalloped. (b) Photomicrograph of a cross section of a rabbit humerus showing primary lamellar bone running in parallel sheets around the periosteal circumference of the bone. Osteonal bone—demonstrating both primary and

structure such that each is surrounded by several circular lamellae, forming a primary osteon with a primary Haversian canal at its center.

Plexiform bone occurs when the rate of formation is greatly increased by continually constructing a trabecular network on the surface and filling in the gaps. The result is a mixture of woven bone (the trabeculae) and lamellar bone (the filled-in spaces). Plexiform bone contains rectilinear residual vascular spaces, which often produce a “brick wall” appearance. Figure 2.8 contains brightfield and polarized light photomicrographs of this structure. It typically occurs in large, fast-growing animals like cows. Racehorses, which put enormous stress on some of their bones, have a similar kind of bone that may have exceptional fatigue resistance (Stover et al. 1992).

Secondary bone results from the resorption of existing bone and its more or less immediate replacement by new, lamellar bone. This process is known as remodeling (see Chap. 3). In compact bone, secondary tissue consists of cylindrical structures known as secondary osteons or Haversian systems (Fig. 2.9a). These are about 200 μm in diameter, although this varies with age in predictable ways (Currey 1964; Jowsey 1966; Martin et al. 1980), and consist of about 16–20 cylindrical concentric lamellae surrounding a central Haversian canal. Osteons are about 1–10 mm long and in the human femur they run at an angle of 11–17° with respect to the long axis of the bone (Cohen and Harris 1958). The boundary between the osteon and the surrounding bone is known as the cement line. It is composed largely of glycosaminoglycans and can be either poorly mineralized (Schaffler et al. 1987) or highly mineralized (Skedros et al. 2005). The cement line can act as a weak boundary that traps cracks to slow crack growth in cortical tissue (Burr et al. 1988). The weak interface between the osteon and the rest of the bone matrix increases tissue toughness and fatigue life. A cement line with low mineralization can be a viscous material that may relieve locally high shear stresses. The low shear stiffness reduces the ability of a cement line to transmit energy to a growing crack. Displacement (creep) of cement lines has been observed in human bone over long periods of torsional loading (Lakes and Saha 1979). In adult humans, most compact bone is entirely composed of secondary bone, which may include whole osteons and the remnants of older osteons that have been partially resorbed (interstitial bone). Remodeling of individual trabeculae within cancellous bone, which occurs frequently in both children and adults, rarely produces a complete osteon because the 200 μm diameter osteon is too large to fit within most trabecula of average diameter 100–150 μm . Instead, hemi-osteons are formed (Fig. 2.9b). One

FIGURE 2.7. (continued) secondary osteons can be found adjacent to the circumferential lamellar bone. (c) Bone from a squirrel monkey in polarized light clearly shows primary lamellar bone, surrounding remodeled secondary osteonal bone.

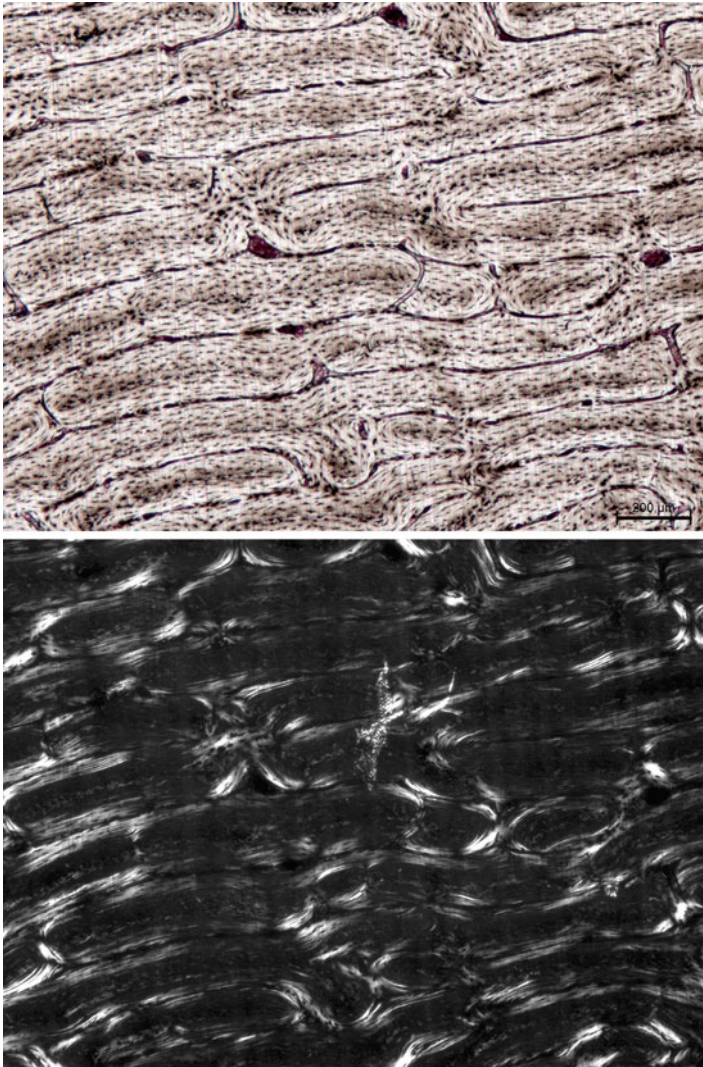


FIGURE 2.8. (a) Photomicrograph of plexiform bone. Vascular channels run between the blocks of bone. (b) In polarized light, one can appreciate the structure of the plexiform bone, with non-lamellar bone forming the core (which is *dark* in this image), surrounded by lighter primary lamellar bone. This bone is from a cow. Orig. mag. $\times 5$.

surface of the hemi-osteon borders the marrow cavity, rather than a Haversian canal, but the remaining surface is separated from older bone within the trabecula by a cement line just as whole osteons are separated from their surroundings in compact bone.

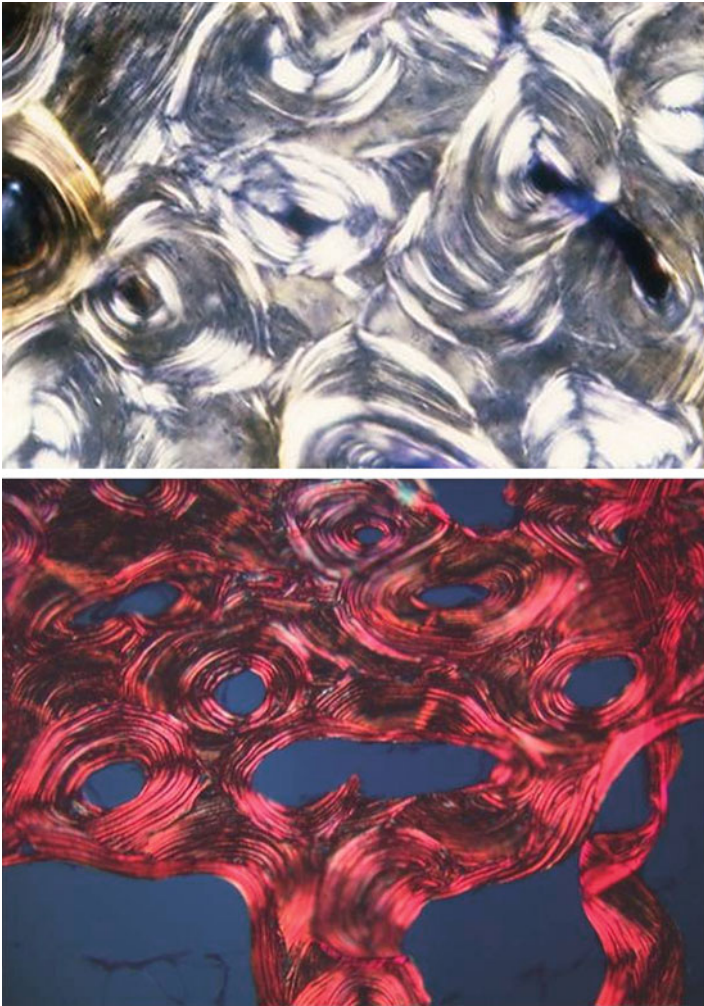


FIGURE 2.9. (a) Higher magnification image in polarized light of remodeled secondary osteonal bone, taken from a chimpanzee femur. (b) In trabecular bone, new packets of bone are often formed from the surface and do not create an entire osteonal structure. The “half osteons” are termed hemi-osteons. One can also see complete secondary osteons within the individual trabeculae.

2.4 The Collagen-Mineral Composition of Bone

At the ultrastructural level, bone is composed of organic and mineralized components, mainly consisting of a matrix of cross-linked type I collagen mineralized with nanocrystalline, carbonated apatite. About 65 % of bone by weight is mineral, and 20–25 % is collagen, mostly Type I. Although bone

is a hard and stiff tissue, water composes about 10 % of its structure and is essential to normal mechanical properties. Non-collagenous proteins make up the remainder, only about 1–2 % of the bone, but have vital regulatory and mechanical functions that belie their relatively minor contribution to bone mass (Table 2.1).

Collagen is a structural protein (molecule), widely distributed throughout the animal kingdom that can spontaneously organize itself into strong fibers. Several dozen types of collagen have now been identified, although only a few of these are found in bone. The predominant fibrillar collagen in bone is Type I, which is also found in tendons, ligaments and skin. Type I collagen is arranged in a triple helix that aggregates into a periodic arrangement of parallel molecules spaced in a quarter-staggered array with separations of ~67 nm between their ends to form collagen fibrils of about 150 nm (Fig. 2.10). The mean diameter of collagen fibrils in osteoporotic bone is

TABLE 2.1. Composition of whole bone by weight dogs

Component	Site or specific molecule	Volume (%)
Water, 10 %	Bound	60–80 ^a
	Free	20–40 ^a
Organic matrix, 25 %	Collagen	90
	Non-collagenous proteins	10
Apatite mineral, 65 %	In gaps between collagen ends	28
	Intrafibrillar	58
	Interfibrillar	14

^aThe precise ratio of bound and unbound water depends on a number of factors, including age

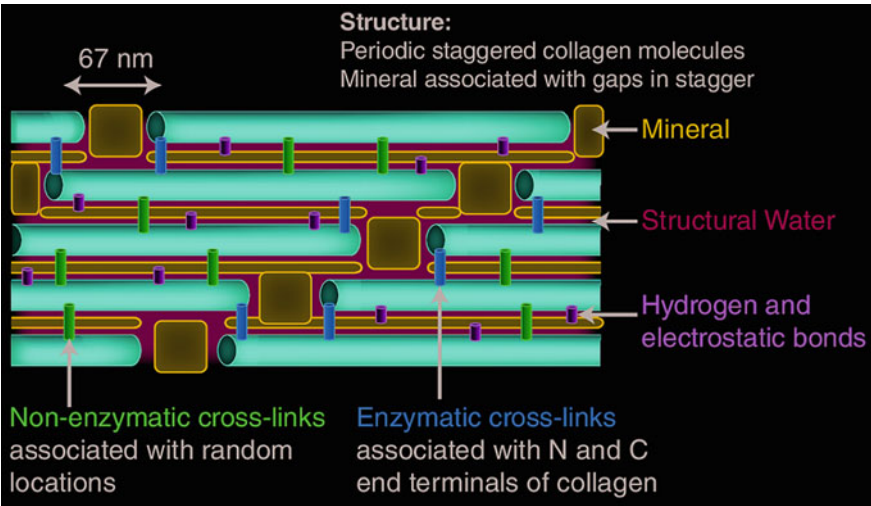


FIGURE 2.10. The structure of a collagen fiber. Note the arrangement of the mineral and water within the fibers.

less than in healthy bone, and this may increase bone's fragility. Variations in spacing may also accompany certain diseases affecting collagen formation, such as osteogenesis imperfect (Wallace et al. 2011).

Chemical bonds called cross-links form bridges between the collagen molecules, stiffening the fibrils and fibers. Immature divalent cross-links are converted fairly rapidly (within weeks of bone deposition) to mature, or irreducible, trivalent cross-links by enzymatic processes. Cross-links created by non-enzymatic processes can also be formed. Non-enzymatic reactions between sugars and proteins are common. After a series of Maillard reactions the products are called advanced glycation endproducts (AGEs). Some of the AGEs create collagen cross-links that stiffen collagen fibers in a manner similar to normal enzymatic cross-links.

Small amounts of Types III and V collagen can be found in the peri-lacunar region of the osteocytes, and possibly contribute to the organization of the extracellular matrix and the construction of the Type I collagen fiber itself (Olsen 1995).

The amount and maturity of the Type I collagen in bone, and the manner in which it is cross-linked, can have a profound effect on strength, stiffness and energy to fracture of bone tissue. Lower collagen content is a good predictor of reduced energy to fracture (Ding et al. 1997). Increased density of mature, trivalent pyridinoline and deoxypyridinoline cross-links increases the compressive strength and stiffness of bone (Bailey et al. 1992; Banse et al. 2002; Lees et al. 1990; Oxlund et al. 1996), but increases in these natural cross-links have not been demonstrated to change bone's fracture energy or ductility (Zioupos et al. 1999; Hernandez et al. 2005; Keaveny et al. 1994; Wang et al. 2002). There is a smaller proportion of immature divalent cross-links in newly formed bone from individuals with osteoporosis (Bailey and Knott 1999), which could increase bone fragility and decrease bending strength.

Non-enzymatic Maillard reactions can form AGEs that cross-link the collagen and reduce bone toughness by reducing ultimate strain and by post-yield deformation. They reduce collagen fibril diameter by dehydration, which could also contribute to increased fragility. AGEs naturally accumulate in bone with age. Bone from individuals over 70 years old has three times more non-enzymatic cross-links than bone from those less than 50 years. AGEs are found in abundance in diabetic bone. Bone from these individuals is more fragile and more likely to break, even in the face of normal or elevated bone mineral density.

Mineral in bone consists almost entirely of highly-substituted, poorly crystalline apatite mineral $(\text{Ca}_{10}\text{PO}_4)_6(\text{OH})_2$ that nucleates within the spaces between the ends of the collagen fibrils, known as hole-zones, as well as in ~35 nm gap zones known as pores that run longitudinally between the fibrils. The individual crystals are rods or plates with hexagonal symmetry, measuring about $50 \times 50 \times 400 \text{ \AA}$ ($1 \text{ }\mu\text{m} = 10,000 \text{ \AA}$). As bone tissue matures, the

mineral crystals grow, become more plate-like and orient themselves parallel to one another and to the collagen fibrils. Therefore the measured average size of mineral crystals is highly dependent on tissue age. In both cancellous and cortical bone, the c-axis, or long axis, of the mineral and collagen aligns with the longitudinal axis of the structure, i.e. the trabecula or the osteon, respectively. When bone is fully mineralized, the crystals are sufficiently close to one another that removal of the organic matrix by ashing or by chemical means leaves behind a solid piece of material rather than a pile of nanocrystalline “dust.” It is unknown whether the crystals actually merge together into a single continuous phase such as in a stony material like granite or are cemented together like sandstone, but their mechanical behavior in fully mineralized tissue seems to be similar to a continuous polycrystalline solid.

The hydroxyapatite found in bone contains many ionic substitutions (e.g. strontium, lead, carbonate, fluoride). These substitutions are governed by the composition of the body fluids and in turn affect the solubility and material behavior of the bone mineral. As bone matures, mineral crystal size and perfection increases due to changes in ion substitutions and mineral stoichiometry (Miller et al. 2001). These can substantially alter the hardness and brittleness of bone in either a positive or a negative manner. Substituting carbonate for the phosphate in hydroxyapatite makes the crystal more soluble, reduces the stability of the mineral lattice and increases fragility. Different cation and anion substitutions in the apatite lattice, can in some cases alter the mechanical properties of bone and in others affect the activity of osteoblasts and osteoclasts. Although not a strict substitution due to size constraints, citrate binding on the surface of the apatite crystal lattice prevents further crystal growth, leading to thinner nanocrystals that may render the overall structure more resistant to crack propagation (Hu et al. 2011).

It is clear that there is a positive correlation between tissue mineralization and stiffness, but also an inverse relationship between these and toughness (Currey 2004). As the mineral in bone increases, the tissue becomes stiffer, stronger and more brittle. The stiffer tissue deforms less under load and when the bone begins to fail, cracks will propagate with less energy than in less mineralized bone. Thus, a higher level of mineralization makes it more difficult to start a crack in the bone (it is stronger) but once a crack starts it is more likely to propagate and break the bone. Lower mineralization may therefore be beneficial, but only to a point; under-mineralized bones, such as occur in rickets, are weak and readily break. Women who have osteoporotic fractures tend to have bones that are either in the highest or the lowest quartile of the mineralization spectrum (Ciarelli et al. 2003).

There are many non-collagenous proteins (NCPs) and polysaccharides (sugar polymers) that have different functions in regulating the assembly and mineralization of the bone matrix. Non-collagenous molecules can be divided into several large groups that include the proteoglycans (versican, heparin sulfate, biglycan, and decorin); the glycoproteins (alkaline

phosphatase, thrombospondin 1 and 2, fibronectin, vitronectin); the SIBLING proteins (osteopontin, bone sialoprotein, Dmp-1, MEPE); and several other ubiquitous and very important proteins, including osteocalcin and osteonectin. The proteoglycans have variable spatial locations with versican expressed highly in osteoid and prevalent during skeletal development. Decorin and biglycan regulate the construction of collagen fibers, and their mineralization, and act in the differentiation of bone forming cells (osteoblasts). The glycoproteins are active in bone formation and mineralization, but are also important during development in the early stages of bone formation and cell proliferation. The SIBLING family of proteins regulate bone mineralization, and several of them (Dmp-1, MEPE) are expressed highly by osteocytes. Osteocalcin and osteonectin (also called SPARC) modify mineral deposition. Osteocalcin is a marker of bone formation (or turnover), whereas osteonectin plays a role in osteoblast proliferation and helps to regulate growth factors in bone.

Box 2.3 The Interaction of Collagen and Mineral Plays a Role in Mechanical Behavior

The interaction of collagen and mineral adds another level of complexity to the mechanical properties of bone. Although collagen supports most of the matrix strain in tension (Gupta et al. 2006), there is a coupled deformation between collagen and mineral. Tensile strain is distributed through an elongation of collagen fibrils and shear deformation of the interfibrillar matrix (Gupta et al. 2006). This protects the hydroxyapatite, which is by itself brittle, and doubles its load-carrying capacity (Gupta et al. 2006) while evenly distributing the forces and reducing stress throughout the matrix (Fratzl et al. 1994; Gao et al. 2003). Smaller collagen fibers oriented around osteocytes alter local strains at the nano- and micrometer level, diverting cracks away from living cells, and preserving them (Ascenzi et al. 2008). The orientation of mineral crystallites also changes the strain environment at the nano-scale (Giri et al. 2008). Thus, the mechanical properties of the collagen and mineral composite, and the viability of bone cells, are enhanced through coupling and load transfer around cellular lacunae and other discontinuities in bone.

Some of the water in the calcified bone matrix is free, and some is bound to other molecules. This water increases ductility and toughness of the bone by allowing sliding between the collagen fibrils and the mineral components (Nyman et al. 2008; Buehler 2007; Fritsch et al. 2009). The mineralization of osteoid (the organic portion of extracellular bone) displaces part of its water, which increases stiffness but decreases the ductility of the tissue. Table 2.1 gives the approximate composition of bone tissue by weight in dogs, measured in primary or secondary lamellar bone several months after formation.

2.4.1 *Quantitative Analyses of Bone Composition*

We have seen that bone contains three primary constituents: mineral, an organic matrix that is largely collagen, and water. Clearly, these three substances have distinctive physical properties that contribute to the overall mechanical behavior of the tissues they comprise. Architectural organization and substrate interactions are also important contributors to skeletal performance but the relative masses of bone's three primary constituents are a good indicator of bone health and enable a reasonable first approximation of its intrinsic strength and stiffness.

Simple gravimetric analysis is a time-honored method of determining the mineral, organic and water content of extracted bones. In this method bone specimens are precisely weighed before and after drying in an oven at 100 °C to determine wet weight, dry weight and percent water. The specimen is then placed in a muffle furnace and heated to over 600 °C which volatilizes all organic constituents and removes any remaining bound water leaving the ash mass which is weighed to determine the mineral component of the tissue. The organic content (predominately collagen) and ash content (mineral) are then determined as a percentage of the original wet weight. Tissue mineralization or ash fraction is calculated as the percentage of ash weight to dry weight. The ash fraction is independent of bone porosity and is an excellent measure of tissue mineralization, which varies by species, skeletal site, age and disease. Collagen content can be measured as the difference between total dry weight and ash weight, divided by dry weight. The gravimetric method can be extended to measure carbonate content by heating specimens to above 900 °C, the temperature at which incorporated carbonate decomposes to CO₂.

Gravimetric analysis yields good quantitative information about the major constituents of bone but it cannot measure the quality of collagen or mineral in the sample, nor can it provide architectural information related to bone density and porosity. The average weight density of the mineralized matrix (ρ_{bone}) can be determined in the laboratory using Archimedes law of buoyancy using a fluid of known weight density ($\rho_{\text{fluid}} = 1 \text{ g/cm}^3$ for distilled water; Eq. 2.1) provided that all marrow elements and fat are removed beforehand.

$$\rho_{\text{bone}} = \rho_{\text{fluid}} (\text{weight in air} / (\text{weight in air} - \text{submerged weight})) \quad (2.1)$$

The density of either cortical or trabecular bone specimens can be measured using this approach, but completely removing marrow elements and fat from highly porous trabecular samples is difficult. Apparent density, the weight of the sample divided by its total overall volume (typically a machined cube or cylinder), including the pores and spaces within, is a far more common measure for trabecular samples. Once the apparent density

is known the porosity (P) of the sample can be determined using the following equation:

$$P = 1 - (\rho_{\text{apparent}}/\rho_{\text{bone}}) \quad (2.2)$$

where ρ_{apparent} is the apparent density of the sample and ρ_{bone} is the density of the bone tissue in the sample as determined directly using Archimedes principle or another technique. A generally accepted value of 2.0 g/ml is frequently substituted for direct measurements of ρ_{bone} , thereby allowing for easily obtained estimates of porosity.

The porosity (P) of a bone sample is simply the percentage of total sample volume not occupied by bone tissue. Conversely, the bone volume fraction (BV/TV) of a sample is the percentage of volume occupied by bone matrix. The material comprising the bone volume fraction is often referred to as bone tissue to distinguish it from the term “bone,” which usually refers to a region large enough to contain Haversian canals, marrow spaces, or other voids. Cancellous bone has been somewhat arbitrarily defined as having porosity greater than 0.5, with free space largely occupied by marrow elements; in cortical bone, porosity is less than 0.5 and space is defined by Haversian and Volkmann’s canals and resorption cavities created by osteoclasts. For practical reasons, the smallest voids in bone, consisting of osteocyte lacunae and canaliculi, have traditionally been included in the bone volume fraction rather than the porosity fraction.

The foregoing describes weight-based measures of bone volume, density and porosity. Another traditional method of analysis employs histologic sections and the principles of stereology. Before the advent of automated procedures this approach typically entailed overlaying a two-dimensional histologic image of a bone section with a series of lines or an array of points and measuring the length of line passing through, or the number of points landing on, bone or void space. When enough sections are sampled the approach yields accurate estimates of bone composition because the ratio of bone tissue to void space remains the same whether assessed by counting points (0D), measuring line lengths (1D), measuring areas (2D) or measuring volumes (3D). Bone volume, bone density and a host of other architectural measures such as bone surface area, trabecular number, thickness and trabecular spacing are now easily made using newer 3D imaging modalities and automated procedures.

2.4.2 Three Dimensional Imaging of Bone Using High Resolution Computed Tomography (microCT)

A bone can change shape both from the outside and the inside as a result of age, diet, disease, fracture and from many other causes. Today, there are many ways to image the shape of a bone in a person or animal without

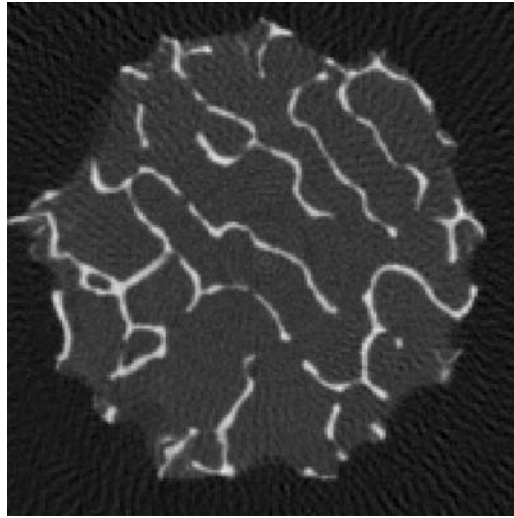


FIGURE 2.11. Micro-computed tomography “slice” from a specimen of human vertebral cancellous bone (50 μm resolution). The *lightest colors* are mineralized bone tissue, the *gray* is bone marrow and the *darkest* is air.

harming them. X-Ray Computed tomography (CT or “Cat Scan”) is often the best method for seeing bones inside the body. The machines used in hospitals ordinarily have the resolution to image clearly objects down to about a millimeter, which is outstanding for many medical purposes but insufficient to see trabeculae, osteons, Volkmann’s canals (blood vessels) and other microanatomy in bones that contribute to its health and strength.

There now are many commercially available CT scanners for small specimens that have better than one millimeter resolution. The highest resolution machines are often called “X-Ray Microcomputed Tomography” scanners or “MicroCT” scanners for short. For resolution on the order of 35–50 μm there are machines that can be used to image the bones of small animals such as mice or rats while they are alive. For finer resolutions on the order of 1–10 μm the required x-ray doses are dangerous to life making high resolution scanning usually restricted to dead or nonliving specimens.

The three dimensional (3D) images obtained from microCT (Fig. 2.11) are used to understand bone microarchitecture in many ways. One method is to use computer assisted image processing to measure bone density or connectivity and to count trabeculae or other features directly from the image (Hildebrand et al. 1999). In the past, measurement of the interior of a bone specimen was done by physically sectioning the bone, mounting the thin histologic bone specimens onto microscope slides and using stereometric analyses (Parfitt et al. 1987; Russ and DeHoff 2000) to estimate 3D

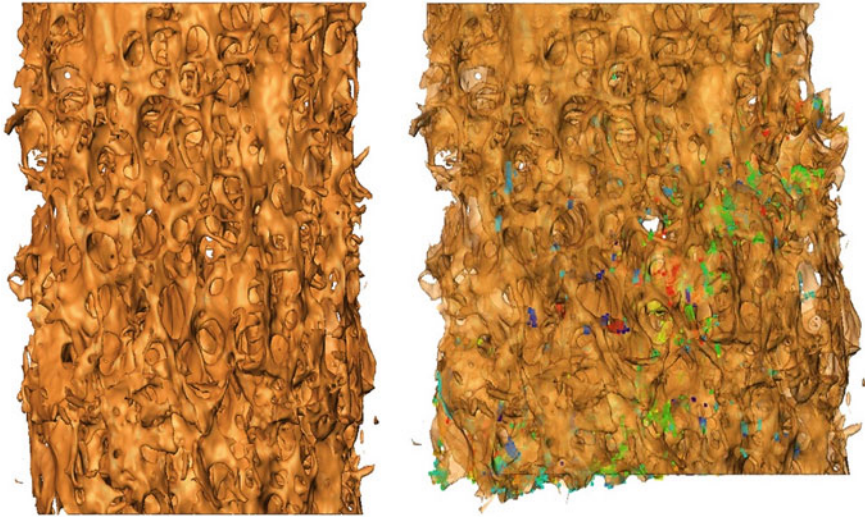


FIGURE 2.12. Three-dimensional rendering of a finite element model created from a microCT image of human vertebral cancellous bone before (*left*) and after (*right*) compression to failure. The *colors* in the *right image* are predicted damage.

information from the 2D microscopic images. MicroCT scanning and image analysis has eliminated much of the need to use bone histology for measuring the interior microanatomy of bone, although histology is still necessary to image bone cells since microCT cannot resolve them adequately. Histology is, of course, still necessary to measure cell-level information such as gene expression and protein distribution, or to quantify dynamic bone remodeling.

A powerful tool for understanding how bone microstructure contributes to tissue strength and stiffness is to convert the 3D microCT image into a finite element stress analysis model (Fig. 2.12). Finite element models of bone tissue built in this way are much better at predicting the mechanical properties of bone tissue than are simpler models that don't include microanatomy (Hou et al. 1998; Homminga et al. 2002; Zauel et al. 2006). Prior to the invention of high resolution microCT finite element modeling of trabecular bone in 3D was extraordinarily time consuming or impossible. Now, however, stress analysis is a normal part of research into what makes bones weak or strong. With continued technological development, higher resolution CT scanners may become available to hospital patients and some of the insights from microCT could become useful to predicting the risk of bone fractures for individual patients.

2.5 Bone Cells

So far we have discussed the extracellular matrix, that part of bone which is outside of living cells and can be mineralized. Formation and maintenance of the matrix is carried out by the cells that make up a small but critical percentage of the bone volume. There are four types of bone cells. These cells fall into two categories: those that resorb bone (osteoclasts), and those that form or have formed bone (osteoblasts, osteocytes, lining cells). The “resorbers” are closely related to macrophages, cells that migrate throughout all tissues of the body to remove debris and pathologic material. The “formers” are closely related to cells like fibroblasts, which produce structural molecules in other tissues. Despite their different lineages, osteoclasts and osteoblasts work together at discrete sites during bone remodeling by forming coordinated assemblies of cells called Bone Multicellular Units, or BMUs.

Osteoclasts are the cells that resorb bone (Fig. 2.13). They are multinucleated cells (with 4–20 nuclei) that are derived from the monocyte/macrophage lineage under the stimulation of Receptor Activator of NF κ B Ligand (RANKL) and Macrophage-Colony Stimulating Factor (M-CSF). The molecule RANKL (RANK ligand) that binds to the RANK receptor is produced by osteoblasts, so it can be seen that to some degree osteoclast differentiation and function is regulated by osteoblasts or osteoblast progenitors. Resorption occurs along a highly invaginated brush or ruffled border of the

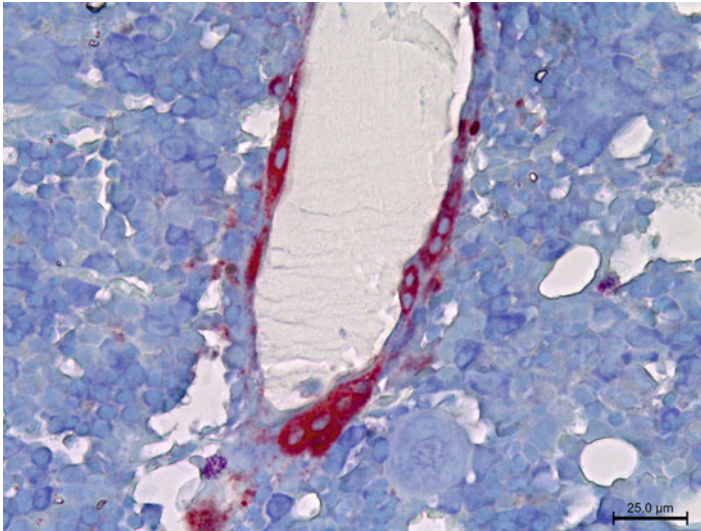


FIGURE 2.13. TRAP (tartrate-resistant acid phosphatase)-stained section of bone showing multi-nucleated osteoclasts (*red*) on the surface of an osteonal remodeling front.

cell, which is sealed to the bone surface by a peripheral clear zone. Osteoclasts erode their way through bone at a rate of tens of micrometers per day by first demineralizing the adjacent bone through the activity of a proton pump (H^+) and chloride ions which leads to low pH and acidification of the bone mineral. This exposes the collagen which is then decomposed by the enzymes cathepsin K and matrix metalloproteinases. These destructive chemicals are manufactured by the cell and transported to the ruffled border in portable intracellular chambers called secretory vesicles. Failure of the ruffled border to form properly, or failure of the polarization of the intracellular actin to form an actin ring around the clear zone, will prevent the osteoclast from resorbing bone. Once resorption is complete, osteoclasts die by nuclear disintegration (apoptosis) and disappear.

Osteoblasts are mononuclear, cuboidal cells that produce osteoid, the organic portion (Type I collagen and non-collagenous proteins) of the bone matrix, at a rate of about $1\text{ }\mu\text{m/day}$ (Fig. 2.14a).

This is called the bone apposition rate. Osteoblasts develop from mesenchymal precursor cells from the bone marrow that differentiate under the early influence of bone morphogenetic proteins, Runx2 and osterix, which regulate osteoblast function by coordinating several other signaling pathways that include Wnt and BMPs. Full differentiation of osteoblasts also requires some level of mechanical stress, one reason that bone formation is impaired in cases of disuse or during spaceflight. Osteoblasts contain large

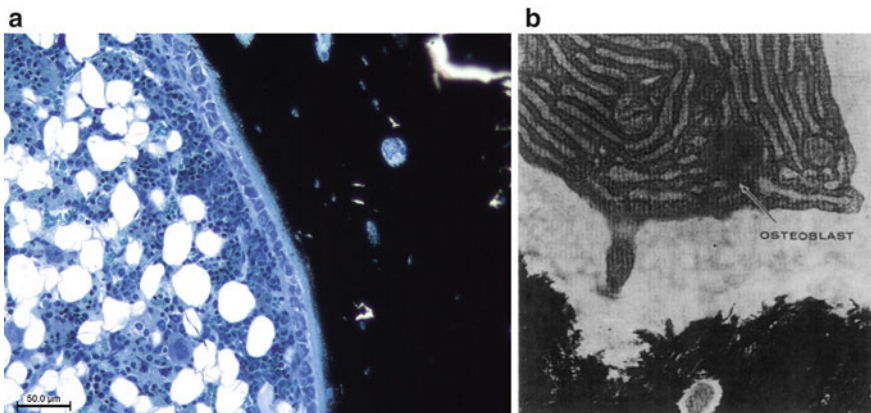


FIGURE 2.14. (a) Bone section showing cuboidal osteoblasts depositing osteoid (blue) on the surface of mineralized bone (black). Marrow cells (blue) and adipocytes (white) can be seen within the marrow cavity. Stained with McNeal's tetrachrome. (b) Electron photomicrograph of osteoblast forming bone. Dark material at bottom is mineralized bone. The lighter, interposed material is osteoid, production of which necessitates the extensive rough endoplasmic reticulum in the cell. A portion of a process protrudes from the cell; another is seen in cross section in the bone matrix (Reddi and Anderson 1976).

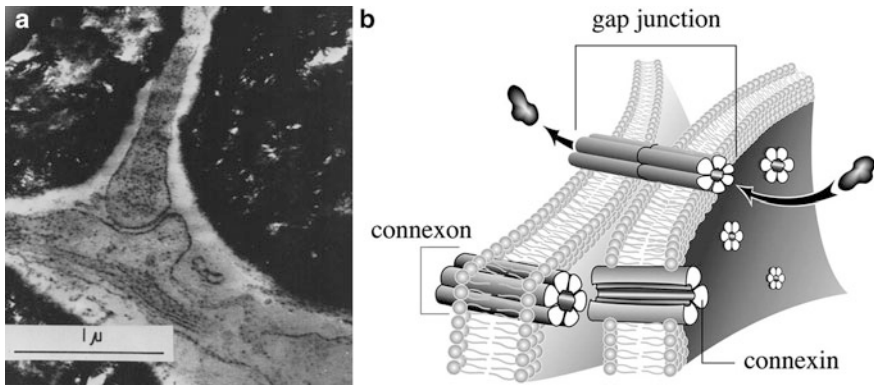


FIGURE 2.15. (a) Electron photomicrograph shows junction between two osteocyte processes (Cooper et al. 1966). (b) Osteocytes and osteoblasts communicate through gap junctions that are spanned by bundles of proteins called connexons, that are made up of individual connexins. Connexin-43 is one of the most important for bone, and regulates bone adaptation to loads and to disuse.

nuclei and Golgi apparatus, and have a large endoplasmic reticulum, where pro-collagen and the other organic matrix molecules are produced (Fig. 2.14b). The boundary between osteoid and calcified bone is called the mineralization or calcification front. Osteoblasts express osteocalcin and alkaline phosphatase which can be measured in blood serum. The concentrations of these molecules are indicative of the rate of bone formation and mineralization. Osteoblasts have extensive communications with other cells on bone surfaces (lining cells), within the bone matrix (osteocytes) and in the bone marrow through cellular processes that connect at gap junctions with processes from neighboring cells (Fig. 2.15). Gap junctions are inter-cellular channels composed of adjoining hemi-channels, each spanning the lipid bilayer of their respective cell and each composed of six identical protein molecules. The most abundant gap junction molecule in osteoblasts is connexin 43. Absence of this communication protein causes dysregulation of osteoblast differentiation and function.

Osteocytes, which make up 90–95 % of all the cells in bone, are former osteoblasts that have become buried in the bone which they and their neighbors have made. Osteocytes sit in cavities called lacunae and communicate with each other and with osteoblasts via processes passing through narrow tunnels called canaliculi (“little canals”) (Fig. 2.16). Processes from adjoining cells are connected by gap junctions, implying that networks of cells buried in bone tissue, resident on bone surfaces or close by in the marrow communicate and exchange substances to a significant degree. There are about 15,000 lacunae per cubic millimeter (mm^3) of bone (Mullender et al. 1996), but because of their small size, the lacunae and canaliculi occupy only about 1 %

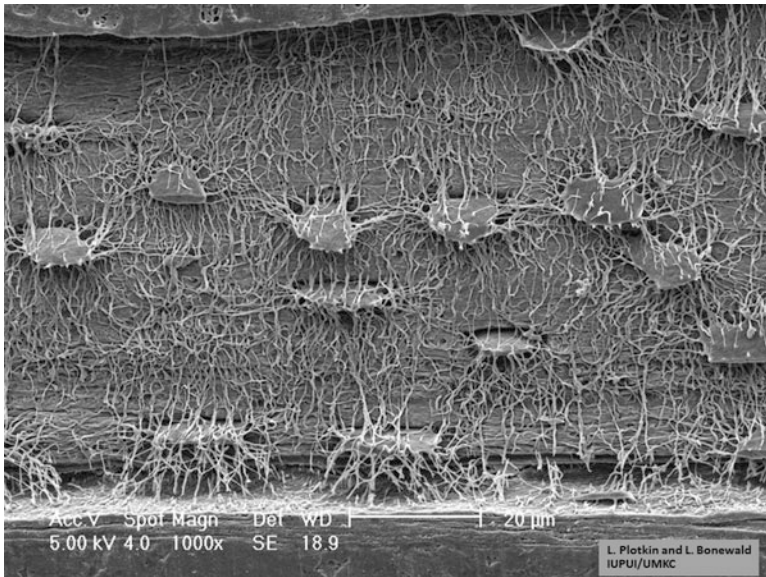


FIGURE 2.16. Photomicrograph shows the network of osteocyte canalicular processes from an acid-etched preparation taken from murine bone (Courtesy of Lilian Plotkin, Indiana University).

of the bone tissue volume. Their surface area, on the other hand, is huge. Johnson (Johnson 1966) estimated the total skeletal surface area of canaliculi in an adult male skeleton to be $1200 \mu\text{m}^2$, compared to $3.2 \mu\text{m}^2$ for Haversian and Volkmann's canals and $9 \mu\text{m}^2$ for cancellous bone surfaces. This intimate contact with virtually every nook and cranny of the skeleton is one reason that osteocytes are thought to be important both in transporting mineral into and out of bone, and in the transduction of mechanical signals (mechanotransduction).

It is now generally appreciated that, because osteocytes form a functional syncytium (a group of cells in which the cytoplasm of one cell is continuous with that of adjoining cells) that connects cells on the bone surface with those in bone marrow and with those in blood vessels, they form a mechanosensory network that is able to detect and respond to changes in mechanical stress. It is likely that they detect load through strain-induced flow of fluid within the canaliculi that deforms the cell processes. Circulating factors, such as calcium ions, may serve to amplify the signal. During normal physical activity, osteocytes produce a protein called sclerostin that inhibits Wnt signaling and suppresses bone formation. When high levels of mechanical loading are detected, sclerostin expression is suppressed, stimulating osteoblast differentiation and increasing bone formation (Robling et al. 2008). In other situations, for example when the bone is damaged or when the bone is not

loaded much at all, osteocytes begin to die by apoptosis. Either the apoptotic osteocytes (Kennedy et al. 2012), or those healthy osteocytes adjacent to the dying ones (Nakashima et al. 2011; Xiong et al. 2011), will begin to produce RANKL, which is important for the differentiation and activation of osteoclasts that are recruited to remove the damaged bone. When osteocytes are removed entirely from bone through targeted ablation, bone resorption leading to significant bone loss occurs (Tatsumi et al. 2007). Thus it appears that the osteocyte syncytium is responsible in large part for orchestrating processes of both resorption and formation. Bone lining cells are, like osteocytes, “retired” (quiescent) osteoblasts. These are the osteoblasts that escaped being buried in newly formed bone and remained on the surface when bone formation ceased. As production of bone matrix stops, bone lining cells become quiescent and flattened against the bone surface, but they do not form a continuous, gap-free barrier over the bone. They maintain communication with osteocytes and each other via gap-junctional processes, and also appear to maintain their receptors for parathyroid hormone, estrogen, and other chemical messengers. Like osteocytes, they are thought to be responsible for transfers of mineral into and out of bone (Parfitt et al. 1987).

Box 2.4 Technical Note: Bone Histology

Microscopic examination of biologic tissues requires that they be infiltrated and embedded in a solid medium to retain proper shape and cellular anatomy during subsequent sectioning. Paraffin, a waxy mixture of hydrocarbons commonly used to manufacture candles, has been the traditional medium of choice because of its low melting point and its miscibility with other organic compounds used to dehydrate and fix the tissue. Manufacturers now add plastic polymers and dimethyl sulfoxide to the paraffin to improve its infiltration and cutting characteristics. In the histology lab, small pieces of tissue, usually about a cubic centimeter in size, are immersed in a mold containing melted paraffin that is allowed to solidify. The embedded “tissue block” is then mounted in a microtome and thin sections (approximately 5 μm thick) are cut, mounted on glass slides, and examined under light microscopy. Skilled histologists can cut and mount several ribbons of serial sections an hour; what they cut is essentially a block of soft wax with cellular and extracellular material of similar softness suspended within.

Because of their hardness, bones and teeth present special problems. Obviously, processing these tissues in the standard fashion would result in broken microtome blades, shabby sections, and haggard histologists. To circumvent the problem, diagnostic histology laboratories first demineralize the sample to make it soft enough for processing using standard embedding and sectioning techniques. Methods of decalcification (the traditional name despite the fact that phosphorous is being liberated along with the calcium) have been in use since the turn of the century, and usually employ solutions of acid or chelating agents to solubilize and extract the mineral. Ethylenediaminetetraacetic acid (EDTA) is the current chelating agent of choice. Although such methods make paraffin embedding possible and are adequate for routine examinations of the cellular elements in tumors and marrow, they do not

render the complete microscopic picture, particularly in terms of bone as a dynamically changing structural tissue. To learn about bone as a tissue, researchers in the field usually use intact mineralized sections that require specialized equipment and procedures to process.

Before the advent of plastics in the 1930s, the only way to prepare acceptable mineralized sections of bone was by grinding down nonembedded rough-cut sections, most often by hand. Wet sandpaper was often used to prepare sections, especially thick sections of 100 μm or more; Frost's seminal discoveries were made using this method. For thinner sections, the specimen was sometimes sandwiched and ground with abrasive powder between two glass plates. Surprisingly beautiful histologic specimens only 5 μm thick can be produced in this way, however the technique is time consuming, tedious, and runs the risk of destroying trabecular architecture. Things changed for the better when ethyl and methyl methacrylates were introduced as embedding media around 1940. Much like paraffin embedding, samples of bone could be immersed and infiltrated with solutions of liquid media that were then polymerized. These hard plastics were a much closer match to the material properties of bone and thin sections (<3–5 μm) could be cut directly from the tissue blocks using specialized heavy-duty microtomes, or thick sections (100 μm) could be prepared without grinding.

In the latter half of the twentieth century, examinations of undecalcified histologic sections of bone have elucidated the cellular machinery responsible for normal bone growth and repair. Histologic manifestations of metabolic bone diseases have been well defined. Dr. Frost and his colleagues developed staining and analysis techniques for mineralized sections of bone that enabled clear identification of pertinent cells, endosteum and periosteum, osteoid, cement lines, and canaliculi, as well as patterns of mineralization. These methods, combined with fluorochrome labeling have given us insight into the complex cellular activity responsible for maintaining the mechanical integrity of our skeletons throughout life.

2.6 Introduction to Cartilage

Unlike bone, cartilage is less hierarchically organized, but not necessarily more homogeneous. Its heterogeneity is determined both by the orientation (or lack of orientation) of the collagen fibers, and by the relative concentration of proteoglycans in different layers of the cartilage matrix. Collagen and proteoglycan content are to some extent inversely related, with the more superficial layers of the cartilage matrix, closer to the articular surface, more highly enriched in collagen, whereas the deeper layers have less collagen and more proteoglycan. This makes the surface layers more resistant to tensile stress, but the deeper layers better able to cope with the large compressive stresses placed on the joint. Because it is composed primarily of water, cartilage is a highly viscoelastic structure, but the movement of water into and out of the cartilage matrix also provides the primary means for cellular nutrition in this avascular tissue. Also unlike bone, cartilage has a low

TABLE 2.2. Cartilage vs. bone

Characteristic	Cartilage	Bone
Collagen (predominantly)	Type II	Type I
Proteoglycan	Aggrecan	Non-aggregated proteoglycans
Composite	Collagen-aggrecan	Collagen-mineral
Organization	Non-lamellar	Lamellar
Cells	Not encapsulated No processes No cell-cell contact	Encapsulated Osteocytic processes Cell-cell contact (gap junctions)
Functional unit	Chondrons	Bone multicellular unit (BMU)
Mechanical properties	Viscoelastic Inhomogeneous Anisotropic Weak in shear Strong in tension	Viscoelastic Inhomogeneous Anisotropic Weak in shear Strong in compression
Growth/ossification	Epiphyseal (endochondral)	Epiphyseal (endochondral) Appositional (intramembranous)
Response to injury	Poor	Excellent
Metabolic rate	Low	High
Remodeling/repair	Single cell (chondrocyte)	Multicellular (osteoclast, osteoblast, osteocyte)

metabolic rate and is somewhat hypoxic. Injury to cartilage is therefore more serious as the cartilage does not repair itself readily. Whereas in bone, turnover and renewal of the tissue is performed by a coordinated set of activities performed by different cells, in cartilage a single cell type performs both catabolic and anabolic functions. This may be one reason for its low capacity for repair. Table 2.2 provides a comparison of the essential similarities and differences between cartilage and bone.

Like bone, cartilage comes in different forms, each suited to a particular application. For the sake of simplicity, we define three types of cartilage.

Hyaline cartilage is the most prevalent type of cartilage in the adult. It is found in the ventral ends of ribs, in the tracheal rings, and covering the joint surfaces of bones, where it is known as articular cartilage. In addition, the growth plates are composed of this cartilage (the growth plate is described in Sect. 2.7).

Elastic cartilage is the variety of cartilage found in the external ear, eustachian tubes, and epiglottis. It has greater opacity, flexibility, and elasticity than hyaline cartilage, and is yellowish rather than white. Its extracellular matrix is permeated with dense, branching elastic fibers unique to this type of cartilage.

Fibrocartilage is the type of cartilage occurring in intervertebral disks, the pubic symphysis, menisci, and in the bony attachments of certain tendons. It also may form when hyaline cartilage is damaged.

The articular cartilage of synovial joints is hyaline, although some joints also contain fibrocartilaginous disks (e.g., the menisci of the human knee). The extracellular matrix macromolecular framework of articular cartilage is composed mostly of type II collagen and proteoglycans manufactured by cartilage cells (called chondrocytes), and then assembled outside the cell into a mesh of collagen interwoven with aggregated proteoglycan molecules called aggrecan. Like bone cells, chondrocytes live in chambers called lacunae. Each chondrocyte lacuna is surrounded by a distinct pericellular matrix (PCM), or glycocalyx, that together with its encapsulated cell is called a chondron. Chondrons often contain more than one chondrocyte.

Unlike bone cells, chondrocytes are encapsulated not in mineralized bone but in a viscoelastic extracellular matrix of collagen fibers and proteoglycan molecules. The pericellular matrix surrounding the chondrocyte within the lacuna is enriched in binding proteins, such as anchorin, that help to anchor the chondrocytes to the surrounding substrate. Immediately adjacent and surrounding the pericellular matrix is a territorial matrix containing large amounts of nonfibrillar Type VI collagen (which may also be found in the pericellular region as well) that may serve as a mechanotransducer of deformations during loading. The cells, their lacunae and the territorial matrix are separated from each other by an interterritorial matrix composed of thicker fibers of Type II collagen, proteoglycans that are relatively enriched in keratan sulfate glycosaminoglycans, molecules that regulate collagen fiber assembly, and various growth factors (e.g. TGF- β and FGF), produced by the chondrocytes (Fig. 2.17). The chondrocytes also produce proteolytic enzymes, which contribute to turnover of the cartilage matrix. There is no direct cell–cell contact among chondrocytes, nor do these cells have processes to help them communicate with each other. The mechanisms for intercellular signal transduction, and how chondrocytes coordinate their activities, if they do, are unknown.

Cartilage matrix is ordinarily not mineralized and is about 70 % (by mass) water. Approximately 60 % of the dry weight of the matrix is collagen and 20 % each is proteoglycans or other non-collagenous proteins. Chondrocytes make up less than 1 % of the volume of articular cartilage in humans, although perhaps slightly more in other animals with smaller volumes of cartilage (Buckwalter and Mankin 1997). Type II collagen is the predominant species (about 90 %) of collagen found in articular cartilage. The assembly of the Type II collagen molecule is very similar to that in bone in that three separate α -chains (but different ones from those in bone) form a triple helix, the procollagen molecule. When the ends of the molecule are cleaved, they organize themselves into a quarter-staggered array. Although Type II collagen is the most abundant in cartilage, there are several other collagen types also present (Table 2.3). Type XI collagen microfibrils form a core for the type II collagen fibril, and control fibril growth. Larger Type XI triple helices help to form the collagenous network of articular cartilage.

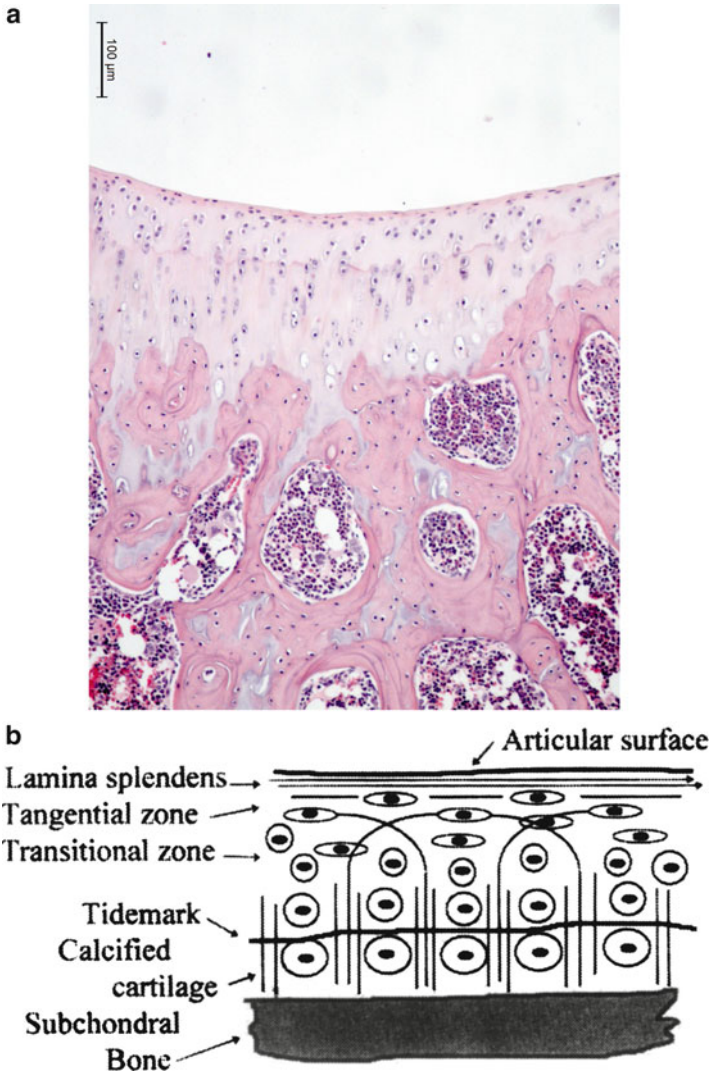


FIGURE 2.17. (a) Histological section showing the structure of adult articular cartilage. The tidemark can be seen as a *lightly stained line* within the cartilage, separating the transitional zone from the calcified cartilage. Rat patellar groove, stained with H+E. (b) Sketch of articular cartilage structure.

Type IX collagen, which is really a form of proteoglycan molecule, is found periodically along the Type II collagen fibril and is covalently cross-linked to it. It is probably important to the structural integrity of the collagen fibril assembly, and it has been suggested that disruption of the Type IX–Type II

TABLE 2.3. Types of collagen in articular cartilage and their functions

Collagen type	Function
II	Structural and mechanical; primary constituent of articular cartilage
VI	Pericellular adhesion molecule; regulates protein-protein interactions
IX	Fibril association; stabilizes type II
X	Hypertrophic zone of growth plate; role in calcification postulated
XI	Core of type II; controls fibril assembly and growth
XII	Cohesion and stability of collagen and the extrafibrillar matrix
XIV	Cohesion and stability of collagen and the extrafibrillar matrix

bond may be one factor leading to the degeneration of articular cartilage and a condition called osteoarthritis.¹

Type VI collagen is found in the territorial region around the chondrocyte lacunae. It is connected by molecular links to Type II collagen and to the larger proteoglycans, providing stability to the cell–extracellular matrix system. This collagen fiber network is stabilized by the formation of cross-links, just as in bone. Cartilage cross-links are excreted in the urine and can be used as markers (though perhaps not very good ones) of cartilage degradation in osteoarthritis. Type X collagen is found in small amounts within Type II collagen fibers. However, it is primarily found forming a network around hypertrophic chondrocytes in growth plates, where it is probably involved in controlling mineralization of the cartilage matrix. Interestingly, it is also produced in greater amounts in osteoarthrotic cartilage during the process of cartilage degeneration. Types XII and XIV have also been recently discovered in cartilage, and may provide cohesion between the extrafibrillar matrix and collagen fibers.

Proteoglycans in aggregate form are the other primary polymer in the matrix of cartilage. A core protein backbone provides the site for the attachment of glycosaminoglycans side chains (GAGs) (Fig. 2.18). The GAG chains consist of repeating disaccharide units in which a sugar is linked to a hexosamine. Chondroitin 4-sulfate, chondroitin 6-sulfate, and keratan sulfate, all of which carry a negative charge, are the three predominant GAGs found in cartilage. These molecules are covalently bound to the core protein to form the aggrecan molecule which, when linked to a hyaluronic acid (hyaluronan) molecule becomes the proteoglycan aggregate. Complete absence of the link protein weakens the bond between proteoglycans and hyaluronan and is perinatally lethal in mice. Cartilage matrix proteins

¹ Osteoarthritis is a term applied to joint degeneration that has a mechanical rather than a primary inflammatory cause. This condition is more commonly referred to as osteoarthritis, even though the “itis” in that term implies that inflammation is the initiating cause.

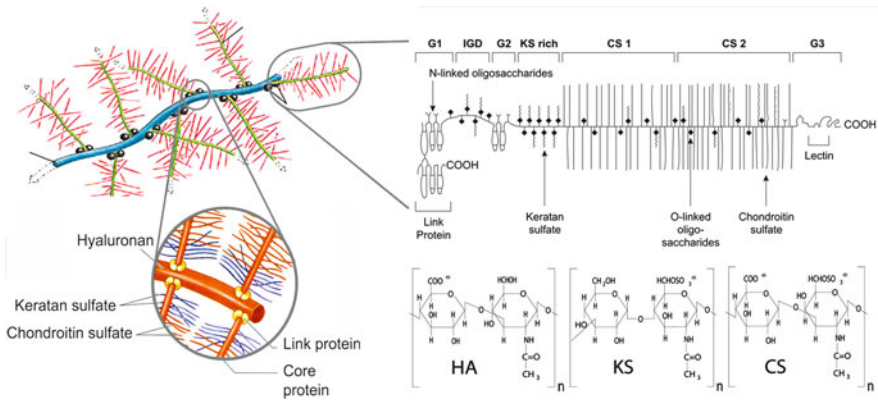


FIGURE 2.18. Two-dimensional diagram of proteoglycan aggregate structure. It should be kept in mind that branching of the core proteins from the hyaluronan backbone and the proteoglycan subunits from the core protein occurs in three dimensions. Reprinted with permission from Springer Science+Business Media (Schulz and Bader 2007).

(CMP), in particular matrilin-1 and 3, bind to the surface of the aggrecan molecule and may control the ability of protein to bind to collagen.

Although Type II collagen fibril size is in part regulated by the minor collagens, other molecules, some of them smaller proteoglycans, contribute to determining fibril size. One of the most prevalent and important is Cartilage Oligomeric Matrix Protein (COMP), which is found in relative abundance in the inter-territorial matrix (0.1 % of wet tissue weight). COMP, or thrombospondin 5, is one of several thrombospondins (including TS-1 and TS-4) found in articular cartilage that all serve to regulate collagen fiber assembly. COMP binds to the surface of Type II and Type XI collagen, and can accelerate the rate of fibril formation. Decorin is a smaller proteoglycan that “decorates” the collagen fibrils and changes fibril assembly. It probably functions to reduce fibril diameter, particularly close to the articular surface, where it is found in greater abundance. Fibromodulin and lumican are found in increasing amounts in the cartilage matrix away from the chondrocyte, binding to the collagen and inhibiting fibril formation. The disruption of fibromodulin may be an early event in the osteoarthrotic process that disrupts the fibrillar “weave” of the cartilage resulting in over-hydration and cartilage swelling. Other proteoglycans (epiphykan, versican, neurocan, mimecan, and perlecan) are found in varying amounts in different kinds of cartilage.

2.6.1 Mechanical Significance of Cartilage

Articular cartilage has unique chemical properties allowing it to serve as a bearing surface. It is able to transfer loads from one bone to another while

simultaneously allowing the load-bearing surfaces to articulate (i.e., roll and slide over one another) with very low friction. Here, however, we may ask, “How can a material that is 70 % water and not mineralized support loads that are several times body weight?” The answer to this question lies in the proteoglycan portion of the cartilage matrix. Specifically, the proteoglycan aggrecan serves this function. The proteoglycan monomers attached to the hyaluronic acid backbone are all highly negatively charged because of the chemical structure of the sulfated GAGs that compose them. This charge causes the aggrecan molecules to repel one another and branch out into something with the appearance of a bottlebrush. It also causes the aggrecan to be very hydrophilic. This attraction for water molecules resists the tendency for water to flow out of the cartilage when it is compressed. As the “bristles” of the brush restrain the water, they are placed in tension, and when the compressive load on the cartilage is removed the aggrecan molecules re-attract water to replace whatever has managed to escape. The large size and negative charge of the aggrecan molecule give cartilage its hydrophilic properties. Thus, compressive loads are borne by a volume of water that cannot escape, rather like a rubber hot water bottle filled with water. In this case, however, the water is constrained not by an outside membrane but by an internal structure that electrically confines the water molecules.

2.6.2 *Organization of Articular Cartilage*

Collagen and proteoglycans are dispersed differently through the thickness of the articular cartilage, with more collagen toward the joint surface and more proteoglycan in the deeper layers of cartilage matrix. This arrangement works well mechanically. Large tensile stresses on the articular surface and at the edges of joint contact areas are resisted primarily by tangentially oriented collagen fibers. The compressive and hydrostatic stresses found in the deeper layers of cartilage are resisted by the incompressibility of water, which is held in place by the hydrophilic aggrecan molecules.

The matrix of collagen fibers in cartilage is not organized randomly. Instead, the structure varies from place to place in response to functional demands, some of which are mechanical. The surface layer, the lamina splendens, consists of closely packed, small diameter (~30-nm) fibrils, tangentially arranged. The lamina splendens acts as a barrier membrane to enzymes and large molecules such as hyaluronic acid. Thus, small molecules like glucose in the synovial fluid can penetrate the lamina to nourish the chondrocytes deeper in the cartilage, but large, lubricating molecules such as hyaluronic acid remain on the surface of the cartilage. This also protects the cartilage to some degree from immune responses. Mechanically, the horizontal fiber arrangement allows the surface to resist the high tensile and shear forces in and around the joint contact area. However, with age, gaps appear in the lamina, allowing synovial fluid to transport enzymes (e.g., collagenase or hyaluronidase) that can break down the cartilage matrix.

The lamina splendens is only a few micrometers thick and merges gradually into the tangential zone in which the collagen fibrils are slightly larger, although still arranged parallel to the articular surface. This layer is more cellular than the lamina, containing small, flattened chondrocytes that are not very metabolically active in synthesizing collagen, but catabolize proteoglycans more quickly than chondrocytes in other layers. They demonstrate a polarity with a smooth but invaginated surface facing the articular surface, and cytoplasmic processes running into the deeper layers of cartilage. Typically, smaller proteoglycan aggregates are found in this layer. The tangential zone also contributes to the tensile stiffness of the cartilage matrix. When this layer becomes disrupted, as it does in the early stages of osteoarthritis, it is less able to dissipate surface tensile stresses, exposing the underlying cartilage layers to increased compressive stresses.

The transitional zone (or radial zone) contains collagen fibers that are larger than those in the tangential zone and also are more widely spaced and obliquely or pseudorandomly arranged. Together with the larger proteoglycan aggregates found in this layer, and which attract water to make the cartilage matrix turgid, the more vertically oriented collagen fibers in this layer help to stiffen the matrix, and reduce large strains (Shirazi et al. 2008). The chondrocytes tend to line up in radial columns (i.e., perpendicular to the surface), particularly in the deeper layers of this zone. These chondrocyte columns are parallel to the surrounding collagen fibers, and the chondrocytes within them actively synthesize structural molecules that replenish the cartilage matrix. This is reflected by the greater amounts of endoplasmic reticulum and Golgi found within the chondrocytes in this layer.

This zone ends at the tidemark, which is the remnant of the original calcification front in the growing bone and which continues to retain high levels of lipid and alkaline phosphatase, as one might expect in a mineralizing front. The calcified layer deep to this is mineralized with hydroxyapatite—the same mineral as found in bone. Although it was thought at one time that the calcified cartilage might serve to make the change in elastic modulus between articular cartilage and bone more gradual, it now appears that the calcified cartilage is more highly mineralized than bone (Fig. 2.19). This would be likely to increase shear stresses at the tidemark and at the osteochondral junction when the joint is loaded. Perhaps as an adaptation to this, collagen fibers from the transitional zone of the articular cartilage cross the tidemark and reinforce against shear between the stiff (calcified) and compliant (uncalcified, articular) cartilage layers.

For many years, investigators have used a simple method to visualize the orientation of collagen fibers in articular cartilage: simply sticking a pin dipped in India ink into the cartilage surface at multiple points over the surface of a joint reveals the pattern because the material splits parallel to the fiber direction. Figure 2.20 illustrates the articular surface at the distal end of the human knee after this has been done.

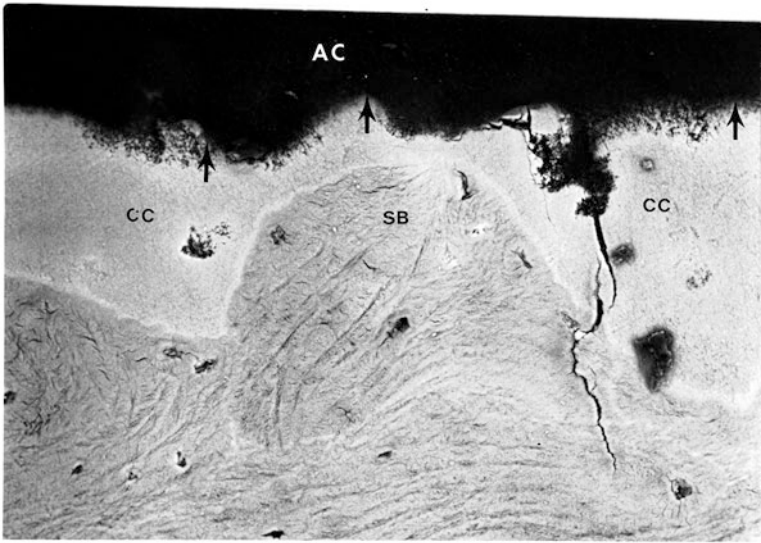


FIGURE 2.19. Higher magnification electron micrograph showing the tidemark (arrows) and osteochondral junction of the human femoral head. Note that the tidemark, which often appears as a well-defined line at lower magnification under the light microscope, is not definite. *Darker areas* within the calcified cartilage (CC) are non-mineralized regions or regions that are undergoing active mineralization. *SB* subchondral bone, *CC* calcified cartilage, *AC* articular cartilage (With permission from Brandt et al., 2003).

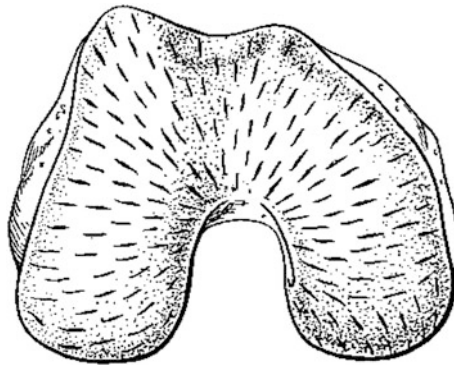


FIGURE 2.20. "Split line map" of the articular surface on the distal end of the human femur.

Noncalcified articular cartilage has no nerves and no blood vessels (calcified cartilage has both). The chondrocytes are nourished by diffusion through the extracellular matrix, which is largely water. This passive transport of nutrients may be aided by flows induced by mechanical loading, but the

dependence on passive transport rather than a blood supply limits the capacity of cartilage to repair large defects. Thus, although chondrocytes are constantly resupplying the extracellular matrix with collagen and proteoglycans as it degrades, the capacity to repair large amounts of damage is apparently absent.

Box 2.5 Bone Remodeling in Osteoarthritis

Historically, subchondral bone sclerosis has been associated with progressive deterioration of articular cartilage in osteoarthritis (OA). It was proposed that the increased subchondral bone density created an overly stiff substrate and internal stress concentrations that caused increased cartilage stresses, and led to its deterioration. However, observations have also demonstrated that there is increased bone remodeling in OA. This increased remodeling would be associated with at least a transient loss of bone, increased porosity in the subchondral region, and reduced density, just the opposite of sclerosis. It has been difficult to rationalize the subchondral sclerosis found in OA with these observations of increased bone remodeling. It is possible, however, that these two observations are not inconsistent, but simply reflect different temporal stages of disease, and/or different spatial locations.

In early OA, the normal mineral apposition rate of $\sim 0.70\text{--}1.00\text{ }\mu\text{m/day}$ increases threefold to fivefold to average values of $>3.50\text{ }\mu\text{m/day}$. There is also an increase in the initiation of new remodeling sites within the subchondral bone. This increased rate of remodeling reduces the thickness of the subchondral plate, and is associated with reduced synthesis of glycosaminoglycans and increased articular cartilage destruction. This has been shown in two different canine models of OA, but similar effects can be shown in early stage OA in young adults. When these bone changes are prevented by preventing the increased remodeling early in the disease process using an agent known to be a potent suppressor of bone remodeling (osteoprotegerin or OPG), bone volume in the subchondral region significantly increases, but cartilage deterioration is largely prevented. Suppression later in the disease process is not effective in either animal models or humans. Therefore, it is clear now that increased rates of remodeling in early stage OA are associated with subsequent deterioration of the overlying cartilage.

The bone resorption and cartilage loss seem to be spatially associated within a single joint, with a sevenfold increase in risk of cartilage loss in regions with subchondral attrition. Changes in subchondral bone mineralization and volume occur only beneath areas of significant cartilage destruction. The increased rates of remodeling in early stage OA may contribute to this by altering joint shape and load transmission that predisposes to progressive cartilage loss. These observations are not yet sufficient to determine whether this association is part of a pathogenetic sequence, or simply represents non-causal and independently occurring changes in bone and cartilage.

Induction of subchondral sclerosis directly, avoiding the early increased remodeling associated with OA, does not cause deterioration of the overlying cartilage nor progressive OA. This suggests that subchondral densification alone, without an intervening period of increased bone remodeling, does not lead inevitably to OA.

Therefore, the progressive development of osteoarthritis appears to involve a series of events which are temporally distinct and spatially localized. In early OA, there is increased bone remodeling, which is associated with a thinner subchondral plate, and bone tissue with a lower modulus that is less stiff. As the disease progresses, the remodeling rate decreases, but with uncoupling between bone resorption and formation leading to a net increase in bone formation. This increases bone volume, and may be associated with an apparent sclerosis. But because bone formation is still relatively elevated, the tissue itself may not be very stiff, which could alter the stress distribution in the joint and alter the shape of the joint surface. It is likely, though by no means proven, that both an early increase in remodeling and a later shift in the resorption-formation balance towards greater formation may be required for progressive OA to develop, and that sclerotic bone, by itself, is not sufficient to induce progressive joint deterioration.

2.6.3 *The Role of Cartilage in Growth*

Suppose that you were given the job of designing the mechanism by which a child's bones are to grow longer. The bones must do this while the child is active and loading them. The simplest solution might be to put osteoblasts on the ends of each bone, and let them form a layer of new bone there each day. The obvious problem with that solution is that the forces and motions between the bones would crush and grind the osteoblasts to bits. An alternative would be to put the osteoblasts inside the bone and let them "push out" new bone matrix, expanding the volume from within. Some early bone scientists thought this was how longitudinal growth worked, but eventually it became clear that the mineralized bone matrix is too rigid for this to happen. Bone can only be formed by cells on a bone surface because cells can only extrude new matrix into an unmineralized, "soft tissue" space. But how can such a soft tissue support loads during growth?

The solution to this dilemma is cartilage. Because it is largely water, cells can "push out" new matrix, expanding its volume, yet it is a tissue that can support very high loads if they are primarily compressive. There is evidence that nature has arranged a sort of pact between cartilage and bone such that the former exists where stresses are small or mainly compressive (hydrostatic) and the latter can bear relatively large compressive and shear (deviatoric) stresses.

This remarkable capacity for growth of a load-bearing material makes cartilage especially suitable as a skeletal substance during development. Indeed, most of the skeleton is first formed in cartilage models, which later are replaced by bone. The occurrence of cartilage is more restricted in post-fetal life, but it continues to play an indispensable role in the longitudinal growth of children's bones and in the maintenance of articular surfaces in adults.

2.7 Longitudinal Growth of Bones

Longitudinal growth of bones occurs in a nonmineralized region of growth near, but not at, each end of the bone. This region is called the growth plate or physis. The physis separates the bony epiphysis from the bony metaphysis (Fig. 2.21).

New cartilage is constantly formed by chondrocytes within the growth plate. It turns out that cartilage is viscous enough that cells within it can extrude more cartilage matrix, increasing its volume. This would increase its thickness except that the cartilage on the metaphyseal side of the growth plate mineralizes and becomes part of the metaphyseal bone, so that the length of the metaphyseal and diaphyseal parts of the bone is increased, while the thickness of the growth plate remains relatively constant. The width of the growth plate region is wider in bones that grow quickly, and narrower in bones that grow more slowly.

The growth plate is divided into zones, each representing a stage in the life cycle of its chondrocytes (Fig. 2.22). The number of these zones can vary, depending on the reference consulted. In the following description, you can imagine that you are traveling down through these zones and into the metaphysis at a given point in time, leaving the upper surface of the growth plate behind. Alternatively, you can imagine that you are sitting at a point which is fixed in relation to the bone as a whole, and watching your surroundings change as time passes, so that the top of the growth plate

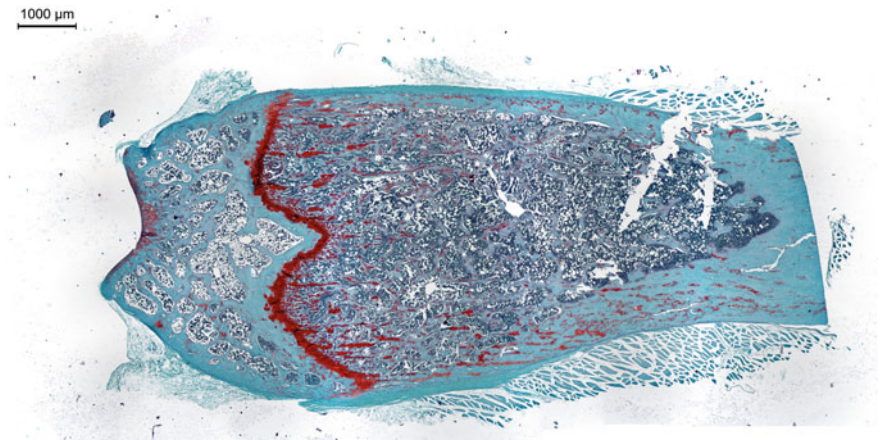


FIGURE 2.21. Distal femur of a rat showing the separation of the epiphysis and the metaphysis by the growth plate (*red*). Stained with Safranin O Fast Green.

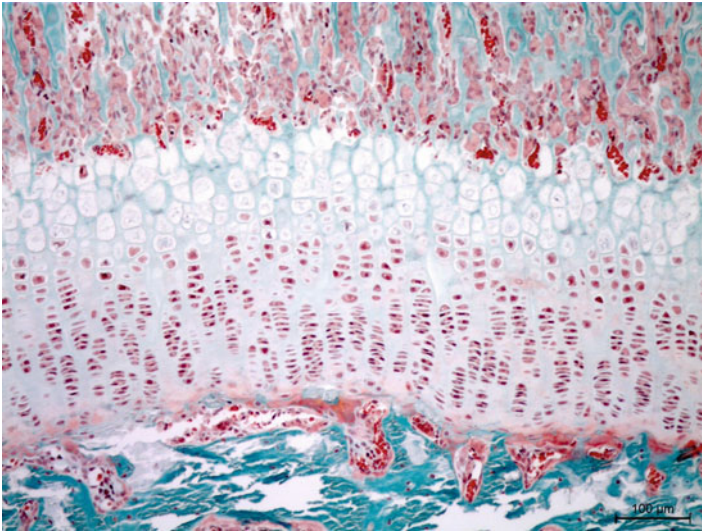


FIGURE 2.22. Higher magnification showing the structure of the different zones of the growth plate. Rat proximal tibia.

leaves you behind.² Forming a clear mental image of this relativity between time and position will help you understand the process by which bones grow longitudinally. The processes that occur during this temporal sequence through the growth plate are largely directed by a protein called SOX 9, perhaps in collaboration with hedgehog proteins (Leung et al. 2011). The hedgehog family of proteins is also important to differentiation of the different tissues, and even the fingers, during limb development. The growth plate's zones are, from the top (i.e., the epiphyseal side), as follows.

Reserve or resting zone: Cells of moderate size are scattered irregularly throughout this zone, which is anchored to the bone of the epiphysis and receives nourishment from epiphyseal blood vessels. Some of these cells may be chondroprogenitor cells, but the expression of Sox9 suggests that many have differentiated to chondrocytes. These cells are not really “resting,” but are dividing slowly to provide chondrocytes for the remainder of the growth plate. Type II collagen content is high in this region, and the fibrils are randomly arranged.

Proliferative zone: Here, the chondrocytes divide repeatedly, are close together, and arrange themselves in columns. They also become disklike, so that they look somewhat like a stack of coins. This is the region where most

² Recall Einstein's famous question, “Does Zurich stop at this train?”

of the growth in length occurs in the bone, primarily due to cell proliferation rather than cellular activity itself (Hunziker et al. 1987). Chondrocyte proliferation and maturation is permitted by suppression of Type X collagen production by SOX 9, but at the same time SOX 9 stimulates production of Type II and Type IX collagen, which allows for the assembly of the fibrillar network. Each cell produces about two times its own volume in new matrix during its lifetime (Hunziker et al. 1987). The collagen fibers are aligned parallel to the cell columns. The chondrocytes produce large amounts of proteoglycan and other molecules needed for the mechanical integrity of the surrounding extracellular matrix.

Hypertrophic zone: As they stop proliferating, the chondrocytes are said to “mature.” They accumulate glycogen in their cytoplasm and secrete copious amounts of matrix, increasing the volume of the surrounding substance. They also hypertrophy, increasing their volume and surface area by 4–10 times (Horton 1993). It is thought that the chondrocytes in this zone are in the early stages of apoptosis (programmed cell death). They stop producing cartilage matrix and begin producing molecules that prepare the adjacent cartilage for calcification.

Zone of provisional calcification: In this zone the degenerating chondrocytes continue to hypertrophy. As they reach the limit of this zone, where the growth plate ends and the metaphysis begins, the chondrocytes die. Simultaneously, the surrounding cartilage matrix is calcified. The calcification mechanism is imperfectly understood, but a working hypothesis for this process is as follows. The hypertrophying chondrocytes synthesize type X collagen and a protein called chondrocalcin, which may help initiate calcification. Proteoglycans may also be disassembled by enzymes released by the degenerating cells. Mineral crystal formation commences within matrix vesicles, which are intact spherical membranes extruded from the hypertrophic chondrocytes. These vesicles contain the enzyme ATPase, which provides energy to transport calcium ions into the vesicle against a concentration gradient. In addition, other enzymes in the vesicles cleave phosphate and calcium from compounds to which they are bound, allowing them to form crystals. The initiation of apatite crystal growth is the most important function of the matrix vesicle. As the levels of calcium and phosphate increase, amorphous calcium–phosphate aggregates form, eventually spreading beyond the vesicle wall and accumulating more mineral by epitaxy (crystal growth that imitates the form of the substrate). At this point, apatite crystals begin to grow in such a fashion that they are integrated with the collagen molecules in the matrix, and water is displaced as the cartilage matrix mineralizes.

The anabolic and the catabolic processes that occur at the growth plate are coupled so that the width of the growth plate is normally controlled within limits. However, if mineralization of the distal regions is delayed, or if chondrocyte division and matrix synthesis in the proliferative region is stimulated, the width of the growth plate will increase. Disturbances to

growth plate width—either narrowing or broadening, or to other aspects of its morphology can be caused by hormonal deficiency, metabolic disorders and genetic disease that can disturb growth plate development resulting in dysplasia (abnormality of development).

2.7.1 *Development of Metaphyseal Trabeculae*

The columns of enlarged lacunae left by the dead chondrocytes form the basis for tunnels, allowing blood vessels and cells from the metaphysis to gain access to the calcified cartilage at the bottom of the growth plate. As each chondrocyte hypertrophies, it dissolves some of the cartilage around it, so that only narrow walls, or septa, are left between the enlarged lacunae. Matrix vesicles which help to control the process of mineralization can be found within these septa. Cells called chondroclasts, another monocyte descendant, but perhaps not identical to osteoclasts, resorb the septa at the bottom of each chondrocyte column. Blood vessels from the metaphysis then invade these tunnel-like spaces, which are separated by columns or “trabeculae” of calcified cartilage. Chondroclasts also resorb some of the calcified cartilage on the walls of these tunnels. Osteoblasts then lay down small amounts of bone on these calcified cartilage trabeculae so that early in growth, the trabecular struts may retain a core of calcified cartilage that can persist in some diseases in which resorption is impaired. This remodeling activity begins to convert calcified cartilage to bone and marks the transition from growth plate to metaphyseal “bone.” At this early stage in the formation of spongy bone the tissue is called primary spongiosa; it still contains considerable calcified cartilage within the trabecular cores. As time goes by, and remodeling continues to replace the cartilaginous portions of the trabeculae with metaphyseal bone, the trabeculae become the secondary spongiosa. Eventually, the trabeculae become entirely bone. Those trabeculae in the center of the metaphysis are eventually entirely resorbed to form the medullary canal, with an arch of trabecular struts remaining to transmit loads from the central portion of the growth plate out to the cortices of the diaphysis.

2.7.2 *Growth of the Pysis*

In the newborn, the pysis is typically a flat, more or less circular plate. As the child grows and the bones become larger, the pysis must increase in diameter. This increase is achieved by cell division at its circumference in a region called the zone of Ranvier. In addition, the pysis loses its flat shape and becomes curved, often with a complex system of ridges and valleys. The pysis, being composed of cartilage, is weak in shear and tension, and physeal injuries in children are relatively common; the epiphysis of the proximal femur sliding off the shaft at the pysis (slipped capital femoral

epiphysis) is a well-known example. One consequence of the development of an irregularly shaped contour to the physis in the growing child is that the complex interlocking geometry of the epiphysis and metaphysis confers some protection against failure of the physis in response to shearing forces.

Most long bones have two physes, one at each end (although some bones, such as the metacarpals and phalanges, have only one functional physis). Usually one of these contributes more than the other to longitudinal bone growth. For example, in the femur the distal physis contributes more to growth, whereas in the tibia the proximal physis is more active. Historically, the relative growth rates of the two physes of a long bone have been measured by comparing bones of different age with respect to some naturally occurring marker in the diaphysis, usually the entrance point of the nutrient artery.

2.7.3 Closure of the Physes

As a child matures, bones reach their adult length and the physis is no longer required. At this point, the physes “close” by ossification, connecting the epiphysis to the metaphysis with bone. The blood circulatory systems of the epiphysis and metaphysis, formerly independent, also unite. This region of bone remodels over many years, but the physeal “scar” or “ghost” (a plate of bone at the site of the former physis) may persist into old age.

Although it is often thought that closure of the epiphyses stops longitudinal growth, available evidence suggests that growth slows down and stops as a function of the increase in sex steroids during maturation, and that closure of the physis follows cessation of growth (Parfitt 2002). Some animals (e.g. rodents such as rats and mice) never completely stop growing, although growth slows markedly, and consequently their growth plates never fuse. In animals that lack an ossified epiphysis (many reptiles, e.g. crocodiles), there is no physeal closure and the bones continue to increase in length indefinitely, so that great longevity becomes associated with great size.

Closure of the physes of different bones, and even the proximal and distal ends of the same bone, occurs at different ages and in a specific sequence. This allows one to determine radiographically the skeletal age of an individual by observing which physes, usually those of the left hand and wrist, have closed. The hands and wrist are most useful because of the number of bones, and the fact that the eight wrist bones, which are derived only from a primary center of ossification and therefore do not have epiphyses, appear radiographically at different ages and grow at different rates (the skeletal elements are present before they appear radiographically, but are completely cartilaginous and are not visible). Skeletal age is not the same as chronological age; the growth plates close earlier in some individuals than others. In girls, physes close several years earlier than in boys; this contributes to the shorter average stature of women compared to men, and probably also to their lower overall bone density.

2.8 Introduction to Ligament and Tendon

Ligaments and tendons, the flexible structures that bind together the musculoskeletal system, are extraordinarily strong in resisting tensile loads. For example, the digital flexor tendon from the foreleg of a horse can bear the weight of two large automobiles without failure. The Anglo-Saxon word for tendon is “sinew,” which also means “strong” or “tough.” As Aristotle realized, without ligaments and tendons to stabilize and animate our skeletons, they would be mechanically useless. We now turn our attention to these important tissues.

Tendons and ligaments differ in morphology and function. Ligaments bind one bone to another to restrict their relative motions. Tendons provide the connecting link from a muscle to a bone. Although textbooks sometimes refer to ligaments as if they were distinct and easily identified anatomical structures, this is often untrue. Many ligaments represent thickenings or specializations within a joint capsule, and their margins may be blurred and indistinct. Moreover, there is considerable individual variation in ligaments, and a ligament that is well defined in one person may be obscure or absent in another. Similarly, textbooks often treat tendon as if it were a uniform generic material, when in fact there is considerable variation in the properties of different tendons that represent adaptations to specialized function. These variations are unlikely to be entirely genetic in origin; many may also derive from the different ways in which each individual habitually uses a particular part of their musculoskeletal system. As with bone, cells may remodel ligaments and tendons to adapt them to the imposed loads.

Tendons and ligaments, along with other non-calcified, non-cartilaginous load bearing structures of the musculoskeletal system (e.g. synovial joint capsules, aponeuroses, and retinacula), are all composed of dense connective tissue. These structures are characterized by an abundance of tightly packed collagen fibers which give the tissue superior tensile strength. In tendons and ligaments the collagen is packaged into bundles of parallel fibers aligned with the predominant force direction. The vast majority of the collagen is type I with trace amounts of types II, III, V, VI, IX, XII, and XIV present at some locations. Ligaments and to a lesser extent tendons, also contain highly extensible elastic fibers constructed from another fibrillar protein called elastin. Small amounts of the proteoglycan molecules biglycan, decorin, fibromodulin, lumican and versican are also present in the extracellular matrix of tendon and ligament, many of which modulate collagen fibril formation and size. Aggrecan, the large hydrophilic molecule found in articular cartilage, is present in regions of the tissue that wrap around bony structures (e.g., tendon of posterior tibialis at the medial malleolus of the ankle; annular ligament that retains the head of the radius) and are routinely subjected to compressive force.

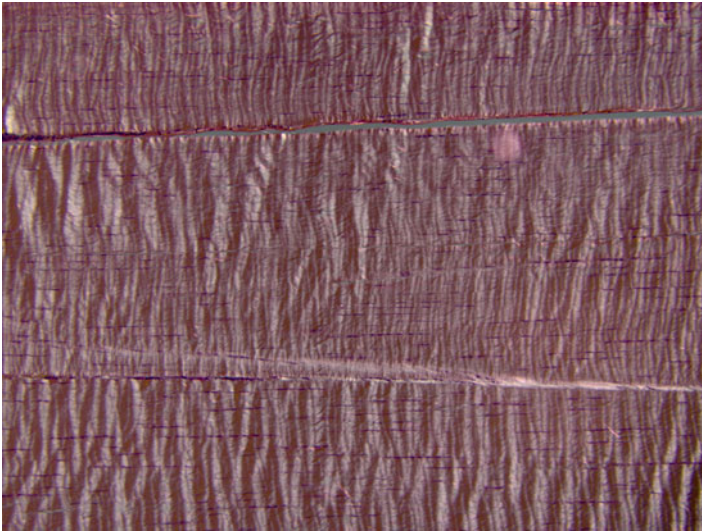


FIGURE 2.23. Photomicrograph of equine tendon as viewed under polarized light illustrating the parallel arrays of collagen fibers interspersed with darker staining fibroblasts. The wavy or “crimp” pattern is characteristic of unloaded tendon and ligament (Image provided courtesy of Dr. Brian Murphy, University of California, Davis). Orig. mag. $\times 100$.

As is the case for bone and cartilage, dense connective tissues originate in the embryonic mesoderm and the cells that populate and manufacture structural components needed to build and maintain these tissues are descended from pluripotent mesenchymal stem cells. Mesenchymal stem cells differentiate into osteocytes, chondrocytes, adipocytes (fat cells), erythrocytes (red blood cells) and leukocytes (white blood cells). They are also the precursors to fibroblasts, the predominant cells found in tendons and ligaments. Fibrocartilage cells, intermediate cells possessing characteristics of both fibroblasts and chondrocytes, can be found in regions where these structures attach to bone and at anatomic sites subjected to compressive forces. Fibroblasts in tendon and ligament tend to lie in longitudinal rows aligned with and interspersed among the collagen fibers of the extracellular matrix (Fig. 2.23). Interestingly, fibroblasts, like osteocytes within bone, display long cellular processes that connect with processes from cells in nearby rows. And like osteocytes, these cells have been shown to communicate through gap junctions composed of connexin 43. It seems highly likely that local networks of interconnected cells sense the magnitudes and modes of loading in the immediate vicinity and respond accordingly to produce and maintain an appropriate extracellular matrix, but the transduction pathway(s) responsible have yet to be worked out.

Blood supply to the dense soft tissues is highly variable but in general less than that found in bone and far less than that in the more metabolically active tissues (e.g. muscle). Vessels and/or capillaries are predominately located on the surface of these structures and to a lesser extent within the midsubstance, or may be absent altogether depending upon structure and location. Vessels within tendon run longitudinally and may originate in the adjoining muscle or bone or permeate the tendon at intermittent locations along its length. Tendons and ligaments have a rich nerve supply that provides proprioceptive feedback for motor control and to prevent damage through overload. Pain receptors and nerve fibers involved in vasoregulation are also present in these tissues.

2.8.1 Structure and Composition

2.8.1.1 Ligament

In its natural state, ligament is 55–65 % water; its other components are shown in Table 2.4 as percentages of dry weight. Collagen comprises approximately 70–80 % of the dry weight and elastin usually comprises another 10–15 %, more in some specialized ligaments (e.g. nuchal (neck) ligaments). In most cases proteoglycans comprise a very small percentage of the extracellular matrix in ligaments and tendons. It should be kept in mind that these proportions are general approximations; exact percentages vary according to ligament location and function.

Collagen, the major organic constituent of bone and cartilage, is by far the most abundant extracellular protein contained in ligament. Fibrillar collagen (mainly type I and small amounts of types, III and V) gives ligament its high tensile strength and is synthesized by fibroblasts. The primary collagen molecule in ligament is Type I, a glycine-rich triple helix, composed of two α -1 chains and one α -2 chain, just as Type I collagen in bone. As with any other protein, collagen synthesis begins with gene transcription in the nucleus followed by translation of the mRNA in the cytoplasm. Translated products are routed to the endoplasmic reticulum where the procollagen molecule is assembled and then extensively modified before being secreted into the extracellular matrix. Even after secretion several enzymes continue to modify the procollagen molecule in preparation for assembly. Finally, the processed molecules, now simple triple-helical subunits, spontaneously self-assemble

TABLE 2.4. Major components of ligament, tendon, and skin as percent dry weight

Component	Ligament	Tendon	Skin
Collagen (mostly as type I)	70–80	75–85	56–70
Elastin	10–15	<3	5–10
Proteoglycans	1–3	1–2	2–4

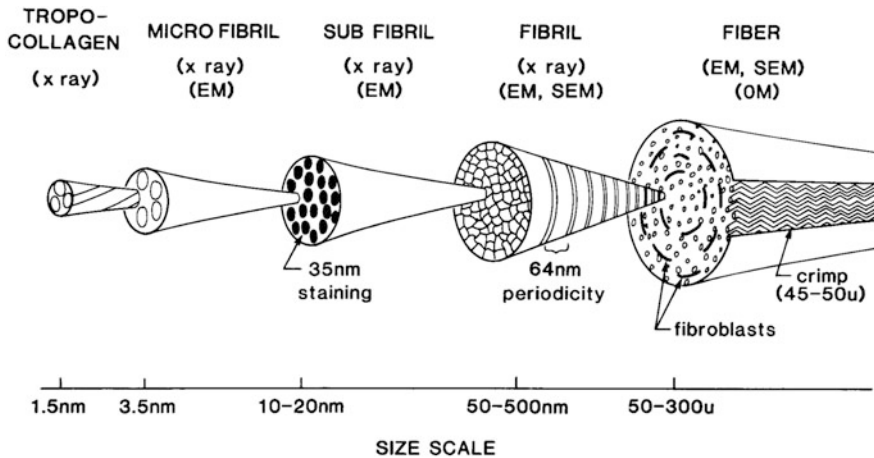


FIGURE 2.24. Structure of a collagen fiber in a ligament. Parenthetical notes at top indicate the kind of imaging that revealed the structure: X-ray diffraction, transmission or scanning electron microscopy (EM, SEM), or optical microscopy (OM) (Kastelic and Baer 1980).

into collagen microfibrils (4 nm in diameter), which in turn assemble into sequentially larger strands of protein called subfibrils (20 nm in diameter) and fibrils (50–500 nm in diameter). Collagen fibrils can be visualized with electron microscopy and have a characteristic striated pattern resulting from the ordered assembly of the individual collagen molecules within the microfibrils. At the top of the structural hierarchy are collagen fibers (100–300 μ m), which can be seen using light microscopy (Fig. 2.24).

Fibroblasts also possess the machinery required to break down and remove old collagen, an important function in response to injury. To do this, fibroblasts synthesize and secrete neutral collagenases, enzymes designed to cleave the collagen triple helix. Collagen turnover is a normal and continual process throughout life with the rate of synthesis and degradation dependent upon the levels of physical activity and imposed load (Kjaer et al. 2006).

In the absence of load, the collagen fibers demonstrate a wavy crimp pattern. As the structure is stretched, the crimp pattern disappears and the fibers become straight. The collagen fibers may be grossly aggregated into bundles that are oriented along the principal axis of stress experienced by the ligament. As a joint progresses through a range of motion, different portions of the ligament are stressed more than other portions at each joint position as a consequence of simple geometry. In many ligaments, the collagen bundles are not strictly parallel; this represents an adaptation by the ligament to provide optimal strength at all joint positions.

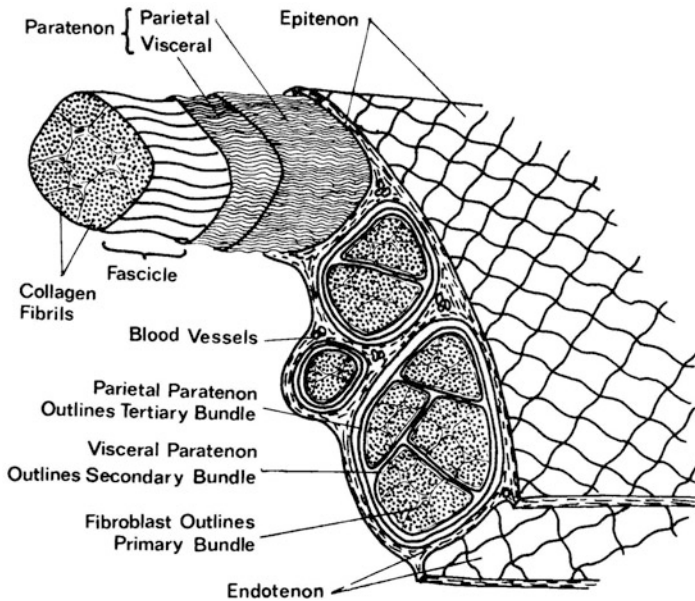


FIGURE 2.25. The structure of tendon, at different length scales.

2.8.1.2 Tendon

The composition of tendon is similar to that of ligament (Table 2.4) Collagen comprises as much as 85 % of tendon dry weight; of this, 95 % is type I and 5 % type III and/or type V (Cetta et al. 1982). Unlike ligament, elastin constitutes less than 3 % of tendon dry weight. In most locations, the proteoglycan concentration is usually less than 2 %; however, tendons that curve around bony surfaces experience compressive forces, and the cells within them respond by producing more proteoglycans. Increasing proteoglycan concentrations give the tendon a more cartilaginous quality that decreases friction and enhances motion.

The structural hierarchy of tendon resembles that of ligament except that collagen fibers are arranged into discrete packets, which in turn are bundled together within fascicles. As shown in Fig. 2.25, primary fiber bundles and fascicles are bound together and supported by the endotenon, a loose connective tissue containing blood vessels, nerves, and lymphatics. The endotenon is continuous with the epitenon, which encases the whole tendon. Long tendons such as the digital flexors are enclosed in a synovial sheath that nourishes the structure and enhances gliding. Fibroblasts are interspersed between collagen bundles and synthesize collagen for structural repair and maintenance. The fibroblasts in tendons are sometimes called tenocytes. On average, tendons contain fewer cells and are less metabolically active than ligaments.

2.8.2 *Injury and Repair*

Ligament and tendon injuries are common and generally entail disruption of the collagen network with associated loss of mechanical integrity. As in other tissues, healing involves a localized inflammatory response typically divided into three phases. Phase I, acute inflammation, lasts from about 48–72 h post-injury and is characterized by pain, swelling, redness, and warmth from vasodilation. During this phase cells are recruited and a hematoma forms. Phase II, proliferation, lasts up to 6 weeks post injury, is characterized by increased numbers of fibroblasts, collagen synthesis, and scar tissue formation. Phase III involves remodeling of the scar and can progress for 12 months or longer. Collagen synthesis decreases, water content drops, and the extracellular matrix becomes more organized, usually meaning a more longitudinal arrangement of fibers. Scar contraction coincides with increasing tensile strength. A number of cytokines activate and guide the healing process. Principal among them, insulin-like growth factor (IGF-1) and transforming growth factor- β (TGF- β), both are involved in the recruitment and migration of fibroblasts and inflammatory cells to the area of injury as well as cell proliferation and synthesis of matrix components. Vascular endothelial growth factor (VEGF) reestablishes vascularization. The robustness of the response to more severe injury is dependent on sound clinical management and the general health of the individual.

2.8.2.1 Ligament

Competitive and recreational athletes are most at risk for ligament damage, with higher rates of injury in adolescents, particularly young women. Grade I sprains entail relatively minor damage to the collagenous matrix with some associated pain and swelling but spontaneously resolve without intervention. Grade II sprains are more severe with tearing of the fibrous matrix, more internal bleeding and swelling and more severe pain. Grade III sprains are complete ruptures resulting in joint instability and often requiring surgical reconstruction. Ligament healing progresses as described above but the affected structure may never regain full strength; primate studies indicate tensile strength to be 80 % of the intact ligament at 1 year (Clancy et al. 1981). Full thickness tears (grade III sprains) cannot be effectively repaired and therefore must be replaced, usually by an autograft (i.e. the patient's own tissue taken from another anatomic site), to restore function. The anterior cruciate ligament (ACL), an important intra-articular stabilizer of the knee, is highly vulnerable to grade III sprains, with over 100,000 ACL reconstructions performed annually in the United States.

2.8.2.2 Tendon

Tendon injuries are most frequently associated with overuse but lacerations and contusions from trauma are also common. Overuse injuries tend to affect a broader and older population. In the case of tendon, damage to the

extra-cellular matrix due to overload is called a strain rather than a sprain. Strains usually occur at tendon-muscle or tendon-bone junctions, which each possess different mechanical properties than the tendon proper; mid-substance, full thickness tears occur much less frequently. The rotator cuff, Achilles, and patellar tendons are most prone to injury due to overuse. Repetitive loading produces sub-rupture fatigue damage leading to pain and disability. The resulting tendinopathy (sometimes called tendinosis) is a failed healing response characterized by the proliferation of abnormal fibroblasts (tenocytes), degenerate collagen fibers, and increased glycoaminoglycan and water content. Some tendinopathies resist current treatment modalities and become chronic conditions. The search for more effective approaches to the management of tendinopathy is an important area of research.

In closing this section, it is important to point out that under normal conditions our dense connective tissues are being constantly modeled, remodeled and renewed, continuously adapting to changing mechanical demands. In most cases protective mechanisms keep imposed loads within appropriate boundaries such that renewal keeps pace with demand, but when these limits are exceeded sprains and strains occur. We will learn more about the functional adaptation of these structures and their mechanical behavior in Chap. 4.

2.9 Summary for Further Reading

Several excellent works can be suggested for further reading. These include *Principles of Bone Biology* (Bilezikian et al. 2002), *Orthopaedic Basic Science* (Simon 1994), and *Osteoporosis* (Marcus et al. 1996). *Rheumatology* (Hochberg et al. 2015) provides excellent chapters reviewing bone, cartilage and tendon/ligament biology. For more in-depth reading, one may turn to B. K. Hall's series of books titled *Bone*. A recent issue of the *Journal of Biomechanics* is devoted to bone as a hierarchical material (Ascenzi and Reilly 2011). For more specifics on cartilage, we recommend *Osteoarthritis* second edition by Brandt (Brandt et al. 2003). An extensive review of all aspects of ligament and tendon biology and mechanics may be found in the book *Injury and Repair of the Musculoskeletal Soft Tissues*, edited by Woo and Buckwalter (Frank et al. 1988). A more recent and excellent review of tendon biology, however, is *Structure-function relationships in tendons: a review* (Benjamin et al. 2008), published in the *Journal of Anatomy*.

2.10 Exercises

The information in this chapter introduced the basic biological structure of the musculoskeletal system. This is a necessary foundation to begin to think quantitatively about the dynamics of growth and tissue renewal, which will

be visited in the next chapter. The exercises here are meant to start you thinking quantitatively about the biological material introduced in this chapter, in preparation for the more analytical chapters to follow:

- 2.1. Estimate the average rate of addition of bone on the diaphyseal surface of your tibia, in micrometers/year, as you grew to skeletal maturity. Palpate your tibia to estimate its present diameter, and assume its diameter was 0.8 cm when you were born. Assume that your tibia's periosteal apposition rate essentially stopped at age 14 if you are a woman or age 16 if you are a man, ignoring any further increase in the diameter of the tibia that occurred after closure of the physes.
- 2.2. Typically, a newborn infant's tibia is about 6.8 cm long (Jeanty et al. 1984). Estimate the present length of your tibia, then estimate the average rate of growth in the tibial growth plates, in millimeters/year and in micrometers/day. Again, assume that longitudinal growth of the tibia stops at age 14 in females and age 16 in males. How does this compare with the bone apposition rate estimated in Exercise 3.1?
- 2.3. The rate of growth of a growth plate is equal to the daily chondrocyte production rate for each column multiplied by the height of the maximally hypertrophied cells (Sissons 1955). Write a paragraph explaining why this is so in a steady-state situation and stating the necessary assumptions.
- 2.4. Consider the relationship between the rate of growth in a growth plate, the average number of cells being produced each day, and the average amount of hypertrophy each cell experiences. Assume the following. The growth plate consists of many parallel cell columns, all growing at the same rate. The stem cell at the top of each column divides at a rate r_s (divisions/year), replacing itself and producing one new chondrocyte. Each resulting chondrocyte divides, producing two new chondrocytes, which in turn divide. This mitotic sequence continues until N generations of chondrocytes are produced, all of which are aligned in a single column. Each chondrocyte in the column hypertrophies to an ultimate cell height, h . Using the principle of Exercise 2.3, show that the physal growth rate is $G = h r_s 2^N$.
- 2.5. If it takes 2 months to refill an osteonal resorption cavity that is initially 250 μm in diameter, what is the average apposition rate in micrometers/day? Assume the Haversian canal diameter is 50 μm after refilling is complete. Compare this apposition rate with that in Exercise 3.1.
- 2.6. Bone can be considered a two-phase, fiber-re-enforced material. Describe the nature of this composite structure at the nano-, micro- and macrostructural levels. Compare and contrast it with existing implant materials.

- 2.7. Compare the functional adaptations of bone, articular cartilage, and tendon/ligament with regard to their structure and their respective roles in skeletal support and load transmission. How do the morphological differences at each level of organization permit them to efficiently carry out their respective functions. Explain how these tissues act together to fulfill their roles and to protect each other from damage during the aging process.

References

- Ascenzi A, Bonucci E. The ultimate tensile strength of single osteons. *Acta Anat.* 1964;58:160–83.
- Ascenzi A, Bonucci E. The tensile properties of single osteons. *Anat Rec.* 1967;158:375–86.
- Ascenzi A, Bonucci E. The compressive properties of single osteons. *Anat Rec.* 1968;161:377–91.
- Ascenzi A, Bonucci E. The shearing properties of single osteons. *Anat Rec.* 1972;172:499–510.
- Ascenzi M-G, Reilly G. Bone tissue: hierarchical simulations for clinical applications. *J Biomech.* 2011;44:211–364.
- Ascenzi M-G, Fill J, Lomovtzev A. Orientation of collagen at the osteocyte lacunae in human secondary osteons. *J Biomech.* 2008;41:3425–35.
- Bailey AJ, Knott L. Molecular changes in bone collagen in osteoporosis and osteoarthritis in the elderly. *Exp Gerontol.* 1999;34:337–51.
- Bailey AJ, Wotton SF, Sims TJ, et al. Post-translational modifications in the collagen of human osteoporotic femoral head. *Biochem Biophys Res Commun.* 1992;185:801–5.
- Banse X, Sims TJ, Bailey AJ. Mechanical properties of adult vertebral cancellous bone: correlation with collagen intermolecular cross-links. *J Bone Miner Res.* 2002;17:1621–8.
- Benjamin M, Kaiser E, Milz S. Structure-function relationships in tendons: a review. *J Anat.* 2008;212:211–28.
- Bilezikian JP, Raisz LG, Rodan GA. *Principles of bone biology.* New York: Academic; 2002.
- Brandt KD, Doherty M, Lohmander LS. *Osteoarthritis.* Oxford: University Press; 2003.
- Buckwalter JA, Mankin HJ. Articular cartilage. Part I: tissue design and chondrocyte-matrix interactions. *J Bone Joint Surg.* 1997;79-A:600–11.
- Buehler MJ. Molecular nanomechanics of nascent bone: fibrillar toughening by mineralization. *Nanotechnology.* 2007;18:295102.
- Burr DB, Schaffler MB, Fredrickson RG. Composition of the cement line and its possible mechanical role as a local interface in human compact bone. *J Biomech.* 1988;21:939–45.
- Carden A, Morris MD. Application of vibrational spectroscopy to the study of mineralized tissues. *J Biomed Opt.* 2000;5:259–68.
- Cetta G, Tenni R, Zanaboni G. Biochemical and morphological modifications in rabbit Achilles tendon during maturation and aging. *J Biochem.* 1982;204:61–7.

- Ciarelli TE, Fyhrie DP, Parfitt AM. Effects of vertebral bone fragility and bone formation rate on the mineralization levels of cancellous bone from white females. *Bone*. 2003;32:311–5.
- Clancy WG, Narechiana RG, Rosenberg TD. Anterior and posterior cruciate ligament reconstruction in rhesus monkeys: a histological, microangiographic, and biomechanical analysis. *J Bone Joint Surg*. 1981;63A:1270–84.
- Cohen J, Harris WH. The three-dimensional anatomy of Haversian systems. *J Bone Joint Surg*. 1958;40A:419–34.
- Cooper RR, Milgram JW, Robinson RA. Morphology of the osteon. *J Bone Joint Surg*. 1966;48A:1239–71.
- Currey JD. Three analogies to explain the mechanical properties of bone. *Biorheology*. 1964;2:1–10.
- Currey JD. The mechanical consequences of variation in the mineral content of bone. *J Biomech*. 1969a;2:1–11.
- Currey JD. The relationship between the stiffness and the mineral content of bone. *J Biomech*. 1969b;2:477–80.
- Currey JD. Incompatible mechanical properties in compact bone. *J Theor Biol*. 2004;231:569–80.
- Ding M, Dalstra M, Danielsen CC, et al. Age variations in the properties of human tibial trabecular bone. *J Bone Joint Surg Br*. 1997;79:995–1002.
- Donnelly E, Baker SP, Boskey AL, et al. Effects of surface roughness and maximum load on the mechanical properties of cancellous bone measured by nanoindentation. *J Biomed Mater Res A*. 2006;77:426–35.
- Frank C, Woo S, Andriacchi T, Brand R, Oakes B, Dahners L, DeHaven K, Lewis J, Sabiston P. Normal ligament: structure, function, and composition. In: Woo SL-Y, Buckwalter JA, editors. *Injury and repair of the musculoskeletal soft tissues*. Park Ridge, IL: American Academy of Orthopaedic Surgeons; 1988. p. 45–128.
- Fratzl P, Roschger P, Eschberger J, et al. Abnormal bone mineralization after fluoride treatment in osteoporosis: a small angle X-ray-scattering study. *J Bone Miner Res*. 1994;9:1541–9.
- Fritsch A, Hellmich C, Dormieux L. Ductile sliding between mineral crystals followed by rupture of collagen crosslinks: experimentally supported micromechanical explanation of bone strength. *J Theor Biol*. 2009;260:230–52.
- Gao H, Ji B, Jager IL, et al. Materials become insensitive to flaws at nanoscale: lessons from nature. *Proc Natl Acad Sci USA*. 2003;100:5597–600.
- Giraud-Guille MM. Twisted plywood architecture of collagen fibrils on human compact bone osteons. *Calcif Tissue Int*. 1988;42:167–80.
- Giri B, Tadano S, Fujisaki K, et al. Understanding site-specific residual strain and architecture in bovine cortical bone. *J Biomech*. 2008;41:3107–15.
- Gourion-Arsiquaud S, Faibish D, Myers E, et al. Use of FTIR spectroscopic imaging to identify parameters associated with fragility fracture. *J Bone Miner Res*. 2009;24:1565–71.
- Gupta HS, Seto J, Wagermaier W, et al. Cooperative deformation of mineral and collagen in bone at the nanoscale. *Proc Natl Acad Sci USA*. 2006;103:17741–6.
- Havers C. *Osteologia nova*, or some new observations of the bones, and the parts belonging to them, with the manner of their accretion, and nutrition, communicated to the royal society in several discourses. London: Samuel Smith; 1691.

- Hengsberger S, Kulik A, Zysset P. Nanoindentation discriminates the elastic properties of individual human bone lamellae under dry and physiological conditions. *Bone*. 2002;30:178–84.
- Hernandez CJ, Tang SY, Baumbach BM, et al. Trabecular microfracture and the influence of pyridinium and non-enzymatic glycation-mediated collagen cross-links. *Bone*. 2005;37:825–32.
- Hildebrand T, Laib A, Muller R, Dequeker J, Ruegsegger P. Direct three-dimensional morphometric analysis of human cancellous bone: microstructural data from spine, femur, iliac crest, and calcaneus. *J Bone Miner Res*. 1999;14:1167–74.
- Hochberg MC, Silman AJ, Smolen JS, Weinblatt ME, Weisman MH. *Rheumatology*, 2-volume set. 6th ed. Philadelphia: Mosby/Elsevier; 2015.
- Homminga J, McCreadie BR, Ciarelli TE, Weinans H, Goldstein SA, Huiskes R. Cancellous bone mechanical properties from normals and patients with hip fractures differ on the structure level, not on the bone hard tissue level. *Bone*. 2002;30:759–64.
- Horton WA. Morphology of connective tissue: cartilage. In: Royce PM, Steinman B, editors. *Connective tissue and its heritable disorders*. New York: Wiley-Liss; 1993.
- Hou FJ, Lang SM, Hoshaw SJ, Reimann DA, Fyhrie DP. Human vertebral body apparent and hard tissue stiffness. *J Biomech*. 1998;31:1009–15.
- Hu Y-Y, Liu XP, Ma X, Rawal A, Prozorov T, Akinc M, Mallapragada SK, Schmidt-Rohr K. Biomimetic self-assembling copolymer–hydroxyapatite nanocomposites with the nanocrystal size controlled by citrate. *Chem Mater*. 2011;23:2481–90.
- Hunziker EB, Schenk RK, Cruz-Orive L-M. Quantitation of chondrocyte performance in growth-plate cartilage during longitudinal growth. *J Bone Joint Surg*. 1987;69-A:162–73.
- Jeanty P, Rodesch F, Delbeke D, Dumont JE. Estimation of gestational age from measurements of fetal long bones. *J Ultrasound Med*. 1984;3:75–9.
- Johnson LC. The kinetics of skeletal remodeling. *Birth Defects Orig Artic Ser*. 1966;2:66–142.
- Jowsey J. Studies of Haversian systems in man and some animals. *J Anat*. 1966;100:857–64.
- Kastelic J, Baer E. Deformation in tendon and collagen. In: Vincent JFV, Currey JD, editors. *The mechanical properties of biological materials*. Cambridge: Cambridge University Press; 1980. p. 397–435.
- Keaveny TM, Guo XE, Wachtel EF, McMahon TA, Hayes WC. Trabecular bone exhibits fully linear elastic behavior and yields at low strain. *J Biomech*. 1994;27:1127–36.
- Kennedy OD, Herman BC, Laudier DM, Majeska RJ, Sun HB, Schaffler MB. Activation of resorption in fatigue-loaded bone involves both apoptosis and active pro-osteoclastogenic signaling by distinct osteocyte populations. *Bone*. 2012;50:1115–22.
- Kjaer M, Magnusson P, Krogsgaard M, Boysen Møller J, Olesen J, Heinemeier K, Hansen M, Haraldsson B, Koskinen S, Esmarck B, Langberg H. Extracellular matrix adaptation of tendon and skeletal muscle. *J Anat*. 2006;208:445–50.
- Lakes R, Saha S. Cement line motion in bone. *Science*. 1979;204:501–3.
- Lees S, Eyre DR, Barnard SM. BAPN dose dependence of mature crosslinking in bone matrix collagen of rabbit compact bone: corresponding variation of sonic velocity and equatorial diffraction spacing. *Connect Tissue Res*. 1990;24:95–105.

- Leung VYL, Gao B, Leung KKH, et al. SOX9 governs differentiation stage-specific gene expression in growth plate chondrocytes via direct concomitant transactivation and repression. *PLoS Genet.* 2011;7(11):1–16.
- Marcus R, Feldman D, Kelsey J. Osteoporosis. New York: Academic; 1996.
- Martin RB, Pickett JC, Zinaich S. Studies of skeletal remodeling in aging men. *Clin Orthop Relat Res.* 1980;149:268–82.
- Miller LM, Vairavamurthy V, Chance MR, et al. In situ analysis of mineral content and crystallinity in bone using infrared micro-spectroscopy of the $\nu(4)PO(4)(3-)$ vibration. *Biochim Biophys Acta.* 2001;1527:11–9.
- Mullender MG, van der Meer DD, Huiskes R, Lips P. Osteocyte density changes in aging and osteoporosis. *Bone.* 1996;18:109–13.
- Nakashima T, Hayashi M, Fukunaga T, et al. Evidence for osteocyte regulation of bone homeostasis through RANKL expression. *Nat Med.* 2011;17:1231–4.
- Neville AC. Cuticle: organization. In: Bereiter-Hahn J, Matoltsy AG, Richards KS, editors. *Biology of the integument. I. Invertebrates.* Berlin: Springer; 1984. p. 611–25.
- Nyman JS, Ni Q, Nicoletta DP, Wang X. Measurements of mobile and bound water by nuclear magnetic resonance correlate with mechanical properties of bone. *Bone.* 2008;42:193–9.
- Olsen BR. New insights into the function of collagens from genetic analysis. *Curr Biol.* 1995;7:720–7.
- Oxlund H, Mosekilde L, Ortoft G. Reduced concentration of collagen reducible cross links in human trabecular bone with respect to age and osteoporosis. *Bone.* 1996;19:479–84.
- Parfitt AM. Misconceptions (1): epiphyseal fusion causes cessation of growth. *Bone.* 2002;30:337–9.
- Parfitt AM, Drezner MK, Glorieux FH, Kanis JA, Malluche H, Meunier PJ, Ott SM, Recker RR. Bone histomorphometry: standardization of nomenclature, symbols, and units. *J Bone Miner Res.* 1987;2(6):595–610.
- Reddi AH, Anderson WA. Collagenous bone matrix-induced endochondral ossification and hemopoiesis. *J Cell Biol.* 1976;69:557–72.
- Reisinger AG, Pahr DH, Zysset PK. Elastic anisotropy of bone lamellae as a function of fibril orientation pattern. *Biomech Model Mechanobiol.* 2011;10:67–77.
- Rho JY, Tsui TY, Pharr GM. Elastic properties of human cortical and trabecular lamellar bone measured by nanoindentation. *Biomaterials.* 1997;18:1325–30.
- Robling AG, Niziolek PJ, Baldridge LA, et al. Mechanical stimulation of bone in vivo reduces osteocyte expression of Sost/Sclerostin. *J Biol Chem.* 2008;283:5866–75.
- Ruppel ME, Miller LM, Burr DB. The effect of the microscopic and nanoscale structure and bone fragility. *Osteoporos Int.* 2008;190:1251–65.
- Russ JC, DeHoff RT. Practical stereology. 2nd ed. New York: Kluwer Academic/Plenum Publishers; 2000.
- Schaffler MB, Burr DB, Frederickson RG. Morphology of the osteonal cement line in human bone. *Anat Rec.* 1987;217:223–8.
- Schulz RM, Bader A. Cartilage tissue engineering and bioreactor systems for the cultivation and stimulation of chondrocytes. *Eur Biophys J.* 2007;36(4).
- Shirazi R, Shirazi-Adl A, Hurtig M. Role of cartilage collagen fibrils networks in knee joint biomechanics under compression. *J Biomech.* 2008;41:3340–8.
- Simon SR. Orthopaedic basic science. Rosemont, IL: American Academy of Orthopaedic Surgeons; 1994.

- Sissons HA. Experimental study of the effect of radiation on bone growth. In: Mitchell JS, Holmes BF, Smith CL, editors. *Progress in radiobiology*. Edinburgh: Oliver and Boyd; 1955. p. 436–52.
- Skedros JG, Holmes JL, Vada EG, et al. Cement lines of secondary osteons in human bone are not mineral-deficient: new data in a historical perspective. *Anat Rec A Discov Mol Cell Evol Biol*. 2005;286:781–803.
- Stock SR, Yuan F, Brinson LC, Almer JD. Internal strain gradients quantified in bone under load using high-energy X-ray scattering. *J Biomech*. 2011;44:291–6.
- Stover SM, Pool RR, Martin RB, Morgan JP. Histological features of the dorsal cortex of the third metacarpal bone mid-diaphysis during postnatal growth in thoroughbred horses. *J Anat*. 1992;181:455–69.
- Tai K, Dao M, Suresh S, Palazoglu A, Ortiz C. Nanoscale heterogeneity promotes energy dissipation in bone. *Nat Mater*. 2007;6:454–62.
- Tatsumi S, Ishii K, Amizuka N, et al. Targeted ablation of osteocytes induces osteoporosis with defective mechanotransduction. *Cell Metab*. 2007;5:464–75.
- Tavassoli M, Yoffey JM. *Bone marrow structure and function*. New York: Liss; 1983.
- Wallace JM, Orr BG, Marini J, Banaszak Holl MM. Nanoscale morphology of type I collagen is altered in the Brl mouse model of osteogenesis imperfect. *J Struct Biol*. 2011;173:146–52.
- Wang X, Shen X, Li X, et al. Age-related changes in the collagen network and toughness of bone. *Bone*. 2002;31:1–7.
- Xiong J, Onal M, Jilka RL, et al. Matrix-embedded cells control osteoclast formation. *Nat Med*. 2011;17:1235–41.
- Xu J, Rho JY, Mishra SR, et al. Atomic force microscopy and nanoindentation characterization of human lamellar bone prepared by microtome sectioning and mechanical polishing technique. *J Biomed Mater Res A*. 2003;67:719–26.
- Zaue R, Yeni YN, Bay BK, Dong XN, Fyhrie DP. Comparison of the linear finite element prediction of deformation and strain of human cancellous bone to 3D digital volume correlation measurements. *J Biomech Eng*. 2006;128:1–6.
- Ziopoulos P, Currey JD, Hamer AJ. The role of collagen in the declining mechanical properties of aging human cortical bone. *J Biomed Mater Res*. 1999;45:108–16.

3

Growth, Modeling and Remodeling of Bone

[H]ow curious the hand of Heaven has been in the Framing and Ordering of this Timber-work of our Bodies.

Osteologia Nova, Clopton Havers (1657–1702) (Havers [1691](#))

...there is a perpetual Waste and Renewal of the Particles which compose the solid Fibres of Bones;...the Addition from the Fluids exceeding the Waste during the Growth of the Bones; the Renewal and waste keeping pretty near par in adult middle Age; and the waste exceeding the Supply from the Liquors in old Age; as is demonstrable from their Weight. . .

Anatomy of the Human Bones, Alexander Monro (primus), 1697–1767 (Monro A. The anatomy of the human bones and nerves and lacteal sac and duct. Dublin [1776](#))

Good mathematical models don't start with the mathematics but with a deep study of certain natural phenomena.

Bull. Amer. Math. Soc., Stephen Smale, 1930—(Smale [1978](#))

Abstract The preceding chapter provided a general description of the structure and biology of bone, cartilage, tendons and ligaments. In this chapter, we will learn how bones grow, are sculpted by modeling and continuously renewed by remodeling, and repair when fractured. We will also learn how these processes in bone can be quantified, and what assumptions we make when attempting to put numbers to dynamic processes.

Keywords Growth • Modeling • Remodeling • Fracture • Histomorphometry

3.1 Introduction

Three processes account for the amount of bone we have, and for where it is found in our skeletons. Growth is defined as the accrual of bone mass through processes of bone formation during development. It does not account for the shapes of the bones, except for their length, and does not

involve resorption of tissue. Modeling is a term used to define either bone formation or bone resorption, but not both, at a given site. Modeling modifies the structural geometry of the whole bone; this is really its primary function. Remodeling, on the other hand, is defined by the sequential resorption and subsequent formation of bone at a single location. Although remodeling may modify the geometry of the bone as well, its primary function is to maintain both the mass and the shape of the bone, and to repair small amounts of damage within the bone tissue.

Remodeling is the only one of these processes that occurs significantly throughout one's lifetime. Growth is essentially restricted to children and modeling slows substantially in adults. Fracture healing occurs occasionally (or never). Because bone remodeling occurs throughout life, it plays a dominant role in determining the structure of most of the tissues in the skeleton, most of the time. Consequently, remodeling is a primary determinant of the mechanical properties of bone, their resistance to fatigue failure, and their ability to function in a changing mechanical environment.

3.1.1 *Growth*

In the previous chapter, we described the way in which most bones grow in length. This kind of bone formation, in which cartilage is formed first, calcified, and then replaced by bone, is known as endochondral ossification. Another mode of bone formation, in which osteoblasts make bone directly, is called intramembranous ossification. So-called “flat bones” (skull, scapulae, pelvis) are formed in this way. The word “intramembranous” refers to the fact that in the embryo, this kind of ossification occurs adjacent to membrane-like layers of mesenchymal cells that differentiate into osteoblasts. The same sort of direct formation of bone by osteoblasts occurs in periosteal modeling, where it is still called “intramembranous bone formation” because the osteoblasts and their precursor cells are found in the deep layer (cambium) of the periosteal membrane. The direct formation of bone onto any surface, whether as part of a membrane or not, is said to be appositional. The four surfaces upon which apposition of bone can occur—periosteal, endosteal, trabecular, and Haversian—are the four skeletal envelopes (Fig. 2.2).

3.1.2 *Modeling Versus Remodeling*

As the long bones develop, their shafts must grow in diameter as well as length. We have seen how growth in length throughout childhood is accomplished by continuation of endochondral ossification in the growth plate. Growth in diameter is accomplished by periosteal intramembranous ossification. However, as growth occurs, it is not enough just to add material to make the bone organs longer and larger. The bones must also be shaped in

various ways. That means that bone must be removed in some places while it is added in others. This sculpting, involving osteoclastic activity (bone resorption) in some places and osteoblastic activity (bone formation) in others, has become known as modeling. In addition, fatigue damage must be repaired inside bones and their internal architecture must be adjusted to varying load conditions, both while the skeleton is developing and throughout the remainder of one's life. This repair involves removal and replacement of bone in particular places by the coupled actions of osteoclasts and osteoblasts working at the same site. This replacement process has become known as remodeling.

Thus, modeling and remodeling refer to the actions of osteoblasts and osteoclasts in reshaping and replacing portions of the skeleton. They are distinguished from one another in several ways:

Modeling involves independent actions of osteoclasts and osteoblasts.

Remodeling involves sequential, "coupled" actions by the two types of cells.

Modeling results in change of the bone's size, shape, or both. Remodeling does not usually affect size and shape.

The rate of modeling is greatly reduced after skeletal maturity. Remodeling occurs throughout life, although it too is substantially reduced after growth stops.

Modeling at a particular site is continuous and prolonged, whereas remodeling is episodic, with each episode having a definite beginning and ending.

These terms were coined by Frost. Be aware, however, that many orthopedists and bone scientists refer to both modeling and remodeling activities, as defined here, as "remodeling."

3.1.2.1 Modeling

Modeling is necessary because the longitudinal growth process does not produce a bone with the correct shape for an adult, and because each growing child loads his or her skeleton in a somewhat different way, requiring each skeleton to be "customized." Modeling may be resorptive or formative. Examples of modeling include the following.

Metaphyseal modeling to reduce bone diameter: As some long bones grow longer, the diameter of the cylinder of bone created beneath the growth plate must be reduced to create the diaphysis. This change is accomplished by osteoclasts working continuously on the periosteal surface of the metaphysis to cut the shaft down to size (Fig. 3.1, top). This action is most important in bones with a widely flaring metaphysis such as the proximal tibia; in other bones, diametric growth of the physis is coordinated with longitudinal growth in such a way that removal of bone in the lower metaphysis is unnecessary (e.g., metatarsals).

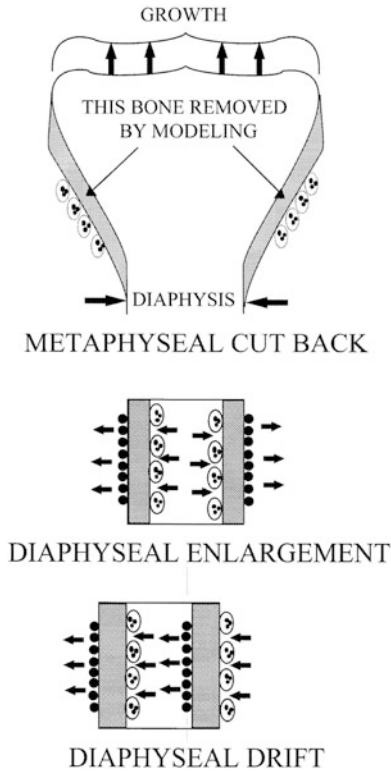


FIGURE 3.1. *Top:* Resorptive modeling beneath the growth plate to form the diaphysis from the metaphysis. *Middle:* Formative (periosteal surface) and resorptive (endosteal surface) modeling to enlarge the diaphysis. *Bottom:* Modeling to “drift” the diaphysis to the *left* (thereby altering diaphyseal curvature).

Diaphyseal modeling to increase bone diameter or alter curvature: As a child continues to grow, long bone shafts become larger in diameter, accomplished by slow, continuous addition of bone to the periosteal surface by osteoblasts. Simultaneously, bone is removed from the endosteal surface (Fig. 3.1, middle). In addition, the curvature of long bones is usually adjusted during growth. This is accomplished by increased formation and resorption on the sides of the bone such that the cross section “drifts” sideways relative to the ends of the bone (see Fig. 3.1, bottom). This phenomenon may increase or decrease bone curvature as required in specific bones. It also serves to correct the curvature resulting from a poorly reduced fracture.

Modeling of a flat bone: As a child grows, the bones of the skull must increase in size to accommodate the increasing size of the brain. This enlargement cannot be accomplished simply by addition of bone at the suture lines between the cranial bones because the plate's radius of curvature must also be adjusted. Modeling of these bones (resorption on the inner surface, formation on the outer surface) is required for proportionally correct increases in cranial size. Simultaneously, modeling transforms the facial features of a child into those of an adult. As one might imagine, similar processes must occur in the ribs to adjust the curvature of the rib cage as the individual's chest expands during growth.

3.1.2.2 Remodeling

Remodeling removes a portion of older bone and replaces it with newly formed bone. This process repairs microscopic damage and prevents accumulation of fatigue damage that could lead to fatigue fracture. It has been hypothesized that remodeling also mechanically "fine tunes" the skeleton to increase its mechanical efficiency.

3.1.3 *The A-R-F Sequence and BMUs*

Remodeling is accomplished by teams of osteoclasts and osteoblasts that work together in basic multicellular units, or BMUs (Fig. 3.2). A BMU consists of about ten osteoclasts and several hundred osteoblasts. There are three principal stages in a BMU's lifetime: activation (A), resorption (R), and formation (F). Activation occurs when a molecular or mechanical signal causes osteoclasts to form (by fusion of monocytes) and begin to remove bone somewhere on or in the skeleton. The osteoclasts resorb a volume of bone in the form of a ditch (on bone surfaces) or tunnel (in the midst of compact bone) about 200 μm in diameter that progresses along the surface or entirely within the cortex, moving at about 40 $\mu\text{m}/\text{day}$ in the longitudinal direction, but only about 5 $\mu\text{m}/\text{day}$ radially. After the tunnel has been excavated to its full diameter, osteoblasts are differentiated from mesenchymal cells over a period of several days and begin to replace the resorbed tissue. Formation is much slower than resorption. Completed BMU's render new osteons (Haversian systems) roughly 3–9 mm in length. If one were embedded in a human bone watching a BMU advance through a particular site, the time required to fully excavate the tunnel to its full diameter (i.e. the resorption period) would be about 3 weeks; while the time required to refill the resorption space with new bone (i.e. the formation or refilling period) would require about 3 months. The total remodeling period for any discreet location along the length of a progressing BMU is about 4–6 months. Because the bone is resorbed and formed in a packet, this has been termed the "quantum concept of bone remodeling" by Parfitt (1979).

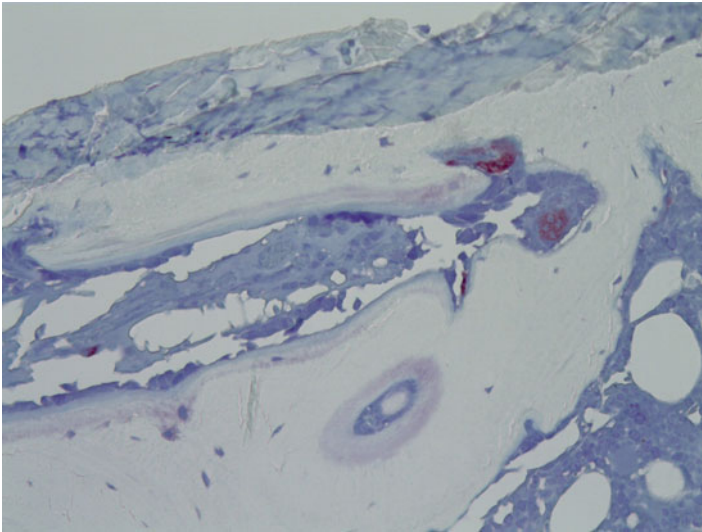


FIGURE 3.2. A photomicrograph of a cutting cone in bone, same orientation as in Fig. 3.3. Because the wall of the tube of new bone laid down within the excavated tunnel is about $80\text{ }\mu\text{m}$ thick, and the average accretion rate is about $1\text{ }\mu\text{m/day}$, the time to refill is roughly 80 days. That is, if it were possible to observe a bone cross section in vivo, it would take about 80 days for the initial resorption cavity to be converted into a completed osteon. Before refilling, about 30 days would have passed between the time when the osteoclasts initially began to open the resorption cavity and the appearance of the osteoblasts. Historically, these resorption and refilling times have been called sigma times because Frost originally used this symbol to represent them. Current nomenclature uses the term remodeling period in place of sigma (or formation period and resorption period in place of sigma-F and sigma-R). The total remodeling period is approximately 4–6 months in human cortical bone [Courtesy of Matthew R. Allen].

3.1.4 *The Six Phases of an Osteon's Lifetime*

The A-R-F process may be further divided into six separate phases, which are always sequential in normal bone remodeling.

3.1.4.1 **Activation**

Differentiated cells must be recruited from precursor cell populations before any bone resorption or bone formation can occur. It is important to recognize that in the purest sense this process occurs only at the point of origin of the BMU. The osteoclasts produced at this time appear to survive for weeks, and perhaps for the entire duration of the tunneling by the BMU. Thus, the “opening up” of a resorption space on a cross section is not actually “activation” because it does not involve the initial recruitment of osteoclasts, but rather the passage of osteoclasts through a particular plane of observation.

The original activation process occurred earlier; in either case, however, the time required is about 3–5 days.

3.1.4.2 Resorption

Newly created osteoclasts begin to resorb bone, moving longitudinally at a rate of about 40 $\mu\text{m}/\text{day}$ in a so-called cutting cone, an ellipsoidal surface about 200 μm in diameter and about 300 μm long. However, it is difficult to be certain where this region stops and the following region begins, so the length of each is debatable.

3.1.4.3 Reversal

The transition from osteoclastic to osteoblastic activity takes several days and results in a cylindrical space lying between the resorptive region and the refilling region. The length of this region may vary considerably, depending on the lag between the resorptive and formative phases. In a completed secondary osteon, the cement line coincides with the location of the bone surface during the reversal period; the cement line is also known as the reversal line. In humans, about 30 days are required for the resorption and reversal periods combined.

3.1.4.4 Formation

Osteoblasts appear around the periphery of the tunnel formed by the osteoclasts, and refilling begins. The osteoblasts lay down concentric lamellae at a decreasing rate as refilling progresses, but the average radial closure rate is 1–2 $\mu\text{m}/\text{day}$. Because the osteoclasts and osteoblasts require nourishment, the osteoblasts do not completely refill the tunnel, but leave a central passageway for a vascular loop to support the metabolism of the BMU and the osteocytes in the bone matrix, and to carry calcium and phosphate to and from the bone when necessary. This passageway is the Haversian canal, and it is typically 40–50 μm in diameter in humans. The formation phase in adult humans averages about 3–4 months.

3.1.4.5 Mineralization

Following deposition of the unmineralized organic matrix of bone, or osteoid, mineral is deposited within and between the collagen fibers. This process is delayed by a period of time known as the mineralization lag time, which is normally about 10 days. This delay is manifested by a layer of osteoid between the osteoblasts and mineralized bone. Once begun, approximately 60–70 % of the mineralization of osteoid occurs during the first 3 weeks (Parfitt 1983; Fuchs et al. 2008); this is called primary mineralization. The remainder of the mineral is added at a decreasing rate for about a year during secondary mineralization (Fuchs et al. 2008). The bone in new, incompletely mineralized osteons can exhibit very different mechanical properties from that in older osteons.

3.1.4.6 Quiescence

After the tunneling and refilling processes are completed, the osteoclasts disappear from the scene, and the osteoblasts become osteocytes or Haversian canal lining cells or disappear. A period of relative inactivity ensues during which the BMU and its associated cells carry on their mechanical, metabolic, and homeostatic functions.

Each of these six phases is location-, magnitude-, and rate specific, so that alterations in the intensity or timing of any one produce an altered histologic morphology characteristic of specific skeletal disorders. For example, an increased lag between the bone formation and bone mineralization stages produces the excessive amounts of osteoid found in osteomalacia (a disorder characterized by decreased mineralization of bone).

3.1.5 Remodeling Cycle Duration

Figure 3.3 depicts these phases in a schematic diagram. This figure also illustrates the relationship between the appearance of an osteonal BMU in longitudinal and cross sections and, in addition, the fact that the duration of events perceived by an observer in a particular bone cross section does not correspond to the duration of events that would be observed by someone moving along the bone with the BMU. For example, the moving observer would say that the resorption phase of the BMU lasts from the time that osteoclasts begin to excavate the tunnel until the tunneling ends, perhaps a centimeter down the bone shaft; this might take 200 days. A stationary observer at some arbitrary cross section along the way would say, however,

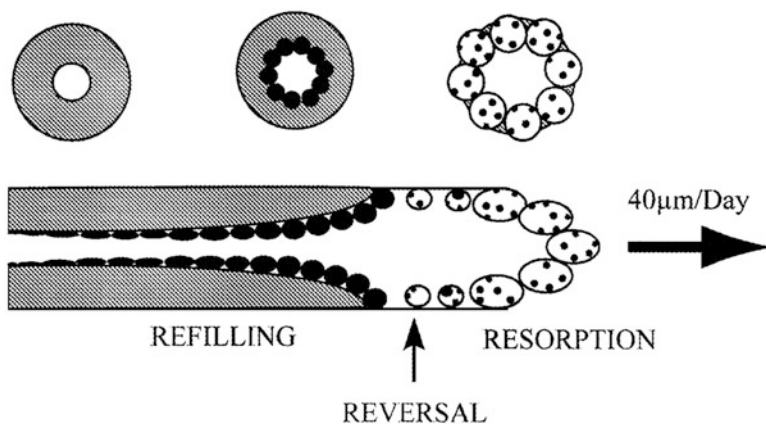


FIGURE 3.3. Schematic diagram of an osteonal BMU. Larger, multinucleated cells to *right* are osteoclasts; smaller cells shown in *black* to *left* are osteoblasts. Cross-sectional views are shown above. Not to scale.

that the time required for resorption is only the few days required for the osteoclasts to pass through the plane of the section and open a hole corresponding to the osteonal cement line. In virtually all discussion of osteonal BMU dynamics, time periods refer to those observed in cross sections, because it is such sections that are routinely observed and analyzed.

As one can see from this, there is a period of time during which bone has been resorbed, but new bone has not formed at that location. This is called the remodeling transient because in healthy bone it is a transitory phase that will eventually be resolved when new bone is formed at that location. This transient also creates a remodeling space, which is all of the space in bone that is undergoing processes of bone remodeling that have not yet been completed. This space has been estimated at 5–10 % of the bone volume. The size of the remodeling space is defined by the frequency with which new remodeling sites are started (the activation frequency, or Ac.f), the rate at which bone is removed or formed at those sites, and the overall length of time it takes for remodeling to be completed at those sites (the remodeling period).

Box 3.1 Biomarkers of Bone Remodeling

It is not always possible to obtain a piece of bone to measure the components of the bone remodeling process histologically. Even for animal studies, it may be desirable to follow changes in remodeling rates longitudinally during the course of an experiment, and beneficial—both for the animal and for the investigator’s budget—not to have to destroy more animals to do that. In these cases, the rates of bone resorption and formation can be estimated from blood or urine by measuring biochemical substances that are released from bone when it is resorbed, or that are produced by osteoblasts or osteoclasts during the bone remodeling process. This is an imperfect measure of bone turnover processes (as is histology as well, for different reasons that are described later) because it provides only a systemic picture of bone resorption and formation and is not location-specific. Moreover, many of the compounds are not specific to bone, but found in tissues and organs other than bone. Their appearance in blood or urine may reflect processes occurring at those sites. Moreover, the measurement does not consider the fact that some substances may be reabsorbed without entering the blood stream or going through the kidneys. Nevertheless, measuring these compounds can provide a non-invasive means to screen for or diagnose disease, or to monitor changes during an experimental challenge.

The most commonly used measure of bone formation is osteocalcin (also sometimes called bone Gla-protein), which is produced by osteoblasts, but also released from bone during the process of resorption by osteoclasts. In part because of this, and because of the coupling between resorption and formation, osteocalcin is really a marker of overall bone turnover. It has recently been proposed that osteocalcin production by osteoblasts enhances insulin secretion and so plays a role in glucose tolerance and the body’s energy metabolism (Ferron et al. 2010; Fuzele et al. 2010). Alkaline phosphatase is another often used marker of bone formation. However, alkaline phosphatase is also produced in the liver, intestine and kidney, and so

bone-specific alkaline phosphatase (BSAP) is a better reflection of bone specific processes. Two pro-peptides of the Type I collagen molecule, one from the C-terminal end (PICP) and one from the Amino-terminal end (PINP) are also now used. These are produced by both osteoblasts and by fibroblasts, and so in part the picture they give may be altered by processes occurring in soft tissues. All these bone formation markers can be measured from blood serum.

The most useful resorption markers are those that are derived from parts of the collagen fiber or its cross-links when it is degraded by osteoclastic resorption. The difficulty with some of these, for example hydroxyproline and the cross-link pyridinoline (PYD), is that they are found not just in bone (Type I collagen), but also in cartilage (Type II collagen) and/or skin and other tissues that contain different types of collagen. Deoxypyridinoline (DPD) is another cross-link, but is more specific to bone (although it is also found in dentin). The most widely used resorption markers, therefore, are those more specific to Type I collagen, and that come from the C- and N-terminal telopeptides of the Type I collagen molecule (CTx and NTx). These can be measured from urine or serum, and although they are found in other tissues and organs that have Type I collagen (e.g., skin), the highest contribution is from bone. Tartrate-resistant acid phosphatase (TRAP) is an enzyme produced by osteoclasts during the process of resorption. It is measured from blood plasma or serum, and is also very commonly used, especially the isoform TRAP5b.

Although attempts have been made to use markers to evaluate the balance between resorption and formation, these attempts have not been very successful, in part because of the inherent variability and non-specificity of the markers, but probably also because of temporal differences in the processes of resorption and formation.

3.1.6 Bone Structural Units Are Produced by Remodeling

The packet of new bone produced by a BMU is called a bone structural unit, whether the remodeling occurs in cortical or trabecular bone or on an endosteal or periosteal surface. Thus, a secondary osteon in cortical bone is a bone structural unit, as is the hemi-osteon in cancellous bone.

3.1.6.1 Coupling and Balance

Because the processes of resorption and formation are tightly linked within the BMU, they are said to be coupled. This concept refers only to the relationship of resorption and formation, and the cross-talk between osteoclasts and osteoblasts that must occur to coordinate the remodeling unit. Coupling between bone formation and bone resorption has made it extremely challenging to direct therapies that alter either side of the resorption-formation equation without affecting both sides simultaneously. Balance, on the other hand, refers to the relative amounts of bone that are resorbed or formed. If more bone is resorbed than formed, the bone balance is negative (i.e., bone is lost) and vice-versa if the balance is in favor of formation. These terms are often confused and incorrectly used

interchangeably, but concern very different aspects of the remodeling scheme. For instance, resorption and formation can be both coupled and unbalanced, as in the example above. However, there are some situations in which it is not clear if the processes are uncoupled or imbalanced. One of these would be the resorption that occurs along the endocortical surface of bone. This bone may eventually be replaced (suggesting coupling), but it may require a longer period than normal to complete.

3.1.7 *Remodeling on the Different Bone Envelopes*

3.1.7.1 **Osteonal Remodeling**

When a BMU tunnels through compact bone, it creates a secondary osteon. Figure 3.2 shows the leading portion of an osteonal BMU in longitudinal section, tunneling to the right. A few osteoclasts may be seen on the resorbing surface. At left, osteoblasts and osteoid line the refilling surfaces. The “head” or front end of a BMU contains a capillary “bud” to supply nutrients and remove byproducts created by working osteoclasts and osteoblasts, and probably to supply additional progenitor cells as required. Because osteonal BMUs become isolated deep within the cortex, this vascular supply must be maintained. Therefore, the tunnel cannot be entirely refilled, and each BMU leaves a new Haversian canal in the bone. Within each Haversian canal are two capillaries: a “supply” and a “return” vessel. These vessels connect with the vasculature in the medullary canal or on the periosteal surface. Haversian canals also contain nerve fibers that produce peptides that likely play a regulatory role in bone homeostasis.

In human adults, osteonal BMUs replace about 3 % of compact bone each year. Figure 3.4 shows how the rate of remodeling changes with age in the human rib. It is very high in children, is reduced in young adults, then rises and falls again in older individuals. (The gender of the individuals in the graph was not reported, but the rise after age 50 probably results from the effects of menopause in women; estrogen inhibits bone remodeling.)

3.1.7.2 **Trabecular Remodeling**

Trabeculae are remodeled similarly, except that the BMUs work on their surfaces, digging and refilling trenches. Because a BMU is about the same diameter as a trabecula, it cannot usually tunnel longitudinally within a trabecular strut without constantly breaking out along its sides. Therefore, BMUs trench out “half-moon” shaped erosion pits, and refill these with new bone, creating a hemi-osteon (Fig. 2.9b). BMUs replace human adult trabecular bone at a relatively high rate, about 25–30 % each year. In fact, the rate of bone turnover varies widely throughout the skeleton. Table 3.1 demonstrates this using data from various skeletal sites in young adult dogs. It can also be seen that the remodeling rates in dogs are much higher than in humans.

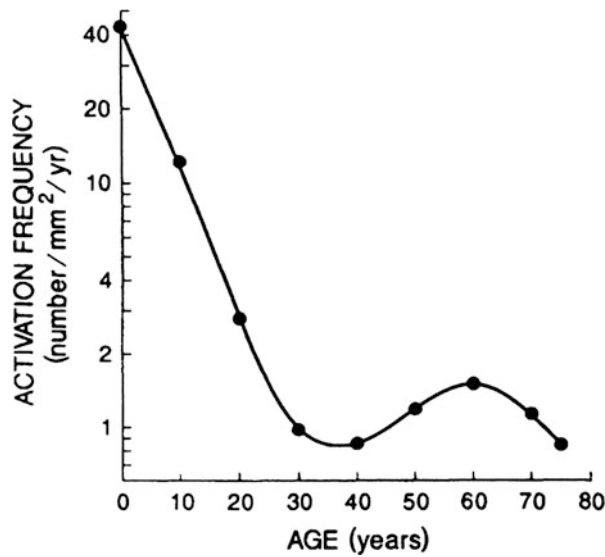


FIGURE 3.4. BMU activation rate vs. age for human ribs (Frost 1964).

TABLE 3.1. Cancellous bone turnover rates in young adult dogs

Site	Turnover rate (%/year)
Lumbar vertebrae	205
Thoracic vertebrae	167
Cervical vertebrae	121
Mandible	105
Skull	60
Calcaneus	120
Proximal humerus	174
Proximal femur	138
Proximal radius	127
Proximal tibia	112
Carpals	31

After Kimmel and Jee (1982). These values are approximate. Much more detailed information can be found in the original reference, in some cases divided into metaphyseal vs. epiphyseal regions, and specified by gender

3.1.7.3 Endosteal and Periosteal Remodeling

Remodeling also can occur on endosteal (endocortical) and periosteal surfaces. Endosteal remodeling is similar to remodeling along trabecular surfaces, although the erosion sites are often not filled completely, or at all,

causing a life-long gradual expansion of the marrow cavity once growth is complete. Long bones expand radially with age in adults as well as children. This expansion is the result of direct bone apposition to this surface (ie, modeling in formation mode). Although periosteal remodeling was dismissed at one time as non-existent or not important, it is now clear that this does occur in the long bones with increasing frequency with age (Orwoll 2003; Bliziotis et al. 2006). In adults, this remodeling always results in scalloping of the periosteal surface and some bone loss.

In some cases in which bone is being treated with anabolic, or bone forming, agents such as parathyroid hormone, the trabeculae can become quite thick. This could eventually cause a problem with fluids and nutrition reaching the cells in the deepest layers of the trabecular strut, but BMUs tend to “tunnel” through overly thick trabeculae, creating two trabeculae where one existed before. This has the additional advantage of creating four surfaces for bone formation instead of two, allowing bone to be formed much more quickly. Besides these advantages, tunneling also alters the architecture of the cancellous structure in a way that increases trabecular connectivity and improves bone strength.

3.1.7.4 Remodeling and Cartilage

Surfaces covered by cartilage that has not calcified cannot be resorbed by osteoclasts, and cannot serve as a working surface for osteoblasts, so they cannot be modeled or remodeled. The same is considered to be true of surfaces covered by osteoid. Consequently, some tuberosities on bone where muscles attach, e.g., the tibial tuberosity, protrude not because of additional bone or cartilage formation on their surfaces, but because of active resorption at their base.

3.1.7.5 Cellular Events in Modeling and Remodeling

Modeling and remodeling involve some combination of four basic cellular events: proliferation of precursor cells, differentiation to the mature phenotype, migration to the surface where resorption or formation will occur, and matrix synthesis (osteoblasts) or matrix resorption (osteoclasts). Some of these events are governed by hormones, neuropeptides or other signaling molecules originating outside the skeleton (e.g., parathyroid hormone [PTH] or neuropeptide Y). At a local level, most of these complex changes are controlled by proteins called cytokines and others called growth factors, which are produced by cells present in and around bone surfaces (e.g., bone morphogenic proteins [BMPs], osteoprotegerin [OPG]). Bone cells produce literally dozens of cytokines, growth factors, and noncollagenous proteins that regulate processes related to the formation, mineralization or dissolution of bone matrix itself. Some cytokines affecting bone remodeling are synthesized by bone marrow cells.

3.1.7.6 Skeletal Envelopes and Bone Loss

Trabecular and endosteal surfaces are bathed in marrow, and remodeling that occurs on them is usually characterized by an excess of resorption relative to formation. This endo-trabecular deficit of bone replacement during remodeling is the reason for normal, age-related bone loss that occurs in men and women after age 30–35, and for the more rapid loss of bone in women after menopause. The periosteal envelope experiences formative modeling which expands the outer diameter of the bone and increases bone strength; such processes partially compensate for involutional (i.e., senile) bone loss. However, as mentioned above, remodeling can also occur on this surface in older people. Like the endosteal and trabecular surfaces, the balance is usually in favor of resorption and bone loss, and in the older individual, contributes to involutional loss. The osteonal, or Haversian, envelope also contributes to involutional bone loss. A Haversian canal must be created for each new osteon, increasing the porosity of compact bone until a level of porosity is attained at which each new BMU removes an existing Haversian canal before creating a new one.

The reason why a deficit in formation occurs during remodeling on surfaces adjacent to marrow is unknown. There is, however, an association between a remodeling deficit and the amount of fat in the marrow (Minaire et al. 1984; Martin et al. 1990). With age, part of the deficit is probably indicative of the smaller number of stem cells available to differentiate into bone forming osteoblasts.

3.1.7.7 Regional Acceleratory Phenomenon

A sudden transient increase in the rate of remodeling sometimes occurs in one or more bones. This is called a regional acceleratory phenomenon, or RAP. The cause is usually trauma to the bone. When this happens, a transient osteoporosis occurs because there is a lag between resorption and formation in each BMU. The osteoporosis will resolve itself over a remodeling period (about 4 months) as the additional BMUs complete their refilling.

3.2 Fracture Healing

Fracture is the most obvious sign that the mechanical functional capacity of the skeleton has been exceeded. It may be argued that much of skeletal physiology is directed at preventing fracture. Even so, fractures are relatively common, and usually occur because of trauma: an accidental overload substantially exceeds the normal range of loading to which the bone has adapted during its growth and development.

Pathologic fractures are caused by normal loading of a bone weakened by disease. The most common pathologic fractures are caused by osteoporosis in the elderly. Osteoporotic fractures are more common in women than men.

Bone tumors are another important cause of pathologic fractures. The tissue in a bone tumor is variable in its mechanical properties, but it will almost always be different than the surrounding bone and therefore serve as a stress concentrator. Removal of a bone tumor obviously also places the patient at risk for fracture. Predicting and preventing pathologic fractures is an important issue in orthopaedic medicine.

Once a fracture occurs, how does it heal? Today, people seek medical attention for the simplest of fractures, but the basic process of fracture healing obviously evolved as a natural phenomenon. Healed fractures are commonly seen in archaeological skeletons of prehistoric people, and in wild animals. Therefore, fracture healing must be a self-controlling process that ordinarily proceeds on its own to reunite the broken bone.

3.2.1 *Basic Concepts*

As we have seen, in the embryo and growing child, bone is formed directly by osteoblasts (intramembranous bone formation) and indirectly by the calcification of cartilage and replacement by bone (endochondral ossification). In the healthy adult, the latter process is absent, and bone is formed only by osteoblasts on existing bone surfaces. When a fracture occurs, however, adults as well as children are able to make bone by both intramembranous and endochondral osteogenesis, leading to much more rapid repair of the damaged bone. Unlike skin and other soft tissues, wound healing in bone does not produce a permanent scar. Eventually, the initial repair tissues are replaced by normal lamellar bone. Therefore, while the woven bone and calcified cartilage in the healing fracture can be considered “scar tissue,” they are temporary, and the tissues in the healed bone are ultimately indistinguishable from those in the original bone.

A fracture that fails to heal is called a non-union. Frequently, non-unions produce a jointlike cartilaginous structure called a pseudoarthrosis (literally, ‘false joint’); these pathologic joints are lubricated with synovial fluid. A fracture that does not heal in the expected length of time, but which is still progressing, is called a delayed union. The length of time required to heal a particular fracture depends primarily on its anatomical site. Ribs, clavicles, and other bones of the torso heal in a few weeks; appendicular bones generally require several months (Table 3.2).

3.2.2 *Important Tissues in Fracture Healing*

Two tissues provide the cells responsible for the healing of a typical fracture: the periosteum and the marrow of the injured bone. The periosteum has two layers: a fibrous outer layer, and a highly cellular cambium or inner layer. Ordinarily, the osteoblasts of the adult periosteum produce lamellar bone at a relatively slow rate, so that long bones expand very slowly with age.

TABLE 3.2. Typical healing times for common fractures

Bone	Typical healing time (weeks)
Distal radius	6
Humeral shaft	12
Tibial shaft	18
Femoral neck	24

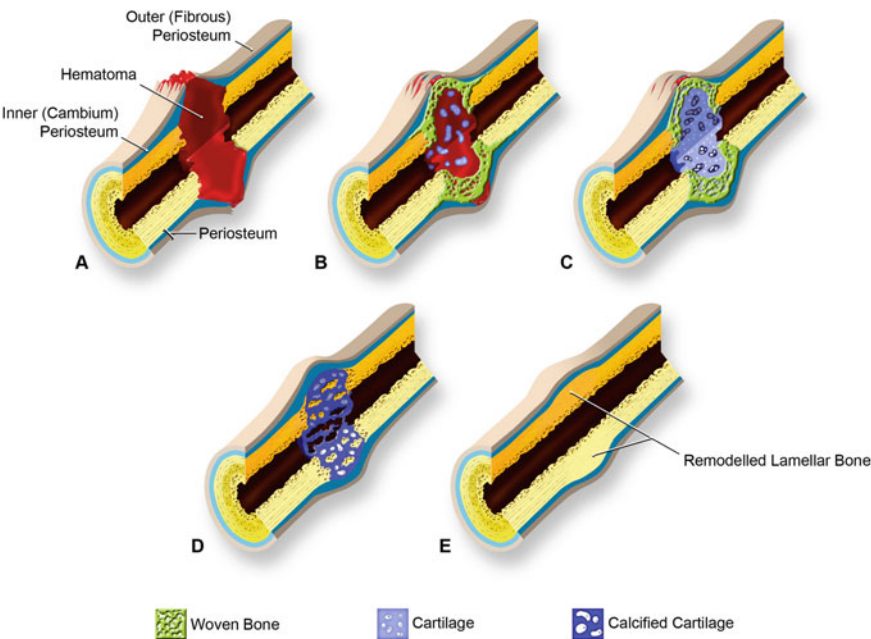


FIGURE 3.5. Stages of a healing fracture. (a) Initially a hematoma forms at the fracture site, bringing cells to the site for repair. (b) These cells rapidly form a bridge across the fracture gap composed of woven bone. (c) A cartilage callus forms to further stabilize the fracture. (d) The cartilage callus will gradually calcify through a process of endochondral ossification, further stabilizing the fracture. (e) The calcified cartilage gradually remodels to re-establish the integrity of the marrow cavity and both periosteal and endocortical surfaces will model to re-establish the original shape of the bone (used with permission from Burr and Guillot 2012).

However, even a slight amount of trauma stimulates the periosteum to produce more osteoblasts, which manufacture large amounts of woven bone in a short amount of time. In the case of a fracture, this periosteal woven bone takes the form of trabeculae, which arch over the fracture line and fuse to bridge the gap and make what is known as the periosteal callus (Fig. 3.5).

The endosteal surfaces of bones are not actually covered by a membrane. Instead, they are lined with bone lining cells; this is true whether or not the endosteal space contains trabecular bone. In places, the endosteal surface may be experiencing ordinary remodeling, but most of the adult endosteal surface is quiescent. The bone lining cells may be stimulated to initiate bone remodeling by various factors, including mechanical ones, but such remodeling does not play a role in the initial fracture healing response. Instead, trauma to the marrow stimulates the formation of the medullary version of the periosteal callus. It has been shown that, even without fracture of the cortex, mechanical disruption of the marrow activates the differentiation of osteoblasts and the formation of woven bone within the medullary canal (Amsel et al. 1969) (this bone formation is a prerequisite to the regeneration of the hematopoietic function of the marrow, and so serves a dual role when fracture occurs). The periosteal callus tends to be more obvious, but the medullary callus is usually considered more important in achieving initial union.

Obviously, the jagged ends of the broken bone, as well as other trauma to the fracture site, can be expected to break many blood vessels. The resulting hematoma was once thought to be the primary source of the cells responsible for fracture healing. It is now thought that most of these cells come from the marrow and periosteum, but under the influence of inflammatory mediators produced in the hematoma.

3.2.3 *Three Biological Phases*

Healing of a fracture may be divided into inflammatory, reparative, and remodeling phases which, taken together, form a self-driven sequence that leads to a healed fracture.

3.2.3.1 **Inflammatory Phase**

This initial phase serves to immobilize the fractured bone and activate the cells responsible for repair. Immobilization is promoted by pain and by swelling, which in many cases serves to hydrostatically splint the fracture. In addition to hematoma formation, this phase is characterized by vasodilation, serum exudation, and infiltration by inflammatory cells. It lasts 3–7 days.

It is important to realize that fracture repair is not accomplished by existing osteoblasts and chondroblasts but depends on the creation of an entirely new work force of cells from a relatively small population of stem cells. The mobilization of this work force requires a variety of chemical mediators and ancillary cell types that are poorly understood. However, it is clear that this process is similar to other biologic cascades. Once set in motion, each step is predicated on what happened earlier, and if a mistake is made, the process may make a wrong turn rather than stopping, with no way

to recover the conditions at the point of the error. It follows that many if not most non-unions or delayed unions are caused by a defect in the initial mobilization of the repair process, and once the problem becomes apparent, there is no way to correct it by manipulating the current group of cells. This is consistent with the fact that the most effective treatment for non-unions is to replace the tissues in the fracture site with a bone graft, which essentially initiates an entirely new repair process. Alternatively, electrical signals (see following) may be able to start a new cascade.

Interruption of the normal vascular supply to the bone by the fracture also results in death of the osteocytes in the bone matrix. The significance of this is not clear, but it may initiate the remodeling activity that is important in the last stage of fracture healing.

3.2.3.2 Reparative Phase

In this phase of healing, which lasts about 1 month, the periosteal and medullary calluses are formed by osteoblasts from the periosteum and marrow. The stem cells for medullary osteoblasts are thought to be pluripotential mesenchymal cells in the marrow stroma. These cells are also stem cells for fibroblasts, chondroblasts, and some marrow cells. Mesenchymal cell proliferation and differentiation are accompanied by intense vascular proliferation. The resulting osteoblasts produce woven bone at a high rate.

The fracture site exhibits some degree of endochondral ossification as well as direct bone formation by osteoblasts. This is thought to be related to inadequate blood supply in some portions of the fracture region. In the absence of adequate vascularization, stem cells are thought to differentiate into chondroblasts rather than osteoblasts. An important contributing factor to inadequate vascularization can be insufficient immobilization of the fracture fragments. Alternatively, the nature of the stresses at a site within the callus may affect the differentiation pathway (Carter et al. 1988). In any event, the chondrogenic regions of the callus, which tend to be near the fracture line, usually calcify and are replaced by lamellar bone, just as the woven bone regions are. If this does not happen, a pseudoarthrosis may develop.

Regardless of how much of the callus is cartilaginous, it takes a week or two for the woven bone or cartilage to accumulate substantial mineral. This early, rather compliant callus is known as the provisional callus. As calcification proceeds, the callus material becomes increasingly rigid; it is then referred to as the bony callus. The rigidity of the callus is enhanced by the fact that it has a larger cross-sectional area (and moment of inertia, see Sect. 7.3) than the original bone cortex. This means that even though the material in the bony callus is less rigid (and strong) than cortical bone, the rigidity (and strength) of the callus structure can equal or exceed that of the intact bone. When this occurs, bony union has been achieved, completing the reparative phase.

3.2.3.3 Remodeling Phase

Once bony union has been achieved, the broken bone is approximately as strong as the intact bone was, but often it has greater mass than the original bone, and is therefore less mechanically efficient. Modeling and remodeling of the fracture site gradually restore the original contour and internal structure of the bone, although radiographic evidence of the fracture may persist for many years. As the bone remodels, it is able to maintain its strength with less material than in the earlier phases of healing, thereby increasing its mechanical efficiency. The medullary and periosteal calluses are removed, and the remaining material (woven bone or calcified cartilage) is replaced by secondary lamellar bone (cortical or trabecular, as the site dictates). If the fracture has healed with incorrect angulation, this may be partially corrected as well by modeling drifts. In general, modeling and remodeling proceed much more rapidly in children than in adults, and this applies to the remodeling phase of fracture healing as well.

In considering the capacity of a malaligned healed fracture to correct itself, there are two important geometric principles to remember. First, an angulation in the plane of motion of the adjacent joint corrects itself better than one out of the joint plane. Second, rotational errors do not correct themselves very well.

Box 3.2 Clinical Assessment of Healing Fractures

It is very important for the physician treating a fracture patient to be able to judge the strength and stiffness of the healing bone. If the cast is removed before the fracture is mechanically competent to bear the loads and moments applied to it, there is risk of refracture or malunion. On the other hand, if the cast is left on too long, there is risk of excessive disuse osteoporosis and contracture of soft tissues about the adjacent joints, not to mention the inconvenience and expense of needless immobilization. The usual clinical methods to assess fracture healing are inspection of radiographs and manipulation to see how stiff the healing fracture feels; both of these are subjective and imprecise. Several groups of investigators have experimented with methods to assess fracture healing in a more quantitative way.

If the fracture has been fixed with an external fixation frame, application of a strain gauge to the fixation frame can allow bone stiffness to be estimated (Burny et al. 1984; Cunningham et al. 1989). A second method for measuring the mechanical characteristics of a healing fracture involves the propagation of sound or stress waves through the tissues. The velocity of sound in an elastic bar is $c = (E/\rho)^{1/2}$ where E is the elastic modulus and ρ is the density of the tissue. The resonant frequency of a vibration in an elastic bar depends on the ratio of c to the length of the bar. When a vibration traveling along the bar reaches a segment having different material properties, a portion of the wave is reflected backward. The magnitudes of the transmitted and reflected waves depend on the ratio of the acoustic impedance, $z = c \rho$, of the two materials. These dependencies on the elastic modulus of the tissues are potential means for probing the stiffness of the fracture callus. For example, the velocity of an ultrasonic wave across the fracture site can be measured (Gerlanc et al. 1975). As the fracture heals, the wave velocity increases; healing can be assessed by comparison with the contralateral,

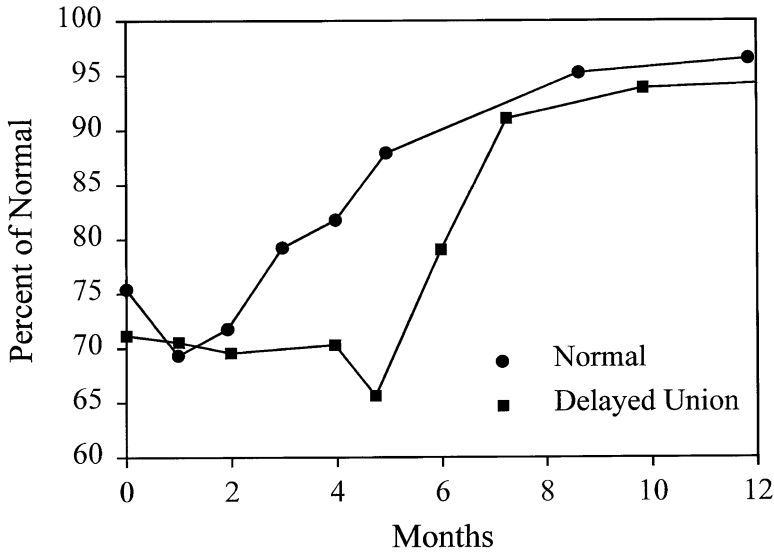


FIGURE 3.6. Sound velocity, normalized by the values in the contralateral bone, measured through the tissues in a healing fracture is plotted against healing time. Examples shown represent human tibias with normal (*circles*) and delayed (*squares*) healing (Gerlanc et al. 1975). With kind permission of Springer Science+Business Media.

uninjured bone (Fig. 3.6). Alternatively, the resonant frequency can be determined by tapping the fractured bone with a hammer (Collier and Donarski 1987), or vibrating it over a range of frequencies with a shaker. Another approach is to measure the attenuation of the impulse as it crosses the fracture site. These methods are most easily applied to the ulna and tibia, which are subcutaneous along at least one margin. Bones that are fully enveloped in muscle (e.g., the femur) are too inaccessible for these vibration methods.

Although each of these methods is capable of probing the mechanical attributes of a healing fracture, in practice they are too cumbersome and imprecise to have come into general clinical use and their applications have been limited to research. A fast, noninvasive, and easy method to measure bone stiffness *in vivo* remains to be invented. However, recent development of a micro-indentation system that can be used *in vivo* in humans may provide a relatively easy method for assessing the stiffness of a healing fracture (Diez-Perez et al. 2010; Bridges et al. 2012; Allen et al. 2015).

3.2.4 Four Biomechanical Stages

A standard way to study fracture healing in experimental animals is to mechanically test the healing bone in torsion and compare the result with the stiffness of an intact bone. When this was done with rabbit tibial fractures (White et al. 1977), four stages of healing were identified (Fig. 3.7).

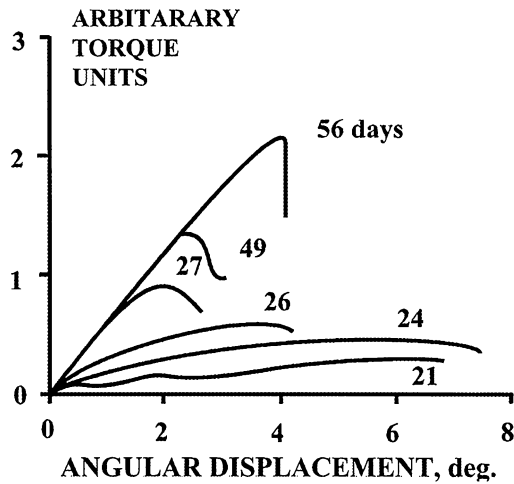


FIGURE 3.7. Torsional stiffness of healing rabbit fractures as a function of days of healing (White et al. 1977).

3.2.4.1 Stage I

Virtually no stiffness is seen, and failure occurs through the original fracture line. This stage corresponds to the inflammatory and early reparative (i.e., provisional callus) phases of healing. (Days 1–26 in the rabbit tibia.)

3.2.4.2 Stage II

Substantial stiffness is now encountered, but failure still occurs through the original fracture line. This stage corresponds to the middle of the reparative phase when the callus has become bony but still is not so large or calcified as it will become. (Days 27–49 in the rabbit tibia.)

3.2.4.3 Stage III

There is no further increase in stiffness, but failure is now partially through the original fracture line and partially through intact bone. (Days 49–56 in the rabbit tibia.)

3.2.4.4 Stage IV

Failure occurs through the intact bone rather than the callus at the fracture site. (After day 56 in the rabbit tibia.)

Stages III and IV correspond to the later portion of the reparative phase and the early remodeling phase, when the full bony callus exists but its mineral content and strength are still increasing. Another interesting observation from these mechanical studies of healing fractures is that there is typically a slight reduction in the strength of the fracture callus during

stage IV of healing. It is reasonable to think that this corresponds to the beginning of the remodeling phase, when the resorptive portion of BMU activity has begun but substantial refilling of the osteonal cavities has yet to occur.

3.2.5 *Stability of Fixation*

The importance of achieving good stability or immobilization of a fracture depends very much on the anatomical site. For example, ribs heal well even with a great deal of motion—rib non-unions are unknown. On the other hand, tibias and other large bones are very prone to non-union as the result of excessive motion of the fragments. This knowledge must be balanced against the fact that some motion is beneficial: there seems to be an inverse relationship between the amount of motion and callus size, up to a point that has not been defined. In the same vein, applying loads across the fracture (i.e., weight-bearing) serves to strengthen the callus, but obviously too much weight can lead to refracture or excessive motion and non-union. There are no clear rules about how much weight-bearing or motion are safe.

Box 3.3 The Rapid Loading Torsion Tester

The torque-deflection curves used to illustrate the four biomechanical stages of fracture healing in Fig. 3.7 were obtained using a device known as the Rapid Loading Torsion Tester (Fig. 3.8). An engineer named Albert Burstein and an orthopaedic surgeon named Victor Frankel developed this apparatus at Case Western Reserve University in the late 1960s. The classic textbook, *Orthopaedic Biomechanics*, which was written by these two investigators was published at about the same time. This book taught a generation of orthopaedic surgeons the application of engineering principles to musculoskeletal function and helped to establish orthopaedic biomechanics as an important discipline.

Burstein and Frankel set out to devise a standard testing apparatus to measure the mechanical strength of whole bone. Knowing that bone strength is dependent on loading rate (a characteristic of viscoelastic materials), the investigators desired an apparatus that could apply a reproducible and rapid loading rate, representative of typical trauma, to a variety of long bones. Other important design criteria included the ability to produce fractures similar to those seen in patients and the ability to subject the entire bone to equally severe loading conditions so as to identify weak points along its length. The device also needed to be inexpensive and easy to use.

Torsional loading best satisfied these requirements. The investigators chose a pendulum system to generate the torsional loads because it could be easily designed to produce a loading rate similar to that estimated for common traumatic fractures. The pendulum traverses the bottom 30° of its swing in 75 ms, similar to the load duration in fractures caused by skiing or motor vehicle accidents. The pendulum has a mass of 14 kg and a length of 23 cm. The pendulum's energy greatly exceeds the

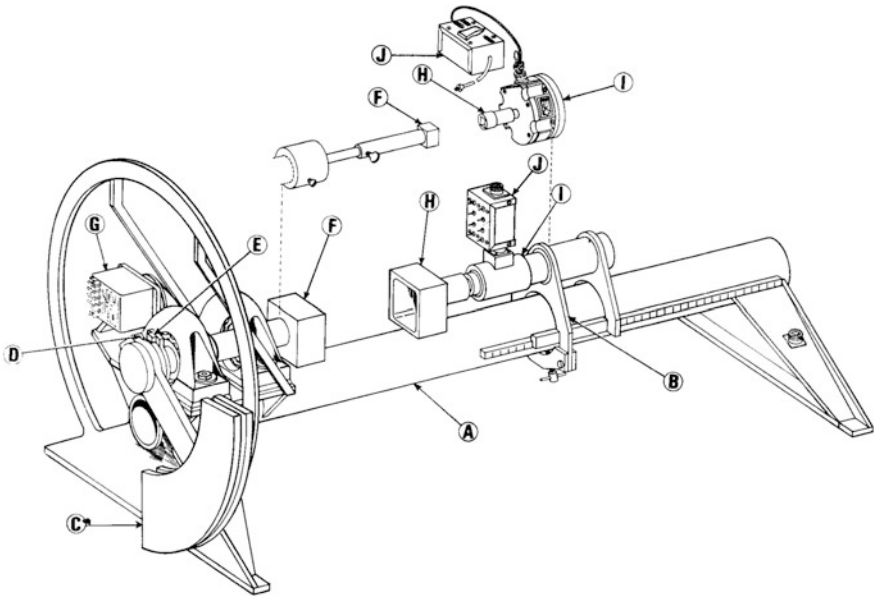


FIGURE 3.8. The torsion testing machine: (A) frame; (B) tail stock; (C) pendulum; (D) and (E) dog clutches on pendulum and rotating shaft, respectively; (F) rotating grips; (G) angular deformation transducer; (H) stationary grips; (I) torque transducer; (J) self-calibration system. *Upper drawing* shows a second subsystem used for smaller bones (Burstein and Frankel 1971).

energy required to fracture most bones, so that minimal changes in pendulum velocity or loading rate occur during the fracture event (less than 2 % in the case of rabbit bones). Measurement instrumentation consists of a reaction torque sensor, a precision angular position transducer, and an oscilloscope to capture and display the voltage outputs from the two transducers.

To conduct a test, the bone is rigidly mounted in alignment with the rotational axis of the device. The pendulum is dropped from a set height and at the bottom of its swing engages a cam on the load shaft; this imparts an angular deformation to the specimen that is measured by the angular position sensor. The resistive torque in the specimen is simultaneously measured by the torque transducer at the opposite end of the bone. A torque-vs.-angular deformation curve is then constructed, which represents the whole bone's failure characteristics under typical traumatic loading conditions. Common variables derived from the curve include the failure torque and deformation and the energy absorption to failure. Because bones are not perfectly straight cylinders, the derivation of material properties from these data is not usually attempted.

The advent of more advanced material testing machines has reduced the popularity of the Burstein–Frankel apparatus, but, thanks to its superb design, it remains a very legitimate means of testing the structural integrity of whole bones.

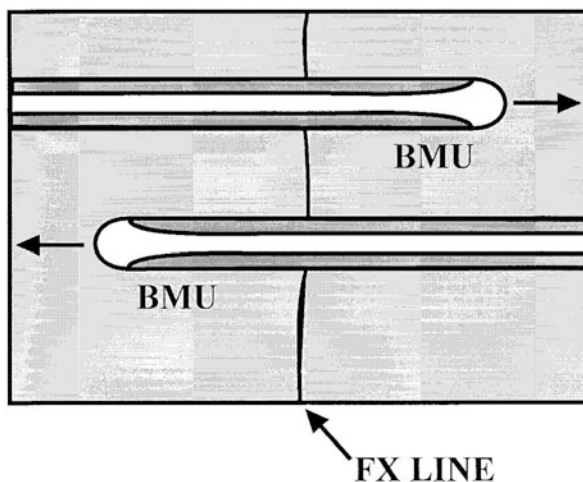


FIGURE 3.9. Sketch illustrating how osteonal BMUs may tunnel across a tightly fixed fracture line or osteotomy (saw cut) interface in cortical bone. This is known as primary healing.

3.2.6 Primary Union

If a fracture is rigidly fixed so that the bone ends are closely approximated, healing may occur by osteonal remodeling across the fracture line (Fig. 3.9). This is known as primary fracture healing. (Therefore, secondary healing is ordinary fracture healing as previously described.) Usually the only way that primary healing can happen is when a plate or screw is used very effectively to internally “fix” (rigidly fasten together the fragments of) the fracture. In these cases, there is almost no periosteal callus.

3.2.7 Posttraumatic Osteoporosis

It is commonly observed clinically that the intact portions of the fractured bone become osteoporotic as healing occurs. This generalized osteopenia of the intact regions, called posttraumatic osteoporosis or posttraumatic bone atrophy, is caused by two factors. First, in addition to the healing response, the fracture causes a remodeling RAP throughout the bone, so that osteonal BMUs riddle the entire cortex with resorption cavities. The second factor is the removal of mechanical loading from the fractured bone. As is seen later, disuse also activates remodeling. If the fracture is well fixed and sufficient load-bearing is resumed, the resorption spaces will refill and the osteopenia will be transient. To the degree that the bone does not resume load-bearing, the atrophy persists.

3.2.8 *Electrical Phenomena and Fracture Healing*

When bones are mechanically deformed, they become electrically polarized via two mechanisms: (1) the ionic fluids in the calcified matrix are forced to move about, creating streaming potentials (Eriksson 1974); and (2) the collagen molecules in the bone tissue are piezoelectric, so that they exhibit a dipole moment when strained (Fukada and Yasuda 1957). The relative roles of these two phenomena in producing stress-generated potentials in bone is unclear, but it is well established, for example, that when a bone is bent, the concave surface becomes negatively charged and the convex surface becomes positively charged. Furthermore, when a bone is fractured, the fracture site becomes negatively charged with respect to the remainder of the bone. Experiments have shown that bone formation is enhanced in the vicinity of a cathode, and the optimal current for this effect is about 20 μA . For a review of electrical phenomena in bone, see Chap. 4 in Martin and Burr (1989).

Since the 1970s, electrical stimulation has been used as a treatment for non-unions. There are two basic methods of producing such currents (Brignton et al. 1979). Thin wires were placed into the fracture site to carry direct current, with an anode pad on the skin nearby. Bassett and co-workers (Bassett et al. 1981) passed currents through external coils to induce currents in the tissues noninvasively. Both methods are reported to be about 70–85 % successful at healing nonunions but they require good immobilization and do not work in cases of true synovial pseudarthrosis. Encouraging studies by these investigators and others led to development of several commercially available devices, federally approved for use in cases of stubborn nonunions. The exact mechanisms responsible for enhanced healing have not yet been fully elucidated and the degree to which electrical stimulation enhances clinical outcomes remains controversial and difficult to define even after several years of clinical use (Mollon et al. 2008).

3.3 Analysis of Bone Remodeling

The dynamic biology of bones during remodeling can be measured directly using microscopy of bone samples collected from animals or people. Bone histomorphometry is used commonly in animal experiments to understand the effects of disease, drugs, mechanical loading, healing, age and other factors on bone metabolism. Bone histomorphometry is also used as a diagnostic and research tool on samples collected from human patients or volunteers, also for the understanding of bone changes in disease, although (of course) there are fewer cases of human experiments because sample collection requires surgery. In the remaining parts of this chapter, some of the methods used for standard histomorphometry will be introduced and then we will explore some mathematical models that help in the understanding of the dynamic remodeling process.

Box 3.4 Bones That Glow

In 1736, John Belchier discovered that developing bone in livestock could be stained *in vivo* by feeding the animals madder, a dye derived from the Eurasian herb *Rubia tinctorum*. This happened when he observed red streaks in the bones of some pork that he was having for dinner. He investigated and found that the pig had been fed waste plant material from a dye shop. During the ensuing two centuries Belchier's discovery provided impetus for several qualitative studies of bone growth and fracture healing. In the late 1700s, using madder-fed pigs, the famous English surgeon-anatomist John Hunter was the first to grasp the significance of what we now call modeling in the growth of the skeleton. Hunter's experiments demonstrated that bones grow radially by endosteal resorption and periosteal apposition. (The prevailing view had been that bones expanded outward as a consequence of new bone deposition within the cortex.) In the 1800s, synthetic alizarin replaced the natural dye and was found to be equally effective as a bone label. These dyes bind to hydroxyapatite as it is being formed, and therefore preferentially stain the interface between osteoid and mineralized bone. More recently, water-soluble versions of such dyes have been administered by injection, producing more sharply defined labels. However, these dyes would not be considered safe to administer to humans.

The modern use of such labels in human bone histomorphometry began in the 1950s when Milch et al. (1958) reported that the common antibiotic tetracycline was incorporated into actively mineralizing bone surfaces and fluoresced yellow when histologic sections were observed in ultraviolet light. Along with Milch, Frost (Frost 1960a), Harris (Harris 1960), and others recognized the significance of this finding: for the first time, the dynamics of human bone modeling and remodeling could be safely studied. Soon, tetracycline-based histologic analysis of bone remodeling was being widely used to study bone physiology. Fluorochrome labeling techniques are now routinely used to study the effectiveness of new drugs aimed at preventing osteoporosis and other metabolic bone diseases, new techniques and treatments intended to increase the rate of fracture healing, and the responses of bone to total joint replacements and other implants.

It turns out that "natural" fluorochrome labeling of our bones has been going on for some time, even without our knowledge! Skeletons excavated from a cemetery in southern Sudan dating from the fourth to sixth centuries A.D. were found to fluoresce yellow-green when placed under UV light. These people were agriculturists who stored grains until they could be eaten. When the stored grains got damp, they provided a rich environment for the growth of *Streptomyces*, a bacterium which produces tetracycline. Thus, by eating the grain, the people were not only ingesting a natural antibiotic (which may have kept them healthier), but were also labeling their bones in a way that can be useful to archeologists and bone biologists today in trying to understand dynamic bone remodeling in prehistoric populations (Bassett et al. 1980).

Alizarin and tetracycline are not the only *in vivo* labels available, although tetracycline remains the only label used in humans. Calcein green, calcein blue, and xylenol orange are other compounds used to label mineralizing bone *in vivo*. It turns out that all these compounds fluoresce, and they are collectively termed fluorochromes. Sequential administration of different fluorochromes, each fluorescing

a different color, allows the researcher to track the time course of a skeletal process such as fracture healing.

Bone specimens intended for fluorescent histomorphometric analysis must be processed with the mineral intact, since the fluorochromes bind to mineral. This is contrary to standard histologic procedures, which demineralize bone so that it can be cut with a knifelike instrument (called a microtome), just as soft tissues are. The simplest (but most tedious) means of making a section of undemineralized bone is to saw off a slab approximately 200–500 μm thick, usually using a diamond-impregnated saw, and then grind it down between two pieces of wet sandpaper until it is about 100 μm thick. Much of Frost's pioneering work was done using this simple method (Frost 1958). More modern methods embed the bone in plastics or resins with material properties similar to the bone tissue. It is then possible for a skilled technician to cut trabecular bone sections a few micrometers thick (or cortical cross sections 30–50 μm thick) using special knives or saws.

Histomorphometric analyses of labeled bones are performed using epifluorescent microscopes. The term "epi-" refers to the fact that the ultraviolet light is reflected off the surface of the section, which renders a sharper plane of focus than would transmitted light. Filters are used to enhance fluorescence and enable polychrome analysis of a single image. Fluorescence decays with time, so measurement techniques must be efficient. As outlined in the text, quantitative analysis requires measuring several histologic features, only a few of which involve fluorescent labels. The bone tissue must be stained to identify features such as osteoid and cement lines. If these stains interfere with the observation of fluorochrome labels, separate sections must be made for quantifying labels.

3.4 Mathematical Notation for Bone Remodeling

Frost (1964, 1969) was the first person to develop mathematical models for bone remodeling. Being the first, he was free to choose mathematical symbols for key variables. For example, he used the symbol "sigma" (σ) to represent the time required by each phase of the BMU to be completed at a particular location, Mf to denote the rate at which bone is added to a forming surface, and μ for the rate at which new BMUs are created in a bone.

Subsequently, Frost's mathematical notation was used by other investigators in reporting their experimental data. Eventually, however, concern arose because various investigators would use the same symbol to denote a variable measured or expressed in somewhat different ways. For example, the amount of bone surface on which formation was occurring could be expressed per unit volume of bone (i.e., in three dimensions) or per unit area of histologic cross section (i.e., two dimensions). To eliminate this problem, the research community developed a standard symbology for referring to bone remodeling variables (Parfitt et al. 1987). For example, the area of bone surface covered with osteoid would be called OS in three dimensions or O.Pm in two dimensions. Here, the ".Pm" refers to the fact that surfaces

TABLE 3.3. Nomenclature used in the previous edition and standard nomenclature

Variable	Suggested units	First edition nomenclature	Standard nomenclature
Resorption period	Days	TR	Rs.P
Reversal period	Days	TI	Rv.P
Formation period	Days	TF	FP
BMU activation frequency	#/mm ² /year	Fa	Ac.f
(Mineral) apposition rate	μm/day	MF	MAR
Mineralization lag time	Days	Tml	Mlt
Osteoid thickness (2-D section)	μm	Tos	O.Wi
BMU mean wall thickness	μm	Tmw	W.Wi
Bone volume fraction	–	Bv	BV/TV
Bone formation rate/bone volume	mm ³ /mm ³ /year	VF	BFR/BV
Bone formation rate/bone surface	μm ³ /μm ² /day		BFR/BS

show up as perimeters in a two-dimensional section. The resorption “sigma” became Rs.P (resorption period), the reversal period became Rv.P, and the formation sigma became FP (formation period). The total of these three periods is the remodeling period (Rm.P). In the first edition of this book a nonstandard nomenclature was used for remodeling parameters. Table 3.3 includes nomenclature used in the first edition along with the standard histomorphometric nomenclature, as defined by Parfitt et al. (1987), introduced here.

There are numerous standard measurements used to understand bone remodeling. Some of these are discussed in the following sections along with some of the assumptions that underlie their validity.

3.5 Histomorphometric Measurement of Remodeling

The concept of bone remodeling in quantum units is important because it recognizes a predictable biologic system that is particularly amenable to quantitative analysis. The utility of the quantitative histologic techniques introduced by Frost, and used today in clinical medicine as well as basic science research, is a direct result of these qualities.

Frost demonstrated that histologic sections could be analyzed to quantify bone remodeling. He originally did this for osteonal remodeling in the cortex of the human rib. Since then, his method has been extensively used in animal experiments and adapted for use on human bone biopsies from the iliac crest. The latter is used in preference to the rib because it is more accessible and contains more cancellous bone, where bone remodeling is more active and

bone loss appears more quickly. We provide a brief overview of the measurement process used in evaluation of bone remodeling for three reasons. First, histomorphometric analysis requires the reduction of the mechanics of remodeling and osteon formation to a few well-defined and measurable steps, which leads to a clearer and more integrated understanding of the osteonal remodeling process. Also, this sort of quantitative analysis is an excellent way to identify points of inconsistency in our theories and frame pertinent questions for future experiments. Finally, an understanding of histomorphometric analysis contributes greatly to a clearer perception of the problems associated with diagnosing pathologic states and defining their etiology.

After reviewing the methods used to quantify bone remodeling in a histologic section, we shall continue to use mathematical modeling as a tool for learning more about how remodeling works and the effects it has on bone structure. Making mathematical models forces us to be quantitative as we test hypotheses about the unknown. Eventually, these models may lead to a much clearer understanding of skeletal biology and mechanics and the relationship between them. You have the distinction of studying this subject in its infancy, and you should view the mathematical analyses in this chapter as “works in progress,” rather than derivations of the kinds of mathematical results that you have encountered in other courses. When asked how they obtained their insights in the world of physics, Newton said “By thinking into them.” And Einstein said “I grope.” In this chapter, we use mathematical models to think our way into bone remodeling. However, we are still very much in the groping stage, and you should view this chapter in that light. Join in the adventure, look skeptically at what follows, and try to find flaws and fix them!

3.5.1 *Assumptions in Histomorphometry*

A number of assumptions are made when attempting to quantify dynamic processes of bone remodeling, and it is important to understand what these are so that the information that derives from them is interpreted properly. The most important assumption is that remodeling is in a steady state, that is, that new BMUs are created at a constant rate, and are functioning in ways that are not changing with time. This implies, for the analysis of remodeling surfaces, that resorption and formation are coupled and in balance as well. Of course, this assumption of a steady state is never strictly true. Dynamic histomorphometry is usually employed to evaluate potential reasons that bone is not in a steady state, whether caused by a disease process or caused by experimental manipulation or genetic modifications of animals. Situations in which histomorphometry is used require the understanding that the investigator is measuring transient remodeling changes. A second assumption is the stereometric assumption that makes it possible to use

2-dimensional (2D) microscope slides to calculate estimates for the 3-dimensional (3D) biology (Mouton 2011).

The accuracy of the estimation of 3D measurements from 2D microscopic images depends on how well the practical measurements meet the mathematical assumptions that underlie unbiased stereology. In the analysis of bone, we regularly violate some of the assumptions of unbiased stereology; for example unbiased stereology commonly assumes that the microscope slides are oriented randomly in the 3D object (the bone). In the practical stereometric analysis of cortical bone it is not possible to randomize the orientation of the slides and they are always cut either directly parallel to, or directly perpendicular to the cortical diaphysis. Always sampling the bone in the same direction violates the random orientation assumption of stereology and creates systematic errors in the conversion of 2D measurements into estimates of the 3D data. In practice, however, this doesn't create a problem as long as one compares "like-to-like," that is, data collected using parallel (or perpendicular) slides is compared to data from slides cut in the same way. A different issue arises in the collection of data from slides of cancellous bone, in which trabecular struts may run at any angle with respect to the plane of section. As a result, 2D data for cancellous bone cut systematically from different sites has different bias (systematic error) when used to estimate 3D information. The mathematical ways to correctly adjust 2D measurements to 3D (Parfitt et al. 1987) are important to understand, especially for cancellous bone; however the details of these methods are beyond the scope of the current discussion. In practice, many labs simply measure several sections and average the values across those sections to estimate volumetric measurements. A third assumption is that the bone structures (or cells) that are measured were created during the period of interest. In actuality, even with sequential fluorochrome labeling, measurements are capturing a single snapshot in time. So, for instance, the mean wall thickness (W.Th) is defined as the average distance from the cement or reversal line to the bone surface. Mean wall thickness can be measured directly from any section containing osteons; however, there isn't any guarantee that the osteonal walls being measured were formed during the same period. More subversively, perhaps they were formed well before the onset of the bone disease or before the experimental manipulation that you want to measure. Since W.Th is used to calculate how long it takes for an osteon to form, not knowing when the tissue was created can lead to errors in evaluating the rate at which new BMUs are formed.

An additional problem for making biological measurements from microscopy slides is that very recently, with the advent of drugs for osteoporosis that create severe suppression of bone turnover, the methods used for histomorphometric analysis sometimes break down. If bone remodeling is suppressed to the degree that double fluorochrome labels cannot be found, then rates of bone formation, and a host of other calculated variables that define the remodeling system, cannot be determined. As the analytical method is

based on the existence of a remodeling steady state in which remodeling is occurring at a measurable rate, many experiments and some clinical trials cannot be accurately assessed using conventional histomorphometric techniques. Overcoming this problem presents an important area for research (Recker et al. 2011).

3.5.2 Histomorphometric Measurements for Cortical Bone

Histomorphometry for cortical bone uses the total area (Tt.Ar) of the cortical bone cross-section as the normalizing reference. Each section is measured for resorption and formation perimeters and areas, as well as labeled osteonal surfaces and distances between double labels and the measurements are normalized by dividing by Tt.Ar (Fig. 3.10). It is also important to measure

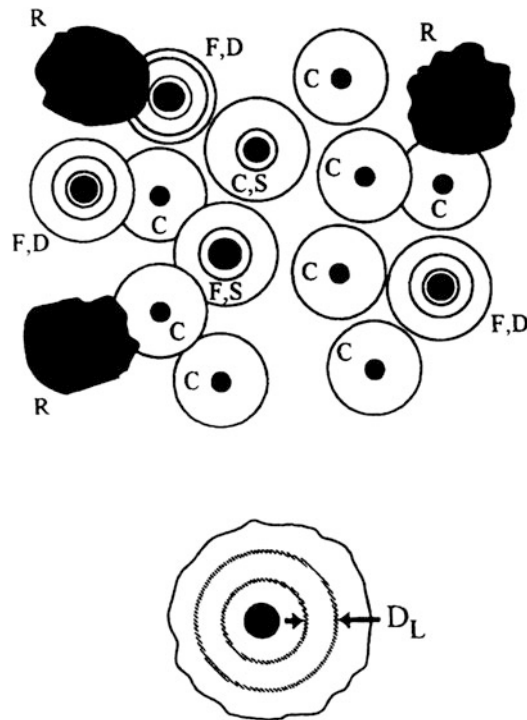


FIGURE 3.10. *Upper*: Schematic diagram of BMUs in a cross section of cortical bone. *Black*, irregularly shaped BMUs are resorption spaces (R). BMUs with large canals are refilling; they have either a single label (F, S) or a double label (F, D). Completed osteons have small canals (C); some of these have a single label (C, S). *Lower*: A single BMU in the last stages of refilling is depicted. It has a double label (D_L); the distance between the labels is measured as shown (Ir.L.Wi).

the number of labeled (forming) osteons per bone area N.O/B.Ar as this can be used in the calculation of bone formation rate (which will be a volume based measure, BFR/BV, rather than a surface based one).

Technically, this notation expresses the number of forming osteons. By assuming that all labeled osteons represent actively forming osteons, the nomenclature can be used. This could also be expressed, however, as N.On.L.

3.5.3 Calculating BMU-Level Results

The mineral apposition rate (MAR) is the rate that mineralized bone matrix is added to a bone and is calculated from the interlabel width (Ir.L.Wi, μm) and the interlabel period (Ir.L.P, days) as:

$$\text{MAR} = \text{Ir.L.Wi}/\text{Ir.L.P} \text{ (}\mu\text{m/day)} \quad (3.1)$$

The word “mineral” is attached to this variable because it is measured from fluorochrome labels, which lie at the junction between the osteoid “seam” (i.e., layer) and the layer of bone experiencing primary mineralization. In a steady-state situation, the mineral apposition rate is the same as the rate at which osteoblasts are laying down osteoid. Therefore, it is common to shorten the term to “apposition rate.”

The area of mineralizing surface per bone surface (MS/BS) is estimated as the simple ratio of the total mineralizing surface ($\text{MS} = \text{sL.S}/2 + \text{dL.S}$) and the bone surface (BS):

$$\text{MS/BS} = (\text{sL.S}/2 + \text{dL.S})/\text{BS} \quad (3.2)$$

The bone formation rate is often expressed per bone surface area. This is calculated from both MAR and from the total perimeter of bone formation per bone surface (MS/BS) as:

$$\text{BFR/BS} = (\text{MAR})(\text{MS/BS}) \quad (3.3)$$

Bone formation rate can also be expressed per bone area as

$$\text{BFR/B.Ar} = (\text{MAR})(\text{MS/BS})(\text{BS/B.Ar}) \quad (3.4)$$

These two examples of bone formation rate show how the histomorphometric parameters are combined to normalize a measurement in different ways.

The time that elapses between the production of osteoid by the osteoblasts and its initial mineralization is the mineralization lag time (Mlt). Histologically, this delay is represented by the thickness of the osteoid seam, O.Th (μm); this would be zero if the new bone mineralized immediately. The mineralization lag time may be determined by measuring the average O.Th

for the refilling BMUs on the section and dividing the result by the apposition rate (MAR):

$$\text{Mlt} = \text{O.Th}/\text{MAR} \text{ (days)} \quad (3.5)$$

The time required to refill a typical BMU (the formation period, FP in days) is then logically assumed to be the mean wall thickness (MWT) divided by the apposition rate:

$$\text{FP} = \text{MWT}/\text{MAR} \text{ (days)} \quad (3.6)$$

As indicated previously, there is a fundamental problem in this part of the analysis. The measurements are supposed to reflect the current situation, but measuring completed osteons means that the results will be skewed toward the conditions that existed when the osteons were formed, perhaps years ago. One way to handle this problem would be to limit the Mlt measurements to completed osteons that microradiographs reveal to be incompletely mineralized. Another method would be to measure Mlt in those few osteons bearing at least one of the labels given for the analysis but no osteoid seam.

The activation frequency is defined as the frequency with which a new BMU will be initiated in a given cross-section of bone. Its calculation relies heavily on the steady-state assumption. The times for each phase of BMU development are assumed to be constant, so that the number of BMUs activated each day is the same as the number entering the refilling phase. Use is also made of the fact that the size of any constant population is equal to the product of the birthrate and the average longevity of its individuals. Applying this to refilling BMUs relates the number of forming osteons per bone area (defined by the number of osteons with osteoid, or N.O) to the activation frequency (Ac.f; number per bone area per time) and the formation period (FP) as

$$\text{N.O}/\text{B.Ar} = (\text{Ac.f})(\text{FP}) \quad (3.7)$$

That is, the population of refilling BMUs per millimeter squared of bone section is equal to the birthrate per unit bone area (activation frequency, Ac.f) of refilling BMUs multiplied by their longevity, or the time required for refilling, FP. Rearranging this equation, one has

$$\text{Ac.f} = [(\text{N.O}/\text{B.Ar})/\text{FP}] \text{ (refilling BMUs/mm}^2\text{/yr)} \quad (3.8)$$

for the BMU activation frequency.

3.5.3.1 Estimating Bone Resorption Rate

The bone formation rate (BFR), is defined as the amount of new bone made per unit time, expressed as a fraction of the surface or volume of bone being considered. For example, if within each cubic millimeter of bone in a human femur the remodeling process formed a volume of bone equal to 0.05 mm^3 in 1 year, the bone formation rate would be $0.05 \text{ mm}^3/\text{mm}^3$ per year. Equivalently, if on a 1 mm^2 area of bone $0.05 \text{ mm}^2/\text{year}$ was formed then the bone formation rate would be $0.05 \text{ mm}^2/\text{mm}^2/\text{year}$. Because fractions of area measured on a bone cross-section are equivalent to volume fractions in the same region, this variable may be computed from the mineral apposition rate (MAR $\mu\text{m}/\text{day}$), osteoid perimeter (O.Pm μm of osteoid perimeter) and osteon number (O.N number of forming osteons per square millimeter) as:

$$\text{BFR} = (\text{MAR})(\text{O.Pm})(\text{O.N}) \quad (\text{mm}^2/\text{mm}^2/\text{yr}) \quad (3.9)$$

In this equation you need to make sure to use the correct units—don't forget that MAR and O.Pm are usually measured in micrometers/day and micrometers, respectively. As the amount of bone resorbed during osteonal remodeling is normally only slightly greater than that formed (the difference in osteonal bone resulting from the creation of new Haversian canals), the bone formation rate is often taken to be the same thing as the bone “turn-over” or replacement rate.

There is no straightforward way to measure the osteoclastic equivalent of MAR, which one may call the erosion rate, BRs.R, but it may be approximated if the rate at which bone is being lost can be determined. Suppose, for example, that biopsies are obtained before and after an experimental treatment, and measurements of porosity indicate that $0.01 \text{ mm}^2/\text{mm}^2$ of bone was lost during the course of a year. This time rate of change of porosity, or bone balance, must equal the difference between the bone resorption rate and the bone formation rate

$$\text{Bone balance} = \text{BRs.R} - \text{BFR} \quad (3.10)$$

If we define the resorption perimeter (Rs.Pm) as the length of surface with active resorption, Rs.N as the number of resorbing osteons per square millimeter formed per year and the linear rate of bone resorption at a surface as MRR (analogous to mineral apposition rate, MAR) then we can write an expression for osteonal resorption that is exactly analogous to that for formation:

$$\text{BRs.R} = (\text{MRR})(\text{Rs.Pm})(\text{Rs.N}) (\text{mm}^2/\text{mm}^2/\text{yr}) . \quad (3.11)$$

The bone balance, therefore, is the difference between this new equation (Eq. 3.11) and the original equation for bone formation rate (Eq. 3.9). If the bone balance is known, then the mineral resorption rate (MRR) can be

TABLE 3.4. Bone remodeling data for rib osteonal BMUs

Variable	Dogs ^a	Rhesus monkeys ^b	Humans by age group (years) ^c		
Age, years	1.5–3.5	“Adolescent”	1–9	30–39	70–89
On.N, #/mm ²	3.09 ± 1.75	4.03 ± 1.64	2.7 ± 1.0	0.21 ± 0.54	0.66 ± 0.50
Rs.N, #/mm ²	0.58 ± .034	1.96 ± 1.15	2.0 ± 0.62	0.25 ± 0.33	0.84 ± 0.46
O.Pm, μm	120 ± 30	200 ± 30	280 ± 40	250 ± 140	300 ± 72
MAR, μm/day	0.95 ± 0.26	1.33 ± 0.23	1.40 ± 0.46	1.1 ± 0.36	0.72 ± 0.20
FP, days	117 ± 110	54 ± 11	51 ± 47	73 ± 58	109 ± 55
Ac.f, #/mm ² /year	15.5 ± 11.8	27.9 ± 13	19 ± 9	1.1 ± 0.5	2.2 ± 1.1
MWT, μm	57 ± 6	63 ± 2	68 ± 4	70 ± 5	66 ± 7
BFR/BV, %/year	11.4 ± 9.4	35.9 ± 18.6	38 ± 24	1.8 ± 0.6	4.4 ± 1.5

BMU basic multicellular unit

^aA. Villanueva, personal communication

^bSchock et al. (1972)

^cFrost (1969)

estimated from BRs.R-BFR, since all of their terms can be directly determined from the cross section except for MRR.

The change in bone volume also can be estimated as the difference between the depth of the resorption cavity (resorption depth, Rs.De), the amount of bone that has refilled it (W.Wi), the perimeter over which resorption occurred (Rs.Pm) and the number of resorption sites per area per time (Ac.F) as

$$\text{BRs.R} = (\text{W.Wi} - \text{Rs.De})(\text{Rs.Pm})(\text{Ac.F}) \quad (3.12)$$

Rs.De. W.Wi and Rs.Pm are averaged over the completed BMU's in the section, first by measuring the distance between the reversal line and the original bone surface (which can be estimated in cancellous bone and the endosteal envelope, although not so easily in intracortical osteons), the second by measuring between the reversal line and the surface of new bone in a complete BMU and third by measuring the total length of the resorption surface. The time rate of bone loss (or gain) for the entire cross-section is the average of these values times the number that occur in the area (Ac.F) over time.

Table 3.4 shows typical values for a number of these histomorphometric variables as they have been measured in the ribs of humans, dogs, and monkeys. Clearly, there is variation between species, between bones (the rib remodels much more quickly than other cortical bone) and with age, so that numbers such as FP=90 days, which we cited earlier as typical in humans, must be used with discretion.

3.5.3.2 Label Escape Error

When two fluorochrome labels are given and the bone is examined afterward, some of the labeled osteons only have one label. One reason for this is

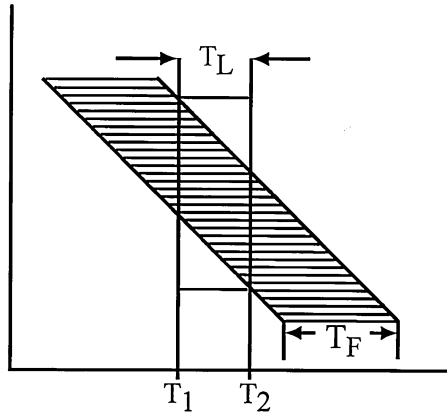


FIGURE 3.11. Frost's "ladder diagram" for deriving the label escape equation (Frost 1983). $T_L = Ir.L.P$, $T_F = FP$.

that some osteons began to refill after the first label was given, and others stopped refilling before the second label was given. Frost (Frost 1983) worked out the correction for this using the graph shown in Fig. 3.11. In this figure, time is represented on the horizontal axis, and the two vertical lines (T_1 and T_2) represent label times. The time between labels is $Ir.L.P = T_2 - T_1$. The diagonally distributed horizontal lines represent BMUs that began refilling at the time indicated by their left end and completed refilling at the time indicated by their right end. Thus, the lengths of the BMU lines are all FP . Doubly labeled BMUs have both vertical lines through them and singly labeled BMUs have only one vertical line. Frost (Frost 1983) showed that the ratio of the number of singly labeled osteons ($sL.N$) to those doubly labeled ($dL.N$) should be

$$sL.N/dL.N = 2(Ir.L.P)/(FP - Ir.L.P) \quad (3.13)$$

and the "true" number of refilling osteons ($N.O$; number of osteons with osteoid) should be taken as

$$N.O = (1/2) sL.N + dL.N \quad (3.14)$$

Demonstration of this relationship using similar triangles is left as Exercise 3.1. There is experimental evidence that the $sL.N/dL.N$ ratio is larger than predicted by Eq. (3.13) because some (or all) osteons pause during refilling. Such pauses may be exacerbated by age or mechanical disuse. This is an important detail of remodeling, which has been further analyzed (Li et al. 1990) but remains to be fully understood.

3.5.4 Histomorphometric Measurements for Cancellous Bone

Histomorphometric variables can generally be divided into those that can be directly measured—primary variables—and those that must be calculated—derived variables. Primary measurements can be broken down into a few essential kinds of measurements that include: area or volume, perimeter or length, widths, and number. Because some bones are larger than others, or have more or less surface area, these measurements must be normalized by using a referent. For example, the amount of mineralizing surface is usually expressed per total bone surface (MS/BS), as are other surface measurements, and the number of cells is also normalized in this way (eg N.Ob/BS for number of osteoblasts per unit of surface).

Using these variables, nearly every aspect of the remodeling process can be quantified, although measuring the rate of bone resorption is challenging because it is very difficult to measure bone that is gone! There are literally dozens of parameters that can be measured or derived in this way, but usually only a few of these are needed to get a clear picture of the remodeling pattern.

For iliac crest bone biopsies only a few primary variables are needed to derive all remaining variables. These measurements can be divided into three groups:

3.5.4.1 Areas

Total area, trabecular area and osteoid area	T.Ar, Tb.Ar, O.Ar

3.5.4.2 Perimeters or Lengths of Surfaces

Trabecular perimeter, erosion perimeter	Tb.Pm, Rs.Pm
Osteoclast perimeter, osteoid perimeter	Oc.Pm, O.Pm
Osteoblast perimeter, wall width	Ob.Pm, W.Wi

3.5.4.3 Perimeters or Lengths of Labels (from Unstained Sections)

Single- and double-labeled perimeter	sL,Pm, dL.Pm
Interlabel width	Ir.L.W

From the primary variables many useful derived variables are calculated to characterize the bone remodeling system, including:

3.5.4.4 Derived Variables

Trabecular number, thickness and separation	Tb.N, Tb.Th, Tb.Sp
Mineralizing surface per bone surface	MS/BS

(continued)

Mineral apposition rate	MAR
Bone formation rate per bone surface or bone volume	BFR/BS or BFR/BV
Mineralization lag time	Mlt
Formation, resorption and total remodeling periods	FP, RP, Rm.P
Activation frequency	Ac.f
Osteoid thickness	O.Th

The formulae for derived parameters for cancellous bone are usually the same as those presented previously for cortical bone. Detailed formulae for the histomorphometric calculations for cancellous bone are found in Parfitt et al. (Parfitt et al. 1987) and Li et al. (Li et al. 1990).

Although histomorphometric measurements can be quite tedious to make, a very few measurements can fully characterize the remodeling system. One drawback is that the derived parameters can be quite variable. Each time you use several primary variables to calculate a derived variable the inherent errors in each measured variable can combine to create even more variability in the calculated number. It is always a good idea to look critically at your result to see whether it makes any sense!

Box 3.5 Estimation of Trabecular Thickness Using the Plate Model

Quantification of the three-dimensional (3D) shape (morphology) of cancellous bone is often done from two-dimensional (2D) histological slides. To estimate 3D trabecular morphology from 2D measurements a model for the 3D structure is needed. It is often assumed that the three-dimensional (3D) structure of cancellous bone is a series of parallel plates (Fig. 3.12). The 2D histological measurements are extrapolated into 3D using stereological principles to estimate trabecular thickness (Tb.Th), trabecular separation (Tb.Sp), and trabecular number (Tb.N). Figure 3.13a is a 2D slice through a set of 3D parallel plates passing into the paper. The yellow is hard tissue and the orange is marrow. Figure 3.13b has the same color scheme and is a perpendicular cross-section through a cylinder of human vertebral cancellous bone that was cored parallel to the axis of the spine. It is the interaction between a test grid (in black) and the image that is used to calculate the 2D parameters which are extrapolated into 3D:

The Grid

Has a total length of solid line of L.

There are P points (intersections of solid and dashed lines) in the grid.

Interaction of the Grid with the Images

The fraction of the points that are on bone (yellow) is:

$$P_p = (\text{Points on yellow})/P$$

The number of intersections of the solid lines with the yellow/orange interface is:

$$I_L = (\text{Number of intersections})/L$$

Find

An equation for the trabecular thickness (Tb.Th).

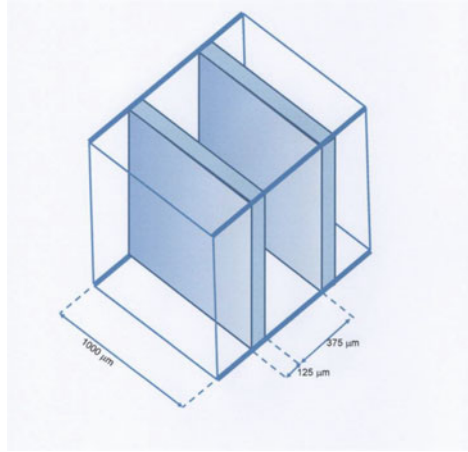


FIGURE 3.12. Diagram to show relationships between calculated three-dimensional quantities. A cube of tissue with sides of length 1 mm contains two parallel plates of bone, each with a surface area of 1 mm² on each side. With the distances shown, $BV/TV = 25\%$, $BS/TV = 4.0 \text{ mm}^2/\text{mm}^3$, $BS/BV = 16 \text{ mm}^2/\text{mm}^3$, $Tb.Th = 125 \text{ μm}$, $Tb.N = 2/\text{mm}$ and $Tb.Sp = 375 \text{ μm}$.

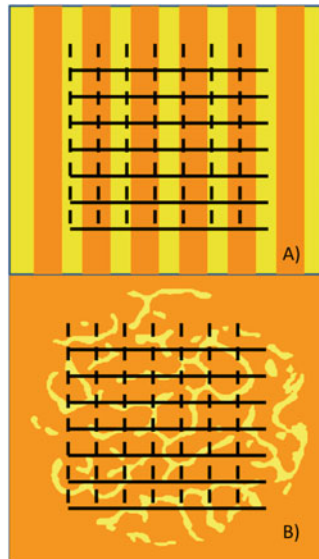


FIGURE 3.13. (a) A measuring grid superposed on a cross-section with bone = yellow and marrow = orange. The top example is artificial, with perfectly parallel strips of bone and marrow. The bottom example (b) was created using a transverse section of a cylindrical specimen of cancellous bone from the anterior-central part of a human T₁₂ vertebra.

Partial Solution

For both images, the average length of solid line over bone per total length of line (L_L) is equal to P_P . (Can you prove this?)

For Image (A) the number of trabeculae ($Tb.N = N_L$) per length of line is $\frac{1}{2}$ of L_L because the solid lines intersect both sides of the trabeculae:

$$N_L = \frac{1}{2}L_L \quad (A)$$

Therefore, the average trabecular thickness for image (A) is:

$$Tb.Th = L_L / N_L = P_P / N_L = 2P_P / L_L. \quad (B)$$

Unfortunately, the number of trabeculae ($Tb.N$) is only calculated correctly using Eq. (A) when the grid lines intersect the trabeculae exactly twice, as is true for (A). As a result of the inaccuracy of Eq. (A) for general images, the trabecular thickness ($Tb.Th$) equation is also not exact for general images (or for B)—but it is often used anyway.

Discussion

Plate model estimates of $Tb.Th$ are smaller than direct measurements of trabecular thickness made from 3D CT images. Furthermore, differences in trabecular morphology between anatomical sites or caused by disease change the underestimate between experiments. Despite these weaknesses, the plate model remains widely used because it is easy to use with 2D sections. It is important to know the limitations of the plate model estimate of $Tb.Th$ (also trabecular number $Tb.N$ and trabecular separation $Tb.Sp$) when reading papers and to determine whether the author is using the plate model or some other measurement (perhaps in 3D) that has the same name. Today it is generally agreed that trabecular numbers, thickness and separation are better calculated directly from 3D images when they are available. This is an advanced topic that is well presented in recent work (Liu et al. 2011; Day et al. 2000).

3.5.5 True vs. Histologic BMU Activation Frequency in Cortical Bone

Activation frequency has two definitions. The first is the number of BMUs that appear on a cross-section in a given time ($Ac.f = BMUs/mm^2/year$). The second is the number of BMUs created per unit volume of bone per unit time (e.g., $BMUs/mm^3/year$). We refer to this definition as the origination frequency ($Or.f$). Typically, BMUs are “born” when bone lining cells are stimulated to recruit osteoclasts to the site. These osteoclasts dig into the adjacent bone matrix, and in a few days create a tunnel or erosion cavity. The BMU continues to tunnel through the bone until some signal causes it to stop. What that signal or those signals may be is not known.

The first definition of activation frequency is the one we previously discussed; it arises when remodeling is quantified histomorphometrically.

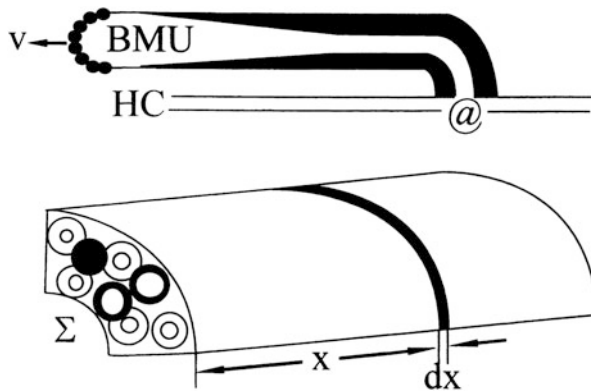


FIGURE 3.14. *Upper:* An osteonal BMU has been activated on the wall of an existing Haversian canal (HC) at the site marked @, and is tunneling to the *left* with velocity v . *Lower:* A quadrant of cortex having cross-sectional area A_r is shown. A histologic section Σ at the *left* end is histomorphometrically analyzed to obtain Ac.f. BMU resorption cavities arrive at Σ after having originated at a distance x , within a zone dx .

In this context, “activation” is not defined as the original creation of a BMU at a specific site along the bone, but as the appearance in the cross section of a functioning BMU in a particular stage of activity (Fig. 3.14).

As we have seen, the standard method of calculating activation frequency is to determine the population of refilling BMUs per mm^2 of section and divide this number by the refilling time (i.e., the “lifetime” of the refilling stage). This yields the “birthrate” of refilling BMUs in the section, that is, the rate at which the “front ends” of refilling regions enter the section. If all BMUs are similar with respect to the lengths of their resorbing and refilling regions, this “refilling region birthrate” is identical to that for the resorbing region, which is in turn identical to the rate at which BMUs themselves arrive at the cross section and begin to create a resorption cavity (Frost 1964). In other words, this definition involves the “creation” of sections of BMUs within the two-dimensional world of a histologic section, as opposed to the creation of the entire BMU in the three-dimensional world of the bone. We refer to the variable measured in this two-dimensional way as the activation frequency, Ac.f. As previously stated, units for Ac.f are typically given as $\text{BMUs}/\text{mm}^2/\text{year}$. How is Ac.f, which is routinely measurable, related to the more biologically relevant three-dimensional Or.f, which cannot be measured?

Consider a region of cortical bone having cross-sectional area A_r and its long axis oriented in the x -direction (Fig. 3.14). If Ac.f is measured in a histologic section, Σ , made perpendicular to the bone at $x=0$, we wish to know its relationship to Or.f in the bone as a whole. Let v be the mean velocity at which BMUs tunnel through the cortex, and let R be the mean

BMU range, that is, the average distance that they dig before ceasing their activity. Now, those BMUs just arriving at Σ originated at various locations, x_i , in the cortex, and at various times, T_i , in the past. The latter times may be called the ages of the BMUs arriving at $x=0$. If all traveled at the same average velocity, ν , then the originating distances and the ages are proportional to one another such that

$$x_i = \nu T_i \quad (3.15)$$

Let $\text{Or.f}(x, t)$ be the rate at which BMUs are being created at time t and position x along the bone. The number of BMUs originating within an infinitesimal section dx at time t in the past per day is

$$dN = \text{Or.f}(x, t) \text{ Ar } dx \quad (3.16)$$

where x is the location of dx and Ar is the area of the cross-section. The number of BMUs that originated in dx arriving at Σ during each unit of time is

$$dN = \text{Or.f}(x, T) \text{ Ar } dx \quad (3.17)$$

where $T = x/\nu$ is the time required for a BMU to travel from dx to Σ . The total number of BMUs/mm² arriving at Σ during each unit of time is obtained by integrating Eq. (3.17) from $x=0$ to $x=R$, the maximum distance from which a BMU could travel to reach Σ :

$$\text{Ac.f} = N/\text{Ar} = \int_0^R \text{Or.f}(x, T) dx \quad (3.18)$$

If Or.f varies in time and space, its histologic analog must be found by integrating Eq. (3.18). Let us now assume that Or.f is the same throughout the bone, varying only with time. Then substituting $dx = \nu dt$, and letting $L = R/\nu$ be the lifetime of a BMU's tunneling activity, one has

$$\text{Ac.f} = \int_0^L \text{Or.f}(t) \nu dt \quad (3.19)$$

If Or.f is also constant in time, one has

$$\text{Ac.f} = (\text{Or.f}) \nu L \quad (3.20)$$

or,

$$\text{Ac.f} = (\text{Or.f}) R \quad (3.21)$$

That is, in a steady-state situation, the two-dimensional, histologically measured activation frequency is equal to the three-dimensional, true activation frequency multiplied by the mean BMU range.

This derivation has only considered BMUs arriving at Σ from one direction. When BMUs arriving from the other direction are considered, it might seem that Eq. (3.19) would yield values of $Ac.f$ that are too small by half. That is not the case, however, because it can be assumed that half the BMUs represented by $Or.f$ will tunnel away from Σ . Thus, the effects of ignoring both these details cancel one another. Another detail should be mentioned as well. As noted, sometimes activation results in a “double-headed” BMU, which arises at one site but forms two cutting cones that tunnel in opposite directions. For the purposes of this analysis, and in terms of the amount of remodeling that result from such incidents, these events should be counted as two activations.

If $Ac.f$ and $Or.f$ are related through the BMU range, R , the next question is, what is the value of R ? This is not an easy question to answer. Because of the difficulty in sorting out anastomoses between Haversian canals and Volkmann’s canals, it is difficult to trace the paths of BMUs from their point of origin to their terminus. However, Cohen and Harris (Cohen and Harris 1958) observed that “In approximately one centimeter of the length of the femur, an osteon and its descendants coursed from periosteal lamellae to endosteum.” Tappen (Tappen 1977) was able to trace forming osteons back 5–6 mm from the cutting cone toward their origin. These observations suggest that a BMU may be expected to tunnel several millimeters before it ends. In that case, the true activation frequency, in BMUs/mm³/year, would be several times smaller than the histologically measured value, in BMUs/mm²/year.

In some early work, activation frequency was measured as just described, obtaining values of $Ac.f$, but the data were reported as $Or.f$, that is, as BMUs activated annually per cubic millimeter rather than per square millimeter. Other authors would simply report $Ac.f$. This may seem to be a small distinction, but it is important for two reasons. First, the origination frequency is probably only a fraction of that measured. This is not important when comparing groups in the same experiment, but it could be very important elsewhere, such as when three-dimensional mathematical models are constructed for remodeling phenomena. Second, ignoring the distinction deemphasizes important details of remodeling dynamics that are rarely discussed. Activation of BMUs at one site is only the beginning of an ongoing process that reaches out in both space and time to generate effects which may appear unrelated to the initial event. If BMUs travel several millimeters after they are activated, the lifetime of a BMU in a bone, as opposed to its lifetime in a section, is considerably longer than Frost’s $Rm.P$. This means that transients caused by changes in activation frequency may take longer to play themselves out than one $Rm.P$. For example, activation of remodeling at a screw hole in the middle of a long bone sends waves of BMUs traveling up

TABLE 3.5. Remodeling periods for cortical bone: cross-sectional and longitudinal values

	Cross-sectional (days)	Longitudinal (days)
Recruitment	5	5
Resorption	15	165
Reversal	0–20	0–20
Formation	90	260
Total ^a	95–125	265–285

^aNote that the reason the numbers for each process do not add up to equal the total remodeling period is that these processes are occurring simultaneously

and down the diaphysis, so that its whole length is affected for months afterward (Martin 1987). Other events that alter activation frequency can be expected to produce effects of similar scope, and it is important to know their scale in time and space. It was estimated by Frost that such transients may require as much as four remodeling periods to resolve, but that is just an estimate, and it might be even longer.

These concepts, which rely on the differences between understanding the bone remodeling system in three dimensions rather than two dimensions, also underscore the reason that the stereological constructs used to measure bone histomorphometrically are only valid when used on the proper cross-sections of bone. The measured periods for resorption, formation and total remodeling are different when viewed in longitudinal section (which might show the entire lifespan of a BMU), than when viewed from a cross-section, or even a series of cross-sections. This is illustrated by comparison of the times for various phases of bone remodeling, shown in Table 3.5.

The relationship between Or.f and Ac.f in cancellous bone is unclear. While investigators describe trabecular BMUs as migrating over trabecular surfaces how far an individual BMU may wander and in what direction before it merges with another, divides into two, or “dies” is poorly understood.

3.6 Remodeling Details

3.6.1 *Activation Is Prime*

Whatever causes activation of osteonal BMUs in a given situation, the body’s ability to vary activation is much greater than the ability to control the numbers of osteoblasts or osteoclasts within the BMU or the rates at which they work. Figure 3.15 shows this by comparing the correlations between bone formation rate and apposition rate (MAR, a measure of osteoblast activity, top graph), and activation frequency (Ac.f, bottom graph). Clearly,

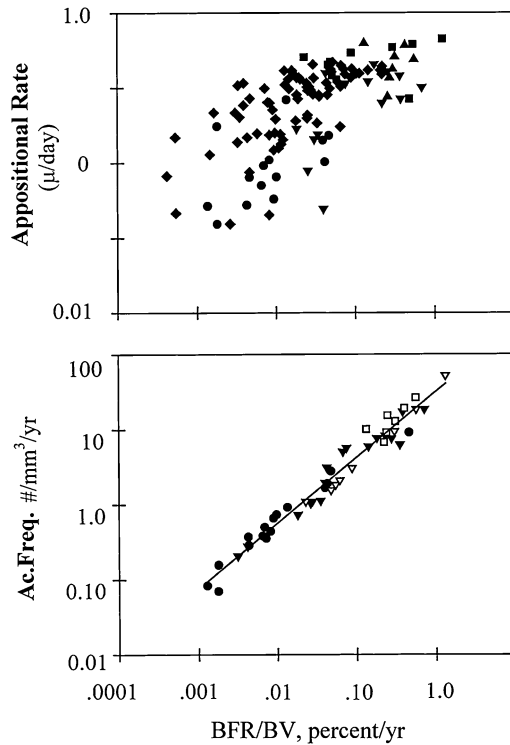


FIGURE 3.15. Graphs of apposition rate (*top*) and activation frequency (*bottom*) vs. bone formation rate (BFR) in human rib specimens. Note the much tighter correlation in the *lower* plot. The different *symbols* represent specimens from various kinds of patients, including normal controls.

the correlation between bone formation rate (BFR) and Ac.f is much higher than that between MAR and BFR. From this we deduce that changes in activation frequency have a greater effect on the amount of bone formed (or turned over) each year. Thus, effective (i.e., reasonably rapid and precise) responses to a situation calling for decreased bone mass, for example, cannot be achieved simply by decreasing osteoblast activity or stimulating osteoclasts; activation of many more BMUs is required. It turns out that activation is both the most important bone remodeling parameter and the one we know the least about.

3.6.2 Resorption: Out with the Old

Bone resorption is primarily the function of osteoclasts, which are mobile, multinucleated cells closely related to macrophages. Osteoclasts attach themselves to a bone surface, form a peripheral seal, and break down the bone

within the sealed area by means of enzymes and other chemicals. The final stages of the dissolution of the bone tissue appear to be carried out by mononuclear cells following along behind the osteoclast. Nevertheless, osteoclasts are the primary resorbing cells within the BMU. Although osteoclasts exhibit considerable mobility on trabecular bone surfaces, apparently ranging over resorption territories several times their contact area, in osteonal cutting cones they seem to dig straight ahead in a tightly bunched configuration.

3.6.2.1 How Fast Is an Osteoclast?

In 1972, Jaworski and Lok (Jaworski and Lok 1972) measured the longitudinal advancement rate of osteonal cutting cones in dog ribs using double tetracycline labels. By measuring the distance between the leading end of the two labels and dividing by the interlabel time, Ir.L.P, they determined that the region where the “refilling cone” began to mineralize advanced longitudinally some $39 \pm 14 \mu\text{m}/\text{day}$. Assuming the BMU’s geometry does not change as it moves through the bone, the cutting cone should have a similar rate of advancement. The rate at which the osteoclasts dig radially can be determined from the longitudinal advancement rate and geometry of the cutting cone. As seen in Fig. 3.16, in longitudinal section the cutting cone is ellipsoidal rather than hemispherical. In dog ribs the major longitudinal radius is about $300 \mu\text{m}$ and the minor radius (i.e., the cement line radius) is about $100 \mu\text{m}$. Thus, the time to open up the resorption space is $300/39$ or about 8 days, and the rate of radial erosion is $100/8$ or $12 \mu\text{m}/\text{day}$. It is not clear just how the osteoclasts move in the process of opening the tunnel. One can imagine each cell starting at the apex of the cutting cone and digging radially, or all the cells digging straight ahead. So little is known of the dynamics of cells in the osteonal BMU that neither possibility has much evidence in its favor, but radial motions would seem to require that the cells either have short lifetimes or recirculate.

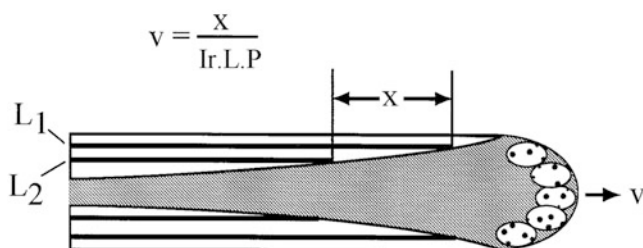


FIGURE 3.16. Schematic diagram shows measurement of osteonal BMU advancement rate (V) in a longitudinal section. X is the distance between the leading edges of labels L_1 and L_2 given Ir.L.P days apart.

3.6.2.2 How Long Does an Osteoclast Live?

The average lifetime of osteoclasts is a rather arbitrary concept, for they appear to consist of an evolving syncytium of mononuclear cells that have fused. This was learned initially as a result of experiments in which dogs were injected with tritiated thymidine and the labeled cell nuclei traced (Jaworski and Hooper 1980; Jaworski et al. 1981). When this was done, the number of labeled osteoclast nuclei increased and then decreased, demonstrating that nuclei enter and leave the cell while it persists in its activity. The average life span of a labeled nucleus in canine osteoclasts is 11 days.

More recently, using live cell microscopy, this process has been observed directly (Ineke et al. 2012). Fusion of mononuclear cells to each other, with multinucleated cells, and fusion of individual multinucleated cells to each other has been observed. Fission has also been identified, with the osteoclast compartmentalizing one or more nuclei, which either split to become a “new” osteoclast, or die by apoptosis. This constant fission and fusion may serve, much like genetic reshuffling during reproduction, to maintain osteoclast vitality, almost like “hybrid vigor,” by mixing different sets of nuclei.

3.6.2.3 The Phagocytic Connection

New cells are apparently added to and removed from the osteoclast as time goes by. These pre- and postosteoclasts are thought to be mononuclear phagocytes (MNP), cells of hematopoietic origin that circulate in the blood and are attracted to sites of injury or infection. There, groups of them fuse to form macrophages that engulf, digest, and eliminate foreign debris and damaged tissue. The osteoclast is also a composite of MNPs. The branching point between macrophages and osteoclasts seems to occur early in the phylogeny of the MNP. There are complex relationships between bone cells and hematopoietic cells (Calvi et al. 2012), and between inflammatory responses and bone remodeling (Takayanagi 2007; Lee et al. 2008; Lorenzo et al. 2008; Baron et al. 1984). This idea recalls the intense activation of BMUs precipitated by trauma or stress, known as a regional acceleratory phenomenon, or RAP. It is possible that this phenomenon is related to inflammatory responses mediated by cells of the immune system.

MNP-like cells are observed in the reversal region, “dressing” the walls of the resorption cavity and possibly creating the cement line over this surface (Baron et al. 1984). Perhaps MNPs in the reversal region are “postosteoclasts” in the sense that they are phagocytes that have reemerged from osteoclasts in the cutting cone. There is evidence that by-products of osteoclasts, such as collagen fragments, lysosomal enzymes, and glycosaminoglycans, are chemotactic for MNPs, causing these cells to approach the cavity wall and begin preparing the surface to receive osteoid. There is also evidence that osteoblast precursors are attracted and induced to differentiate and begin laying down bone by similar chemotactic factors near the cement line.

3.6.3 Refilling: *In with the New*

In some respects, the refilling process is more amenable to analysis than the resorption process, partly because newly mineralizing bone surfaces can be labeled with fluorochrome markers. The basic processes of bone formation within BMUs are clear, but many questions remain about the details.

3.6.3.1 Osteoblast Origins

Osteoblasts are mononuclear cells that are smaller and less mobile than osteoclasts. As previously noted, they apparently originate from stem cells in the mesenchyme (the embryonic tissue from which all connective tissues are derived). Thus, although bone formation follows resorption in bone remodeling, labeling BMU cell nuclei has excluded the possibility that osteoblasts are differentiated from osteoclasts or their remnants. Cells from many tissues (e.g., spleen) may be induced to differentiate into osteoblasts and form bone by continuous exposure to certain cytokines or tissues (e.g., urinary bladder epithelium), but only cells from bone marrow and the periosteum spontaneously produce osteoblasts.

An interesting observation by Coccia et al. (Coccia et al. 1980) underscores the different origins of osteoblasts and osteoclasts. After a man's marrow was transplanted into a female osteopetrosis¹ patient, her osteoclasts were found to contain the man's Y chromosome, but her osteoblasts did not. This implies that the two cell types have different origins; it also suggests that the osteoblast precursors in marrow do not survive transplantation. There are two principal types of cells in marrow: hematopoietic and mesenchymal stem. Owen (Owen 1980) proposed that hematopoietic cells are the source of osteoclasts, and that the mesenchymal stem cells which are the source of marrow's stromal cells are the source of osteoblasts. Presumably, mesenchymal cells are also required for growth of the vascular system within the osteon. As mesenchymal cells do not circulate, it seems that osteoblast precursors must come from a self-perpetuating pool of mesenchymal cells that follows the cutting cone from its point of origin.

In any event, assuming that osteoblast precursor cells are present in the reversal region, there must be mitogenic and chemotactic factors to orchestrate their conversion to functioning osteoblasts on the refilling surface. The products of bone dissolution by osteoclasts could perform this function and ultimately attract differentiated osteoblasts to the bone surface. Both monocytes and osteoblast-like cells are attracted to portions of the disassembled collagen molecule.

¹ Osteopetrosis (literally, "stone-bones") is a disease caused by an inability to produce sufficient osteoclasts. Thus, normal resorption and remodeling cannot occur, and the patient's bones are much harder and more dense than usual. It can be treated by transplanting marrow containing osteoclast pre-cursor cells from a normal person to the patient.

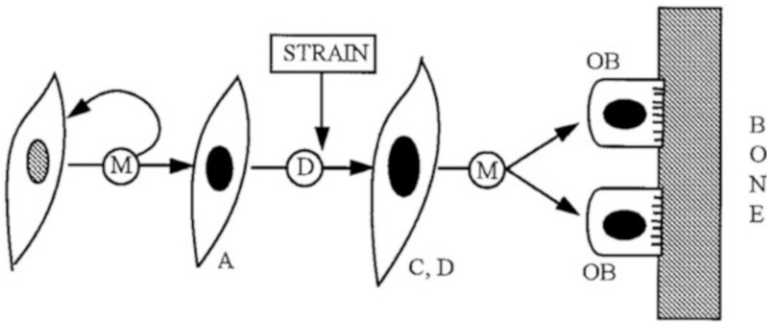


FIGURE 3.17. Diagram shows osteoblast precursor cells differentiating through intermediate stages. *M* mitosis, *D* differentiation, *OB* osteoblasts (After Roberts et al. 1984).

Mechanical loading appears to be required for osteoblasts to differentiate from their precursors. Using a simple yet elegant experimental approach, Roberts et al. (Roberts et al. 1982) characterized the stages of osteoblast differentiation in terms of nuclear size and used mathematical models to reveal a differentiation pathway through four precursor cell types to the functional osteoblast (Fig. 3.17). They found that, in the rat periodontal ligament,² the differentiation process is keyed to five light–dark cycles, so that the entire process requires 60 h or 2.5 days. Furthermore, the transition from stem cells to committed osteoprogenitors (type A cells) to pre-osteoblasts (type C and D cells) requires the presence of normal levels of mechanical stress (or strain) in the tissues; this has been confirmed in rats subjected to either simulated or actual weightlessness (Barou et al. 1996). In an osteonal BMU advancing at 40 $\mu\text{m}/\text{day}$, the cutting cone would move forward 100 μm during the time required for osteoblasts to go through the differentiation process. Allowing some time for events preliminary to those observed by Roberts, these observations are entirely consistent with the fact that reversal regions are typically 100–200 μm long. Abnormally long reversal regions could mean that the osteoblast differentiation process has been compromised, but observations in dogs subjected to disuse suggest that global bone formation may increase with disuse—the number of forming osteons with osteoid actually increases—even though the balance at each individual BMU favors resorption (Li et al. 2005). The effects of disuse on osteoblast differentiation in osteonal BMUs have not yet been observed.

Labeling with tritiated thymidine has revealed that osteonal osteoblasts derive from spindle-shaped cells located near the vascular sinus in the

² The periodontal ligament is the tissue that supports each tooth in its socket. It is the interface between tooth and bone, and is the source of cells that excavate or fill in bone when teeth are moved naturally or by an orthodontist.

resorption cavity, behind the advancing osteoclasts (Jaworski and Hooper 1980). These spindle cells are believed to constantly divide to maintain their population while half the new cells differentiate into osteoblasts, a process consistent with Roberts' observations in periodontal ligament and trabecular bone. Furthermore, the differentiation process appears to carry each subsequent group of precursors further from the capillary bud. Once created, osteoblasts begin work as cube-shaped cells occupying an area of bone surface about $15\text{ }\mu\text{m}^2$, and become shorter, flatter, and extend over more bone surface as they work. Ultimately, some of them end up embedded in bone as osteocytes, and others as lining cells, each covering perhaps $300\text{ }\mu\text{m}^2$ of the finished bone surface.

Jaworski and Hooper (Jaworski and Hooper 1980) also showed that osteoblasts do not move along the filling cone, but remain stationary with respect to the surrounding bone, falling further and further behind the cutting cone. Because the BMU is advancing about $40\text{ }\mu\text{m}/\text{day}$, and each osteoblast occupies about $15\text{ }\mu\text{m}$ of the closing cone's length, simple geometry suggests that about three new rings of cells must be created each day. If the radius of the resorption cavity is initially $100\text{ }\mu\text{m}$, its circumference is about $600\text{ }\mu\text{m}$, and each new ring could contain some 40 osteoblasts. Thus, about 120 osteoblasts must be produced daily to maintain the filling cone.

3.6.3.2 One or Many Generations?

Is the initial ring of osteoblasts responsible for the complete refilling job at its section of the osteon, or do additional generations of osteoblasts follow to complete the work? If the former possibility obtains, then the lifetime of some osteoblasts must reach FP, or about 3 months. Alternatively, several generations of osteoblasts could do the job, each living a fraction of this time.

Frost (Frost 1960b) described "resting" osteonal osteoid seams based on a qualitative change in their histologic appearance when stained with basic fuchsin. He associated these characteristics with lack of tetracycline uptake but normal osteoid thickness, suggesting that both the elaboration of osteoid by osteoblasts and its mineralization ceased simultaneously. This could occur because one generation of osteoblasts has reached the end of its lifetime and another was being recruited to resume the work. If multiple generations of osteoblasts refill the osteon, both resting osteoid seams and lamellar structure could result from the transition from one generation of osteoblasts to another. The exact nature of lamellar structure is controversial, but if one merely accepts the view that completed Haversian systems contain a number of discrete layers of bone between the cement line and the canal, then the boundaries between these layers may somehow be associated with interruptions or pauses in the refilling process.

Johnson (Johnson 1964) suggested that refilling discontinuities involves the replacement of generations of osteoblasts rather than pauses in the activity of the original cells; existing data may be used to test this hypothesis. Using polarized light, Kellin and Frost (Kellin and Frost 1964) counted the

numbers of lamellae in 350 osteons in normal human ribs; they counted bright layers, which were separated by dark layers. They found 7.9 ± 0.5 lamellar pairs per osteon. The mean wall thickness of an osteon is typically $80 \mu\text{m}$. If this wall contains eight lamellar pairs, the average pair thickness is $10 \mu\text{m}$. For a mean apposition rate of $1 \mu\text{m}/\text{day}$, the time required to form the first lamellar pair is 10 days; this can be regarded as a maximum time, as we show that the initial accretion rate is normally much greater than $1 \mu\text{m}/\text{day}$. Thus, the 11-day period of the Jaworski–Hooper experiment should have seen the labeled osteoblasts well into formation of the second lamellar pair, suggesting that lamellar structure is not correlated with separate generations of osteoblasts. On the other hand, this observation certainly does not prove that refilling is completed by only one generation of cells, and it leaves open the possibility that lamellar structure is associated with mere pauses in osteoblastic activity.

Tam and his co-workers (Tam et al. 1980) showed that apposition rate measurements in rats yield results that are inversely proportional to the time between labels when this time is greater than 72 h. This result could be explained by the occurrence of 3-day rest periods and longer work periods during refilling. As shown in Fig. 3.18, interlabel times longer than 3 days (e.g., A and B) could produce double-labeled BMUs that had rested between labels. These BMUs would underestimate apposition rate because they would overestimate the time actually devoted to making bone. Double-labeled BMUs obtained with interlabel times shorter than 3 days (e.g., C and D in Fig. 3.18) could not have rested, and would thus give accurate results. A 3-day rest would be long enough to include Roberts' 60-h osteoblast differentiation time. Thus, the "rest period" could be either a "sabbatical" for a persistent population of osteoblasts, or a time in which a new generation of osteoblasts is produced.

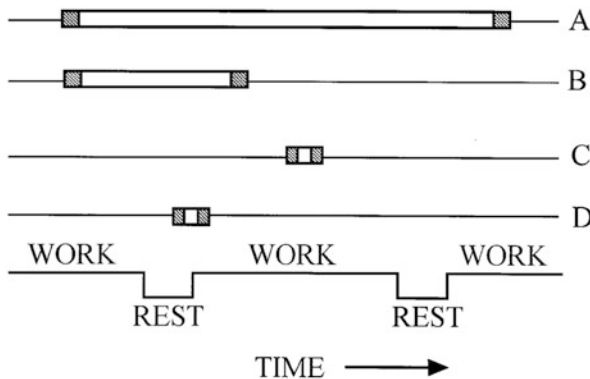


FIGURE 3.18. Schematic diagram shows possible relationships between fluorochrome labels and putative osteoblast work-rest cycles in a refilling BMU. Labels of varying intervals are shown as *shaded blocks* at ends of label intervals on time lines A–D. Assumed work and rest periods for bone formation are shown at *bottom*.

3.6.3.3 Apposition Rate Slows During Refilling

There has been some controversy regarding the uniformity of the rate at which osteonal BMUs refill. As refilling is initiated, it appears that there is rapid bone formation that “plasters over” the cusps of Howship’s lacunae and fills in other eccentricities in the resorption cavity surface. Vincent (Vincent 1955) originally observed that the rate of refilling slows down in proportion to the diameter of the remaining cavity. Other investigators suggested that the apposition rate was constant over most of the refilling period, or depended on the diameter of the original osteonal resorption cavity rather than on the age of the osteoblasts or the time since refilling began.

Manson and Waters (Manson and Waters 1965) performed an analysis that resolved the issue. They postulated that the apposition rate, MAR, is not constant and directly proportional to the radius, Rd, of the refilling cavity. That is,

$$\text{MAR} = d(\text{Rd})/dt = -k \text{ Rd} \quad (3.22)$$

where k is a constant and the minus sign occurs because $d(\text{Rd})/dt$ is directed opposite to the radius vector. The solution to this equation is

$$\text{Rd} = (\text{Cl.Rd}) e^{-kt} \quad (3.23)$$

where Cl.Rd is the cement line radius. Figure 3.19 shows this exponential model’s prediction about how refilling would proceed. Formation period, FP, can be estimated by solving for t when Rd is the Haversian canal radius, H.Ca.Rd. One obtains

$$\text{FP} = [\ln(\text{Cl.Rd}) - \ln(\text{H.Ca.Rd})]/k \quad (3.24)$$

This expression is quite different from the conventional histomorphometric formula which, for a diaphyseal cross-section takes the form

$$\text{FP} = (\text{Cl.Rd} - \text{H.Ca.Rd})/\text{MAR} \quad (3.25)$$

which assumes a constant apposition rate.

To test the model and obtain values for the constant k , Manson and Waters planned an experiment in which fluorochrome labels would be given at times T1 and T2, producing rings of radii R1 and R2, respectively, in those osteons that were doubly labeled. Consider now the result of plotting R2 vs. R1. If the apposition rate were constant, one would have

$$\text{R2} = \text{R1} - (\text{MAR})(\text{Ir.L.P}) \quad (3.26)$$

where $\text{Ir.L.P} = \text{T2} - \text{T1}$ is the interlabel interval. The plot of R2 vs. R1 should produce a straight line with a slope of 1.0 and an intercept proportional to

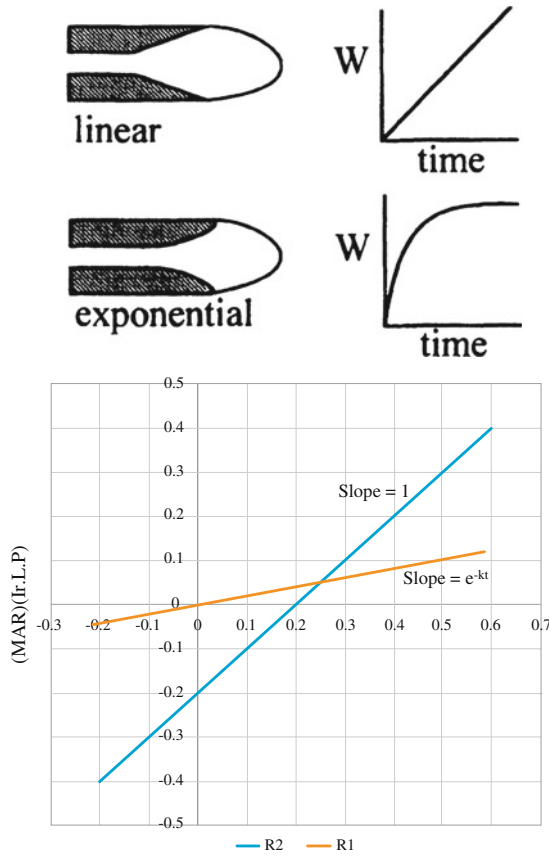


FIGURE 3.19. Linear and exponential models for refilling compared. *Upper left:* The shapes of the refilling surfaces in longitudinal profile. *Upper right:* Osteon wall thickness, W , as a function of time at a single longitudinal location. *Bottom:* Plots of the radii of double fluorochrome labels have different slopes and intercepts in the two models.

MAR and the interlabel time, Ir.L.P From their exponential model, Manson and Waters obtained a distinctly different expression:

$$R2 = R1 \exp(-k[\text{Ir.L.P}]) \quad (3.27)$$

which indicates the plot should pass through the origin and have a slope determined by k and Ir.L.P. Upon executing their experiment in cats of two different ages, and analyzing several different bones, they found that their data clearly fit Eq. (3.27) much better than Eq. (3.26) (Table 3.6). The intercepts of the R2 vs. R1 plots were nearly zero, and the slopes were significantly different from unity. Thus, the data confirmed the hypothesis that the apposition rate declines exponentially during refilling.

TABLE 3.6. Apposition rate experiments

Age of cat (years)	0.75	2.5
Osteons analyzed	101	169
Interlabel time (days)	14	28
Mean slope	0.539	0.436
Mean correlation coefficient	0.933	0.943
k, days ⁻¹	0.0441	0.0296

Data from Manson and Waters (1965)

TABLE 3.7. Apposition rate modeling: effect of k on other variables

k, day ⁻¹	MAR (μm/day)			
	Initial	Final	Mean	FP (days)
0.01	1.0	0.2	0.5	160
0.02	2.0	0.4	1.0	80
0.03	3.0	0.6	1.5	53
0.04	4.0	0.8	2.0	40

3.6.4 Computation of FP

Although the formula for computing FP that arises from the Manson–Waters theory is different from the conventional histomorphometric formula, if the mean apposition rate is used in the conventional formula (Eqs. 3.6 and 3.25) the correct result is nevertheless obtained. The mean apposition rate is related to k by the expression

$$\text{MAR} = k(\text{Cl.Rd} - \text{H.Ca.Rd}) / [\ln(\text{Cl.Rd}) - \ln(\text{H.Ca.Rd})] \quad (3.28)$$

Table 3.7 shows values of initial, final, and mean apposition rate, and FP for four possible values of k, assuming that the cement line and Haversian canal radii are 100 and 20 μm, respectively.

3.6.5 The Equal Lamellar Work Areas Principle

The fact that osteonal refilling proceeds at an exponentially decreasing rate leads to an important geometric characteristic of the filling cone. As we have seen, there are typically eight lamellar pairs in the osteonal wall, and their thickness is remarkably uniform. In an idealized filling cone, a nonlinear refilling rate leads to a horn-shaped rather than a truly conical surface (Fig. 3.20). Note that the appositional surface of each lamella increases as one proceeds from right to left in the figure (e.g., $x_2 > x_1$ and $ds_2 > ds_1$).

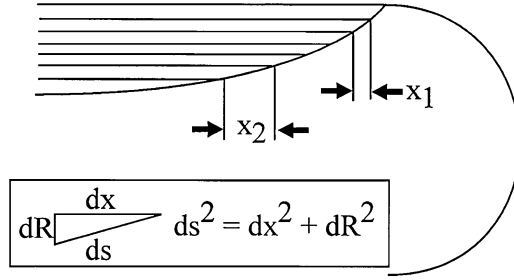


FIGURE 3.20. Effect of nonlinear refilling on exposed surface of individual lamellae. Inner lamellae have a broader surface on which formation can occur, but osteoblasts at this point lay down bone more slowly. *Inset*: Relationship between ds , dx , and dR .

Because the apposition rate declines during refilling, the inner lamellae not only take longer to form than the outer lamellae, but they also present more surface on which the cells can work. In fact, it can easily be shown that this longitudinal extension of the lamellar work surface just compensates the loss of work area resulting from the contracting circumference. To see this, consider the calculation of the surface, dS , of a single lamellar work region bound by longitudinal positions x_1 and x_2 . Because $dS = 2\pi R(x) ds$ and $ds^2 = dx^2 + dR^2$, the surface integral is

$$S = 2\pi \int_{x_1}^{x_2} R(x) \sqrt{dx^2 + dR^2} = 2\pi \int_{x_1}^{x_2} R(x) \sqrt{1 + (dR/dx)^2} dx \quad (3.29)$$

where x is the longitudinal coordinate and R is the radius of the filling cavity. For exponentially declining refilling where we use R as the current radius

$$R(t) = (Cl \cdot Rd) e^{-kt} \quad (3.30)$$

This expression can be converted to one for $R(x)$ by using the longitudinal advancement rate, ν , of the osteon. That is, substituting

$$t = x/\nu \quad (3.31)$$

into Eq. (3.30) gives

$$R(x) = (Cl \cdot Rd) e^{-kx/\nu} \quad d(R(x))/dx = -(k(Cl \cdot Rd)/\nu) e^{-kx/\nu} \quad (3.32)$$

and substituting these expressions into the integral yields

$$(Cl \cdot Rd) \exp\left(-k \cdot \frac{x}{\nu}\right) \cdot \sqrt{1 + \left(\frac{Cl \cdot Rd \cdot k}{\nu}\right)^2 \cdot e^{-\frac{2kx}{\nu}}} \quad (3.33)$$

It can be shown that the square root term is almost equal to 1 for any value of x and typical values of k (0.02/day), $Cl.Rd$ (100 μm), and ν (40 $\mu\text{m/day}$). Making such an approximation and performing the integration,

$$S = 2Cl.Rd(\nu/k)[\exp(-kx_1/\nu) - \exp(-kx_2/\nu)] \quad (3.34)$$

Note that because $R1 = Cl.Rd \exp(-kx_1/\nu)$ and $R2 = Cl.Rd \exp(-kx_2/\nu)$, the term in brackets can be replaced by $(R1 - R2)/Cl.Rd$. Therefore

$$S = 2\nu(R_1 - R_2)/k \quad (3.35)$$

Now $(R1 - R2)$ is just the lamellar pair thickness, which is nearly constant for all lamellae. Thus, the lamellar work surface, S , turns out to be equal to a constant function of k , ν , and lamellar thickness.

Equation 3.35 may be used to estimate the number of osteoblasts that work on each lamellar pair. Letting $FP = 90$ days, $Cl.Rd = 100 \mu\text{m}$, and $H.Ca.Rd = 20 \mu\text{m}$, Eq. (3.24) yields $k = 0.0179/\text{day}$. Then if $\nu = 40 \mu\text{m/day}$, and the lamellar pair thickness is $10 \mu\text{m}$, $S = 2 \times 3.14 \times 40 \times 10 / 0.0179 = 140,000 \mu\text{m}^2$. Osteoblasts that are rapidly forming bone have a “footprint” on the bone surface which is about $10 \mu\text{m}$ in diameter; those which are slowly forming bone are about $20 \mu\text{m}$ across (Holtrop 1990). Assuming, then, that the average osteoblast occupies approximately $15 \times 15 = 225 \mu\text{m}^2$, the number of cells working on each lamella is $140,000/225 = 622$.

3.6.6 Why Does the Apposition Rate Slow Down?

We might postulate four mechanisms for declining apposition rates during refilling: (1) A single generation of osteoblasts refills the cavity and its cells slow down, or spend more time resting as the refilling proceeds, from either internal or external factors. (2) Several generations of osteoblasts refill the cavity, and each generation becomes more indolent, owing to genetic or environmental factors. (3) Several generations of osteoblasts refill the cavity, and the time required for differentiation of each new generation is increased, causing increasingly longer “rest periods.” (4) Both (2) and (3) occur.

Manson and Waters asserted that osteoblastic activity is enhanced when the cells are further from the blood vessels in the osteonal cavity. They did not offer a rationale for this assertion, which seems contrary to a straightforward notion of cell nutrition via diffusion from a capillary. It is, however, consistent with another observation by Roberts et al. (Roberts et al. 1987) concerning osteoblast differentiation. When a periodontal blood vessel is too close to the bone-forming surface, osteoblast precursors are obliged to differentiate along a path that initially takes them away from the bone. They subsequently circle back to the bone surface. On the basis of this behavior, it could be postulated that as the osteonal resorption cavity is refilled,

osteoblast precursors find it increasingly difficult to get enough “elbow room” between the blood vessel and the bone surface, and this affects their rate of differentiation or their subsequent ability to function.

This mechanism, based on spatial factors within the osteonal cavity, is external to the osteoblast; alternatively, it is possible that the decreased rate of osteoid synthesis is caused by factors internal to the cell. This distinction is important because internal factors would be likely to affect osteoblasts in osteonal and surface (i.e., periosteal, endosteal, or trabecular) BMUs alike, but external factors such as cell crowding may be peculiar to osteonal osteoblasts.

Martin and co-workers (Martin et al. 1987) developed a model that assumes that variations in apposition rate are caused by factors within the osteoblast, and compared the abilities of the three available models (linear refilling, slowing from external factors, slowing from internal factors) to describe apposition rates in osteonal and trabecular bone. For osteons, the internal effects model yields essentially the same result as the Manson–Walters model. However, applying the internal effects model to trabecular bone by measuring R1 and R2 from the cement line,

$$R2 = R1 \exp(-k \text{ Ir.L.P}) + C[1 - \exp(-k \text{ Ir.L.P})] \quad (3.36)$$

where k and C are constants. In this case, the slope term is similar to that in the Manson–Walters theory, but now there is also an intercept term. Thus, the three apposition rate models yield distinctly different predictions about R1 vs. R2 plots, as shown in Table 3.8.

These theoretical models were tested against R1–R2 data from both osteonal and trabecular BMUs in adult dogs. As shown in Table 3.9, it was found that osteonal apposition rates had slopes significantly less than 1 and intercepts close to 0, but trabecular BMUs had slopes very close to 1 and intercepts nearly equal to $(\text{MAR})(\text{Ir.L.P})$. Thus, apposition rates were found to be constant in trabecular BMUs but not in osteons. If internal factors (i.e., internal to the osteoblast) were responsible for the nonlinear refilling in osteons, trabecular osteoblasts should be similarly affected; as they were not, it appears that external factors are responsible for reducing the rate of bone apposition as osteons refill.

TABLE 3.8. Apposition rate models: predictions about R1 vs. R2 plots

	Slope	Intercept
Linear refilling	1	$-(\text{MAR})(\text{Ir.L.P})$
External factors	<1	0
Internal factors		
Osteonal BMUs	<1	0
Trabecular BMUs	<1	$C [1 - \exp(-k \text{ Ir.L.P})]$

TABLE 3.9. Apposition rates in osteonal and trabecular canine BMUs

	No. of dogs	No. of BMUs	Correlation coefficient	Slope	Intercept	(MAR) (Ir.L.P)
Osteonal BMUs						
Rib	7	88	0.97 ± 0.02	0.77 ± 0.06	-5.3 ± 1.4	15
Humerus	5	140	0.95 ± 0.02	0.70 ± 0.12	-1.5 ± 3.3	14
Trabecular BMUs						
Ilium	5	67	0.99 ± 0.03	1.02 ± 0.05	9.6 ± 1.7	8

Data from Martin et al. (1987)

In thinking about the external factors that could cause this effect, we may reiterate our previous suggestion that decreasing space between the cavity wall and the blood vessel interferes with osteoblast differentiation. Another possibility has to do with Roberts’ observation that the differentiation of osteoblasts is stress dependent. One would expect to find a stress concentration at the front end of a BMU, with diminishing stresses along the refilling region. Perhaps this diminishment of bone strain interferes with osteoblast production (see Chap. 6), or with the regulation of bone formation by the osteoblasts. We now know that sclerostin, a negative regulator of bone formation produced by osteocytes, is inhibited in response to increased levels of strain (Robling et al. 2008). The suppression of sclerostin at the head of the cutting cone would be permissive for new bone formation. But as those strains declined along the refilling region of the BMU, the production of sclerostin by osteocytes would return to normal and reduce the amount of bone formation. This is just one regulator of bone formation; it is likely that there are others that might have similar effects (e.g. TGF β). This idea is explored further in Chap. 6, where we describe a theory relating wall strains in the osteonal BMU to cell function and directional control.

3.6.6.1 Osteoblast Accounting

Another aspect of refilling that appears to be related to both lamellar structure and cell lifetime is the conversion of osteoblasts to osteocytes. Osteocytes are formed when osteoblasts bury themselves in osteoid. Jaworski and Hooper found that the first osteocytes appeared 9 days into the refilling period; this places the first osteocytes near the junction between the first and second lamellar pairs.

It has been argued that because the working surface contracts as osteonal refilling proceeds, some osteoblasts should be forced out into the new bone matrix or back into the resorption cavity (Parfitt 1983). Pursuing this assumption quantitatively, it appears that the number of displaced osteoblasts would be much greater than the number of osteocytes and bone lining cells in a unit length of completed osteon. We first consider what fraction of the osteoblasts become osteocytes using the scheme of cell dynamics

presented here. During the formation of each lamellar pair, 1 of every 28 cells gets trapped and becomes an osteocyte. This figure is deduced as follows. Consider the i th lamellar pair, which consists of a “pipe” of bone having an outer radius R_o and an inner radius R_i . Segments of the length of this pipe may be thought of as being formed by successive rings of osteoblasts arranged around the BMU cavity. If $b = 15 \mu\text{m}$ is the breadth of the work area of an individual osteoblast, then the volume of such a segment of bone formed by a single ring of osteoblasts is

$$V = \pi b (R_o^2 - R_i^2) \quad (3.37)$$

The average density of osteocytes in cortical bone is 12,000–20,000/mm³ (Boyde 1972; Frost 1960c; Mullender et al. 1996). Let Ot.Dn be such an osteocyte density expressed as cells/ μm^3 and let its average value be 16×10^{-6} . Then the expected number of osteocytes ($N.\text{Ot}$) in the pipe segment of volume V is

$$N.\text{Ot} = (\text{Ot.Dn})(V) = \pi b (\text{Ot.Dn})(R_o^2 - R_i^2) \quad (3.38)$$

The average circumference of the work surface during formation of the pipe segment is

$$C = 2\pi[(R_o + R_i)/2] \quad (3.39)$$

Assuming there is an osteoblast every $b \mu\text{m}$ along this circumference, the average number of cells in the osteoblast ring would be

$$N.\text{Ob} = \pi(R_o + R_i)/b \quad (3.40)$$

Combining and simplifying Eqs. 3.38 and 3.40, the ratio of the number of osteocytes buried in the pipe segment to the number of osteoblasts that formed the segment is

$$N.\text{Ot}/N.\text{Ob} = (\text{Ot.Dn}) (R_o - R_i)b^2 \quad (3.41)$$

Note that $R_o - R_i$ is just the lamellar pair width, which can be assumed to be about $10 \mu\text{m}$ in all cases. Therefore, the fraction of osteoblasts converted to osteocytes is also constant for all lamellae, and works out to be $N.\text{Ot}/N.\text{Ob} = (16 \times 10^{-6})(15)^2(10) = 0.036 = 1/28$.

Assuming the cement line radius was $100 \mu\text{m}$, and the osteoblasts averaged $10 \mu\text{m}$ in diameter at the start of refilling, the original ring of osteoblasts would have contained about $2\pi \times 100/10 = 63$ cells. Then only a couple of these cells could be expected to end up as osteocytes. If the radius of the completed Haversian canal is $20 \mu\text{m}$, and it is covered by bone lining cells having a diameter of about $20 \mu\text{m}$, there would be about 2π or 6 lining

cells left from the original ring of 63 osteoblasts. That leaves about 55 osteoblasts unaccounted for, assuming that a single generation of cells accomplishes the refilling. One interesting possibility is that these cells are “surfing” on the advancing wave of new bone, that is, translating forward in the wake of the advancing resorption cavity, which would reduce the numbers of new osteoblasts that have to be produced. Alternatively, some of these cells may be forced off the shrinking refilling surface and die, or revert to another type of cell. What actually happens is unknown.

3.7 Long-Term Effects of Osteonal Remodeling

3.7.1 *Implications for the Aging Skeleton*

Activation frequency in the human rib declines from very high levels in childhood to a minimum at about age 35 years, and then rises and falls again at the end of an average lifespan (see Fig. 3.4). Unfortunately, such data are not available for long bones such as the femur or for trabecular bone. It is assumed, however, that this curve is generally descriptive of $Ac.f$ vs. age, although the magnitudes of the values would be different in other bones.

Anyone who has studied dynamic systems theory will be struck by the similarity of this curve to that describing the behavior of a damped oscillation. One way that such oscillations can occur is when a feedback system is “overdriven,” or operated with too much feedback. Whether the variation in activation rate has anything to do with an overdriven regulatory system remains to be learned, but it is certain that the shape and scale of this curve have a number of important consequences.

One of the most obvious of these consequences is that the rate of addition of new osteons, and the rate of turnover of cortical bone, is much higher in children than in adults. Most cortical bone is created by modeling processes associated with growth. The high values of $Ac.f$ in children ensure that most of this bone is rapidly converted to secondary osteonal bone. One may postulate that later, during adulthood, $Ac.f$ is maintained at some minimal level by systemic factors and is increased only under conditions of mechanical or metabolic need.

It appears that the difference between the values of $Ac.f$ in children and adults is greater than the difference in levels of activity or mechanical usage in these two groups. The elevated activation rate is generally assumed to be a growth hormone-driven phenomenon, but part of the elevation may be caused by age-related variations in material properties of the bone matrix. The rapid growth and high rate of remodeling of children’s bone causes it to be undermineralized. This deficit results in a relatively low elastic modulus, producing higher strains for a given load or level of stress. Nunamaker

et al. (Nunamaker et al. 1990) have shown that peak strains in racehorses approach 5000 microstrain but are reduced to less than 3000 microstrain after maturity. Strain levels this high may cause increased fatigue damage and thus increased activation of new BMUs.

After declining to a minimum at about age 35, the BMU activation rate increases again to about twice the minimal value in men and women in their 60s. This increase tends to amplify any deficit in remodeling at the BMU level that occurs in adults as well as that which may be associated with aging itself. The final decline in Ac.f in people in their 80s and 90s tends to reduce the effects of senile remodeling deficits.

3.7.2 *Osteonal Overlapping*

Osteon accumulation is limited by the fact that osteons soon begin to overlap one another as more and more are added. When an osteon overlaps another, it is not prevented from “obliterating” (i.e., refilling) its predecessor’s Haversian canal. While in general new osteons create new canals, they may also destroy an old canal. Age-related changes in porosity therefore are affected by the probability that old canals will be obliterated.

The analysis of osteon accumulation and overlapping is a geometric problem not unlike the one astrogeologists have encountered in trying to estimate planetary age from the numbers and overlapping of craters seen on the surface. Physical anthropologists have used osteon population density to estimate the age at death of skeletal remains. A mathematical theory can help to clarify the nature of senile osteoporosis in cortical bone. It can also allow one to use a simple census of osteons to estimate activation and bone turnover rates—information that normally can only be obtained from vital labeling of living animals.

3.7.3 *The Random Remodeling Assumption*

Because osteons are quite long in comparison to their diameter, their overlapping must be determined as much by the “steering” of the BMU as by its point of origin. It has been hypothesized that osteonal orientation is controlled by the principal strain direction, but it is also reasonable to think that it is affected by variations in the material properties or chemical composition of the bone which appears before the advancing osteoclasts. That is, regions of increased mineral content, patches of osteoid, volumes of damaged tissue and canals could conceivably alter the path. As all these factors may be expected to be at least a little different for the new BMU than for its predecessors, one would expect the steering of an osteon to be different from its nearby older neighbors, leading to quasirandom positioning of each new osteon. This “local randomness” does not imply that new osteons are randomly positioned with respect to the whole cross section but only with

respect to local organization. Such local randomness is an underlying assumption in most analyses of the effects of osteonal remodeling on such tissue properties as porosity and mean tissue age.

3.7.4 Mathematical Theories

The problem of determining the average age of a region of bone (tissue age, or mean skeletal age) is very closely linked to that of osteon accumulation, and has been discussed in some detail by Hattner and Frost (Hattner and Frost 1963). Tissue age depends on the rate of remodeling and the degree to which new osteons obliterate those previously formed. We now examine the effects of the accumulation of osteons inside a cortical bone cross section using probability theory.

Suppose that within an area $Tt.Ar$ of cross section a new resorption space is equally likely to appear at any location (Fig. 3.21). The area of existing Haversian canals that this resorption space can be expected to obliterate is, on the average,

$$Ar = (Rs.Ar)(Po) \quad (3.42)$$

where $Rs.Ar$ is the area of the resorption space (which will be equivalent to the area of the final osteon, $On.Ar$) and Po is the porosity (total area of Haversian canals/total area, $Po = H.Ca.Ar/Tt.Ar$) of the region. The resorption space will refill to leave a Haversian canal of area $H.Ca.Ar$. The net change in void area to be expected is, when averaged over many such incidents,

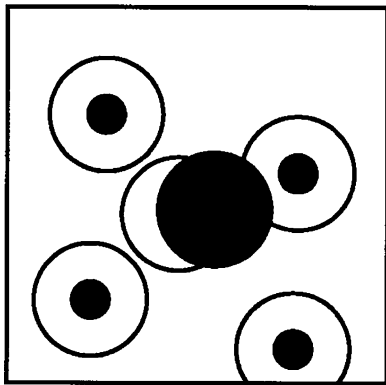


FIGURE 3.21. A new resorption space (*black circle*) appears at a random location in a field of osteons in a cross section having total area $Tt.Ar$.

$$d(\text{H.Ca.Ar}) = \text{H.Ca.Ar} - (\text{On.Ar})(\text{Po}) \quad (3.43)$$

The number of new osteons added to the section per square millimeter per year is just the activation frequency, Ac.f . If one defines an osteon area ratio, f , as $\text{H.Ca.Ar}/\text{On.Ar}$ then in an interval of time dt one has

$$d[(\text{N.On})(\text{H.Ca.Ar})] = (f - \text{Po}) (\text{Ac.f}) (\text{On.Ar})(\text{Tt.Ar})dt \quad (3.44)$$

where Tt.Ar is the total area of the region and N.On is the total number of osteons in the region. Dividing by $(\text{Tt.Ar})dt$, one has

$$d(\text{Po})/dt = (f - \text{Po})(\text{Ac.f})(\text{On.Ar}) \quad (3.45)$$

where $d(\text{Po})/dt = (d/dt)[d(\text{N.On})(\text{H.Ca.Ar})/\text{Tt.Ar}dt]$ is the time rate of change of porosity resulting from Haversian canals. If f , Ac.f , and On.Ar can be assumed to be constant, Eq. (3.45) has the simple exponential solution

$$\text{Po}(t) = f - (f - \text{Po}_0)\exp[-(\text{Ac.f})(\text{On.Ar})t] \quad (3.46)$$

where Po_0 is the initial porosity. When this equation is plotted, it is apparent that as t becomes large, the porosity of the bone does not increase indefinitely, but asymptotically approaches f , which may be viewed as the porosity of each osteon. The equation may be used to study the ways in which osteon dimensions and activation frequency influence the rate of increase of porosity.

Equation 3.46 can be converted to a formula for the number of Haversian canals per unit area ($\text{H.Ca.N}/\text{Tt.Ar}$) by substituting $f = \text{H.Ca.Ar}/\text{On.Ar}$ and $\text{Po} = [(\text{N.On})(\text{H.Ca.Ar})/\text{Tt.Ar}]$. The result is

$$\begin{aligned} \text{H.Ca.N}/\text{Tt.Ar} &= 1/(\text{On.Ar}) \\ &- (1/\text{On.Ar} - [\text{N.On}/\text{Tt.Ar}]_0)\exp[-(\text{Ac.f})(\text{On.Ar})t] \end{aligned} \quad (3.47)$$

or, if the initial osteon density, $[\text{N.On}/\text{Tt.Ar}]_0$, is negligible,

$$\text{H.Ca.N}/\text{Tt.Ar} = [1 - \exp(-(\text{Ac.f})(\text{On.Ar})t)]/\text{On.Ar} \quad (3.48)$$

This variable also is limited in its growth, and one can see that the maximum number of Haversian canals (equal to the number of osteons) that can be fit into total area Tt.Ar is equal to $\text{Tt.Ar}/\text{On.Ar}$ —exactly the number of times the osteon area divides into the total area. (In Chap. 8 it will be seen that this reciprocity between osteon size and packing density has significance in the resistance of bone to crack propagation.) The initial rate at which osteon density increases turns out to be a function only of activation frequency; subsequently, smaller osteons cause $\text{N.On}/\text{Tt.Ar}$ to grow at a faster rate.

Also of interest is the fractional area of a section converted to secondary osteons or their fragments after a given period of time. This fraction,

On.Ar/Tt.Ar, can be calculated as follows. Suppose an increment of resorption space of area $d(Rs.Ar)$ appears in a field of area Tt.Ar of which On.Ar/Tt.Ar has already been remodeled. The expected portion of On.Ar/Tt.Ar that would fall on the already remodeled bone and be removed by the new resorption space is

$$dA = d(Rs.Ar)(On.Ar/Tt.Ar) \quad (3.49)$$

The expected increase in On.Ar is then the fraction of the new space that falls onto the unremodeled portion of the tissue

$$d(On.Ar) = (1 - On.Ar/Tt.Ar)[d(Rs.Ar)] \quad (3.50)$$

In an increment of time dt the number of new resorption spaces appearing on the section would be $(Ac.f)(Tt.Ar)dt$ and the area of each would be $On.Ar_1$, so that $d(Rs.Ar) = (Ac.f)(Tt.Ar)(On.Ar)dt$. Over this period,

$$d[On.Ar] = (1 - [On.Ar/Tt.Ar])(Ac.f)(Tt.Ar)(On.Ar_1)dt \quad (3.51)$$

Reorganizing the equation, one may rewrite this as

$$d[On.Ar/Tt.Ar]/dt = (1 - [On.Ar/Tt.Ar])(Ac.f)(On.Ar_1) \quad (3.52)$$

Assuming that On.Ar/Tt.Ar was initially zero, this differential equation has the solution

$$On.Ar/Tt.Ar = 1 - \exp^{-(Ac.f)(On.Ar_1)t} \quad (3.53)$$

where t is the time during which the remodeling has occurred. On.Ar/Tt.Ar asymptotically approaches 100 %, with the initial rate of increase being equal to the product of activation frequency and osteon size.

The fraction of the cross-sectional area occupied by primary (i.e., unremodeled) bone would be

$$1 - On.Ar/Tt.Ar = \exp^{-(Ac.f)(On.Ar_1)t} \quad (3.54)$$

Naturally, this decreases at the same rate that On.Ar/Tt.Ar increases, eventually reaching zero.

This kind of analysis shows the essential ways in which mechanically important variables such as porosity and osteon density depend on remodeling parameters such as osteon dimensions and activation frequency. If the models can be substantiated by comparison with experimental data, they have the advantage that, unlike empirical correlations between the variables (e.g., polynomial curves), they lead to an understanding of the mechanisms underlying the phenomena.

The derivations of these equations assumed that $On.Ar_1$, $H.Ca.Ar$ and $Ac.f$ are constant. Are these assumptions valid? $H.Ca.Ar$ and $On.Ar_1$ do not vary

much over the course of a lifetime, but we have seen that the activation frequency can vary greatly, contrary to the assumption made in deriving the equations. Figure 3.4 showed that $Ac.f$ may be 20–30 times greater in childhood than in adults, and a substantial increase also occurs as adults age. This problem can be treated by using computer models that incorporate age-corrected values of $Ac.f$ into the aforementioned integrals. On the other hand, Eqs. 3.46, 3.47, 3.53, and 3.54 are useful in showing the basic exponential nature of osteonal accumulation, and they may be reasonably accurate in those regions of cortical bone that are created in late childhood, so that the variations in $Ac.f$ are relatively small over an extended period of time. In general, the outer cortex of most long bones is formed in the latter stages of growth, and many bones “drift” in a particular direction so that much of the adult cross section contains little or no bone from early childhood. For example, most of the adult rib cross section is created after age 12 years (Wu et al. 1970). By comparing the foregoing equations to experimental data, one can find out how this kind of theory holds up in the face of age variations in $Ac.f$ and $On.Ar$.

There are two important applications for these relationships. First, the increase in osteon numbers with age means that age-at-death may be predictable on the basis of the osteon population density or the fraction of the cortex occupied by secondary osteons. Second, the activation frequency may be estimated from the rate at which these variables increase with age. One may also compare the values of these variables with those determined by histomorphometric methods (which have their own limitations, as we have discussed.)

3.7.4.1 Comparison of the Theoretical Results with Experimental Data

Kerley (Kerley 1965) pioneered the use of osteon population density measurements to determine age from skeletal remains. Figure 3.22 contains osteon density data that he published for the midshaft of the femur. These data are means for four sites evenly spaced around the outer half of the cortex. The best least-squares regression fit of these data to Eq. (3.48) yields the curve shown in Fig. 3.22. The form of the theoretical curve matches the data very well, and the parameters of the curve allow one to deduce the mean values of $On.Ar$ and $Ac.f$. (The radius of the osteon, $On.Rd$, is calculated from $On.Ar$ assuming the osteons are circular.) The results of this analysis are shown in the upper part of Table 3.10, along with results for the tibia and fibula obtained from other graphs in Kerley’s paper.³ Note that the fibula is

³ In his 1965 paper, Kerley underestimated his microscopic field area (Kerley and Ubelaker 1978). We have corrected for this in Table 3.10, but not in Fig. 3.22, where the ordinate values should be multiplied by 0.60. Figure 3.23 needs no correction as it represents fractional area.

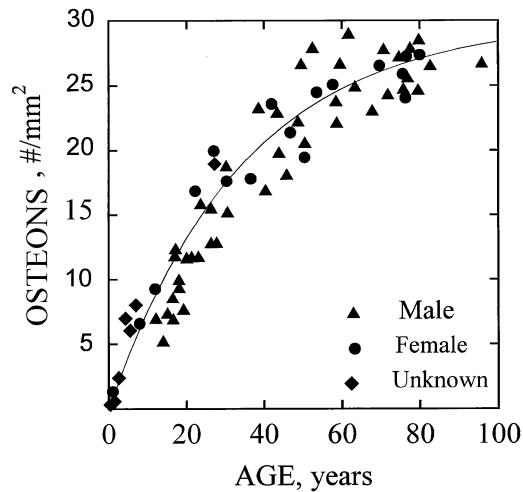


FIGURE 3.22. Osteon density vs. age for a sample of human femurs (Kerley 1965).

TABLE 3.10. Bone remodeling parameters calculated from experimental data

Parameter	Femur	Fibula	Tibia
Osteon density data			
Ac.f(On.Ar), #/year	0.029	0.026	0.018
Ac.f, #/mm²/year	0.53	0.46	0.38
On.Rd, μm	133	137	122
Fraction primary bone data			
Ac.f(On.Ar), #/year	0.031	0.036	0.026
Initial primary bone fraction	0.70	0.55	0.60

From Kerley (1965). Corrected according to data by Kerley and Ubelaker (1978)

similar to the femur with respect to osteon size and activation frequency, but the tibia seems to have smaller osteons and a reduced rate of activation.

Kerley also published graphs relating age to the percent of cross-sectional area occupied by primary bone. Fitting Eq. (3.53) to these graphs yields a second estimate of $(\text{Ac.f})(\text{On.Ar})$ for the same set of samples. Figure 3.23 shows the femoral example, and the second set of $(\text{Ac.f})(\text{On.Ar})$ values appears in the bottom of Table 3.10. The form of Eq. (3.53) again matches the data, and the two estimates of $(\text{Ac.f})(\text{On.Ar})$ for the femur are quite similar. The separate estimates of $(\text{Ac.f})(\text{On.Ar})$ for the other two bones do not agree as well, but this may be partly because they are based on less than half as many data points. Also, while the points in the percent of primary bone vs. age graphs suggest an initiation value of 1, the regression concludes that the cross section started with a significant amount of secondary bone.

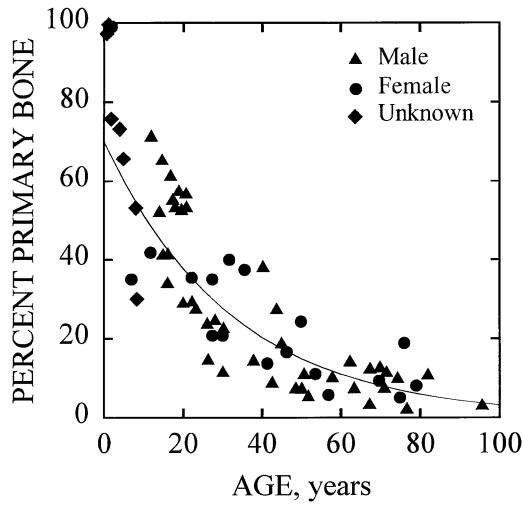


FIGURE 3.23. Percent of cortex that is primary bone vs. age for a sample of human femurs (Kerley 1965).

This conclusion probably reflects the inaccuracy resulting from assuming $Ac.f$ is constant when it is not. The rate of osteon density increase with time was lower in the tibia compared to the femur and fibula, suggesting that $(Ac.f)(On.Ar)$ is different for the tibia compared to the other bones (ie, the tibia remodels differently than the femur and fibula).

Others have independently reached similar conclusions. Evans and Bang (Evans and Bang 1967) correlated mechanical properties with microscopic structure in the same three bones. They, too, found that the fibula has the largest osteons and the tibia has the smallest. As one would expect from this, the fibula had the fewest osteons/mm², but the femur rather than the tibia has the greatest osteon density. The tibia, however, has the greatest number of osteon fragments/mm², suggesting a difference in the way osteons are positioned in these two bones. That is, new osteons in the tibia seem to be more likely to overlap existing osteons than those in the femur, creating more and smaller osteon fragments between whole osteons. With aging, the femur exhibited both the greatest increase in osteon density and the highest $Ac.f$ estimate (see Table 3.10).

As we have noted, Kerley obtained the data in Figs. 3.22 and 3.23 for the purpose of finding the age-at-death for archeological bone specimens. Others have pursued this goal, sometimes correlating age with several different histomorphometric variables (Ahlqvist and Damsten 1969; Bouvier and Ubelaker 1977; Thompson 1979; Stout and Gehlert 1980; Ubelaker 1989). The diminishing slope of the graphs in these figures suggests that determining the age-at-death from the amount of remodeling that has occurred will be less accurate in older individuals, and this is what Walker et al. (Walker

et al. 1994) found. In addition to the slope problem, one cannot be sure that the bone age reflects the person's age-at-death, as was noted. Because the diaphyses of adults expand radially with age, the oldest bone near the endosteal surface is preferentially removed; thus, the average age of the remaining bone may be considerably younger than the person. These limitations serve to illustrate the complex relationship between aging and remodeling in cortical bone, which depends on many other factors as well, including nutrition, activity, genetics, body size, and personal behavior (e.g., smoking and drinking).

The activation frequencies in Table 3.10 are all lower than those for adult rib shown in Table 3.4. There is a paucity of histomorphometrically measured values (i.e., using tetracycline labels) for the activation frequency of osteonal BMUs in bones other than the rib. One reason for this is that $Ac.f$ is so low in cortical bone of the extremities of adults that a great many sections must be examined to find enough remodeling sites for a valid measurement. The theory presented here offers a way to estimate $Ac.f$ for such low-turnover regions and to compare remodeling in one bone with that in another.

3.7.5 *Another Theoretical Approach to Bone Remodeling*

Polig and Jee (Polig and Jee 1990) applied stochastic renewal theory to the problem of determining the relationship between bone turnover and mean skeletal age. Mathematicians have developed the theory of stochastic processes in such a way that it is directly applicable to bone remodeling. By taking advantage of this, some very useful insights and relationships are quickly established. For example, it can be shown that mean skeletal age must be greater than half the average lifetime of any particular modicum of bone tissue. How much greater depends on the degree to which remodeling is random. The authors also showed how, if a single fluorochrome label is given and the amount of this label remaining in the bone is plotted as a function of time, the parameters of bone turnover and mean skeletal age can be obtained from the graph.

Box 3.6 Calculating Activation Frequency Without In Vivo Labels

Suppose that you dug up a mastodon bone in your garden and wanted to know what its average bone formation rate had been. Could you estimate this without benefit of double tetracycline labeling? It is possible to do this by counting the number of osteons/mm² in a section and assuming that the total number of osteons created is the product of the activation frequency and the age of the bone (Wu et al. 1970). However, you have to find a way to estimate the number of osteons that were created but have since been obliterated by more recent remodeling, so that you can add them

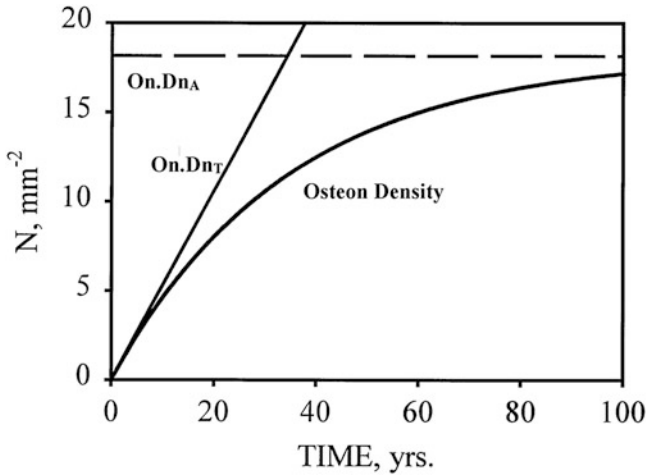


FIGURE 3.24. Theoretical plot of On.Dn (osteons/ mm^2) vs. t (time) for a bone cross section having a constant activation frequency. The curve $\text{On.Dn}(t)$ is that obtained from Eq. (3.48). It has an asymptote at $\text{On.Dn} = \text{On.Dn}_A = 1/\text{On.Ar}$. The line $\text{On.Dn}_T(t)$ represents the total number of osteons/ mm^2 that have been created and would be seen if new osteons never obliterated old ones. Its slope is the activation frequency, Ac.f . The number of observed osteons, On.Dn , is always less than On.Dn_T and On.Dn_A .

to the currently visible osteons to get the grand total. This is shown in Fig. 3.24 of osteon density (On.Dn) vs. bone age (t). As in Fig. 3.23, the curved line represents the visible osteons at any t ; it has an asymptote at On.Dn_A . The straight line represents the total number of osteons created, $\text{On.Dn}_T(t)$. The difference between the two lines is the number of osteons obliterated. Frost (1987) developed a simple algorithm for finding N.On_T when you have measured On.Dn and know the asymptotic value, On.Dn_A . First, calculate $\alpha = \text{On.Dn}/\text{On.Dn}_A$. Then compute $\beta = (1 - \alpha^{3.5}) - 1$, and $\text{On.Dn}_T = \beta \text{On.Dn}$. If you know the age of the bone, t , the activation frequency is estimated as $\text{Ac.f} = \text{On.Dn}_T/t$ and the bone formation rate as $\text{BFR} = (\text{Ac.f})(\text{On.Ar})$, where On.Ar is the mean osteon area. Stout and Paine (Stout and Paine 1994) compared this method's results to measurements made with double tetracycline labels in human ribs. They found that Frost's method gave similar but somewhat lower results, especially in people more than 40 years old. Similar techniques have been used to estimate remodeling rates in human patients (Wu et al. 1970), nonhuman primates (Burr 1992), and modern archeological populations (Burr et al. 1990), as well as in fossils of extinct animals (Wu et al. 1970) or of Neandertal and early Homo sapiens (Abbott et al. 1996; Streeter et al. 2010).

Frost's algorithm is empirical, based on the asymptotic nature of $\text{On.Dn}(t)$ as seen in graphs like those of Kerley (Kerley 1965). One may derive an alternative procedure on the basis of our analysis of randomly overlapping osteonal remodeling. According to Eq. (3.48), the asymptote of the $\text{On.Dn}(t)$ curve is $\text{On.Dn}_A = 1/\text{On.Ar}$. Thus, On.Dn_A can be actually be obtained by measuring On.Ar in the specimen of interest, assuming

that this area does not change appreciably with age. Solving Eq. (3.48) for $Ac.f$, one obtains

$$Ac.f = -\ln[1 - (On.Ar)(On.Dn)]/[(On.Ar)(t)] \quad (3.55)$$

This equation can be solved for the activation frequency if $On.Ar$ and $On.Dn$ are measured directly from the section and the age of the bone, t , is known. Frost's algorithm and this approach give similar results (see Exercise 3.13).

An important difficulty in calculating activation frequency using either of these methods is that the chronological age of an individual often exceeds the age of the bone (t) because of drifts that occur during growth. Based on fluorochrome labeling studies, Wu et al. (Wu et al. 1970) calculated an adjustment of 12.5 years for the middle of the sixth rib to account for the effective "birth" date of the compact bone seen in an adult; i.e., $t = \text{person's age} - 12.5$ years. The correction for other human bones is unknown, and the data necessary to calculate it are not available. Therefore, studies using these techniques often assume the bone age is the subject's age or some other constant; this makes comparison between studies difficult, but maintains consistency within a study.

3.8 Summary and Further Reading

In this chapter we have tried to gain a deeper understanding of bone remodeling. In the process, we have introduced you to several attempts to mathematically model this phenomenon and its effects. We have seen how Frost originally formulated the remodeling process in cortical bone so that it could be histomorphometrically quantified. We have probed deeper into some nooks and crannies of remodeling, trying to pin down details of the refilling process, for example. We concluded by examining the ways in which osteons could accumulate and change bone structure over a lifetime.

We must remember that our mathematical approach was based on the assumption that the morphology and physiology of the osteonal BMU can be expressed quantitatively, even though the histologic picture is one of considerable variation. The reason that this variability does not negate the usefulness of mathematical analysis is that, to be useful, the result must only correctly describe the average behavior of the system. The truth of this is aptly illustrated in the case of the histomorphometric measurement of remodeling dynamics and in the work of Manson and Waters (Manson and Waters 1965). In both cases, conceptual models of osteonal remodeling that assume simple geometric and physiologic features are found to be very successful in describing the average behavior of large numbers of BMUs. It does not matter, for example, that the osteonal apposition rate is not exactly an exponential function, but seems to have an especially rapid early phase. This phenomenon is not apparent when the experimental data are compared to an exponentially varying apposition rate, so it may be regarded as a second-order effect. Second-order effects and histologic variability mark

the limitations of remodeling theories, but do not prevent them from being useful tools for understanding physiology and disease.

Further reading should begin with Frost's 1964 book (Frost 1964), *Mathematical Elements of Lamellar Bone Remodeling*, which laid the groundwork for this subject. *Bone Histomorphometry: Techniques and Interpretation*, edited by Recker (Parfitt 1983), contains several chapters on the details of quantitative histomorphometric analysis, as well as a chapter by Parfitt that remains an excellent resource for the details of bone remodeling. The recent update by Recker et al. (Recker et al. 2011) that deals with problematic cases of very low bone turnover is also worth consulting. The standard reference for histomorphometric technique and nomenclature remains another work by Parfitt, who chaired the ASBMR Task Force to standardize nomenclature (Parfitt et al. 1987). This was recently updated by Dempster et al. (2013). These also contain valuable information on using the plate model for cancellous architecture. Two mathematical analyses of bone remodeling by Polig and Jee (Polig and Jee 1990; Polig and Jee 1987) illustrate how a physicist and an anatomist can very productively collaborate in this area. In addition, several investigators have published computer simulations of bone remodeling: Krumme et al. (Krumme et al. 1984), Kimmel (Kimmel 1985), Martin (Martin 1985), and Reeve (Reeve 1986). Turner et al. (Turner et al. 1993) mathematically analyzed the effects of remodeling on fluoride uptake by the skeleton. Weinhold et al. (Weinhold et al. 1994) examined bone remodeling as a dynamic system capable of oscillation. Finally, more details and examples of the approaches found in this chapter may be found in papers by Martin (Martin 1989; Martin 1984; Martin 1986; Martin 1991).

3.9 Exercises

- 3.1. Derive Eq. (3.13) using the principle of similar triangles in Fig. 3.11. Explain how Eq. (3.14) is represented on the diagram.

The following exercises are intended to start you thinking in terms that can be used to mathematically model the consequences of remodeling processes which are varying in time. To do this, it is necessary to keep track of how many BMUs are currently in each stage of remodeling. These populations depend not on the current activation frequency but on what $Ac.f$ was at some time in the past. For example, BMUs currently in the resorbing phase were activated or “born” between the present time and RP days ago. BMUs in the reversal phase were born before that, and refilling BMUs were born even earlier than that. Also, keep in mind that we are considering the remodeling going on in a representative, two-dimensional cross-section of cortical bone, e.g., Fig. 3.10. In this context, “activation” of a BMU occurs when its leading osteoclasts first break the plane of the section, creating a new resorption cavity. It helps to imagine that you can

observe such a cross section in your own femur as BMUs tunnel across it, excavating and refilling their tunnels.

- 3.2. If a BMU tunnels through cortical bone at a rate of $40 \mu\text{m}/\text{day}$, how long would it take to tunnel 2 cm through your tibia?
- 3.3. Imagine a cross section through your femur. Assume that 5 % of the bone in this cross section is replaced annually by osteonal remodeling. The osteons average $250 \mu\text{m}$ in diameter. What would the osteonal BMU activation rate (new osteons/ mm^2 per year) have to be to produce the 5 % turnover rate?
- 3.4. In Exercise 3.3, if each osteonal BMU produced a new Haversian canal that was $50 \mu\text{m}$ in diameter, what would be the rate of increase in bone porosity (mm^2 of canals/ mm^2 of cross section/year)? Assume that the new osteons do not obliterate preexisting Haversian canals.
- 3.5. Draw a horizontal line representing time. Place a tick mark representing the present time at the left end of the line, so that past time is measured to the right. Let the time scale be 1 cm/week. Assuming the BMU resorption period is $\text{RP} = 21 \text{ days} = 3 \text{ weeks}$, mark the region of the line containing the “birthdays” (activation days) of all BMUs that are currently resorbing bone. Assuming the reversal period is $\text{RvP} = 7 \text{ days} = 1 \text{ week}$, mark the time span during which all BMUs presently in the reversal phase were activated. Finally, assuming the refilling phase lasts $\text{FP} = 56 \text{ days} = 8 \text{ weeks}$, show the time span during which all currently refilling BMUs were “born.”
- 3.6. Using the time line of Exercise 3.5, construct an ordinate axis to represent activation frequency, Ac.f. Plot the following BMU activation history: constant at 2 BMU per mm^2/week until 8 weeks ago, when it suddenly increased to 4 BMU per mm^2/week . It remained at that level until 2 weeks ago, when it rapidly decreased back to the original value. Show graphically the numbers of BMUs/ mm^2 in the three phases of activity (N.Rs, N.Rv, and N.F, respectively).
- 3.7. Using the notation given in Exercises 3.5 and 3.6, write integral equations to represent N.Rs, N.Rv, and N.F as functions of $\text{Ac.f.}(t)$, where t is time. Be sure to specify the integration limits, and continue to let past time be represented by positive numbers.
- 3.8. Now think about the cross-sectional view of a single osteon. Assume that the mean diameter of an osteonal cement line is $200 \mu\text{m}$ and that of the mean Haversian canal is $40 \mu\text{m}$. Using these numbers, calculate Rs.R, the mean rate at which bone is resorbed in each osteon during its resorption phase, in $\mu\text{m}^2/\text{day}$, and F.R, the mean rate at which bone is replaced during refilling of each osteon, in $\mu\text{m}^2/\text{day}$. You will use these

values in the next exercise. (B and C stand for ‘blasts and clasts’, respectively.)

- 3.9. Calculate N.Rs and N.F assuming the activation frequency has been decreasing linearly at a rate of 5 BMU per mm^2/year and is presently 1 BMU per mm^2/year . Use the integration methods you practiced in the earlier exercises, and the same values for RP and FP as before. Then calculate something called the “bone balance”:

$$\text{B.BMU} = \text{F.R}(\text{N.F}) - \text{Rs.R}(\text{N.Rs})$$

using F.R and Rs.R from Exercise 3.8. Based on the equation, what are the units and meaning of F.R(N.F), and Rs.R(N.Rs), and B.BMU? Compare this bone balance with that which exists when the activation frequency is constant at 1 BMU per mm^2/year , assuming everything else remains the same. Why is one number positive and the other negative? Discuss the relevance of this calculation to the porosity of cortical bone in children and adults.

- 3.10. Suppose a tetracycline label is given for 1 week. Write an equation for the volume fraction of the bone that has been labeled by the end of the week. (Remember, volume fraction is the same as the area fraction in our representative cross section. Also, assume that none of the labeled bone has been resorbed during the week.) Use data for a 30- to 39-year-old human from Table 3.4 to calculate the volume fraction. Now, derive an equation for the time rate of decrease of the labeled bone thereafter, assuming it is only removed by remodeling and that the new osteons are randomly placed in the section. Use the concept of Fig. 3.21 and Eq. (3.42), only now let the Haversian canal areas become labeled bone. Calculate the time it would take for half the labeled bone to be removed, assuming again the subject is a 30- to 39-year-old human. What would this “half-life” be for a child?
- 3.11. Discuss the implications of Exercise 3.10 for an adult who has been caught in a nuclear reactor accident and ingested a significant quantity of strontium-90, which is radioactive with a half-life of 28 years and is incorporated into bone mineral in a manner similar to tetracycline. Quantitatively explore the possibilities for speeding up the elimination of the strontium from the victim’s skeleton using what you have learned in this chapter.
- 3.12. Show that the initial slope of the N vs. t curve represented by Eq. (3.48) and Fig. 3.22 is the activation frequency, Ac.f, and that the asymptote is $1/\text{On.Ar}$.
- 3.13. In Box 3.6, two methods were described for determining average activation frequency (Ac.f) from the numbers of secondary osteons visible in a cortical cross section from a bone of known age. Compare

the results of these two methods using the following assumptions. Let the age of the bone be $t = 40$ years and the observable osteon density be $\text{On.Dn} = 40 \text{ mm}^{-2}$. Let the asymptotic value of $\text{On.Dn} = 50 \text{ mm}^{-2}$. Calculate α , β , and Ac.f using Frost's algorithm. Then calculate Ac.f using Eq. (3.48) as described in the box. For consistency, assume that $\text{On.Ar} = 1/\text{On.Dn}$. Stout and Paine (Stout and Paine 1994) found that Frost's method underestimated Ac.f compared to conventional histomorphometric measurements. Does it also yield a lower Ac.f than Equation 3.48?

- 3.14. Differentiate the following features of the bone remodeling system from each other, and describe how you could measure each from bone tissue taken from an experimental animal:

Activation frequency vs. origination frequency

Uncoupling vs. bone balance

An anabolic response of bone to a mechanical stimulus

- 3.15. From a stereological standpoint, measurement of bone remodeling processes from sections of bone cut longitudinally will give different answers from those made on cross-sections. Why (there may be both biological and mathematical answers to this)? How does this affect measurements made of cancellous bone? How would you expect this to alter the comparison of remodeling rates in healthy vs. osteoporotic cancellous bone?

References

- Abbott S, Trinkaus E, Burr DB. Dynamic bone remodeling in later Pleistocene fossil hominids. *Am J Phys Anthropol.* 1996;99:585–601.
- Ahlqvist J, Damsten O. A modification of Kerley's method for the microscopic determination of age in human bone. *J Forensic Sci.* 1969;14:205–13.
- Allen MR, McNerny EMB, Organ JM, Wallace JM. True gold or pyrite: a review of reference point indentation for assessing bone mechanical properties in vivo. *J Bone Miner Res.* 2015;30:1539–50.
- Amsel S, Maniatis A, Tavassoli M, Crosby WH. The significance of intramedullary cancellous bone in the repair of bone marrow tissue. *Anat Rec.* 1969;164:101–11.
- Baron R, Tross R, Vignery A. Evidence of sequential remodeling in rat trabecular bone morphology, dynamic histomorphometry, and changes during skeletal maturation. *Anat Rec.* 1984;208:137–45.
- Barou O, Lafage-Proust MH, Palle S, Vico L, Alexandre C. Effects of immobilization on preosteoblast proliferation assessed histomorphometrically using BRDU uptake on epon-embedded rat bone (abstract). *Bone.* 1996;19:131S.
- Bassett CAL, Mitchell SN, Gaston SR. Treatment of ununited tibial diaphyseal fractures with pulsing electromagnetic fields. *J Bone Joint Surg.* 1981;63A:511–23.
- Bassett EJ, Keith MS, Armelagos GJ, Martin DL, Villanueva AR. Tetracycline-labeled human bone from ancient Sudanese Nubia (A.D. 350). *Science.* 1980;209:1532–4.

- Bliziotes M, Sibonga JD, Turner RT, Orwoll E. Periosteal remodeling at the femoral neck in nonhuman primates. *J Bone Miner Res.* 2006;21:1060–7.
- Bouvier M, Ubelaker DH. A comparison of two methods for the microscopic determination of age at death. *Am J Phys Anthropol.* 1977;46:391–4.
- Boyde A. Scanning electron microscope studies of bone. In: Bourne GH, editor. *The biochemistry and physiology of bone.* New York: Academic Press; 1972.
- Bridges D, Randall C, Hansma PK. A new device for performing reference point indentation without a reference probe. *Rev Sci Instrum.* 2012;83(4):044301.
- Brighton CT, Friedenberg ZB, Black J. Evaluation of the use of constant current in the treatment of nonunion. In: Brighton CT, Black J, Pollack SR, editors. *Electrical properties of bone and cartilage. Experimental effects and clinical applications.* New York: Grune & Stratton; 1979. p. 519–46.
- Burny F, Donkerwolcke M, Bourgois R, Domb M, Saric O. Twenty years experience in fracture healing measurements with strain gauges. *Orthopedics.* 1984;7:1823–9.
- Burr DB. Estimated intracortical bone turnover in the femur of growing macaques: implications for their use as models in skeletal pathology. *Anat Rec.* 1992;232:180–9.
- Burr DB, Guillot III G. Almost invisible, often ignored: periosteum, the living lace of bone. *Medicographia.* 2012;34:221–7.
- Burr DB, Ruff CB, Thompson DD. Patterns of skeletal histologic change through time: comparison of an archaic native American population with modern populations. *Anat Rec.* 1990;226:307–13.
- Burstein AH, Frankel VH. A standard test for laboratory animal bone. *J Biomech.* 1971;4:155–8.
- Calvi LM, Bromberg O, Rhee Y, Lee R, Weber JM, Smith JN, Basil M, Frisch BJ, Bellido T. Osteoblastic expansion induced by parathyroid hormone receptor signaling in murine osteocytes is not sufficient to increase hematopoietic stem cells. *Blood.* 2012;119:2489–99.
- Carter DR, Blenman PR, Beaupre GS. Correlations between mechanical stress history and tissue differentiation in initial fracture healing. *J Orthop Res.* 1988;6:736–48.
- Coccia PF, Krivit W, Cervenka J, Clawson CC, Kersey JH, Kim TH, Nesbit ME, Ramsay NKC. Successful bone marrow transplantation for infantile malignant osteopetrosis. *N Engl J Med.* 1980;302:701–8.
- Cohen J, Harris WH. The three-dimensional anatomy of Haversian systems. *J Bone Joint Surg.* 1958;40A:419–34.
- Collier RJ, Donarski R. Non-invasive method of measuring resonant frequency of a human tibia in vivo. *J Biomed Eng.* 1987;9:329–31.
- Cunningham JL, Evans M, Kenwright J. Measurement of fracture movement in patients treated with unilateral external skeletal fixation. *J Biomed Eng.* 1989;11:118–22.
- Day JS, Ding M, Odgaard A, et al. Parallel plate model for trabecular bone exhibits volume fraction-dependent bias. *Bone.* 2000;27(5):715–20.
- Dempster DW, Compston JE, Drezner MK, Glorieux FH, Kanis JA, Malluche H, et al. Standardized nomenclature, symbols and units for bone histomorphometry: a 2012 update on the report of the ASBMR Histomorphometry Nomenclature Committee. *J Bone Miner Res.* 2013;28:1–16.

- Diez-Perez A, Güerri R, Nogues X, et al. Microindentation for in vivo measurement of bone tissue mechanical properties in humans. *J Bone Miner Res.* 2010;25:1877–85.
- Eriksson C. Streaming potentials and other water dependent effects in mineralized tissues. *Ann N Y Acad Sci.* 1974;238:321–38.
- Evans FG, Bang S. Differences and relationships between the physical properties and microscopic structure of human femoral, tibial and fibular cortical bone. *Am J Anat.* 1967;120:79–88.
- Ferron M, Wei J, Yoshizawa T, Del Fattore A, DePinho RA, Teti A, Ducy P, Karsenty G. Insulin signaling in osteoblasts integrates bone remodeling and energy metabolism. *Cell.* 2010;142:296–308.
- Frost HM. Preparation of thin undecalcified bone sections by a rapid manual method. *Stain Technol.* 1958;33:272.
- Frost HM. Measurement of bone formation in a 57-year-old man by means of tetracycline. *Henry Ford Hosp Med Bull.* 1960a;8:212–9.
- Frost HM. Observations on osteoid seams: the existence of a resting state. *Henry Ford Hosp Med Bull.* 1960b;8:220–4.
- Frost HM. Measurement of osteocytes per unit volume and volume components of osteocytes and canaliculae in man. *Henry Ford Hosp Med Bull.* 1960c;8:208–11.
- Frost HM. Mathematical elements of lamellar bone remodelling. Springfield: Thomas; 1964.
- Frost HM. Tetracycline-based histological analysis of bone remodeling. *Calcif Tissue Res.* 1969;3:211–37.
- Frost HM. Bone histomorphometry: correction of the labeling “escape error”. In: Recker RR, editor. *Bone histomorphometry: techniques and interpretation.* Boca Raton: CRC Press; 1983.
- Frost HM. Secondary osteon populations: an algorithm for estimating the missing osteons. *Yearb Phys Anthropol.* 1987;30:239–54.
- Fuchs RK, Allen MR, Ruppel ME, Dab T, Phipps RJ, Miller LM, Burr DB. In situ examination of the time-course for secondary mineralization of Haversian bone using synchrotron Fourier transform infrared microspectroscopy. *Matrix Biol.* 2008;27:34–41.
- Fukada E, Yasuda I. On the piezoelectric effect of bone. *J Physical Soc Japan.* 1957;10:1158–69.
- Fuzzele K, Riddle RC, DiGirolamo DJ, Cao X, Wan C, Chen D, Faugere M-C, Aja S, Hussain MA, Brüning JC, Clemens TL. Insulin receptor signaling in osteoblasts regulates postnatal bone acquisition and body composition. *Cell.* 2010;142:309–19.
- Gerlanc M, Haddad D, Hyatt GW, Lanlo HT, Hilaire P. Ultrasonic study of normal and fractured bone. *Clin Orthop Relat Res.* 1975;111:175–80.
- Harris WH. A microscopic method of determining rates of bone growth. *Nature.* 1960;188:1038–9.
- Hattner R, Frost HM. Mean skeletal age: its calculation and theoretical effects on skeletal tracer physiology and on the physical characteristics of bone. *Henry Ford Hosp Med Bull.* 1963;11:201–16.
- Havers C. *Osteologia nova, or some new observations of the bones, and the parts belonging to them, with the manner of their accretion, and nutrition, communicated to the royal society in several discourses.* London: Samuel Smith; 1691.

- Holtrop ME. Light and electron microscopic structure of bone-forming cells. In: Hall BK, editor. *Bone*, vol. 1. The osteoblast and osteocyte. Caldwell, NJ: Telford Press; 1990. p. 1–39.
- Ineke DCJ, Vermeer JAF, Bloemen V, Stap JM, Everts V. Osteoclast fusion and fission. *Calcif Tissue Int*. 2012;90:515–22.
- Jaworski JFG, Hooper C. Study of cell kinetics within evolving secondary Haversian systems. *J Anat*. 1980;131:91–102.
- Jaworski ZF, Lok E. The rate of osteoclastic bone erosion in Haversian remodeling sites of adult dog's rib. *Calcif Tissue Res*. 1972;10:103–12.
- Jaworski ZFG, Duck B, Sekaly G. Kinetics of osteoclasts and their nuclei in evolving secondary Haversian systems. *J Anat*. 1981;133:397–405.
- Johnson LC. Morphologic analysis in pathology: the kinetics of disease and general biology in bone. In: Frost HM, editor. *Bone biodynamics*. Boston: Little, Brown; 1964.
- Kelin M, Frost HM. Evidence of periodic changes in the rate of formation of individual osteons in human bone. *Henry Ford Hosp Med Bull*. 1964;12:565–72.
- Kerley ER. The microscopic determination of age in human bone. *Am J Phys Anthropol*. 1965;23:149–64.
- Kerley ER, Ubelaker DH. Revision in the microscopic method of estimating age at death in human cortical bone. *Am J Phys Anthropol*. 1978;49:545–6.
- Kimmel DB. A computer simulation of the mature skeleton. *Bone*. 1985;6:369–72.
- Kimmel DB, Jee WSS. A quantitative histologic study of bone turnover in young adult beagles. *Anat Rec*. 1982;203:31–45.
- Krumme H, Guenay M, Nagel H-H, Delling G. Computerized simulation of bone remodeling: graphic demonstration of dynamic processes. *Metab Bone Dis Relat Res*. 1984;5:253–7.
- Lee S-H, Kim TS, Choi Y, Lorenzo J. Osteoimmunology: cytokines and the skeletal system. *BMB Rep*. 2008;41:495–510.
- Li XJ, Jee WSS, Chow SY, Woodbury DM. Adaptation of cancellous bone to aging and immobilization: a single photon absorptiometry and histomorphometric study. *Anat Rec*. 1990;227:12–24.
- Li CY, Price C, Delisser K, Nasser P, Laudier D, Clement M, Jepsen KJ, Schaffler MB. Long-term disuse osteoporosis seems less sensitive to bisphosphonate treatment than other osteoporosis. *J Bone Miner Res*. 2005;20:117–24.
- Liu XS, Shane E, McMahon DJ, Guo XE. Individual trabecula segmentation (ITS)-based morphological analysis of microscale images of human tibial trabecular bone at limited spatial resolution. *J Bone Miner Res*. 2011;26:2184–93.
- Lorenzo J, Horowitz M, Choi Y. Osteoimmunology: interactions of the bone and immune system. *Endocr Rev*. 2008;29:403–40.
- Manson JD, Waters NE. Observations on the rate of maturation of the cat osteon. *J Anat*. 1965;99:539–49.
- Martin RB. Porosity and specific surface of bone. *Crit Rev Biomed Eng*. 1984;10:179–222.
- Martin RB. The usefulness of mathematical models for bone remodeling. *Yearb Phys Anthropol*. 1985;28:227–36.
- Martin RB. Interpretation of fluorochrome labeling results in dynamic remodeling experiments. *Trans Orthop Res Soc*. 1986;11:247.

- Martin RB. Osteonal remodeling in response to screw implantation in canine femora. *J Orthop Res.* 1987;5:445–52.
- Martin RB. Label escape theory revisited: the effects of resting periods and section thickness. *Bone.* 1989;10:255–64.
- Martin RB. On the significance of remodeling space and activation rate changes in bone remodeling. *Bone.* 1991;12:391–400.
- Martin RB, Burr DB. The structure, function, and adaptation of compact bone. New York: Raven; 1989.
- Martin RB, Dannucci GA, Hood SJ. Bone apposition rate differences in osteonal and trabecular bone. *Trans Orthop Res Soc.* 1987;12:178.
- Martin RB, Chow BD, Lucas PA. Bone marrow fat content in relation to bone remodeling and serum chemistry in intact and ovariectomized dogs. *Calcif Tissue Int.* 1990;46:189–94.
- Milch RA, Rall DP, Tobie JE, Albrecht JM, Trivers G. Florescence of tetracycline antibiotics in bone. *J Bone Joint Surg.* 1958;40A:897–910.
- Minaire P, Edouard C, Arlot M, Meunier P. Marrow changes in paraplegic patients. *Calcif Tissue Int.* 1984;36:338–40.
- Mollon B, da Silva V, Busse JW, Einhorn TA, Bhandari M. Electrical stimulation for long-bone fracture-healing: a meta-analysis of randomized controlled trials. *J Bone Joint Surg.* 2008;90-A:2322–30.
- Monro A. The anatomy of the human bones, nerves and lacteal sac and duct. Dublin; 1776.
- Mouton PR. Unbiased stereology: a concise guide. Baltimore: Johns Hopkins University Press; 2011.
- Mullender MG, van der Meer DD, Huiskes R, Lips P. Osteocyte density changes in aging and osteoporosis. *Bone.* 1996;18:109–13.
- Nunamaker DM, Butterweck DM, Provost MT. Fatigue fractures in thoroughbred racehorses: relationships with age, peak bone strain, and training. *J Orthop Res.* 1990;8:604–11.
- Orwoll ES. Perspective: toward an expanded understanding of the role of the periosteum in skeletal health. *J Bone Miner Res.* 2003;18:949–54.
- Owen M. The origin of bone cells in the post-natal organism. *Arthritis Rheum.* 1980;23:1073–9.
- Parfitt AM. Quantum concept of bone remodeling and turnover: implications for the pathogenesis of osteoporosis. *Calcif Tissue Int.* 1979;28:1–5.
- Parfitt AM. The physiologic and clinical significance of bone histomorphometric data. In: Recker RR, editor. *Bone histomorphometry techniques and interpretation.* Boca Raton: CRC Press; 1983. p. 143–223.
- Parfitt AM, Drezner MK, Glorieux FH, Kanis JA, Malluche H, Meunier PJ, Ott SM, Recker RR. Bone histomorphometry: standardization of nomenclature, symbols, and units. *J Bone Miner Res.* 1987;2(6):595–610.
- Polig E, Jee WSS. Bone age and remodeling: a mathematical treatise. *Calcif Tissue Int.* 1987;41:130–6.
- Polig E, Jee WSS. A model of osteon closure in cortical bone. *Calcif Tissue Int.* 1990;47:261–9.
- Recker RR, Kimmel DB, Dempster D, Weinstein RS, Wronski TJ, Burr DB. Issues in modern bone histomorphometry. *Bone.* 2011;49:955–64.
- Reeve J. A stochastic analysis of iliac trabecular bone dynamics. *Clin Orthop Relat Res.* 1986;213:264–78.

- Roberts WE, Mozsary PG, Klingler E. Nuclear size as a cell-kinetic marker for osteoblast differentiation. *Am J Anat.* 1982;165:373–84.
- Roberts WE, Morey-Holton E, Gonsalves MR. Sensitivity of bone cell populations to weightlessness and simulated weightlessness. In: *The gravity relevance in bone mineralisation processes*. Paris: European Space Agency; 1984.
- Roberts WE, Wood HB, Burk DT. Vascularly oriented differentiation gradient of osteoblast precursor cells in rat periodontal ligament. *J Periodontal Res.* 1987;22:461–7.
- Robling AG, Niziolek PJ, Baldridge LA, et al. Mechanical stimulation of bone in vivo reduces osteocyte expression of Sost/Sclerostin. *J Biol Chem.* 2008;283:5866–75.
- Schock CC, Noyes FR, Villanueva AR. Measurement of Haversian bone remodeling by means of tetracycline labeling in rib of rhesus monkeys. *Henry Ford Hosp Med Bull.* 1972;20:131–44.
- Smale S. Book review: catastrophe theory: selected papers. *Bull Am Math Soc.* 1978;84:1360–8.
- Stout SD, Gehlert SJ. The relative accuracy and reliability of histological aging methods. *Forensic Sci Int.* 1980;15:181–90.
- Stout SD, Paine RR. Brief communication: bone remodeling rates: a test of an algorithm for estimating missing osteons. *Am J Phys Anthropol.* 1994;93:123–9.
- Streeter M, Stout S, Trinkaus E, Burr D. Brief communication: bone remodeling rates in Pleistocene humans are not slower than the rates observed in modern populations: a reexamination of Abbott et al. (1996). *Am J Phys Anthropol.* 2010;141:315–8.
- Takayanagi H. Osteoimmunology: shared mechanisms and crosstalk between the immune and bone systems. *Nat Rev Rheumatol.* 2007;7:292–304.
- Tam CS, Harrison JE, Heersche JNM, Jones G, Wilson DR, Parsons JA, Murray TM. Short term variation in the rate of apposition of mineralized bone matrix in small animals. In: Jee WSS, Parfitt AM, editors. *Bone histomorphometry. Third international workshop*. Paris: Societe Nouvelle Publications Medicales et Dentaires; 1980.
- Tappen NC. Three dimensional studies of resorption spaces and developing osteons. *Am J Anat.* 1977;149:301–32.
- Thompson DD. The core technique in the determination of age at death in skeletons. *J Forensic Sci.* 1979;44:902–15.
- Turner CH, Boivin G, Meunier PJ. A mathematical model for fluoride uptake by the skeleton. *Calcif Tissue Int.* 1993;52:130–8.
- Ubelaker DH. *Human skeletal remains*. Washington, DC: Smithsonian Institution Press; 1989.
- Vincent J. *Recherches sur la Constitution de l'Os Adult*. Brussels: Arscia; 1955.
- Walker RA, Lovejoy CO, Meindl RS. Histomorphological and geometric properties of human femoral cortex in individuals over 50: implications for histomorphological determination of age-at-death. *Am J Hum Biol.* 1994;6:659–67.
- Weinhold PS, Gilbery JA, Woodard JC. The significance of transient changes in trabecular bone remodeling activation. *Bone.* 1994;15:577–84.
- White AA, Panjabi MM, Southwick WO. The four biomechanical stages of fracture repair. *J Bone Joint Surg.* 1977;59A:188–92.
- Wu K, Schubeck KE, Frost HM, Villanueva AR. Haversian bone formation rates determined by a new method in a mastodon, and in human diabetes mellitus and osteoporosis. *Calcif Tissue Int.* 1970;22:204–19.

4

Mechanical Properties of Ligament and Tendon

The movement of animals is like that of automatic puppets...For they have functioning parts that are of the same kind: the sinews and bones. The latter are like the pegs and the iron in our example, the sinews like the cables.

Aristotle (384–322 B.C.), *De Motu Animalium*

Abstract In Chap. 2 we explored the architecture and composition of ligaments and tendons, non-calcified but nonetheless extremely tough skeletal tissues responsible for binding bones together, transmitting forces from muscles and constraining motion within normal limits. To briefly review, ligaments and tendons are composed of linearly arranged collagen molecules assembled in a hierarchal fashion into subfibrils, fibrils and fibers. Collagen fibers within ligaments are architecturally orientated to effectively control and constrain joint motion; in tendon they are grouped into distinct but parallel fascicles.

Keywords Elastic modulus • Residual strain • Aponeurosis • Viscoelastic

4.1 Quasi-static Tensile Properties

In Chap. 2 we explored the architecture and composition of ligaments and tendons, non-calcified but nonetheless extremely tough skeletal tissues responsible for binding bones together, transmitting forces from muscles and constraining motion within normal limits. To briefly review, ligaments and tendons are composed of linearly arranged collagen molecules assembled in a hierarchal fashion into subfibrils, fibrils and fibers. Collagen fibers within ligaments are architecturally orientated to effectively control and constrain joint motion; in tendon they are grouped into distinct but parallel fascicles (Figs. 2.24 and 2.25). In this chapter we will take an in-depth look at the mechanical behavior of these skeletal soft tissues and some of the specialized laboratory and in-vivo approaches used to determine their behaviors. We will also briefly examine some common sites of injury, repair and assessment.

When loaded in the laboratory, ligaments and tendons produce characteristic load-deformation patterns that reflect their structural architecture. Materials such as steel and aluminum display linear stress-strain characteristics because of their crystalline nature. In these materials, the modulus of elasticity is constant throughout the elastic region and is representative of the atomic forces at work within their crystalline architecture. If the elastic limit is exceeded, crystalline materials will yield and eventually fail. In contrast, noncrystalline materials display what is termed *rubbery elasticity*. Tensile tests of these materials produce continuous upwardly concave load-displacement and stress-strain curves. In the noncrystalline materials, the elastic modulus is dependent on load magnitude because of intra- and intermolecular forces and molecular cross-linking. Failure of these materials is often abrupt and can occur without a well-defined yield point. Tensile tests of tendon and ligament produce curves that display both crystalline and rubbery characteristics, reflecting their complex structural hierarchy and molecular interactions.

A typical load-deformation (or stress-strain) curve for tendon or ligament (Fig. 4.1) can be divided into three regions. In the *toe region*, collagen crimps are removed by elongation. Initially minimal force is required for this elongation, but as the collagen fibers straighten, more and more force is required and the curve swings upward. The toe region typically ends at about 1.5–3.0 % strain. It is a matter of some controversy whether the toe region is physiologically or functionally important in vivo. It may actually be artifactual and the result of assuming that the normal resting length of the tendon or

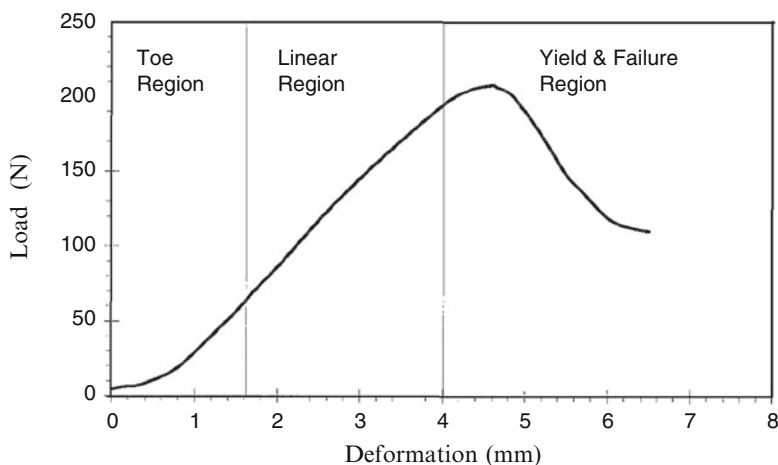


FIGURE 4.1. Typical load-deformation curve from a tendon or ligament showing the toe, linear, and yield and failure regions. The stress-strain curve would have a similar shape. Reprinted with permission from Simon SR (ed): *Orthopaedic Basic Science*. Rosemont, IL, American Academy of Orthopaedic Surgeons, 1994 (Woo et al. 1994).

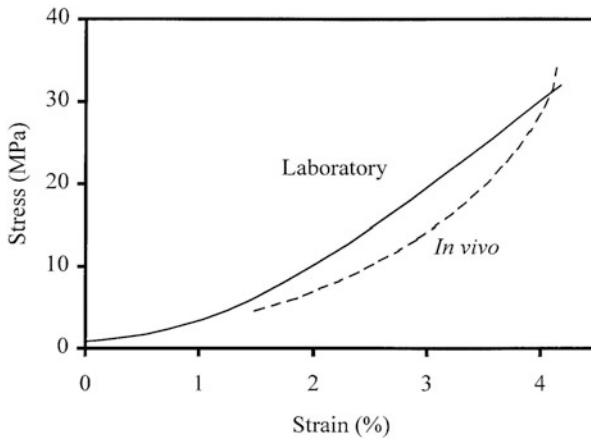


FIGURE 4.2. Different stress-strain curves for equine superficial flexor tendon are obtained when measured *in vivo* (*dashed line*) and on specimens removed from the animal and loaded in the laboratory (*solid line*). Previously published as Fig. 3 from Gillis, C., Sharkey, N.A., Stover, S.M., Pool, R.R., and Meagher, D.M. The effect of maturation and aging on the material and ultrasonographic properties of equine superficial digital flexor tendon. *American Journal of Veterinary Research* 1995; 56:1345–1350 (Gillis et al. 1995).

ligament *in vivo* occurs at zero load. Gillis et al. (1995) showed that in the forelimb of the standing horse the superficial flexor tendon has a strain of about 1.5 % compared to the extracted tendon under zero load (Fig. 4.2). Therefore during standing, a resting position for the horse, the structure is pre-strained beyond the toe region seen in laboratory tests of extracted tendons.

Further loading produces a *linear region* in which the molecular crosslinks of the collagen are stressed (see Fig. 4.1). As for bone, the slope of this region of the load-deformation curve represents the stiffness of the tendon or ligament, and the slope of the stress-strain curve is a measure of the tissue's modulus of elasticity (see Fig. 4.6). In practice, the slopes of these curves are usually determined at a standardized load or stress, close to the limits of the most linear portion of these curves. Elastic modulus values reported for tendon and ligament range between 1.0 and 2.0 GPa; the modulus of highly specialized tissues (such as the *ligamentum nuchae*, described further in Sect. 4.6) can be substantially less.

In the final region of the curve, a reduction in slope marks the *yield point* or onset of cross-link or fiber damage. If loading is continued, the tendon or ligament will eventually fail. Tendons and ligaments can be strained to between 5 and 7 % without damage. The maximum strain that a ligament or tendon can endure before failure is generally in the neighborhood of 12–15 %; however, substantial damage to the collagen network can occur well

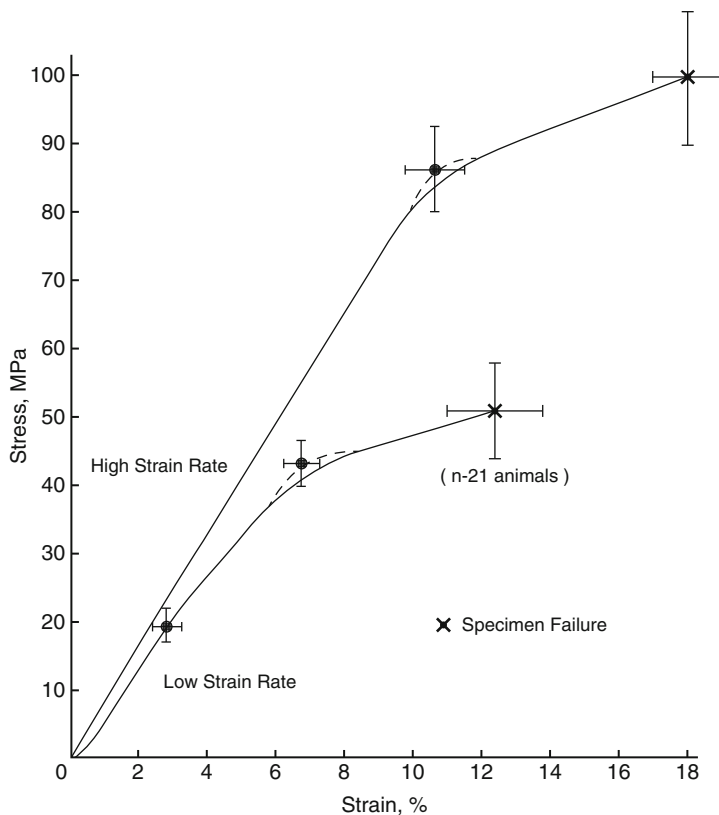


FIGURE 4.3. Effect of strain rate on stress-strain curve for rat tail tendon. At high strain rates the linear portion of the curve is extended. Previously published as Fig. 3 from Haut, R.C. Age-dependent influence of strain rate on the tensile failure of rat-tail tendon. *Journal of Biomechanical Engineering* 1983; 110:296–299 (Haut 1983).

before complete rupture. Ligaments with unusually high elastin content can be strained to 30 % or more without damage. Ultimate tensile strengths for tendons and ligaments range from 50 to 150 MPa (for comparison, recall that the tensile strength of bone is about 150 MPa).

4.2 Viscoelastic Properties

The mechanical behavior of ligament and tendon is *viscoelastic*; thus, their load-displacement and stress-strain relationships are **rate** and **history dependent**. Testing parameters must always be carefully considered when establishing or comparing material properties such as ultimate strength, ultimate strain, and modulus of elasticity. Two representative stress strain curves for rat tail tendon are plotted in Fig. 4.3. The longer and shorter curves

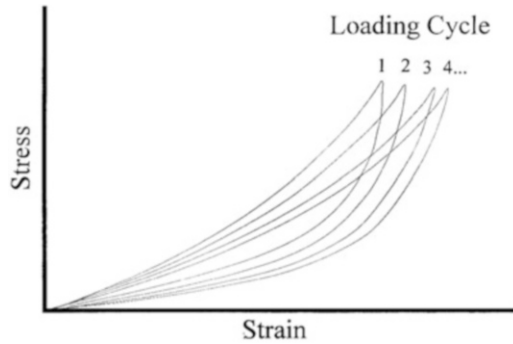


FIGURE 4.4. Energy is lost each time a tendon is repeatedly loaded and unloaded, a process called hysteresis. The consequence of this over time is reduced stiffness and greater deformation for a given load or stress. Generally, a steady state develops after 10–20 cycles.

were produced using strain rates of 720 % per second and 3.6 % per second, respectively. As can be seen, the higher strain rate resulted in higher stress and strain at both failure and yield, and a slightly increased elastic modulus. Compare this behavior with that of bone in Fig. 7.13. Note that the ductility (energy absorption) of bone is decreased at higher strain rates, but that of tendon is increased.

Tendon and ligament display *hysteresis*, another behavior caused, at least in part, by viscoelasticity. Hysteresis refers to loss of energy during a load-deformation cycle (Fig. 4.4). If a purely elastic material is loaded in its elastic range and then immediately unloaded, the load-deformation curve for the loading phase is identical to the load-deformation curve for the unloading phase. The energy absorbed during loading and returned to the system during unloading will be the same, and the system will be 100 % efficient. The loading-unloading profiles of ligaments and tendons are not identical, however. The area under the loading curve is greater than the area under the unloading curve, and the difference between the two represents the energy lost during the load cycle, which is referred to as hysteresis. The amount of hysteresis observed depends on the particular ligament or tendon and on the testing conditions. Tendons involved in gait tend to be efficient stores of elastic energy (i.e., have minimal hysteresis). For example, the digital flexor tendons of horses studied by Gillis et al. (1995) lost only about 5 % of their strain energy during a loading-unloading cycle.

Where the lost energy goes is difficult to determine. Frictional forces at work within the sample cause some energy to be converted to heat, some of which flows to the surroundings while the remainder remains within the sample, raising its temperature and altering its mechanical properties. When the load is removed, the accumulated thermal energy slowly dissipates

and the tendon or ligament may shorten further with time. However, *residual strain* may remain in the tissue for some time or indefinitely. Some of the lost strain energy has probably been used to break molecular cross-links or otherwise disrupt the molecular structure of the tissue.

Figure 4.4 demonstrates the effect of loading history. The mechanical behavior in cycle four is clearly different than that in cycle one. The shift to the right reflects a decrease in tendon or ligament stiffness and may be accompanied by a decrease in ultimate strength. For most tendons and ligaments a steady state is achieved after 10–20 cycles of loading. However, if the peak load is increased on a subsequent cycle, the steady state is disrupted and additional cycles of loading will be required to reestablish stability. Conditioning effects have been measured in-vivo and are an inherent time-dependent property reflective of actual physiologic behavior and not just an artifact of laboratory testing (Maganaris 2003). Tendons and ligaments are also subjected to constant loads or deformations, and, like bone and cartilage, they exhibit *creep* under constant load and *stress relaxation* under constant deformation (Fig. 4.5).

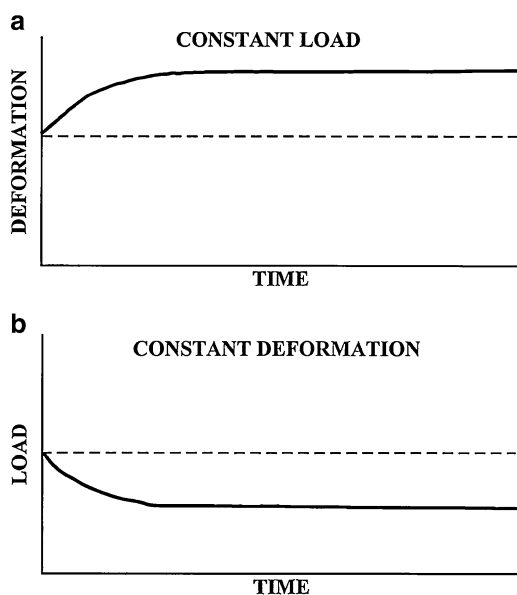


FIGURE 4.5. (a) When a constant load is applied to a tendon or ligament, its deformation gradually “creeps” to a stable value. (b) When a constant deformation is applied, the load (material stress) “relaxes” to a stable value. The times required for these processes are of the order of minutes or hours.

4.3 Mathematical Modeling

When modeling musculotendon dynamics, many investigators have chosen to simulate tendon as a purely elastic component for simplicity's sake. Zajac (1989) assumed pure tendon elasticity in developing a generic musculotendon force-strain curve using two parameters, the *tendon slack length*, L_S , and the *peak isometric muscle force*, F_0^M . L_S is the length at which the tendon just begins to exhibit some resistance to lengthening. Tendon strain is defined as $\epsilon = (L - L_S)/L_S$. Suppose that the force-strain curve for a tendon has the shape shown in Fig. 4.6. The peak isometric force is measured for the tendon's muscle and used to normalize the tendon force values (as shown on the right-hand axis). By dividing the force values by tendon cross-sectional area, the normalized force-strain curve, can be turned into a stress-strain curve which defines the response of a generic tendon. This relationship can then be scaled to obtain the force-length response of a musculotendon unit having a different area, F_0^M , and L_S .

Actually, of course, tendon tissue is viscoelastic and nonlinear. Lieber and associates (1992) developed an empirical model for muscle–tendon units by isolating and measuring the individual mechanical behavior of the frog semitendinosus muscle and its tendon and *aponeurosis* (the connection between the two; see following). They obtained the following regression equations for tendon and aponeurosis material behavior:

$$F_T = 10^{(\epsilon + 0.633)/1.35} \quad (4.1)$$

$$F_A = 10^{(\epsilon + 3.63)/5.66} \quad (4.2)$$

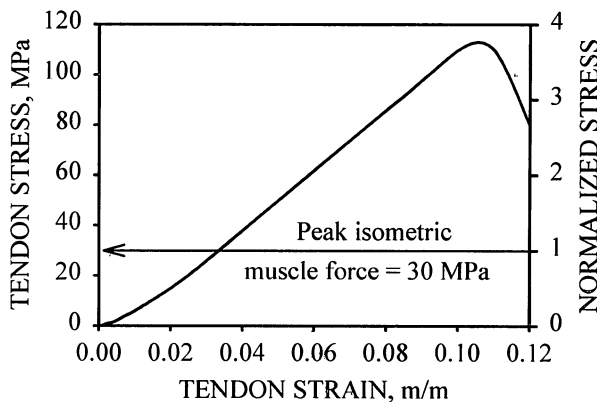


FIGURE 4.6. The stress-strain curve for tendon may be normalized by the peak isometric muscle force, F_0^M (expressed per unit cross-sectional area of the muscle), to obtain a generic tendon response shown by the scale on the right side of the graph (Zajac 1989).

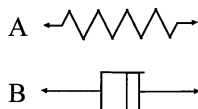


FIGURE 4.7. Symbols used for the two basic rheological model components: a spring [A] and a dashpot [B].

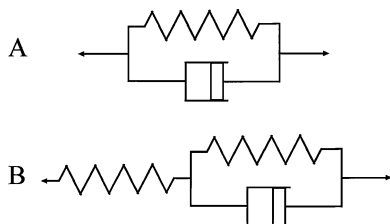


FIGURE 4.8. Two rheological models: a Kelvin solid [A] and a three-parameter solid [B].

where ϵ is percent tendon or aponeurosis strain and F_T and F_A are the corresponding tensions in the tendon or aponeurosis, respectively, expressed as percent of the maximum tetanic muscle tension.

The viscoelastic properties of tendon or ligament are often represented using rheological models composed of two basic components: springs and dashpots (Fig. 4.7). The elastic component of their behavior is represented by the spring element, in which strain is linearly proportional to the applied stress. The dashpot represents the viscous behavior, in which the strain rate is linearly proportional to the applied stress. (The dashpot is a “leaky” piston in a cylinder; under load the fluid slowly flows past the piston, allowing it to move at a velocity proportional to the load.) In practice, combinations of elements are connected to simulate various behaviors. Two common models are the Kelvin solid and the three-parameter solid (Fig. 4.8). It is interesting to compare the behaviors of these models to those seen in Fig. 4.5. For the Kelvin solid, the strain, ϵ , is the same in each element, and the total stress supported by the combination is

$$\sigma = E\epsilon + \eta\dot{\epsilon} \quad (4.3)$$

where $\dot{\epsilon}$ is strain rate, E is the elastic modulus of the spring, and η is the viscosity of the dashpot. In a creep test, one has a constant load, σ_0 , and Eq. (4.3) can be integrated to obtain

$$\epsilon = (\sigma_0/E) \left[1 - e^{-Et/\eta} \right] \quad (4.4)$$

As t becomes very large, the strain asymptotically approaches σ_0/E . Initially, however, $t=0$ and $\epsilon=0$; contrary to Fig. 4.5, there is no instantaneous

elastic response in a creep test. When a constant strain, ϵ_0 , is applied, the second term in Eq. (4.3) is zero, and the stress is constant at $E\epsilon_0$. Thus, there is no relaxation effect for this model, either.

For the three-parameter model, a single spring is in series with a Kelvin solid. The equation for the series spring is $\sigma = E'\epsilon'$ and that for the Kelvin element is $\sigma = E''\epsilon'' + \eta\dot{\epsilon}''$. Using the LaPlace transform method (Flügge 1967), these equations may be combined to obtain

$$(E' + E'') \sigma + \eta \dot{\sigma} = E' E'' \epsilon + E' \eta \dot{\epsilon} \quad (4.5)$$

Using the LaPlace method, one may obtain for the creep solution

$$\epsilon = \left(\sigma_0 / E' \right) [E^* (1 - e^{-\lambda t}) + e^{-\lambda t}] \quad (4.6)$$

where $E^* = (E' + E'')/E''$ and $\lambda = E''/\eta$. In this case, at $t = 0$ there is an instantaneously produced strain σ_0/E' , and the strain subsequently approaches $\sigma_0 E^*/E'$. For the relaxation solution, one obtains

$$\sigma = E' \epsilon_0 [(1 - e^{-\lambda t})/E^* + e^{-\lambda t}] \quad (4.7)$$

where ϵ_0 is the applied strain and $\gamma = (E' + E'')/\eta$. Again, there is a stress $E'\epsilon_0$ at $t = 0$, and this subsequently relaxes asymptotically to $E'\epsilon_0/E^*$. Thus, this model exhibits the general behavior observed in tendon and ligament. By fitting its equations to experimental data, values for E' , E'' , and η can be obtained. These results may be useful for predictive purposes, but the model is still phenomenological rather than mechanistic. That is, we have not learned much about how the viscoelastic behavior is caused by the tissue structure. What parts of the tendon have the attributes associated with E' and E'' ? Presumably η is associated with a flow of water, but where, exactly? These questions remain open.

A more elaborate and popular nonlinear viscoelastic theory was described by Fung (1972). This quasi-linear viscoelastic (QVL) model mathematically describes the stress relaxation of collagenous tissues as the product of independent functions of time and strain:

$$\sigma[\epsilon(t); t] = G(t)^* \sigma^e(\epsilon) \quad (4.8)$$

where σ^e is the *elastic response*, a function of strain (ϵ) only, and G is the *reduced relaxation function*, which is dependent on time (t) only. Versions of this theory have been successfully applied by other investigators as well (Haut and Little 1972; Jenkins and Little 1974; Woo 1982). However, the form of this model assumes time and strain to be separable functions, i.e. time dependent relaxation behavior ($G(t)$) is the same at any given strain (ϵ). Experimental evidence demonstrates that this is not the case (Hingorani et al. 2004; Duenwald et al. 2010) indicating that even more elaborate nonlinear models are required to accurately describe the strain dependent

relaxation behavior of collagenous tissue. A specialized transversely isotropic viscoelastic constitutive model developed by Davis and De Vita (2012) considers finite strain, tissue anisotropy and the strain dependent nature of stress relaxation. The model was able to represent the increase in stress relaxation with strain for rat tail tendon fascicles strained to 0.98, 2.00, and 3.28 %. Their phenomenological model remains in development, but is an improvement for tendon over the QLV model where the relaxation kernel is a function of strain.

4.4 Measuring Structural and Material Properties

A load-deformation curve, such as that shown in Fig. 4.1, is obtained by securely mounting each end of a sample in a material testing apparatus and then elongating the sample until it ruptures. Stress and strain behavior can then be generated by normalizing load to cross-sectional area and elongation to original length. However, unlike metals, plastics, or even bone, tendons and ligaments cannot easily be machined into standardized testing specimens. To further compound matters, they are difficult to grip, perhaps even more sensitive to moisture and temperature than bone, and have complex internal structures that make data interpretation difficult. For these reasons and others accurate and reproducible laboratory measurement of tendon and ligament tensile properties is notoriously difficult.

4.4.1 *Storage and Testing Environments*

The treatment and storage of ligament and tendon specimens and the environment in which testing is actually conducted are important practical considerations. Because it is not always possible to test tissue immediately after death, the issue of storage by freezing is particularly crucial. Investigations as early as 1847 reported the effects of various storage methods on the mechanical properties of tendons and ligaments, and although the data are somewhat conflicting, it would appear that freezing at -20°C or below has little effect on their mechanical behavior (Wertheim 1847; Barad et al. 1982; Noyes and Grood 1976). Later, Woo et al. (1986) evaluated the effects of freezing rabbit medial collateral ligaments for up to 3 months. The only parameter significantly changed by freezing was the area of hysteresis in the first few cycles; by cycle 10 this difference became insignificant. Cyclic stress relaxation, load-deformation characteristics, and energy-absorbing capacity were unchanged. Jung et al. (2011) subsequently examined the consequences of up to eight freeze (-20°C)-thaw (22°C) cycles on human bone-patellar tendon-bone preparations and found little measurable effect on structural properties of these constructs.

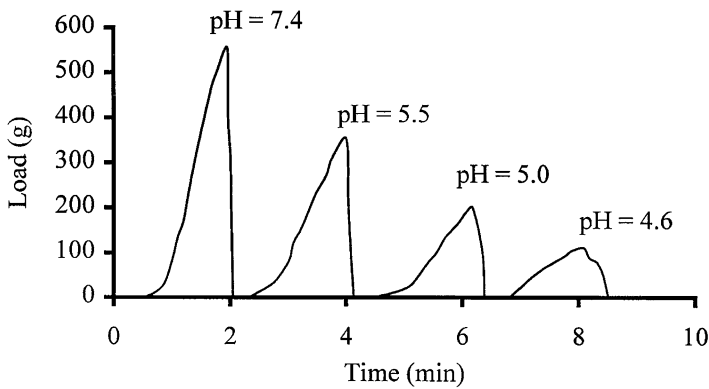


FIGURE 4.9. Rat tail tendon strength declines with diminishing pH. Previously published as Fig. 4 in Davison, P.F. The contribution of labile crosslinks to the tensile behavior of tendons. *Connective Tissue Research* 1989; 18:293–305 (Davison 1989).

Carefully controlled studies have shown that storage of specimens at room temperature has time-dependent effects that appear to be amplified in small specimens. Davison (1989) noted that the breaking strength of rat tail tendon decreased with time. Because this could be prevented by immediately washing the harvested rat tail tendon with enzyme inhibitors and keeping them on ice, Davison concluded that the collagen molecules were being subjected to gradual enzymatic degradation. An acidic environment (pH below 6) can also induce changes in the collagen molecule and lead to decreased material strength. Figure 4.9 shows the breaking strengths of rat tail tendon after being soaked in buffers of varying pH for 30 min at 0 °C.

As we have learned, biologic tissues are viscoelastic with mechanical behavior dependent on fluid pressures and flow. Not surprisingly, numerous studies have shown the mechanical behavior of these tissues to vary as a function of hydration (Chimich et al. 1992; Haut and Haut 1997; Thornton et al. 2001; Hoffman et al. 2005). It is therefore important to keep test specimens immersed, or at least saturated, in physiologic testing fluids throughout experimentation. The fluids most frequently used are physiologic saline, phosphate buffered saline and Ringer's solution. The former is a 0.9 % (g of solute/ml of solvent) solution of sodium chloride. This ionic concentration of salt is termed isotonic because its osmotic pressure is the same as that within the intracellular fluid of most cells. Cells remain intact and normally shaped when bathed in this fluid. Cells bathed in hypotonic solutions of lower ionic concentration take on water, swell, and eventually burst, releasing hydrolases and other potentially harmful enzymes into the extracellular matrix. Ringer's is a more elaborate isotonic electrolyte solution that closely matches the electrolyte concentration of the blood.

It contains small amounts of potassium and calcium in addition to sodium and chloride.

Because the viscoelastic behavior of the tendon and ligament is temperature dependent, good experimental technique dictates careful control of temperature as well as hydration during laboratory testing. Body temperature best simulates the in-vivo situation. Substantial heat can be generated in cyclically loaded tendon. In the laboratory situation, the amount of energy stored as strain energy or released to the environment as heat is dependent on the heat conduction characteristics and temperature of the test system. It is interesting to note that Wilson and Goodship (1994) measured temperatures in the superficial flexor tendons of racehorses greater than 44 °C and postulated heat-induced cell death as a possible cause of tendon degeneration, a frequent problem in the equine athlete.

It should be clear from the foregoing discussion that testing conditions are important and should be closely examined when reviewing previous work. Discrepancies between experimental results can often be explained by the different conditions under which the materials were evaluated. Obviously, conditions closely matched to the biologic environment are best able to reproduce the normal behavior of these tissues within the living body.

4.4.2 *Measuring Cross-Sectional Area*

Derivation of stress-strain behavior from load-elongation data requires knowledge of specimen cross-sectional area (CSA) and resting length (L_o). In testing rigid materials these parameters are easily determined and controlled, but the viscoelastic nature of tendon and ligament makes accurate measurement of CSA and L_o difficult. A variety of techniques have been developed to establish the cross-sectional area of ligaments or tendons, including measurement of specimen weight or volume per unit length, physical measurement using a micrometer, shadow amplitude contour reconstruction, laser reflectance, ultrasonographic measurement, and computer digitization of cross sections, photomicrographs, or inkblots. Significant systematic differences between these methodologies have been demonstrated (Ellis 1969; Woo et al. 1990; Moon et al. 2006). These systematic differences are often responsible for the considerable variation in absolute tissue stress and modulus values reported across studies.

If the density of the tissue is known, one can estimate the specimen's mean CSA by dividing its weight by its length and density. Alternatively, the specimen's volume can be determined by water displacement and divided by length to obtain mean cross-sectional area. These gravitational methods can be very reproducible but are generally useful only with long specimens of uniform dimensions. If a circular or elliptical cross section can be assumed, an alternative approach is to simply measure its diameter(s) with calipers at the site of interest and calculate the area. As might be expected, this technique

produces more variable results. A more reproducible approach deforms the cross-sectional shape of the sample into a rectangle by pushing the segment of interest into a slot of known width and closing the slot with a plunger conforming to the dimensions of the slot. A standardized force is applied through the plunger and over time the specimen conforms to the rectangular geometry of the slot, after which height is measured and the area calculated as height $\times$ slot width. This method of deriving the cross-sectional area of a specimen yields reproducible values but may introduce substantial systematic error. Clearly, the result will depend on the force used to deform the ligament or tendon into a rectangular cross-section, the tissue's normal architecture, and the time between force application and height measurement. If these variables are carefully controlled, the method is convenient and can produce excellent results.

Shadowing amplitude contour reconstruction, a fairly involved technique that requires a collimated light source, a detector, and a mechanical apparatus to rotate the sample, may give a more precise measure of specimen CSA; however, it is less reproducible than other techniques, and the apparatus is expensive to fabricate. Laser micrometer systems (Woo et al. 1990) and newer laser reflectance systems (Moon et al. 2006; Pokhai et al. 2009) offer the advantage of nondestructive measurement on intact specimens, and can also be used to measure CSA while the structure is loaded. Ultrasonography should be included in the list of reliable cross-sectional area measurement techniques and can be used for both in-vitro and in-vivo measurement (Gillis et al. 1995; Gellhorn and Carlson 2013). And finally, computer digitization of histologic cross sections taken from the test specimen after testing is accurate and reproducible but labor intensive and incompatible with destructive testing.

4.4.3 *Gripping Slippery Viscoelastic Tissues*

The mechanical testing of tendons or ligaments requires that they be grasped at the ends for mounting in the testing apparatus; this is particularly difficult because these tissues are so slippery. In addition, the viscoelastic nature of these tissues causes the specimen to deform over time and loosen within clamps, leading to slippage at the interface when loads are applied. Adhesive compounds, cyanoacrylates, have been found to work well for small slips of rat tendon but do not work for larger samples. A variety of methods have been employed for tensile testing larger tendons and ligaments. These include: looping the specimen around a post with or without suture augmentation, use of grips with roughened or serrated surfaces, use of bone–muscle–tendon–bone preparations, and freezing the tendon into the clamps. We have found the latter two methods work best.

Bone–muscle–tendon–bone (or bone–ligament–bone) preparations are an excellent choice if sub-failure characteristics are to be determined. The bone

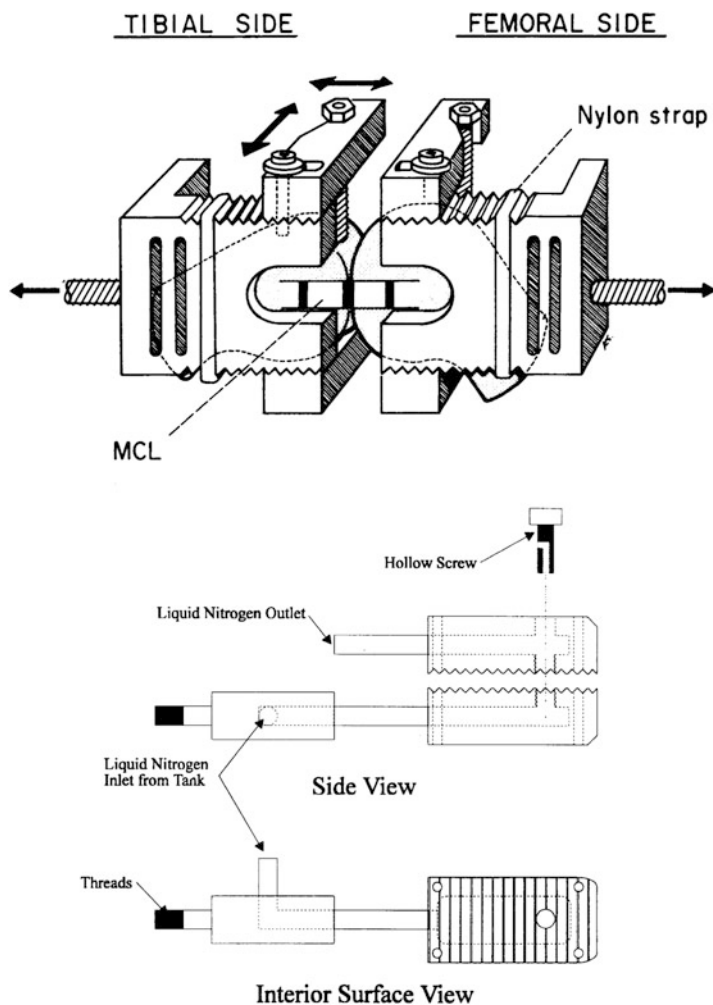


FIGURE 4.10. *Top*: Grips for holding tibia and femur to test the medial collateral ligament (MCL) (Woo et al. 1983). *Bottom*: Apparatus for gripping tendon or ligament tissue by freezing it within the clamp (Sharkey et al. 1995).

affords a reliable grip site that can accommodate high loads. Much has been learned about ligamentous behavior using the rabbit medial collateral ligament (MCL) and methods developed by Woo et al. (1983); see upper panel of Fig. 4.10. However, the method is obviously limited to bone-ligament-bone preparations and failures often occur near the insertions rather than in the mid-substance of ligament. When testing muscle-tendon-bone preparations, the muscle fibers will fail at tendon strains greater than 4 % or 5 %, which precludes complete recordings of the tendon stress-strain relationship.

In either case, appropriate orientation of the bony ends of the specimen is critical to achieve a physiologic application of stress.

Riemersma and Schamhardt (1982) were the first to use cryogenic clamps to grip these tissues during testing. In our modification of this method (Sharkey et al. 1995) (lower panel of Fig. 4.10), liquid nitrogen flows through hollow chambers in the clamp, causing the specimen to freeze thereby forming a rigid interlock with the serrations of the clamp. Because the mechanical properties of tendon and ligament are sensitive to temperature, it is important to have a long specimen so that the test region is well away from the clamps. We have been able to apply supra-physiological loads to a variety of human tendons using this approach. Other investigators have employed other versions of this technique with good success (Matthews et al. 1996; Bowser et al. 2011).

4.4.4 *Determining Resting Length*

When you read in a research paper that a tendon or ligament was stretched 5 %, it is important to also know how the unstretched length was designated. Extracted collagenous tissues such as ligament and tendon do not have an obvious gauge length, i.e. resting length prior to loading that is subsequently used to determine strain. The reason is that their length depends not only on current loads but also on past loading history. Several conventions are used to handle this problem. One is to apply a small load to the structure for a fixed period of time and to arbitrarily call the resulting length the gauge length. An alternative method is to extrapolate from the linear portion of the load-deformation curve to the point at which the load goes to zero and refer to the length intercept at that point as the gauge length.

A closely related problem is the determination of tendon or ligament lengths in vivo and in situ. For both ligaments and tendons, this depends on body position, joint angles, and muscle activity. For example, strains in the ulnar collateral ligament of the elbow joint are diminished when the joint is near full extension if the flexor muscles are active. In tendons, the relationships between body position, muscle activity, tendon compliance, and the architecture of the muscle–tendon construct are extremely complex. This is particularly true for tendons that traverse two or more joints. In addition, because of the viscoelastic features of these structures, cyclic loading alters resting length. Finally, we must keep in mind that there is considerable individual variability in the innate tightness of ligaments; some individuals have very lax ligaments and hypermobile joints, whereas others have very tight ligaments. Women generally have more compliant tendons and ligaments than men, and this difference can be further increased during pregnancy. The viscoelastic properties of the cervix and other tissues in the birth canal are greatly altered by hormonal factors when mammals give birth (see pp. 97–99 in Alexander 1968), and this effect “spills over” into other connective tissues as well.

4.4.5 *Measuring Specimen Elongation*

Methodologies to measure sample elongation can be grouped into two broad categories: measuring the displacement of the testing machine crosshead and measuring changes in a defined gauge length on the test specimen itself. The latter can be done mechanically or by using optical, radiographic or other imaging techniques.

Using crosshead displacement as a measure of sample elongation is the least preferred. Such experiments usually yield erroneously high strain values because of extraneous displacement (i.e., slipping) at the grips. Bone–ligament–bone preparations can reduce but not always eliminate this testing artifact. Because the error is systematic, methods that use machine displacement may be adequate for comparative experiments, but the error may increase non-linearly with strain.

Measurement of elongation within the central region of the specimen, far from the grip sites, is much more likely to reflect the actual material properties of the tendon or ligament. Mechanical techniques for local measurement of elongation use *extensometers* supplied by the testing apparatus manufacturer or more generic linear displacement transducers. In either case the instrument is coupled to the sample directly, which can present problems. Many tendons and ligaments have complex fiber geometry or a helical fiber arrangement; pins inserted through the body of these structures for instrument attachment often rotate and produce out-of-plane motions that hamper axial elongation measurement. Using suture for attachment can reduce this problem but may produce results that reflect the behavior of the collagen fibers in the direct vicinity of the suture and not the tendon or ligament as a whole.

Preferred methods for elongation measurement employ imaging techniques. Cinematography or sequential still frame photographs were used in the past (Holmes et al. 1991), but digitized imaging is now the more common approach. An early schematic of this technique as presented by Woo et al. (1983) is provided in Fig. 4.11. An alternative method used for in-situ measurement employs radiography and radiopaque spheres placed within the substance of the tendon or ligament (Bay et al. 1997). Newer methods employ speckle tracking and digital image correlation to measure local strain and strain fields (Davis and De Vita 2012; Andarawis-Puri et al. 2010). One shortcoming of optical techniques is that they only measure surface strain; internal behavior may be different. Another is that stain or dye markings tend to disperse under load, which makes distance measurements difficult. Recorded behavior may also be influenced by surrounding fascia. Newer texture correlation techniques make it possible to avoid markers altogether and measure strain by tracking the displacements of natural features within the structure (Bey et al. 2002; Doebling et al. 2009). Texture correlation can also be applied to ultrasound images, which allows for in-vivo strain measurements at different depths within structures of interest (Kim et al. 2011; van Doesburg et al. 2012).

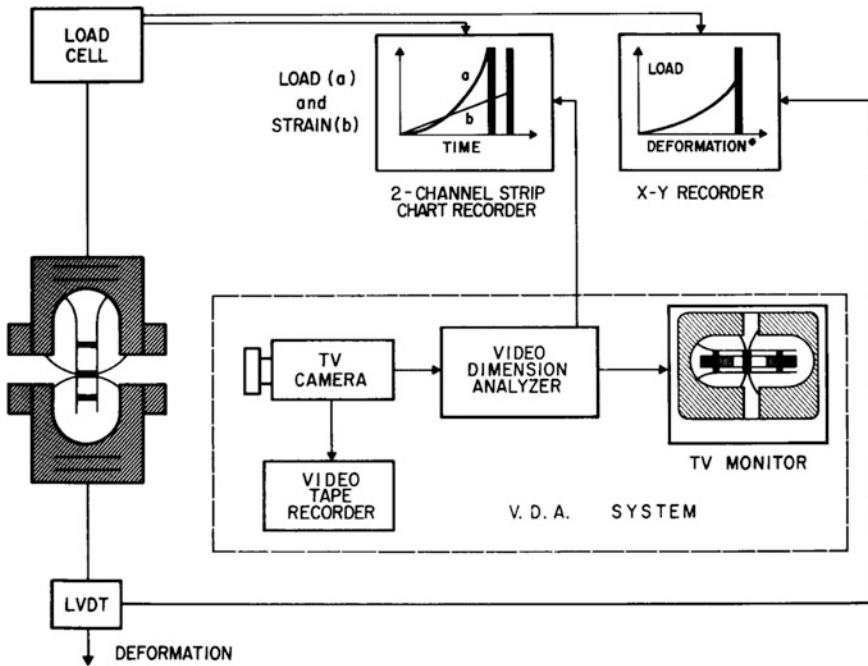


FIGURE 4.11. Video dimension analyzer (VDA) system used by Woo et al. (1983) to videographically record tendon deformations and plot load-deformation curves.

4.4.5.1 Architectural Implications

Although the collagen fibers in tendons typically are parallel, those in ligaments often are not. The pathways from origin to insertion of particular fibers are such that at any given angle of joint rotation certain fibers are stressed more than others. In testing a bone–ligament–bone preparation, orientation of the bones or the bone plugs to which the ligament is attached becomes critical. If only one particular joint angle is simulated, the characteristic behavior of the ligament at other angles of the joint will not be seen. Moreover, introduction of any extraneous twist or malalignment of the tissue may alter the measured mechanical properties. For testing to failure of either ligament or tendon, failure at the bony insertion always raises the concern that slight malpositioning of the bony block in the apparatus has resulted in nonphysiologic stresses at the bone–tendon or bone–ligament junction.

Because of ligament nonuniformity in relation to joint angles, elongation in loaded ligaments is best in two or even three dimensions using speckle techniques or image correlation so that local variations in strain can be appreciated.

4.4.5.2 In-Vivo Approaches

As indicated above, it is now possible to determine the displacements and strains present within the collagenous soft-tissues of live human subjects performing various activities. Ultrasound, the safest and most common technique, has been successfully used to track tendon displacements (van Doesburg et al. 2012) and measure strain at various depths within tissues (Kim et al. 2011). These studies are indeed informative but without simultaneous force measurement structural and material behaviors cannot be determined. Maganaris overcame this limitation by measuring strain from sonographic images of the tendons (and aponeuroses) of the tibialis posterior and gastrocnemius muscles taken while maximally stimulating the respective muscle and simultaneously measuring ankle joint moments with a dynamometer. Force was then indirectly derived by dividing the measured joint moments through by the tendon moment arms measured on magnetic resonance images. Sonography was used to measure tendon cross-sectional area. Force-displacement and stress-strain relationships were thus derived (Maganaris 2002; Maganaris and Paul 2002). The in-vivo results calculated by Maganaris and colleagues are in close agreement with data obtained in bench top experiments, with functional strains for tendon in the 3–5 % range, Youngs modulus values in the neighborhood of 0.9–1.4 GPa, and hysteresis values of 14–22 %.

Several different in-vivo transducer designs have been employed to measure the forces transmitted by the skeletal soft-tissues during normal activities. All designs work on a similar principal in that they deform as axial load is applied to the tendon or ligament to which they are attached. Buckle transducers, implantable force transducers and fiberoptic cables have all been used to measure tendon and ligament forces in humans. Of course all of these devices require surgical implantation with attendant risk to subjects. And once implanted the transducers require calibration, which is usually achieved indirectly by measuring joint moments and calibrating the transducer to the theoretic forces derived by dividing the moments obtained by the moment arm through which the tendon works. Once calibrated, forces within the structure during normal activities can be measured.

Forces in the anterior cruciate ligament (Roberts et al. 1994; Howe et al. 1990) and patellar tendon (Finni et al. 2000) have been successfully measured within implanted transducers, but functional forces in the Achilles tendon have been the primary focus of in-vivo work because of its important functional role in locomotion and relatively easy accessibility (Gregor et al. 1987; Komi et al. 1992, 1996; Fukashiro et al. 1995). Although not without technical challenges, the fiberoptic transducers have been used in several such studies with good success; they are minimally invasive, easy to place, and inexpensive. Finni and colleagues (1998) used fiberoptics to study eight normal subjects and reported peak Achilles tendon forces during walking of 750–2360 N.

4.5 Functions at Junctions

Suppose that you were given the daunting task of designing the interfaces between muscles and tendons and between tendons and bones. How would you design an interface between two soft tissues of very different tensile stiffnesses so that it could resist forces of hundreds of Newtons? Or one between flexible and rigid tissues where the angle of pull would vary considerably? We now examine the solutions found in the skeleton.

4.5.1 *The Myotendinous Junction*

The transition from muscle to tendon occurs at the *aponeurosis*, a flat thin sheet of collagenous material that gradually narrows and thickens to form tendon. Individual muscle fibers attach to these tendinous extensions both proximally and distally. At each junction the force developed by the intracellular contractile apparatus of the fiber is transferred to the extracellular matrix of the tendon. The junction site is characterized by extensive membrane folding, which increases the surface area available for force transmission and thus decreases the junctional stress. The inter-digitation also enables more effective load transmission because the interface is subjected primarily to shear stress rather than tensile stress. In fact, the angles at the junctions between muscle fibers and tendon collagen are usually close to zero, which greatly enhances junctional strength. Increases in the junctional angle may be associated with disuse atrophy and junction failure (Tidball 1983).

Binding of the muscle fiber to the tendon collagen apparently occurs at discrete sites along the interface (Trotter et al. 1981; Tidball and Daniel 1986). Subsarcolemmal “densities” have been identified within skeletal muscle cells at these sites. The cytoskeletal proteins *alpha-actinin*, *vinculin*, and *talin* and the adhesive proteins *integrin* and *vitronectin* have all been implicated in the load transmission process, but our understanding of myotendinous junctions is far from complete.

Several studies of musculotendinous preparations have shown that muscle strain injuries usually occur at or near the junction of muscle and its tendon of origin. Lieber and associates (1991), using frog semitendinosus muscle–tendon–bone preparations, demonstrated that during passive stretch the aponeurosis was significantly more compliant than both the tendon and the bone–tendon junction (Fig. 4.12, Left) Maganaris (2002) measured very similar behavior in the tibialis anterior aponeurosis of live human subjects (Fig. 4.12, Right).

In another study using frog semitendinosus units, Tidball et al. (1993) reported large increases in failure load (30 %) and energy absorption (112 %) in stimulated (actively contracting) preparations as compared to nonstimulated controls. If the muscle–tendon unit is strained to failure with the muscle stimulated, most fibers fail at the myotendinous junction, whereas

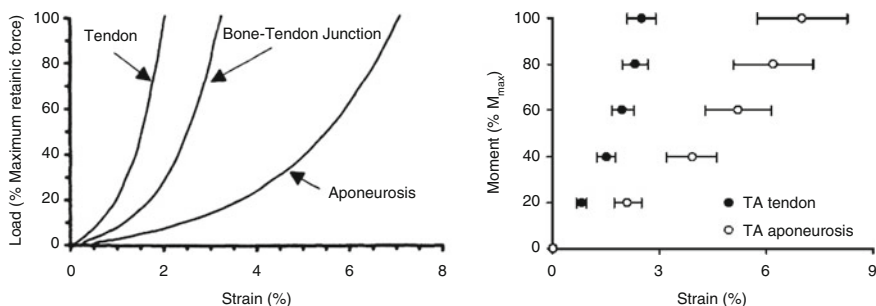


FIGURE 4.12. *Left*: Variations in compliance of the three junctional components of frog semitendinosus muscle-tendon-bone preparations. Previously published as Fig. 3D in Lieber, R.L., Leonard, M.E., Brown, C.G., and Trestik, C.L. Frog semitendinosus tendon load-strain and stress-strain properties during passive loading. American Journal of Physiology 1991; 261:C86–C92 (Lieber et al. 1991). *Right*: Compliance of the tibialis anterior (TA) tendon and aponeurosis measured in live human subjects (Maganaris and Paul 2002).

failure in nonstimulated muscle–tendon preparations occurs within the inactive fibers a short distance from the junction.

Most studies of aponeurosis mechanics, whether under passive, active, in-vivo or in-vitro conditions, have only characterized strains in the longitudinal direction, i.e. along the trajectory of its emanating tendon. But it has been shown that the aponeurosis also expands in the transverse directions, which makes sense given that muscle fibers surround and attach on opposite sides of these tendinous sheets and muscles bulge upon contraction (Scott and Loeb 1995; Donkelaar et al. 1999; Maganaris et al. 2001). Azizi and Roberts (2009) used high-speed fluoroscopy to measure biaxial strains within the aponeuroses of actively contracting gastrocnemius muscles extracted from turkeys. Transverse strain within the aponeurosis was found to be 4X greater than longitudinal strains. More important in terms of functional significance was the fact that longitudinal stiffness varied as a function of transverse strain, thereby providing a mechanism to modulate overall stiffness of the muscle–tendon unit, which in turn modulates energy storage and recovery capacity, factors we briefly explore in forthcoming sections.

4.5.2 Ligament and Tendon Insertions to Bone

Tendons and ligaments attach to bone at *insertions* or *entheses*. Insertions are generally classified as *direct* or *indirect*, but they are anatomically and functionally diverse, constituting a continuum of morphology with each site uniquely suited to its location and function. Direct insertions, also referred to as *fibrocartilagenous entheses* (Benjamin et al. 2006), often meet the adjoining bone at right angles, and have four microscopic transitional

zones of varying thicknesses. Zone 1 is represented by the normal collagen and fibroblasts of tendon or ligament; zone 2 is made of fibrocartilage and contains cells with morphology similar to chondrocytes; zone 3 is mineralized fibrocartilage; and zone 4 is bone. Between zones 2 and 3 is a mineralization tidemark (Fig. 4.13). The transition from tendon or ligament to bone is 300–700 μm across.

Collagen fibers at indirect or *fibrous* insertion sites approach the bone at an acute angle and do not have a fibrocartilaginous zone of transition. However, there is a mineralization tidemark or demarcation between the tendon or ligament and the bone. At indirect insertions, more of the fibers blend with the periosteum, which is firmly attached to the underlying bone. Fibrous insertions into thick cortical bone often display *Sharpey's fibers*—collagen fibers originating in the periosteum and extending into peripheral lamellar bone. These fibers eventually become buried in bone matrix during new subperiosteal bone formation and form a radial collagen network within the underlying bone (Fig. 4.14). Irrespective of classification, tendons and ligaments tend to broaden and flare out at their insertions, often blending with other tendons, ligaments and collagenous structures, which serves to distribute load, reduce stress and increase overall capacity. It is also worth noting that tendon insertions are non-uniform and a single insertion (e.g. Achilles) can in fact be purely fibrous in one region and fibrocartilaginous in another. As will be explored in the following section, regional differences in tissue structure and composition are strongly indicative of regional differences in the local stresses.

Failures near insertion sites due to mechanical overload are common clinical injuries. More susceptible sites of injuries include the combined *entheses* of the extensor carpi radialis brevis muscle and the lateral collateral ligament of the elbow (producing a condition known as tennis elbow), the anterior talofibular ligament (the most common site of ankle sprain) and the Achilles tendon. Stouffer et al. (1985), in a study using human patellar tendon-bone units, concluded that bone-tendon strains were much larger than those measured within the substance of the tendon. Butler and colleagues (1990) also reported high tissue strains near the insertion sites of human patellar tendon and anterior and posterior cruciate ligaments. Damage at insertion sites usually occurs in the underlying bone (*avulsion fractures*) or in the tendon tissue adjacent to the junction rather than at the junction itself. Woo et al. (1986), using rabbit femur—medial collateral ligament—tibia complexes, showed that the specific failure location within these sites is age dependent. Young growing animals were much more susceptible to avulsion of the ligament from the tibia, whereas older animals were more likely to fail by mid-substance tearing of the ligament (Table 4.1). In still older animals, femur fractures often occurred before any type of observable ligamentous failure. Generally speaking, the cross-linking of collagen within connective tissues increases with age; this may explain these experimental results.

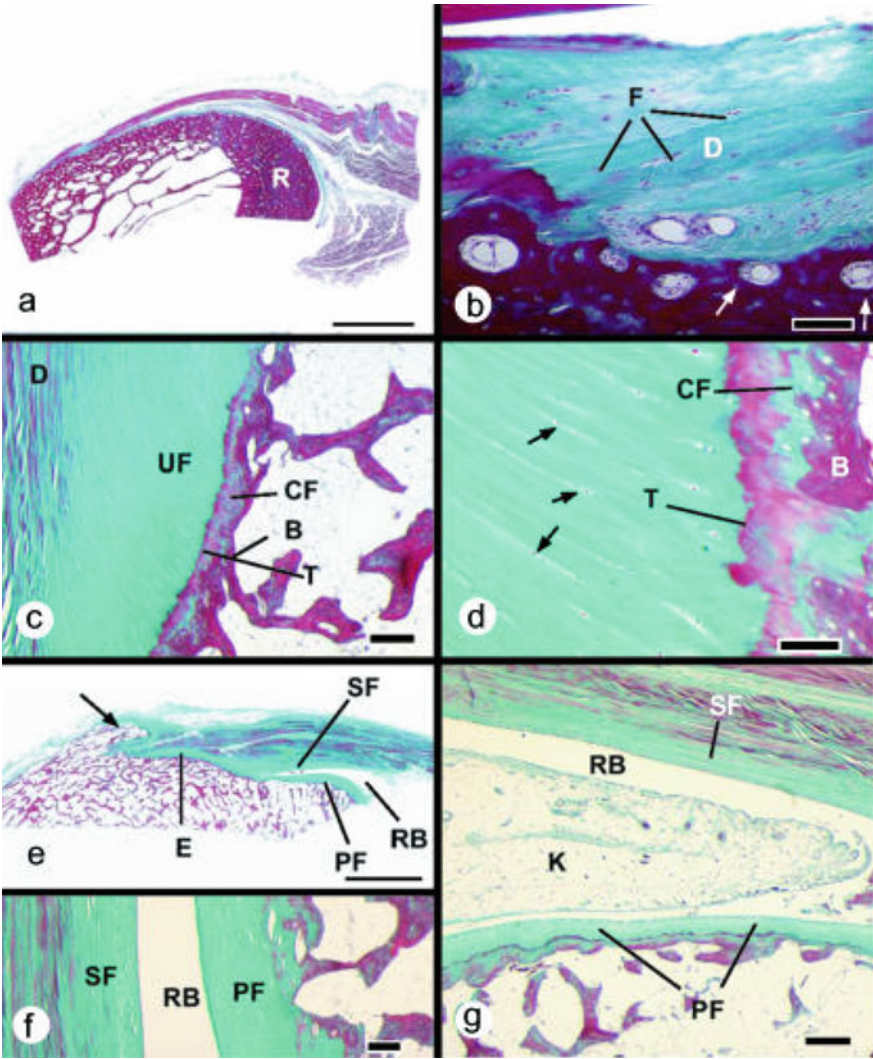


FIGURE 4.13. Low (*Left*) and high (*Right*) magnification photomicrographs of a direct (fibrocartilagenous) insertion for the Achilles tendon. The four zones of tissue include *D* dense fibrous connective tissue, *UF* uncalcified fibrocartilage, *CF* calcified fibrocartilage and *B* bone. The irregular mineralization tidemark (*T*) is shown (Benjamin et al. 2006).

It is often asserted, and some studies have confirmed, that failure occurs within the substance of the ligament at high strain rates, whereas at lower strain rates bone avulsion at the attachment site is the usual mode of failure. However, other studies have failed to confirm these findings. Several studies

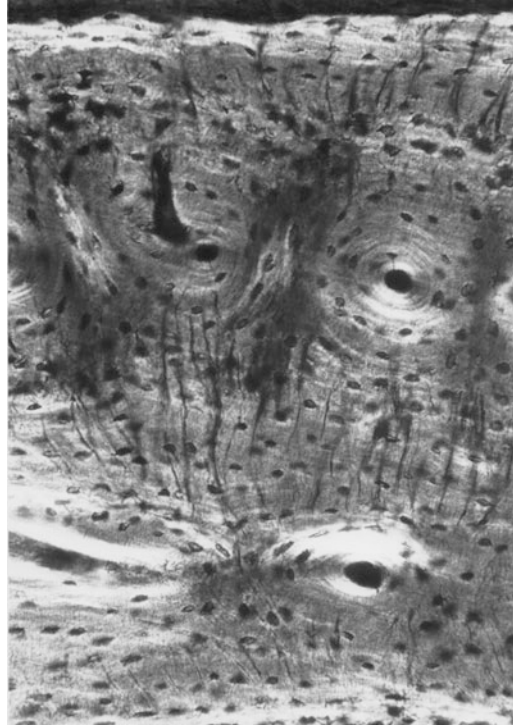


FIGURE 4.14. Sharpey's fibers appear as *dark lines* radiating inward from a tendonous attachment at *top* of this photomicrograph of a canine radius. These fibers have become embedded in the bone matrix as a result of ongoing periosteal apposition during growth. Note how they have also been interrupted by the formation of secondary osteons. Basic fuchsin-stained, undemineralized cross section; field width, approximately 0.5 mm.

TABLE 4.1. Failure modes of the medial collateral ligament as a function of age

	Age, months	No.	Failure modes, %		
			Tibial avulsion	Mid-substance	Femur fracture
Growing animals	1.5	11	100	0	0
	4–5	20	100	0	0
Adult animals	6–7	18	22	67	11
	12–15	16	0	56	44

From Woo et al. (1986)

have shown that joint angle and loading direction have profound influences on failure mechanics at insertion sites. This experimental consideration is important when evaluating the biomechanical properties of tendon and ligament in the laboratory.

4.6 Functional Adaptation and Specialization

Ligaments and tendons demonstrate tremendous variation in mechanical properties that reflect the functional demands placed upon them. For example, the ligamentum flavum connecting adjacent vertebra of the human spine functions to protect the spinal cord and keep it from protruding dorsally. This structure functions over a large range of intervertebral angles and must be highly extensible to effectively stretch during intervertebral flexion while not bulging and impinging on the spinal cord during extension. This is indeed the case; Nachemson and Evans (1968) showed that the ligament can withstand strains exceeding 60 % without rupture. Figure 4.15 presents load-deformation curves for human anterior cruciate (top) and ligamentum flavum (bottom) ligaments. The ligamentum flavum has rubbery elasticity

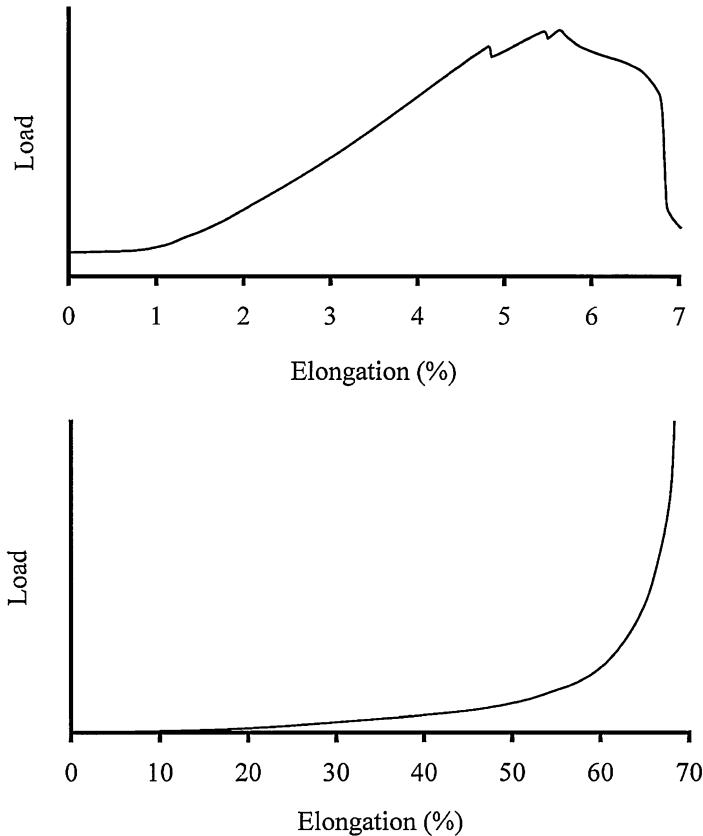


FIGURE 4.15. Load-deformation curves for human anterior cruciate (*Top*) and ligamentum flavum (*Bottom*) ligaments. Note ten-fold difference in abscissa scales.

and a very low elastic modulus (0.1 GPa) compared to most other ligamentous structures. The ligamentum nuchae is another highly elastic ligament that runs along the dorsal aspect of the spine. In the neck of a horse or cow, this ligament can store energy over prolonged periods of time during grazing and then use it to help raise the animal's head. At the other extreme, the elastic modulus of calcified turkey leg tendon approaches that of bone (10 GPa), presumably because it is under constant high stress from the continually flexed position of the turkey's lower limbs.

Specialized function is achieved by physiologic optimization of matrix composition, collagen fiber orientation, and myotendinous and insertional arrangements. Genetic coding undoubtedly controls skeletal tissue development; however, it is clear that the expression of this information is affected by local mechanical conditions. Fibroblasts have been shown to be responsive to strain *in vitro*, and many studies have demonstrated enhanced healing in tendons subjected to active or passive motion. It has been postulated that, in addition to stimulating fibroblastic activity, mechanical factors may also play a role in regulating the extracellular assembly and orientation of collagen fibrils in the matrix.

The mechanical characteristics of tendon and ligament are dictated by the relative concentrations of collagen, elastin, proteoglycan, and, occasionally, hydroxyapatite (mineral), which may not be uniform along the structure's length. By way of example, the elastin content in the ligamentum flavum described earlier approaches 70 %, compared to only 10 % in most ligaments, and even less in tendons. As shown in Fig. 4.15, the ratio of collagen to elastin is a strong determinate of mechanical behavior. Calcification also correlates strongly with increases in structural stiffness, as shown by Leonard et al. (1976) in turkey tendon.

The importance of achieving the right functional balance of compositional constituents cannot be overemphasized. Short-term energy storage and return within the Achilles during running is perhaps the most widely cited example of tendon's important mechanical contributions to efficient musculoskeletal function. But given their passive nature, it is easy to overlook other critical, albeit more subtle, contributions tendons make to efficient muscle function during everyday activities. Two examples are the tendons of flexor hallucis longus and flexor digitorum longus, both of which originate from muscles deep in the posterior shank, traverse several joints along their path and function to flex the toes and counterbalance ground reaction forces during the push-off phase gait. Experimental simulations of normal walking (Kirane et al. 2008; Hofmann et al. 2013) strongly suggest that the material properties of these two tendons are precisely tuned to transmit appropriate time-varying resistive forces to the toes while under constant highly efficient isometric contraction, thereby minimizing energy expenditure (Fig. 4.16). In this case, force development in the tendon under constant muscle length is purely a consequence of the rotations in the joints these tendons traverse, making proper force transmission to the toes entirely dependent on tendon stiffness.

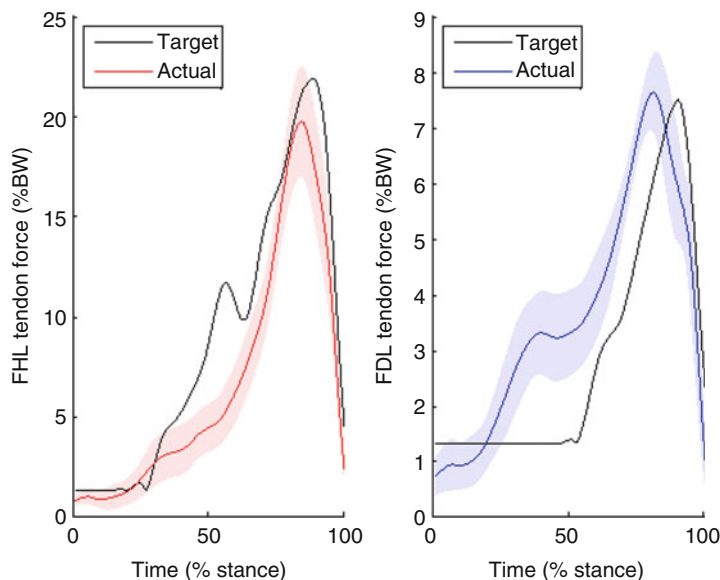


FIGURE 4.16. Data from cadaver simulations of normal gait showing time varying forces in the tendons of flexor hallucis longus (*red/left*) and flexor digitorum longus (*blue/right*) under constant isometric muscle contraction. Muscle forces estimated from electromyographic recordings of normal subjects (*black lines*) are included for comparison (*BW* body weight) (Hofmann et al. 2013).

4.6.1 Determinants of Tendon Architecture

As evidenced above, tendon architecture is determined by considerations of stiffness as well as strength. In thinking about the various factors that govern tendon architecture, it is useful to consider a simple model of tendon action: a loop consisting of a bone, a hinge joint, a second bone, a muscle attached to the first bone, and the tendon of the muscle attached to the second bone so that contraction of the muscle flexes the joint (Fig. 4.17).

Contraction of the muscle exerts a force (F_1) on the tendon that causes rotation of the joint and in turn enables the effector end of the second bone to exert a force of some kind on the environment (F_2). The relationship of the two forces, F_1 and F_2 , depends in part on the relative moment arms of the tendon insertion and the effector point (with respect to the joint center of rotation). The relationship of the two forces also depends on the compliance of the tendon as a structure, which is in turn determined by its elastic modulus, cross-sectional area, and length. If we assume for the moment that the tendon is perfectly elastic, then if it were twice as thick it would stretch only half as much at a given force, resulting in greater motion at the joint for a given level of muscle contraction. Similarly, if it were half as long,

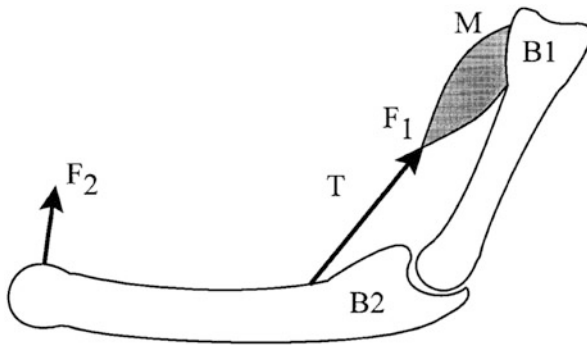


FIGURE 4.17. Sketch of a simple model for tendon mechanics. Bones B1 and B2 are articulated at a hinge joint; muscle M pulls on tendon T with to force F_1 , flexing the joint. The distal end of bone B2 produces force F_2 against an object in the environment.

a given level of muscle motion would result in greater motion at the joint. From the perspective of an organism that must adapt efficiently, the variables which can be manipulated include muscle architecture, tendon cross-sectional area, tendon length, tendon moment arm, and moment arm of the force effector. From an evolutionary perspective, these variables may change together, with different evolutionary “problems” giving rise to different adaptive solutions. For example, if you have a long, compliant tendon, you could make it shorter or thicker (if greater forcefulness were sought), or you could increase fiber length (excursion) of the muscle if precision were more important, or you could increase both muscle fiber length and tendon stiffness while decreasing the tendon insertion moment arm if speed were the critical factor. In some cases, adaptation has led to tendons that are very robust compared to the size of their muscles. An extreme example is the massive suspensory ligament of the horse, which is not, strictly speaking, a ligament, but the tendon for a muscle so tiny that it appears as a reddish fuzz of rudimentary fibers. This tendon functions principally for energy storage during gait; it is loaded primarily by gravity rather than by its muscle. At the other extreme, some of the human finger flexor muscles have extremely long and compliant tendons; these are adapted to maximize precision of finger motion at the expense of force or elastic energy storage (which is not a factor in the fingers).

4.6.2 Sites of Tendon Compression

Proteoglycan is another component of tendon and ligament that increases under special conditions. High proteoglycan content in tendon or ligament produces a fibrocartilaginous tissue. Localized areas of *fibrocartilage* are found in regions of tendon that wrap around and press against bone. Recall

that proteoglycan is a major constituent of cartilage and is the molecule most responsible for this tissue’s ability to withstand compression. Fibrocartilaginous tissues contain specialized cells that show a gradient of phenotypic expression between that of a fully differentiated chondrocyte and that of a fully differentiated fibroblast. These cells produce large amounts of proteoglycan and collagen. (The mix of different collagen types is also varied.) Cell shape is reflective of extracellular matrix composition and a “*fibrochondrocyte*” is more spindle shaped in areas of low proteoglycan (such as tendon or ligament) and more rounded in areas containing large amounts of proteoglycan (like cartilage).

An interesting mechanical theory of tissue differentiation based on localized stress and strain conditions has been proposed by Giori and coauthors (1993). This theory is related to the Carter–Wong theory of cartilage calcification discussed in Sect. 6.7. Here again, it is postulated that cells produce an extracellular matrix that varies according to the balance between distortional and hydrostatic stresses imposed on them. In this case, however, the theory addresses the balance between cartilage and tendon rather than the calcification of cartilage. A graphical representation of the theory is presented in Fig. 4.18. The horizontal axis represents the amount of compressive or

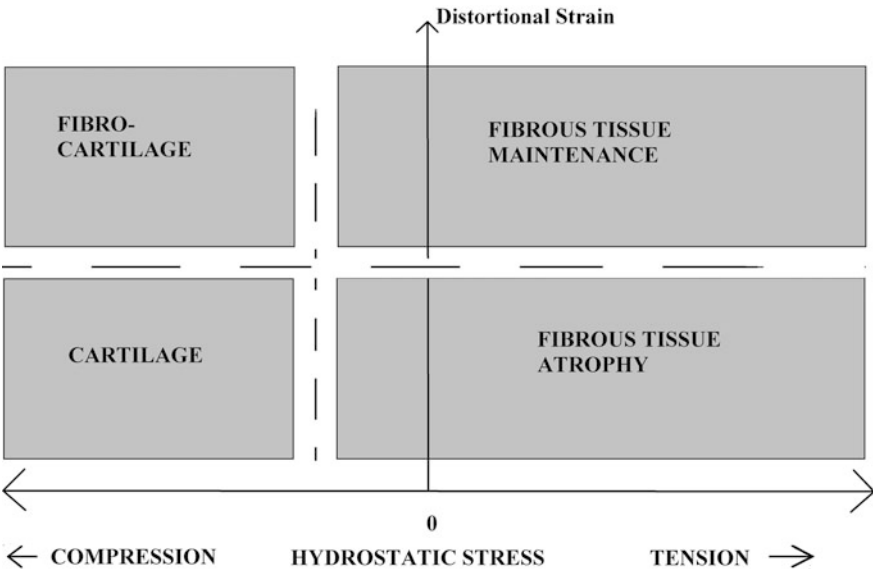


FIGURE 4.18. Hypothesis for the effects of stress and strain on tissues produced by cells in a tendon. The hydrostatic portion of the stress is plotted on the abscissa, with compression to the *left* and tension to the *right*. Distortional (or shear) strain is plotted on the ordinate. Depending on the balance between these hydrostatic and distortional influences, the cells are postulated to produce the kinds of tissue shown (Giori et al. 1993).

tensile hydrostatic stress, and the vertical axis denotes the amount of distortional strain, at a point in the tendon. It is postulated that cartilaginous tissue is produced in an environment dominated by compressive hydrostatic stress and that fibrous tissue occurs in an environment dominated by tensile hydrostatic stress. Fibrocartilage, such as that seen in tendons that pass around and are forced against a bone, is produced in an environment with high compressive hydrostatic stress combined with high distortional strain. Fibrous tissue is maintained in a situation that combines high distortional strain with tensile hydrostatic stress.

To find out which of these situations obtains in various regions of a tendon wrapped around a bone, the authors developed a two-dimensional finite element model. They assumed tendon tissue to be a two-phase, linear elastic fiber-reinforced composite consisting of longitudinally oriented collagen fibers and an extrafibrillar matrix. Their model did not iteratively vary the tissue properties according to the local stress history, as the Carter–Wong models did, but simply studied the distribution of the stresses in a tendon wrapped around and forced against a cylindrical bone. Figure 4.19 shows typical distributions of hydrostatic stress (left) and distortional strain (right) in the finite element model. The distortional strain is relatively high throughout the tendon and greatest along the surface contacting the bone. Thus, the

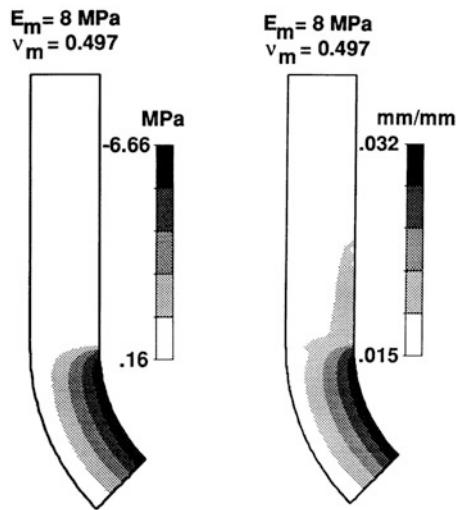


FIGURE 4.19. Hydrostatic stress (*Left*) and distortional strain (*Right*) in a finite element model of a tendon passing around a cylindrical bone. The elastic modulus and Poisson's ratio values used in the model are shown *above* and the scales of shading used to depict stress results to the *right* of each figure. The bone, which would be adjacent to the *dark region* in the tendon, is not shown, and only the upper half of the tendon appears because the model utilizes the symmetry of the situation (Giori et al. 1993).

tendon's mechanical situation is entirely in the upper half of Fig. 4.18. The hydrostatic stress is high and compressive (negative) where the tendon presses against the bone; it is slightly positive elsewhere. The region in contact with the bone is in the "fibrocartilage" zone because its hydrostatic stress is negative. The other portions of the tendon are considered to be in the "fibrous tissue maintenance" zone because their hydrostatic stresses are slightly positive. Therefore, the mechanical situation in the tendon and the distribution of tissue types correspond to the biologic effects postulated in Fig. 4.18.

Matyas et al. (1995) proposed a mechanical hypothesis for tissue differentiation at direct ligamentous insertions (recall that direct insertions contain a band of fibrocartilage). They examined histologic sections of the femoral insertions of rabbit medial collateral ligaments and quantified cell shape as an index of tissue phenotype. They assumed that rounded cells produce a more cartilaginous material and spindle-shaped cells a more fibrous material. To perform their study, the investigators used automated image processing and analysis techniques that differentiated the margins of dark-staining cell nuclei. A one-to-one correspondence of nuclear shape to cytoplasmic shape was assumed based on the authors' preliminary work. High-magnification images of histologic slides were captured with a CCD (charge-coupled device) camera and digitized. The software automatically measured the location, area, and perimeter of each nucleus and calculated a roundness coefficient for each cell as follows:

$$\text{Roundness} = (4\pi \times \text{area})/\text{perimeter}^2$$

Using this equation, a perfect circle would have a roundness coefficient of 1. Finite element techniques were used to characterize the local mechanical environments in and about the insertion site, and the results were correlated to the histologic maps of cell roundness. Corresponding gradients of cell shape and principal stress were noted in the fibrocartilaginous zones at insertion sites. Several mechanical properties correlated with tissue phenotype (cell roundness) (Table 4.2), providing further evidence that functional adaptation is controlled, at least in part, by the mechanical environment.

4.6.3 *Flexors and Extensors*

The long flexor and extensor tendons that contribute to finger motion in humans and quadrupedal gait in animals are also highly specialized structures. The flexors usually have different mechanical properties than the extensors. Shadwick (1990) showed that porcine flexor tendons were consistently stronger, stiffer, and more energy conserving than extensor tendons (Fig. 4.20). Thus, flexor tendons were more effective biologic springs than the extensors, which makes good sense biologically in terms of reducing the amount of muscular energy and the metabolic cost of locomotion. As the

TABLE 4.2. Mechanical correlations to cell shape

Mean nuclear roundness vs.	Pearson correlation
Principal stresses	
Compression	−0.68
Tension	−0.53
Derivatives of principal stress	
Hydrostatic stress	−0.67
von Mises stress	−0.31
Strain energy density	−0.41
Principal strains	
ϵ_1	−0.40
ϵ_2	+0.69
ϵ_3	−0.20
Derivatives of principal strain	
Volumetric strain	−0.67
Distortional strain	−0.31

From Matyas et al. (1995)

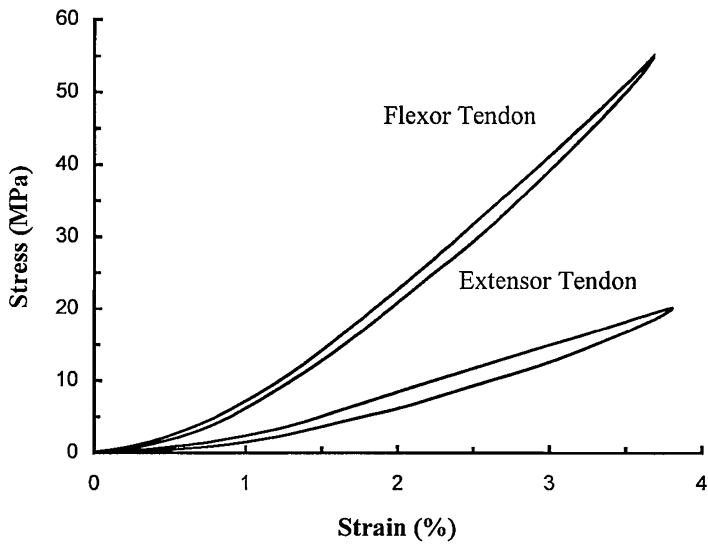


FIGURE 4.20. Comparison of load-deformation curves of flexor and extensor tendons from pigs. Previously published as Fig. 4 in Shadwick, R.E. Elastic energy storage in tendons: mechanical differences related to function and age. *Journal of Applied Physiology* 1990; 68:1033–1040 (Shadwick 1990).

body decelerates when the foot contacts the ground, kinetic and potential energy are converted to strain energy stored in the flexor tendons of the limb. Elastic recoil of the stretched tendons converts this momentarily stored energy back to kinetic and potential energy at push-off. When the tendons are more efficient at storing energy, less work is required for locomotion. Extensor tendons function differently and are not elongated during deceleration, thus efficient energy management is less crucial in these structures. Interestingly, Shadwick also showed that the two types of tendon possess identical mechanical characteristics at birth, providing further evidence that functional demand regulates structure and mechanical behavior (Shadwick 1990).

Tendon stiffness has been shown to increase with increasing load, largely because of the straightening of crimps in collagen fibers. Stiffness increases dramatically as the toe region of the load-deformation curve is surpassed (see Fig. 4.1); when crimp is removed, stiffness remains reasonably constant. Typically, relative muscle force at the point of transition is less than 25 % of maximum. This behavior is amplified in human finger flexors because they are so long. In these structures, high compliance when muscle force is low serves to damp contractile variations, producing smooth and finely controlled motion at the fingers. When higher muscle force is applied, this damping effect is diminished, but high-force tasks usually do not require fine motor control. The greater tendon stiffness during high-load tasks (such as tight gripping) provides more efficient force transfer.

The synovial sheath that surrounds the long flexors and extensors is another important functional adaptation of these tendons. It produces and retains synovial fluid that bathes the tendon and facilitates smooth excursion with minimal friction. Without the synovial membrane these tendons would be subjected to substantial drag because of their length.

4.7 Age and Mechanical Behavior

A number of animal studies have shown that the mechanical properties of tendons and ligaments change with age. During maturation, the size of the toe region of the stress-strain curve decreases and the elastic modulus and ultimate strength increase (Woo et al. 1990; Tipton et al. 1978). Evidence suggests that three processes are responsible for these changes in behavior. The first appears to be an increase in collagen content per unit area, as reflected by a steady decline in water content during maturation. The second is a gradual decrease in collagen crimp, which would account for the reduced toe region. The third is a gradual increase in collagen cross-links, which renders the molecules stiffer. Actually, these changes may be two sides of the same coin. The reduced crimping may just be a manifestation of the

altered conformation of the collagen molecules caused by their additional cross-linking, and the straightening of the collagen may promote additional cross-linking. Increasing tensile strength and stiffness plateau after maturity and apparently deteriorate during old age (Noyes and Grood 1976; Neumann et al. 1994).

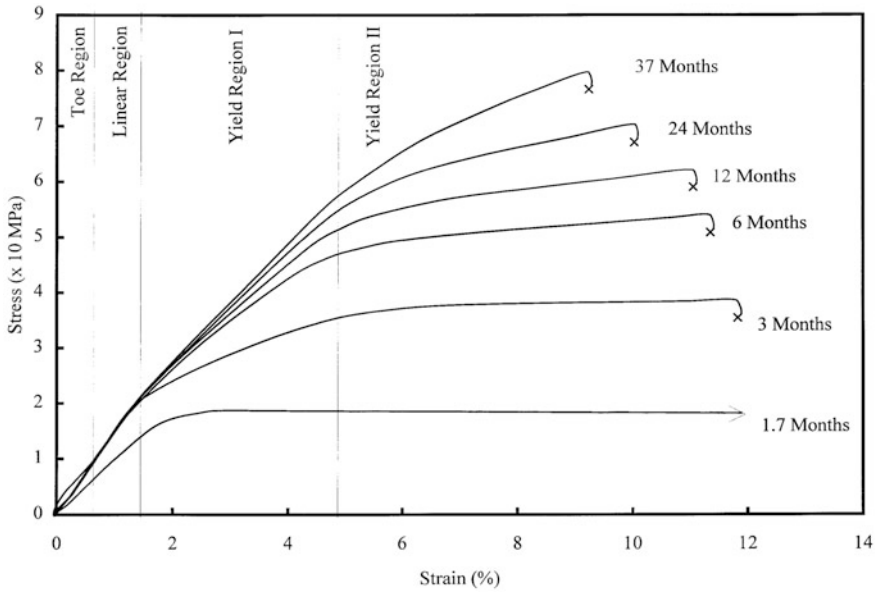


FIGURE 4.21. Effects of age on stress-strain curves for rat tail tendon (Kastelic and Baer 1980).

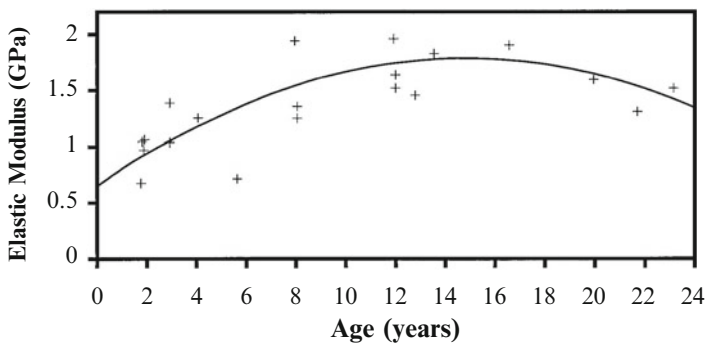


FIGURE 4.22. Effects of age on elastic modulus of equine flexor tendon. Previously published as Fig. 2 in Gillis, C., Sharkey, N.A., Stover, S.M., Pool, R.R., and Meagher, D.M. The effect of maturation and aging on the material and ultrasonographic properties of equine superficial digital flexor tendon. *American Journal of Veterinary Research* 1995; 56:1345–1350 (Gillis et al. 1995).

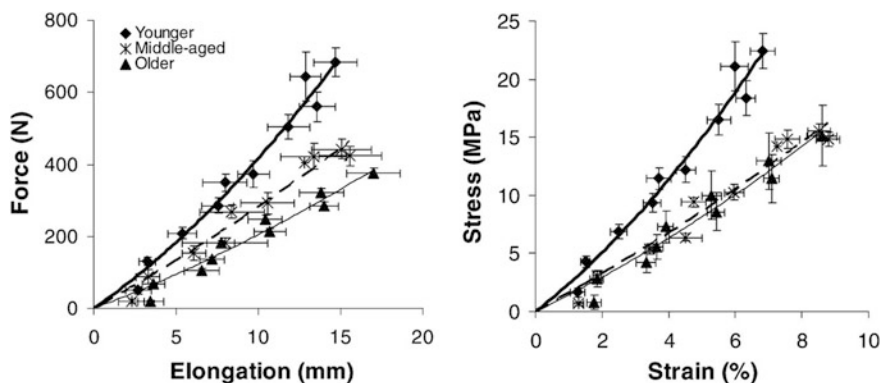


FIGURE 4.23. Force-elongation and stress-strain plots for the human gastrocnemius tendon as measured in 24 young (24 ± 1 year), 10 middle-aged (46 ± 1 year) and 36 older (68 ± 1 year) subjects (Onambele et al. 2006).

Figure 4.21 demonstrates the increasing strength and modulus of rat tail tendon during maturation (Kastelic and Baer 1980). Figure 4.22 shows how the elastic modulus of equine flexor tendon changes with age (Gillis et al. 1995). Note that, in the horse, tendon stiffness gradually increases and then begins to decline in the older animals. Rats may simply not live long enough for this to happen to their tendons.

Data from human studies also show declining stiffness and modulus of elasticity with age and disuse (Karamanidis and Arampatzis 2005; Reeves et al. 2005; Maganaris et al. 2006). Using ultrasonography, Onambele et al. (2006) measured the stiffness and modulus of the gastrocnemius tendon in three cohorts of healthy individuals tightly grouped about mean ages of 68, 46 and 24 years. The Young's modulus of tendon in the older group was found to be 28 % less than that in the younger cohort (Fig. 4.23). Similarly, Stenroth and colleagues (2012) measured a 17 % decline in stiffness and 32 % decline in Young's modulus when comparing subjects in their seventies to those in their mid-thirties. Cross-sectional area of the tendon in the older subjects was 16 % larger than that in the young, which partially compensated for the reduction in modulus. Interestingly, tendon stiffness was highly correlated with maximum voluntary contraction of the triceps surae and there was no difference in stiffness or modulus when comparing old and young subjects having similar muscle strength. These findings suggest that the Achilles tendon adapts to the mechanical performance of the triceps surae, which was significantly smaller and weaker in the older subjects. Thus it could be that sarcopenia is the prime driver of the biomechanical and compositional changes noted in tendons and ligaments as a function of aging.

4.8 Exercise and Immobilization

We have learned that skeletal tissues, including bone, tendon, and ligament, respond in predictable ways to the mechanical demands placed upon them. Specifically, it has been demonstrated that the mass and strength of bone rise and fall according to mechanical demand. To some degree, tendons and ligaments appear to behave similarly. Woo and associates (1980) have shown that exercise increases the tensile strength of extensor tendons and flexor tendon insertions but not flexor tendons themselves. Figure 4.24 compares the mechanical behavior of digital **extensor** tendons extracted from exercised (40 km/week for 1 year) and non-exercised swine. The improved performance of the exercised tendons is obvious. The lack of similar strength changes in the **flexor** tendons may be because tendons and ligaments that are regularly subjected to high loads by normal activities of daily living are already optimally designed. Perhaps the only way to increase the load-bearing capacity of these structures is to increase cross-sectional area, which, unlike bone, would be naturally self-limiting because of geometric and kinematic constraints. This reasoning may also explain why bone-tendon and bone-ligament complexes respond to increased demand by improving strength at insertion sites rather than by architectural modifications of the tendon or ligament itself (Woo et al. 1980; Tipton et al. 1967). Conversely, Laros et al. (1971) and Noyes et al. (1974) demonstrated in dogs and primates, respectively, that the mechanical strength at insertion sites

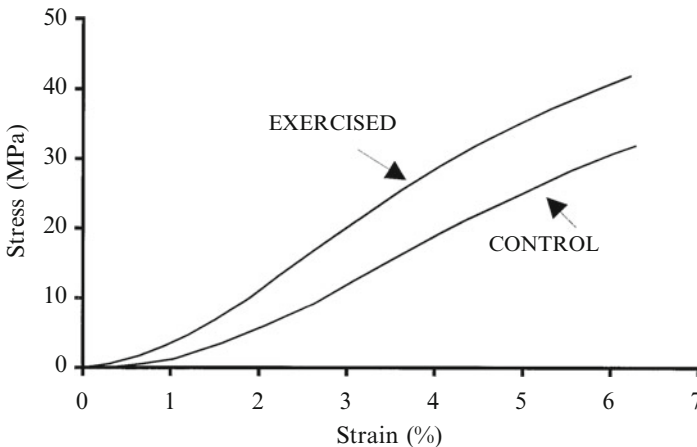


FIGURE 4.24. Extensor tendon load-deformation curves for exercised and control swine. Previously published as Fig. 6 in S. L.-Y. Woo, M. A. Ritter, D. Amiel, T. M. Sanders, M. A. Gomez, S. C. Kuei, S. R. Garfin, W. H. Akeson. *The Biomechanical and Biochemical Properties of Swine Tendons—Long Term Effects of Exercise on the Digital Extensors*. *Connective Tissue Research*, 1980, Vol. 7, No. 3: Pages 177–183 (Woo et al. 1980).

TABLE 4.3. Mechanical properties of normal, immobilized, and reconditioned rhesus monkey anterior cruciate ligaments

Parameter	Normal (n = 30)	Immobilized (n = 18)	Five-month reconditioning (n = 22)	Twelve-month reconditioning (n = 20)
Maximum load (kg)	87.33 ^a (19.40)	53.14 ^b (12.79)	68.68 ^c (18.35)	79.23 ^{a,c} (17.01)
Strain at maximum load (%)	46.66 ^a (8.63)	43.88 ^{a,b} (9.08)	38.53 ^b (7.80)	42.91 ^{a,b} (6.22)
Energy at failure (cm-kg)	34.02 ^a (10.72)	23.16 ^b (5.27)	26.53 ^{b,c} (7.19)	31.42 ^{a,c} (8.16)
Stiffness (kg/mm)	18.49 ^a (3.01)	12.76 ^b (3.29)	17.13 ^b (3.01)	18.10 ^b (2.77)

SD in parentheses. Values without any superscripts in common are significantly different ($p < 0.05$)

From Table 1 of Noyes F.R. Functional properties of knee ligaments and alterations induced by immobilization. A correlative biomechanical and histological study in primates. Clin Orthop Rel Res 1977; 123:210–242 (Noyes 1977)

quickly diminishes in response to immobilization. In the primate study, recovery was still incomplete after 5 months of reconditioning exercise (Table 4.3). Woo and co-workers (1987) also have shown inferior mechanical properties in the rabbit medial collateral ligament after immobilization. All these studies indicate that the insertion sites may be more prone to disuse atrophy, and take longer to regain their original strength, than the ligaments themselves.

Several human studies employing ultrasound and/or MRI imaging to measure tendon elongation and cross-sectional area have confirmed the anabolic effects of training on the structural stiffness and elastic modulus of the patellar tendon (Reeves et al. 2003; Kubo et al. 2005; Malliaras et al. 2013) and Achilles tendon (Kubo et al. 2007; Arampatzis et al. 2010). In most studies, changes in cross-sectional area remain small or non-existent while strain for a given load is reduced, suggesting that internal remodeling of the tissue rather than gross addition of material is responsible for the observed mechanical adaptations. These same in-vivo methodologies also have also been used to show that human tendon stiffness and modulus decline with disuse (Karamanidis and Arampatzis 2005; Maganaris et al. 2006).

4.9 Pathology and Healing

4.9.1 Ligament

Ligament failure occurs by mid-substance tear or bony avulsion. Mid-substance failure is more common in adults and avulsions are seen more frequently in children. *Sprains*, stretching the ligament beyond its

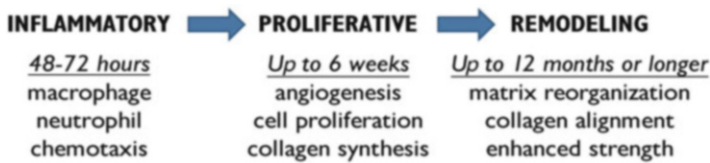


FIGURE 4.25. The normal progression of wound healing in ligament and tendon.

elastic limit, elicit an inflammatory response that will eventually repair the compromised tissue. Ligament healing has been divided into the same three phases as bone fracture healing, but requires a much longer time. Phase I (48–72 h), *acute inflammation*, is characterized by pain, swelling, redness, and warmth from vasodilation. During this phase cells are recruited and a hematoma forms. Phase II (up to 6 weeks post injury), *proliferation*, is characterized by increased numbers of fibroblasts, collagen synthesis, and scar tissue formation. Phase III (12 months or longer) involves *remodeling* of the scar (Fig. 4.25). Collagen synthesis decreases, water content drops, and the extracellular matrix becomes more organized. (Usually this means a more longitudinal arrangement of fibers.) Scar contraction coincides with increasing tensile strength. Unlike bone, the injured ligament may never regain full strength; primate studies indicate tensile strength to be 80 % of the intact ligament at 1 year (Clancy et al. 1981).

4.9.2 Tendon

Like ligament, tendon can fail in mid-substance, at or near the myotendinous junction or at its bony insertion. Avulsion is less common for tendons than ligaments but does occur in young athletes. Tendon lacerations, particularly of the finger flexors and extensors, are common injuries. Overloading of tendon can cause damage to the collagen network, as can repetitive motion and abrasion. Diseases that compromise blood flow, such as diabetes, can cause ischemia-induced degeneration, particularly in areas with little collateral blood supply. Steroid injection has also been associated with tendon degeneration and rupture. Tendon healing follows the same pathway as described earlier for ligaments, but is more problematic. The length and quality of the repair process can vary considerably as the result of major differences in blood supply from site to site. In addition, *adhesions* after surgical repair are a significant clinical problem, particularly in the digital flexors/extensors. Scar formation extends into surrounding tissues, tacks the tendon down, and prevents sliding. Rehabilitation protocols employing both active and passive motions, judiciously applied after surgery, can reduce this and improve clinical outcomes.

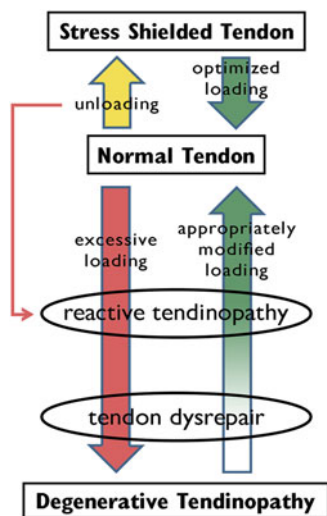


FIGURE 4.26. Schematic illustration of the dependence of the healing response in tendon to imposed load. Too much or too little load can disrupt the normal progression of healing. Loading can be modified during rehabilitation to optimize healing (*right green arrows*) but as tendinopathy progresses due to overload or underload the chances of reversal toward normal recovery lessen, eventually resulting in permanent disability (Cook and Purdam 2009).

Injuries to mechanically regulated tissues can present difficult clinical problems. Too little loading after surgical repair of a tendon or ligament leads to inadequate tissue remodeling, low tensile strength at the repair site, and diminished overall structural strength; too much loading will pull the repair apart or elongate the structure resulting in dysfunction. Absence of motion can permit runaway scar formation, leading to adhesions and a poor functional outcome (Fig. 4.26).

Active motion whereby loading is gradually increased to pre-injury levels is perhaps the best postoperative treatment after ligament reconstruction or repair of partial tendon laceration. Vailas and associates (1981) demonstrated increased collagen synthesis and strength in rat medial collateral ligaments subjected to progressive exercise during healing as compared to non-exercised controls. In cases of complete tendon laceration, many studies have advocated repair techniques that place more strands of suture across the laceration site. These newer techniques increase the postoperative tensile strength of the repair so that active motion rehabilitation can begin soon after surgery. Nevertheless, overload of the repair site is a significant concern, making *passive motion* of the involved joint a popular alternative that subjects the healing tendon to modest stress and lowers the risks of gap formation at the repair site. Freehan and Beauchene (1990), Takai et al. (1991) and

others have shown that passive motion improves healing and increases wound strength compared to immobilized controls. Passive motion has also proven useful in healing cartilagenous injuries in experimental animal models (Salter et al. 1980), suggesting that the treatment might be useful for joint injuries affecting ligament and cartilage as well as tendon. Although its clinical utility in humans remains controversial, a systematic review of 19 studies examining the ability of continuous passive motion to promote and accelerate healing in the rabbit model pioneered by Salter yielded convincing evidence of improved outcomes (Knapik et al. 2013). Regardless of tissue type, most animal studies and some human studies indicate improved tissue generation and organization with reduced inflammation and adhesion formation.

Flexor tendon adhesions and joint contractures continue to be problematic despite advances in surgical techniques and postoperative rehabilitation. In many instances *tenolysis*, surgically freeing the tendon from surrounding tissue, is required to achieve acceptable function. Patients not committed to extensive postoperative rehabilitation may never regain adequate function.

4.9.3 Surgical Repair

Surgical repair of damaged tendons and ligaments is often a difficult task, in large part because of the problems involved in replacing tissues with very specialized properties. Learning something about these surgical procedures helps to clarify the practical significance of the material presented in this chapter.

4.9.4 Anterior Cruciate Ligament

The anterior cruciate ligament (ACL) of the knee is a site of frequent injury, particularly in athletes. The ACL is an intra-articular ligament bathed in synovial fluid; it originates on the posteromedial aspect of the lateral femoral condyle and inserts centrally into the joint surface of the anterior tibia. The ligament restricts anterior translation of the tibia. Figure 4.27 shows that the ACL restrains translation of the tibia relative to the femoral condyles. The posterior horns of the menisci offer a second line of defense in case of ACL rupture, as does active contraction of the hamstrings muscle. Many patients with ACL rupture do well with conservative treatment, which usually consists of hamstring strengthening exercises and bracing, but 30–50 % of all patients choose reconstruction; more than 100,000 such reconstructions are performed annually. No method of reconstruction is perfect, which in part accounts for the extremely large body of literature focused on the ACL and its repair. In the late 1970s and 1980s synthetic grafts were a popular

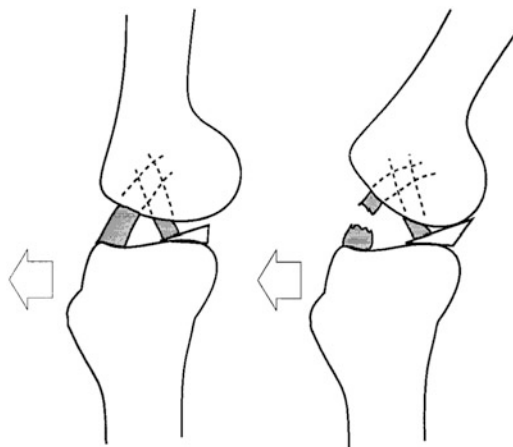


FIGURE 4.27. The anterior cruciate ligament (ACL) of the human knee restrains anterior translation of the tibia (*at bottom*) relative to the distal femur (*above*).

alternative to autogenous (the patient's own tissue) graft materials, but it was soon realized that synthetic materials rapidly wear out and degrade in the environment of the joint. Fatigue damage was a common finding, emphasizing the need for continual maintenance and renewal of collagen fibers within the normal ligament.

Biologic grafts have proven to be much more effective because they are able to be repopulated with host cells (i.e., the patient's own cells) and subsequently remodeled into reasonable facsimiles of the original structure. The most popular and successful techniques currently use *autografts*, the patient's own tissue, for ACL replacement. Table 4.4 compares the strengths of many of the autograft tissues that have been used for ACL reconstruction with that of the normal ACL. The central one-third of the inferior patellar tendon is a popular graft because it can be harvested as a bone–tendon–bone preparation, which can be rigidly attached at the repair site. Table 4.4 shows that the patellar tendon is also a very strong choice.

Allografts, tissue taken from another human, is also used for reconstruction. Allografts eliminate the pain and risk associate with harvesting an autograft but also carry the extra risk of disease transmission. There is also evidence suggesting allografts have a higher failure rate than autografts. Nonetheless, good results have been reported following allograft reconstruction using patellar tendon or Achilles tendon. It must be borne in mind, however, that any graft will try to alter its mechanical properties to suit its new location and function. The question is whether it will be able to do so (in a joint initially inflamed by surgery) while it bears the loads applied by the recovering patient.

TABLE 4.4. Failure loads of tissues used for anterior cruciate ligament (ACL) reconstruction

Tissue	Maximum load (N)	Percent of normal ACL
Normal anterior cruciate ligament ($n = 6$)	1725	100
Bone—patellar tendon—bone		
Central third ($n = 7$)	2900	168
Medial third ($n = 7$)	2734	159
Semitendinosis ($n = 11$)	1216	70
Gracilis ($n = 17$)	838	49
Distal iliotibial tract (18-mm width; $n = 10$)	769	44
Fascia lata (16-mm width; $n = 18$)	628	36
Retinaculum—patellar tendon		
Medial ($n = 7$)	371	21
Central ($n = 6$)	266	15
Lateral ($n = 7$)	249	14

From Noyes et al. (1984)

4.9.5 Supraspinatus Tendon (Rotator Cuff)

The *rotator cuff* of the shoulder joint is another site of injury frequently treated surgically. The rotator cuff comprises the supraspinatus, infraspinatus, subscapularis, and teres minor muscles. All these muscles originate on the scapula and insert on the proximal humerus close to the glenohumeral articulation. They assist in arm elevation and rotation, and also stabilize the humeral head against the upward shear forces generated by deltoid contraction. The supraspinatus is the tendon most frequently affected, and tears of the rotator cuff are usually synonymous with tears or rupture of the supraspinatus tendon. It is the most superiorly oriented tendon of the rotator, passing within a narrow space bordered superiorly by the acromion and coracoacromial ligament and inferiorly by the humeral head. The etiology of supraspinatus damage is not well understood and more than one mechanism of injury is likely. Some investigators believe that *impingement* is the predominate cause of injury to the tendon. In this scenario the tendon incurs damage by being repetitively compressed against the undersurface of the acromion or coracoacromial ligament (Fig. 4.28). For this mechanism of injury to occur, either the humeral head must translate superiorly or the superior structures must invade the space of supraspinatus travel. Either way, the tendon becomes chronically irritated and painful, making arm elevation difficult or impossible. Complete rupture is not uncommon. Cadaveric studies suggest that many people have non-symptomatic damage to the supraspinatus tendon.

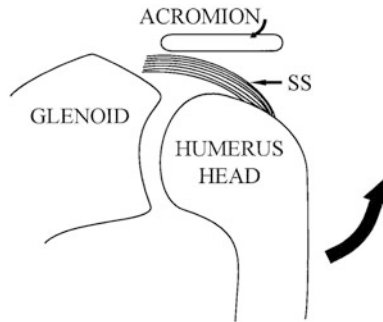


FIGURE 4.28. Cross section through the shoulder joint shows the precarious position of the supraspinatus tendon (SS). Abduction of the *humerus* (as indicated by the arrow) may pinch the tendon against the *acromion* as the humeral head rotates.

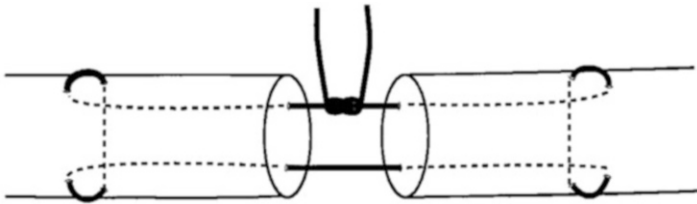


FIGURE 4.29. Typical suturing technique for repairing lacerated flexor tendons in the hand.

Surgical techniques are designed to repair tendon damage and relieve or *decompress* the tendon. Surgery once consisted of complete acromion removal, but this compromised normal deltoid function and was cosmetically damaging. Over the years the procedure has become progressively less radical. The current treatment uses either open or arthroscopic techniques to remove the undersurface of the acromion and repair the tendon. Surgical intervention is not always successful and is indicated only after conservative treatment, consisting of rest and exercise designed to strengthen the inferior rotator muscles, has failed.

4.9.6 Flexor Tendons of the Hand

The flexor tendons are often cut in two by lacerations of the hand. Such injuries demand immediate surgical repair for successful restoration of function. Several suture techniques are available for reconnecting these tendons. Most techniques are aimed at minimizing the amount of suture within the repair site to maximize the area available for collagen fiber deposition and fusion (Fig. 4.29). Evidence indicates that suturing to provide strength

sufficient to allow active flexion postoperatively may be a more important consideration because disuse during the postoperative recovery period leads to adhesions between the tendon and surrounding tissues. Autograft tendon from several different anatomical locations may be used to reconstruct missing segments of flexor tendon. Autografts taken from tendons within synovial sheaths, like those of the flexors they will replace, yield a better functional outcome than tendons taken from sites without an adjoining synovium.

As previously noted, scar formation and tendon adhesion are complicating factors in flexor tendon repair. Experimental evaluations of flexor tendon repair should include tendon gliding and measurement of the resulting range of joint motion, in addition to the material properties of the tendon itself. Achieving superior tendon strength without proper function is useless. Unfortunately, these two factors tend to be inversely related, so that a repair technique and biologic response that yield a strong tendon may also cause significant scar deposition and adhesion to surrounding tissues.

Knowledge concerning the molecular, intra- and extra-cellular signaling involved in tissue development, morphogenesis and repair is expanding in sync with the technical ability to manufacture highly customized biomaterials. These complementary advances in developmental biology and materials science are paving the way for new approaches to accelerate healing and rapidly restore the mechanical performance of structural tissues. New regenerative techniques are employing cell seeding, introducing various combinations of bioactive agents and incorporating scaffolds composed of various substrates. Progress has been made but clinically significant advances using any single approach remain elusive.

4.10 Summary and Further Reading

Tendons transmit muscle forces to bones, storing and releasing energy in the process; ligaments are passive joint stabilizers. Both these structures provide proprioceptive information that is vital to all kinds of motor activity. Both are composed largely of type I collagen, with ligaments containing somewhat more elastin. Both tendon and ligament readily exhibit disuse atrophy, and their compositions and mechanical properties seem to vary in accordance with their function; for example, flexor tendons are generally stiffer than their corresponding extensor tendons. It is hypothesized that this is controlled by cellular responses to mechanical signals. Viscoelasticity is an important feature of the mechanical properties of these tissues, so that they are sensitive to strain rate and load history. They also creep, relax, and exhibit hysteresis. Mechanical testing is complicated by their soft, slippery, wet, and viscoelastic nature, making gripping, dimensional analysis, and strain measurement problematic. When lacerated or torn, ligaments and

tendons will heal, but the healed tissue may never regain full strength and tendon motion is frequently constrained by adhesions to adjacent tissues.

For further reading about structure relationships in tendons, ligaments and their insertions we recommend two review articles by Benjamin and colleagues: *Structure-function relationships in tendons: a review* (Journal of Anatomy 212:211–228, 2008) and *Where tendons and ligaments meet bone: attachment sites ('entheses') in relation to exercise and/or mechanical load* (Journal of Anatomy. 208, 2006) (Benjamin et al. 2006, 2008). An informative review article by Wang entitled: *Mechanobiology of tendon* (Journal of Biomechanics 39:1563–1582, 2006) (Wang 2006) expands upon the limited amount of tendon biology presented here. To further explore biological aspects of tendon and ligament development and repair we refer readers to an article by Yang and co-authors entitled: *Tendon and ligament regeneration and repair: clinical relevance and developmental paradigm* (Birth Defects Research, 2013) (Yang et al. 2013). And although somewhat dated, the book *Injury and Repair of the Musculoskeletal Soft Tissues*, edited by Woo and Buckwalter (Woo and Buckwalter 1988) still provides an excellent overview of ligament and tendon tissue biology and mechanics.

4.11 Exercises

- 4.1. Show that Zajac's peak isometric muscle force-normalized generic muscle model implicitly assumes that the cross-sectional areas of tendons are proportional to the physiologic cross-sectional areas of their muscles.
- 4.2. Equations (4.1) and (4.2) represent empirical force-strain relationships for the tendon and aponeurosis of a frog muscle–tendon unit. In these, force is expressed as percent of the maximum tetanic muscle force. Calculate the tendon and aponeurosis strains when the muscle is exerting this maximum tetanic force. Which tissue has the greater strain? In general, what is the relationship between the tendon and aponeurosis strains for this muscle–tendon unit?
- 4.3. Show that Eq. (4.4) is the solution to Eq. (4.3).
- 4.4. Use the LaPlace transform method to show that Eq. (4.6) is the creep solution to Eq. (4.5).
- 4.5. Your collaborator, a small animal surgeon, has recently taken an interest in the canine medial collateral ligament of the elbow and is exploring different surgical techniques for its reconstruction. As a first step you have decided to completely define the mechanical behavior of the ligament including its failure characteristics. Describe your approach, including the parameters that will be measured, your

experimental methods and instrumentation, and problems which you anticipate.

- 4.6. You are testing anterior cruciate ligament (ACL) reconstructions with Dr. Kneely, an orthopaedic surgeon. You recorded the following load-deformation data for an experimental femur–ACL–tibia reconstruction in an animal model. Failure was by bony avulsion at the tibia. The ACL cross-sectional area was 150 mm^2 and the bone-to-bone resting length of the ligament was 3.5 cm. Calculate the important structural and material properties of the reconstruction.

Displacement, mm	Load, N	Displacement, mm	Load, N	Displacement, mm	Load, N
0.00	0	0.12	214	0.30	1908
0.01	0	0.14	379	0.32	1915
0.02	0	0.16	583	0.34	1940
0.03	0	0.18	796	0.36	2111
0.04	0	0.20	1000	0.37	2037
0.05	0	0.22	1202	0.38	1540
0.06	10	0.24	1406	0.39	1120
0.08	37	0.26	1611	0.40	821
0.10	96	0.28	1792	0.41	698

- 4.7. It was noted in Sect. 4.4.1 that tendon temperatures in racehorses may exceed 44°C . Suppose a typical tendon has an elastic modulus of 1 GPa and cycles between 0.5 and 2 % strain during running. If the stress-strain relationship is linear in this range, what is the fluctuation in strain energy density as the tendon stress rises and falls? If that energy were expressed as heat, how much would it be in calories/ml of tendon/cycle? (The mechanical equivalent of heat is 4.19 J/cal .) If the specific gravity and specific heat capacity of tendon tissue are similar to those of water (1 g/ml and $1 \text{ cal/g per }^\circ\text{C}$, respectively), what percent of the strain energy must be converted to heat to raise the tissue temperature 1°C in 1000 load cycles, assuming adiabatic conditions (i.e., no heat leaves the site)? Is it plausible that tendons could be heated by loading during running?
- 4.8. In Chap. 1 we learned that the skeleton is arranged as a system of levers that amplify the small displacements produced by contracting muscles. Instruments known as micromanipulators use levers to demagnify finger motions. In this chapter you have read that the flexor tendons of the fingers themselves have characteristics that serve to increase fine motor control; that is, they demagnify somewhat the motions of the

fingers which would otherwise be produced by the muscle contractions. Using the simple joint shown in Fig. 4.17 as a model, derive an equation to show how this can be achieved by making flexor tendons longer and more compliant. How would the tendon's length, diameter, and elastic modulus affect the resulting motion?

- 4.9. In Exercise 4.8, how would “pulleys” (loops of tissue to hold the tendon against the bones) near the joint contribute further to fine motor control? What other function would such pulleys have?
- 4.10. When a kangaroo switches from “pentapedal” (four limbs and tail) locomotion to hopping, its oxygen consumption drops, presumably because it then stores more energy in elastic tissues (Dawson and Taylor 1973). One of these tissue “springs” in kangaroos and other animals is the Achilles tendon. A kangaroo's Achilles tendon was found to be 1.5 cm in diameter and 35 cm long. *If both Achilles tendons had an elastic modulus of 1 GPa and were loaded to 2 % strain*, how much strain energy would they possess? Based strictly on energy considerations, how high could this amount of energy lift a 40-kg kangaroo?

References

- Alexander RM. Animal mechanics. Seattle, WA: University of Washington Press; 1968.
- Andarawis-Puri N, Kuntz AF, Kim SY, Soslowsky LJ. Effect of anterior supraspinatus tendon partial-thickness tears on infraspinatus tendon strainthrough a range of joint rotation angles. *J Shoulder Elbow Surg.* 2010;19(4):617–23.
- Arampatzis A, Peper A, Bierbaum S, Albracht K. Plasticity of human Achilles tendon mechanical and morphological properties in response to cyclic strain. *J Biomech.* 2010;43(16):3073–9.
- Azizi E, Roberts TJ. Biaxial strain and variable stiffness in aponeuroses. *J Physiol.* 2009;587(17):4309–18.
- Barad S, Cabaud HE, Rodrigo JJ. Effects of storage at -80°C as compared to 4°C on the strength of rhesus monkey anteriorcruciate ligaments. *Trans Orthop Res Soc.* 1982;7:378.
- Bay BK, Sharkey NA, Szabo RM. Biomechanical behavior of the median nerve in and about the carpal tunnel. *J Hand Surg.* 1997;22A:621–7.
- Benjamin M, Toumi H, Ralphs JR, Bydder G, Best TM, Milz S. Where tendons and ligaments meet bone: attachment sites (‘entheses’) in relation to exercise and/or mechanical load. *J Anat.* 2006;208(4):471–90.
- Benjamin M, Kaiswr E, Milz S. Structure-function relationships in tendons: a review. *J Anat.* 2008;212(3):211–28.
- Bey MJ, Song HK, Wehrli FW, Soslowsky LJ. A noncontact, nondestructive method for quantifying intratissue deformations and strains. *J Biomech Eng.* 2002;124(2):253–8.

- Bowser JE, Elder SH, Rashmir-Raven AM, Swiderski CE. A cryogenic clamping technique that facilitates ultimate tensile strength determinations in tendons and ligaments. *Vet Comp Orthop Traumatol*. 2011;24(5):370–3.
- Butler DL, Sheh MY, Stouffer DC, Samaranayake VA, Levy MS. Surface strain variation in human patellar tendon and knee cruciate ligaments. *J Biomech Eng*. 1990;112(1):38–45.
- Chimich D, Shrive N, Frank C, Marchuk L, Bray R. Water content alters viscoelastic behavior of the normal adolescent rabbit medial collateral ligament. *J Biomech*. 1992;25:831–7.
- Clancy WG, Narechania RG, Rosenberg TD. Anterior and posterior cruciate ligament reconstruction in rhesus monkeys: a histological, microangiographic, and biomechanical analysis. *J Bone Joint Surg*. 1981;63A:1270–84.
- Cook JL, Purdam CR. Is tendon pathology a continuum? A pathology model to explain the clinical presentation of load-induced tendinopathy. *Br J Sports Med*. 2009;43(6):409–16.
- Davis FM, De Vita R. A nonlinear constitutive model for stress relaxation in ligaments and tendons. *Ann Biomed Eng*. 2012;40(12):2541–50.
- Davison PF. The contribution of labile crosslinks to the tensile behavior of tendons. *Connect Tissue Res*. 1989;18:293–305.
- Dawson TJ, Taylor CR. Energetic cost of locomotion in kangaroos. *Nature (London)*. 1973;246:313–4.
- Doehring TC, Kahelin M, Vesely I. Direct measurement of nonuniform large deformations in soft tissues during uniaxial extension. *J Biomech Eng*. 2009;131(6):061001.
- Donkelaar CCV, Willems PJB, Muijtjens AMM, Drost MR. Skeletal muscle transverse strain during isometric contraction at different lengths. *J Biomech*. 1999;32:755–62.
- Duenwald SE, Vanderby R, Lakes RS. Stress relaxation and recovery in tendon and ligament: experiment and modeling. *Biorheology*. 2010;47:1–14.
- Ellis DG. Cross-sectional area measurements for tendon specimens: a comparison of several methods. *J Biomech*. 1969;2:175–86.
- Finni T, Komi PV, Lukkariniemi J. Achilles tendon loading during walking: application of a novel optic technique. *Eur J Appl Physiol*. 1998;77:289–91.
- Finni T, Komi PV, Lukkariniemi J. In vivo human triceps surae and quadriceps femoris muscle function in a squat jump and counter movement jump. *Eur J Appl Physiol*. 2000;83:416–26.
- Flügge W. *Viscoelasticity*. Waltham, MA: Blaisdell; 1967.
- Freehan LM, Beauchene JG. Early tensile properties of healing chicken flexor tendons: early controlled passive motion versus postoperative immobilization. *J Hand Surg*. 1990;15:63–8.
- Fukashiro S, Komi PV, Järvinen M, Miyashita M. In vivo achilles tendon loading during jumping in humans. *Eur J Appl Physiol*. 1995;71:453–8.
- Fung YCB. Stress-strain-history relations of soft tissues in simple elongation. In: Fung YCB, Perrone N, Anliker M, editors. *Biomechanics: its foundations and objectives*. Englewood Cliffs: Prentice-Hall; 1972.
- Gellhorn AC, Carlson MJ. Inter-rater, intra-rater, and inter-machine reliability of quantitative ultrasound measurements of the patellar tendon. *Ultrasound Med Biol*. 2013;39(5):791–6.

- Gillis C, Sharkey NA, Stover SM, Pool RR, Meagher DM. The effect of maturation and aging on the material and ultrasonographic properties of equine superficial digital flexor tendon. *Am J Vet Res.* 1995;56:1345–50.
- Giori NJ, Beaupre GS, Carter DR. Cellular shape and pressure may mediate mechanical control of tissue composition in tendons. *J Orthop Res.* 1993;11:581–91.
- Gregor RJ, Komi PV, JaÈrvinen M. Achilles tendon forces during cycling. *Int J Sports Med.* 1987;8:9–14.
- Haut RC. Age-dependent influence of strain rate on the tensile failure of rat-tail tendon. *J Biomech Eng.* 1983;105:296–9.
- Haut TL, Haut RC. The state of tissue hydration determines the strain rate sensitive stiffness of human patellar tendon. *J Biomech.* 1997;30:79–81.
- Haut RC, Little RW. A constitutive equation for collagen. *J Biomech.* 1972;5:423–30.
- Hingorani RV, Provenzano PP, Lakes RS, Escarcega A, Vanderby R. Nonlinear viscoelasticity in rabbit medial collateral ligament. *Ann Biomed Eng.* 2004;32:306–12.
- Hoffman AH, Robichaud DR, Duquette JJ, Grigg P. Determining the effect of hydration upon the properties of ligaments using pseudo Gaussian stress stimuli. *J Biomech.* 2005;38(8):1636–42.
- Hofmann CL, Okita N, Sharkey NA. Experimental evidence supporting isometric functioning of the extrinsic toe flexors during gait. *Clin Biomech.* 2013;28:686–91.
- Holmes G, Sharkey NA, Syftestad G, Reiser K. Changes in collagen crosslinks, fibril diameter, and mechanical strength in aging chicken tendons. *Trans Orthop Res Soc.* 1991;16:620.
- Howe JG, Wertheimer C, Johnson RJ, Nichols CE, Pope MH, Beynnon B. Arthroscopic strain gauge measurement of the normal anterior cruciate ligament. *Arthroscopy.* 1990;6:198–204.
- Jenkins RB, Little RW. A constitutive equation for parallel-fibered elastic tissue. *J Biomech.* 1974;7:397–402.
- Jung HJ, Vangipuram G, Fisher MB, Yang G, Hsu S, Bianchi J, Ronholdt C, Woo SL. The effects of multiple freeze-thaw cycles on the biomechanical properties of the human bone-patellar tendon-bone allograft. *J Orthop Res.* 2011;29(8):1193–8.
- Karamanidis K, Arampatzis A. Mechanical and morphological properties of different muscle-tendon units in the lower extremity and running mechanics: effect of aging and physical activity. *J Exp Biol.* 2005;208(Pt 20):3907–23.
- Kastelic J, Baer E. Deformation in tendon and collagen. In: Vincent JFV, Currey JD, editors. *The mechanical properties of biological materials.* Cambridge: Cambridge University Press; 1980. p. 397–435.
- Kim YS, Kim JM, Bigliani LU, Kim HJ, Jung HW. In vivo strain analysis of the intact supraspinatus tendon by ultrasound speckles tracking imaging. *J Orthop Res.* 2011;29(12):1931–7.
- Kirane YM, Michdelson JD, Sharkey NA. Evidence of isometric function of the flexor hallucis longus muscle in normal gait. *J Biomech.* 2008;41:1919–28.
- Knapik DM, Harris JD, Pangrazzi G, Griesser MJ, Siston RA, Agarwal S, Flanigan DC. The basic science of continuous passive motion in promoting knee health: a systematic review of studies in a rabbit model. *Arthroscopy.* 2013;29(10):1722–31.
- Komi PV, Fukashiro S, JaÈrvinen M. Biomechanical loading of Achilles tendon during normal locomotion. *Clin Sports Med.* 1992;11:521–31.

- Komi PV, Belli A, Huttunen V, Bonnefoy R, Geyssant A, Lacour JR. Optic as a transducer of tendomuscular forces. *Eur J Appl Physiol.* 1996;72:278–80.
- Kubo K, Ohgo K, Takeishi R, Yoshinaga K, Tsunoda N, Kanehisa H, Fukunaga T. Effects of isometric training at different knee angles on the muscle-tendon complex in vivo. *Scand J Med Sci Sports.* 2005;16(3):159–67.
- Kubo K, Morimoto M, Komuro T, Yata H, Tsunoda N, Kanehisa H, Fukunaga T. Effects of plyometric and weight training on muscle-tendon complex and jump performance. *Med Sci Sports Exerc.* 2007;39(10):1801–10.
- Laros GS, Tipton CM, Cooper RR. Influence of physical activity on ligament insertions in the knees of dogs. *J Bone Joint Surg.* 1971;53A:275–86.
- Leonard F, Moscovitz P, Hodge JW, Adams JP. Age-related Ca-Mg content and strength in turkey tendon. *Calcif Tissue Res.* 1976;19:331–6.
- Lieber RL, Leonard ME, Brown CG, Trestik CL. Frog semitendinosus tendon load-strain and stress-strain properties during passive loading. *Am J Physiol.* 1991;261:C86–92.
- Lieber RL, Brown CG, Trestik CL. Model of muscle-tendon interaction during frog semitendinosus fixed-end contractions. *J Biomech.* 1992;25:421–8.
- Maganaris CN. Tensile properties of in-vivo human tendinous tissue. *J Biomech.* 2002;35(8):1019–27.
- Maganaris CN. Tendon conditioning: artefact or property? *Proc Biol Sci.* 2003;270 Suppl 1:S39–42.
- Maganaris CN, Paul JP. Tensile properties of the in vivo human gastrocnemius tendon. *J Biomech.* 2002;35(12):1639–46.
- Maganaris DN, Kawakami Y, Fukunaga T. Changes in aponeurotic dimensions upon muscle shortening: in vivo observations in man. *J Anat.* 2001;199(Pt 4):449–56.
- Maganaris CN, Reeves ND, Rittweger J, Sargeant AJ, Jones DA, Gerrits K, De Haan A. Adaptive response of human tendon to paralysis. *Muscle Nerve.* 2006;33:85–92.
- Malliaras P, Kamal B, Nowell A, Farley T, Dhamu H, Simpson V, Morrissey D, Langberg H, Maffulli N, Reeves ND. Patellar tendon adaptation in relation to load-intensity and contraction type. *J Biomech.* 2013;46(11):1893–9.
- Matthews GL, Keegan KG, Graham HL. Effects of tendon grip technique (frozen versus unfrozen) on in vitro surface strain measurements of the equine deep digital flexor tendon. *Am J Vet Res.* 1996;57(1):111–5.
- Matyas JR, Anton MG, Shrive NG, Frank CB. Stress governs tissue phenotype at the femoral insertion of the rabbit MCL. *J Biomech.* 1995;28:147–57.
- Moon DK, Abramowitch SD, Woo SL. The development and validation of a charge-coupled device laser reflectance system to measure the complex cross-sectional shape and area of soft tissues. *J Biomech.* 2006;39(16):3071–5.
- Nachemson AL, Evans JH. Some mechanical properties of the third human lumbar interlaminar ligament (ligamentum flavum). *J Biomech.* 1968;1:211–20.
- Neumann P, Ekstrom LA, Kellar TS, Perry L, Hansson TH. Aging, vertebral density, and disk degeneration alter the stress-strain characteristics of the human anterior longitudinal ligament. *J Orthop Res.* 1994;12:103–12.
- Noyes FR. Functional properties of knee ligaments and alterations induced by immobilization. A correlative biomechanical and histological study in primates. *Clin Orthop Rel Res.* 1977;123:210–42.

- Noyes FR, Grood ES. The strength of the anterior cruciate ligament in humans and rhesus monkeys. Age-related and species-related changes. *J Bone Joint Surg.* 1976;58:1074–82.
- Noyes FR, Delucas JL, Torvik PJ. Biomechanics of anterior cruciate ligament failure: an analysis of strain-rate sensitivity and mechanisms of failure in primates. *J Bone Joint Surg.* 1974;56:236–53.
- Noyes FR, Butler DL, Grood ES, Zernicke RF, Hefzy MS. Biomechanical analysis of human ligament grafts used in knee-ligament repairs and reconstructions. *J Bone Joint Surg.* 1984;66:344–52.
- Onambele GL, Narici MV, Maganaris CN. Calf muscle-tendon properties and postural balance in old age. *J Appl Physiol* 1985. 2006;100(6):2048–56.
- Pokhai GG, Oliver ML, Gordon KD. A new laser reflectance system capable of measuring changing cross-sectional area of soft tissues during tensile testing. *J Biomech Eng.* 2009;131(9):094504.
- Reeves ND, Narici MV, Maganaris CN. Strength training alters the viscoelastic properties of tendons in elderly humans. *Muscle Nerve.* 2003;28:74–81.
- Reeves ND, Maganaris CN, Narici MV. Plasticity of dynamic muscle performance with strength training in elderly humans. *Muscle Nerve.* 2005;31:355–64.
- Riemersma DJ, Schamhardt HC. The cryo-jaw, a clamp designed for in vitro rheology studies of horse digital flexor tendons. *J Biomech.* 1982;15:619–20.
- Roberts CS, Cummings JF, Grood ES, Noyes FR. In-vivo measurement of human anterior cruciate ligament forces during knee extension exercises. 40th Annual meeting, orthopaedic research society, New Orleans, LA; 1994. p. 84.
- Salter RB, Simmonds DF, Malcolm BW, Rumble EJ, MacMichael D, Clements ND. The biological effect of continuous passive motion on the healing of full-thickness defects in articular cartilage. An experimental investigation in the rabbit. *J Bone Joint Surg Am.* 1980;62(8):1232–51.
- Scott SH, Loeb GE. Mechanical properties of aponeurosis and tendon of the cat soleus muscle during whole muscle isometric contractions. *J Morphol.* 1995;224:73–86.
- Shadwick RE. Elastic energy storage in tendons: mechanical differences related to function and age. *J Appl Physiol.* 1990;68:1033–40.
- Sharkey NA, Smith TS, Lundmark DC. Freeze clamping musculo-tendinous junctions for in-vitro simulation of joint mechanics. *J Biomech.* 1995;28(5):631–5.
- Stenroth L, Peltonen J, Cronin NJ, Sipilä S, Finni T. Age-related differences in Achilles tendon properties and triceps surae muscle architecture in vivo. *J Appl Physiol.* 2012;113(10):1537–44.
- Stouffer DC, Butler DL, Hosney D. The relationship between crimp pattern and mechanical response of human patellar tendon-bone units. *J Biomech Eng.* 1985;107:158–65.
- Takai S, Woo SL, Horibe S, Tung DK, Gelberman RH. The effects of frequency and duration of controlled passive mobilization on tendon healing. *J Orthop Res.* 1991;9:705–13.
- Thornton GM, Shrive NG, Frank CB. Altering ligament water content affects ligament pre-stress and creep behavior. *J Orthop Res.* 2001;19:845–51.
- Tidball JG. The geometry of actin filament-membrane associations can modify adhesive strength of the myotendinous junction. *Cell Motility.* 1983;3:439–47.
- Tidball JG, Daniel TL. Myotendinous junctions of tonic muscle cells: structure and loading. *Cell Tissue Res.* 1986;245:315–22.

- Tidball JG, Salem G, Zernicke R. Site and mechanical conditions for failure of skeletal muscle in experimental strain injuries. *J Appl Physiol.* 1993;74:1280–6.
- Tipton CM, Schild RJ, Tomanek RJ. Influence of physical activity on the strength of knee ligaments in rats. *Am J Physiol.* 1967;212:783–7.
- Tipton CM, Matthes RD, Martin RK. Influence of age and sex on the strength of bone-ligament junctions in knee joints of rats. *J Bone Joint Surg.* 1978;60A:230–4.
- Trotter JA, Corbett K, Avner BP. Structure and function of the murine muscle-tendon junction. *Anat Rec.* 1981;201:293–302.
- Vailas AC, Tipton CM, Mattes RD, Gart M. Physical activity and its influence on the repair process of medial collateral ligaments. *Connect Tissue Res.* 1981;9:25–31.
- van Doesburg MH, Yoshii Y, Henderson J, Villarraga HR, Moran SL, Amadio PC. Speckle-tracking sonographic assessment of longitudinal motion of the flexor tendon and subynovial tissue in carpal tunnel syndrome. *J Ultrasound Med.* 2012;31(7):1091–8.
- Wang JHC. Mechanobiology of tendon. *J Biomech.* 2006;39:1563–82.
- Wertheim MG. Memoirs sur l'élasticité et la cohésion des principaux tissus des corps humain. *Anales de Química y Física.* 1847;21:385–414.
- Wilson AM, Goodship AE. Exercise-induced hyperthermia as a possible mechanism for tendon degeneration. *J Biomech.* 1994;27:899–905.
- Woo SL-Y. Mechanical properties of tendons and ligaments: quasistatic and nonlinear properties. *Biorheology.* 1982;19:385–96.
- Woo SL-Y, Buckwalter JA. Injury and repair of the musculoskeletal soft tissues. Park Ridge, IL: American Academy of Orthopaedic Surgeons; 1988. p. 1988.
- Woo SL-Y, Ritter MA, Amiel D, Sanders TM, Gomez MA, Kuei SC, Garfin SR, Akeson WH. The biomechanical and biochemical properties of swine tendons—long term effects of exercise on the digital extensors. *Connect Tissue Res.* 1980;7(3):177–83.
- Woo SLY, Gomez MA, Seguchi Y, Endo CM, Akeson WA. Measurement of mechanical properties of ligament substance from a bone-ligament-bone preparation. *J Orthop Res.* 1983;1:22–9.
- Woo SLY, Orlando CA, Gomez MA, Frank CB, Akeson WH. Tensile properties of the medial collateral ligament as a function of age. *J Orthop Res.* 1986;4:133–41.
- Woo SL, Gomez MA, Sites TJ, Newton PO, Orlando CA, Akeson WH. The biomechanical and morphological changes in the medial collateral ligament of the rabbit after immobilization and remobilization. *J Bone Joint Surg.* 1987;69A:1200–11.
- Woo SLY, Peterson RH, Ohland KJ, Sites TJ, Danto MI. The effects of strain rate on the properties of the medial collateral ligament in skeletally immature and mature rabbits: a biomechanical and histological study. *J Orthop Res.* 1990;8:712–21.
- Woo SL-Y, An K-N, Arnoczky SP, Wayne JS, Fithian DC, Myers BS. Anatomy, biology, and biomechanics of tendon, ligaments, and meniscus. In: Simon SR, editor. *Orthopaedic basic science.* Park Ridge, IL: American Academy of Orthopaedic Surgeons; 1994. p. 45–87.
- Yang G, Rothrauff BB, Tuan RS. Tendon and ligament regeneration and repair: clinical relevance and developmental paradigm. *Birth Defects Res C Embryo Today.* 2013;99(3):203–22.
- Zajac FE. Muscle and tendon: properties, models, scaling, and application to biomechanics and motor control. *Crit Rev Biomed Eng.* 1989;17:359–411.

5

Synovial Joint Mechanics

[The bone ends] are covered with a smooth elastic crust, to prevent mutual abrasion; connected with string ligaments, to prevent dislocation; and enclosed in a bag that contains a proper fluid deposited there for lubricating the two contiguous surfaces (Hunter 1743).

Abstract Most large joints, for example the knee, have very low friction under dynamic loading. The viscoelastic properties, fluid flow through, and surface decorating molecules of articular cartilage are keys to the efficient transmission of load through this soft tissue.

Keywords Cartilage • Viscoelasticity • Fluid Flow • Permeability • Electrostatics • Polymer Mechanics • Tribology • Wear • Friction • Lubrication

5.1 Introduction

Now that we have explored the mechanical behaviors of tendons and ligaments, constructs that transmit the forces needed to produce and constrain relative motion between bones, we will turn our attention to the tissue that directly transmits forces from one bone to another while at the same time allowing for smooth almost frictionless motion between them. These motions occur within our synovial joints and the predominant tissue responsible is cartilage. In this chapter we explore the mechanics of synovial joints; however, we must not forget what we have learned about bone, tendon and ligament because their presence heavily influences the characteristics of the new tissues we are going to study.

Synovial or diarthroidal joints are those that move freely and contain synovial fluid, like the knee and elbow joints, as opposed to joints that do not, such as the “tether” joints between the vertebral bodies in the spine. Figure 5.1 is a schematic diagram of a typical synovial joint. The synovial fluid is manufactured by cells in the synovial membrane, which lines the interior of the fibrous joint capsule. As explored previously, the joint capsule is aided by ligaments (not shown) and the surrounding musculotendinous

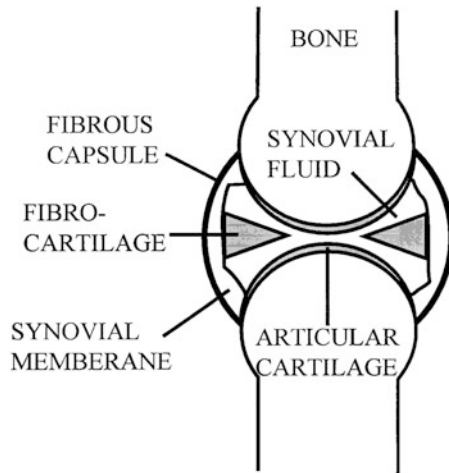


FIGURE 5.1. Schematic diagram of a typical synovial joint.

tissue in holding the two bones of the joint together (i.e., “stabilizing” the joint). The bone surfaces within the joint capsule are covered by articular cartilage, also called hyaline cartilage. The purpose of this tissue is not to serve as a shock absorber, as some textbooks state, but to provide a suitable surface for lubrication and wear prevention. Some joints also contain fibrocartilage disks that help maintain the other joint components in appropriate positions and reduce stress by extending the load-bearing surfaces. Such disks are generally found in joints where there is substantial rotation or torsion but incongruous joint surfaces (e.g., the acromioclavicular joint). In the knee, these disks are called menisci. An important compositional difference between articular cartilage and meniscus is that the latter contains type I instead of type II collagen (Mow et al. 1992).

5.1.1 Functions of a Synovial Joint

Synovial joints provide essentially frictionless motion between limb segments while transmitting relatively high loads between them. Joints in animals are the equivalent of “bearings” in machines. Of course, these bearings do not support continuously rotating shafts, but components that start, stop, and oscillate. A similar engineering example might be the ball joints in the steering mechanism of an automobile.

We saw in Chap. 1 that the forces acting in the joints between bones are quite large relative to the external loads being supported; the stresses in the bones, tendons, and ligaments are therefore relatively high as well. The same may be said of the stresses in cartilage. The normal contact pressures exerted

on the cartilage in various normal joints across species typically range between 0.5 and 5.0 MPa (Brand 2005).

5.1.2 *Joint Diseases*

Joint disease and osteoporosis are the two most common skeletal pathologies in modern society. One or the other seems to give almost everyone problems eventually. Their human and economic costs are enormous, and are increasing as other diseases are controlled and people live longer. Although these conditions do not kill directly, they do so indirectly in most painful and debilitating ways. (Immobilization itself, from the pain of arthritis, can be life-threatening to an elderly person.) The causes of both osteoporosis and arthritis are both poorly understood, but the difficulties with understanding them are different. Mechanically, osteoporosis is rather simple: there is not enough bone tissue, and therefore the bone structure is weak and fractures occur. The complexity is in the bone remodeling biology that leads to diminished bone structure and altered bone quality. In the case of an arthritic joint, on the other hand, both the mechanics of joint function and the cell-level biology are poorly understood. However, it seems clear that most kinds of arthritis involve feedback loops between mechanical damage processes and inflammatory processes (Fig. 5.2). For example, mechanical overuse may create fatigue damage in cartilage. This damage may incite an inflammatory process that activates bone and cartilage cells not only to repair the damage, but also to attempt to reshape the tissues to improve the mechanical situation. If too much begins to happen at once, however, the circumstances become chaotic. Enzymes intended to get rid of a damaged tissue also destroy healthy tissue. The size and shape of a joint can be altered rather easily in young people by bone and cartilage modeling, but after growth stops, these processes become sluggish and, apparently, are not well matched to mechanical needs. New cartilage and bone production formed by the joint result in misshapen “spurs” (called osteophytes) around the periphery of the joint, which often get in the way and cause pain, or cause damage to adjacent muscle (e.g., in the case of some rotator cuff tears in the shoulder).

Alternatively, an infection or a defect in the patient’s immune system may cause an inflammatory response in a joint when mechanical loading is normal. This response can lead directly to tissue destruction, which so jeopardizes mechanical properties that normal loads produce damage-level strains. Thus, there are many possibilities for the development of “vicious circles” in which mechanical and cellular processes in the bone and cartilage tissues, which normally operate in a controlled way at modest levels, become rampant and unstable.

Understanding joint mechanics is essential to understanding various forms of arthritis. The way synovial joints work is not entirely clear, but much is known and it is a fascinating story.

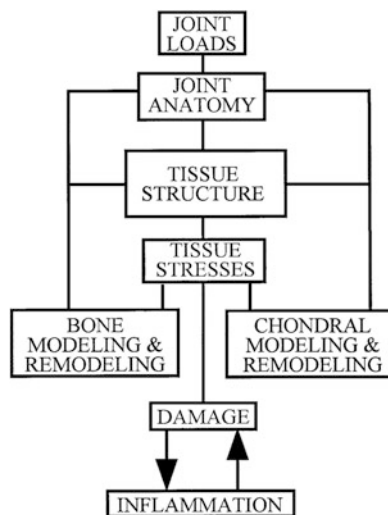


FIGURE 5.2. Block diagram of potential feedback loops in synovial joint physiology. Joint loads and anatomy determine tissue stresses, which produce tissue property-dependent strains that influence bone and chondral modeling and remodeling. If tissues are damaged, inflammatory responses as well as modeling and remodeling of bone and cartilage, may be activated, influencing joint anatomy and tissue properties. In some instances, the system may be stable and self-correcting; in others, it may become unstable and self-destructive.

5.2 Mechanical Properties of Cartilage

5.2.1 Initial Points

The general structure and composition of cartilage were introduced in Chap. 2. There are three important initial points to be made about the importance of this structure to the mechanical role that cartilage plays in joint mechanics.

5.2.1.1 Cartilage Is 70 % Water

The extracellular matrices (ECM) of all connective tissues (including bone and cartilage) are three-dimensional crosslinked networks of insoluble collagen fibers with other more soluble components enmeshed in, integrated with or chemically attached to the collagen fibers. The void dimensions of cartilage ECM are about 50 Å, much smaller than the 50–300 μm void size of bone. The very small distances between molecular components in cartilage very strongly affects mechanical properties. Cartilage ECM is about 95 % of the total tissue volume and the remaining 5 % is filled with cartilage cells called chondrocytes. Water completely fills the entire tissue volume and is in very intimate association with the matrix and, of course, the chondrocytes.

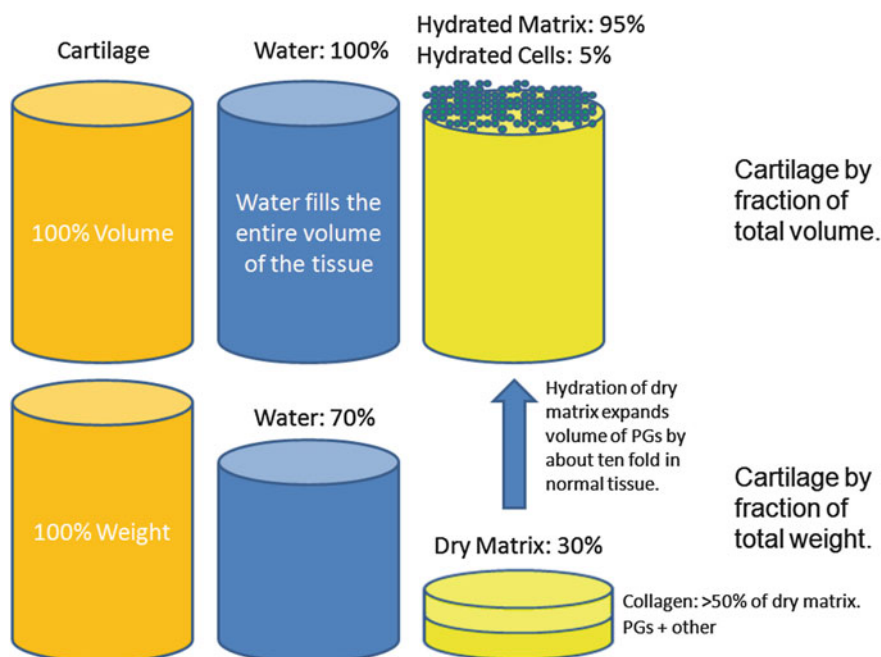


FIGURE 5.3. In the simplest sense, the three components of cartilage are water, extracellular matrix and chondrocytes.

The chondrocytes create and assemble the matrix initially and maintain the framework of collagen and proteoglycans. Cartilage is essentially an extracellular matrix inflated with water that is held in place by strong electrochemical interactions with the proteoglycans. If anything interferes with the proteoglycan content (e.g., the release of destructive enzymes by the chondrocytes), or with the ability of chondrocytes to make normal ECM (e.g., a genetic defect), the mechanical functioning of the cartilage will deteriorate (Kuettner 1992) (Fig. 5.3).

The material properties of articular cartilage depend on the detailed relationships between water and the macro-molecular components of the ECM. Fibrillar type II collagen and proteoglycans are the primary macro-molecules of adult articular cartilage ECM, however, water is 70 % (by weight) of the tissue. The water is very tightly bound to the ECM by the physical chemical properties of the macromolecules. Collagen is more than 50 % of the dry weight (or 10–20 % of the wet weight) of articular cartilage. In adults, approximately 90 % of cartilage collagen is of type II which is found in only a few other tissues such as the vitreous body of the eye. Type II is a fibril-forming collagen and the molecules are made of three identical $\alpha_1(\text{II})$ polypeptide chains. Each of the $\alpha_1(\text{II})$ molecules is itself a

300 nm-long triple helix initially formed with short telopeptides at each end that are enzymatically removed before assembly. Once the collagen molecules leave the cell they assemble into fibrils extracellularly. The resulting network of type II collagen fibrils gives the ECM tensile strength and is essential to maintaining the tissue's volume and shape. The tensile strength of the collagenous matrix is increased by enzymatically produced covalent, intermolecular crosslinks.

Proteoglycans (PGs) are highly abundant in the ECM of all connective tissues but especially in cartilages. In articular cartilage these highly hydrophilic macromolecules are embedded within the fibrillar network in an under-hydrated (compressed) form. They give articular cartilage the ability to undergo large, non-damaging, reversible compression. The predominant PG of articular cartilage is aggrecan. It contains a protein core (215 kDa) to which glycosaminoglycans and oligosaccharides are covalently attached. Ninety percent of the total mass of aggrecan is from the glycosaminoglycans keratan sulfate and chondroitin sulfate.

Aggrecan attaches noncovalently to the very high molecular weight biopolymer hyaluronan. A link protein acts to stabilize this interaction between aggrecan and hyaluronan making it nearly as stable as a covalent bond. The combination of aggrecan and hyaluronan forms an extraordinarily large space-filling molecule called aggregate. In free solution, aggrecan swells to about 50 times its dry volume. However, in healthy cartilage the collagen network constrains swelling to about ten times the dry volume (i.e., 20 % of free swelling). This permits the accumulation of high concentrations of compressed aggrecan in the tissue. Aggrecan is a highly negatively charged molecule. The highly concentrated negative charges repel each other and also strongly attract water molecules creating a swelling pressure. Under compressive load, water is pushed out of aggrecan and the whole tissue. Compression of the proteoglycans increases the negative charge density in the tissue which, in turn, increases electrical repulsion between the negative charges and the attraction of water into the tissue. The process of water expression continues until there is equilibrium between loading forces and the electrical and osmotic swelling stresses. Removing the load permits water to return to the tissue and it expands until the expansion force of the proteoglycans is balanced by tension in the fibrillar network. The numbers of negatively charged groups in the proteoglycans will rise and fall in proportion to the pH of the synovial fluid. Cations in the synovial fluid (e.g., Na⁺ and Ca⁺⁺) also are attracted to and shield the negative proteoglycan charges reducing the effective charge density; divalent cations having effects much more than twice as strong as monovalent cations (June et al. 2009). For these and other reasons, the chemistry of cartilage has substantial effects on its mechanical properties.

In addition to the swelling properties of the PG's, their high negative charge concentration has strong effects on the diffusion of solutes. The negative charges of aggrecan are fixed to the molecule and don't diffuse

independently. As a result, they create a high density of fixed negative charge that can retard the diffusion of negatively charged (anionic) macromolecules into the tissue while attracting and concentrating positively charged (cationic) molecules (Nimer et al. 2003). The effect of the fixed negative charges on the diffusion of neutral macromolecules with a mixture of positive and negative regions (zwitterions) is more complex.

5.2.1.2 Cartilage Is Viscoelastic

Cartilage is a water-filled biopolymer that has strong interactions between its constituent components. As a result, its mechanical properties are viscoelastic due to the flow of water in the matrix and time dependent reconfiguration of the biopolymers under mechanical loading. The elastic modulus of the tissue varies greatly with the rate of loading, demonstrating a broad spectrum of interactions between water and the biopolymers comprising the tissue. At very slow rates of loading, the resulting modulus is only a few MPa, a characteristic that can be reasonably modeled as the slow percolation of water through an elastic matrix. At physiologic rates of dynamic loading the modulus of cartilage is much larger, approaching 500 MPa (Higginson and Snaith 1979; Radin et al. 1970; Unsworth 1981; Yamada 1973). The mechanical behavior of cartilage under these more typical dynamic conditions cannot be explained using simple fluid flow models; additional interactions between the constituent biopolymers are needed to explain this more extreme viscoelastic behavior. This range of behaviors should be borne in mind when reading the rest of this chapter, because many of the descriptions of cartilage stiffness measurements refer to behavior at very slow rates of loading. For example, indentation tests are usually quasistatic (extremely slow loading rate) and the “equilibrium” or “aggregate” modulus used in many models refers to the stiffness associated with an infinitely slow rate of loading.

5.2.1.3 Cartilage Is Not a Shock Absorber

Finally, while it is much less stiff than cortical bone, cartilage does not serve to “cushion” or reduce impulsive forces in joints. It cannot do this because it is so thin that its capacity to absorb energy is insignificant compared to (a) eccentric contractions of muscles and (b) energy absorption in the bones on either side of the joint. (Exercise 5.1 asks you to demonstrate this to yourself). Instead, the role of cartilage is to provide a self-renewing, well-lubricated, compliant, load-bearing surface.

5.2.2 Structure of Articular Cartilage

Figure 5.4 is a schematic diagram of the structure of articular cartilage in relation to its foundation of subchondral bone. The collagen fibers in the cartilage have been described as forming arches (the “arcades of

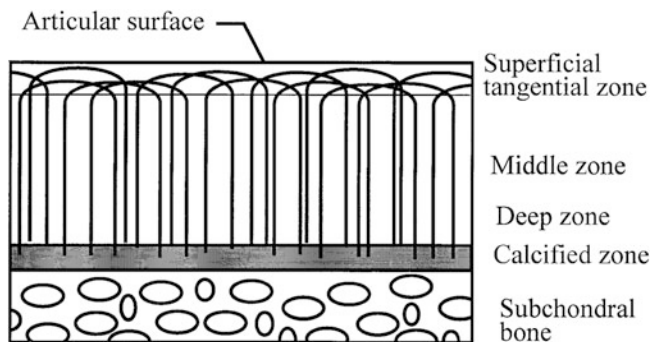


FIGURE 5.4. Schematic diagram of collagen fiber orientations in articular cartilage.

Benninghoff”) (Benninghoff 1925) such that their orientation is parallel to the surface in the superficial tangential zone and perpendicular to it in the deeper zones. This description of the collagen orientation in articular cartilage is not really correct, but rather the collagen appears as arches because the orientation of the collagen fibrils gradually change throughout the cartilage matrix, being more parallel to the surface superficially, and orthogonal to the surface in the deeper cartilage layers. The chondrocytes tend to line up in columns parallel to the collagen fibers in the middle and deep zones. As in the growth plate, the deepest layer of cartilage calcifies.¹ The junction between calcified and uncalcified cartilage is called the tidemark and shows up as a line in stained histologic sections. The calcified cartilage layer may reduce the stress concentration that would exist at the bone–cartilage interface by virtue of their different stiffnesses. Beyond these general observations, the structure of cartilage remains poorly understood, but even this degree of understanding reveals much about the origins of the anisotropy and inhomogeneity in cartilage’s mechanical properties.

5.2.3 Permeability

Because cartilage consists of an organic biopolymeric matrix and water, it can be useful to think of it as a multiphasic system. The simplest multiphasic system is biphasic, with water being one phase and the collagen–proteoglycan biopolymer matrix the other. The volume fraction of the water phase can be called the porosity of the cartilage. Although the water molecules are restrained to a significant degree by the matrix molecules, they may

¹ In fact, one may think of articular cartilage as a miniature growth plate that exists throughout one’s lifetime. Its chondrocytes manufacture cartilage very slowly, however, so that any added length to the parent bone is very small.

TABLE 5.1. Properties of lateral femoral condyle cartilage in several different species

Variable	Human	Bovine	Canine	Monkey	Rabbit
Aggregate modulus (MPa)	0.70	0.89	0.60	0.78	0.54
Poisson's ratio	0.10	0.40	0.30	0.24	0.34
Permeability, $10^{-15} \text{ m}^4/\text{N s}$	1.18	0.43	0.77	4.19	1.81

From Mow et al. (1991)

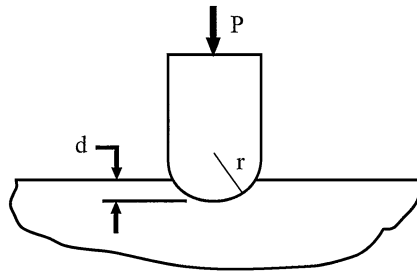


FIGURE 5.5. Diagram of indentation experiment. An indenter having a hemispherical end of radius r is pressed into the surface of the tissue for a distance d by a load P .

nevertheless slowly migrate through the tissue under the influence of a pressure gradient. This may be modeled using Darcy's law:

$$Q = -k(dP/dx) \quad (5.1)$$

where Q is the flow rate per unit cross-sectional area, dP/dx is the pressure gradient, and k is the hydraulic permeability coefficient (sometimes called the apparent permeability). Typical values of k are shown in Table 5.1. Permeability diminishes in the deepest layers of the articular cartilage, and also varies with the composition of the liquid phase. Clearly, the resistance of the tissue to deformation depends to some extent on its permeability.

5.2.4 Indentation Testing

It is difficult to form mechanical test specimens of cartilage because it is only found in thin (2–3 mm) layers. Therefore, many of the early mechanical tests involved indentation of the surface as a means of gaining compressive stiffness data. Such experiments were based on the equation derived by Boussinesq (1885) for a hemispherical indenter applied to an elastic half-space (Fig. 5.5). If an indenter of radius r penetrates to a depth d under load P , the elastic modulus is given by

$$E = P(1 - \nu^2)/2dr \quad (5.2)$$

where ν is Poisson's ratio of the indented material. Assuming that the cartilage was indeed of infinite depth and that it was incompressible (so that $\nu = 0.5$), Sokoloff (1966) obtained $E = 2.3$ MPa for the elastic modulus of articular cartilage. Later, when a mathematical solution was obtained for indentation of an elastic layer on a rigid foundation (i.e., cartilage on relatively rigid bone) by a plane-ended cylinder, Hayes et al. (1972) found that E was only 1.6 MPa. As more work was done, it became clear that this value, too, was probably inaccurate because cartilage is not truly incompressible and fails to meet the assumptions of the tests in other ways as well (e.g., it is anisotropic, inhomogeneous, and viscoelastic). Table 5.1 shows data on the compressive modulus and Poisson's ratio of articular cartilage for various species. These data were obtained from compression tests on machined specimens using techniques similar to those described later in this section.

Box 5.1 The Different Moduli for Simple and Poroelastic Materials (Fig. 5.6)

An isotropic linear elastic material such as discussed in the previous section has two independent material properties, the Young's modulus (E) and the Poisson's ratio (ν). These are measured in the "usual way" in unconfined compression of a specimen as $E = \sigma/\epsilon_y$ and $\nu = \epsilon_x/\epsilon_y$. A material filled with fluid (like cartilage) has similar properties, but they depend on whether the fluid is in or out of the solid matrix (Korhonen

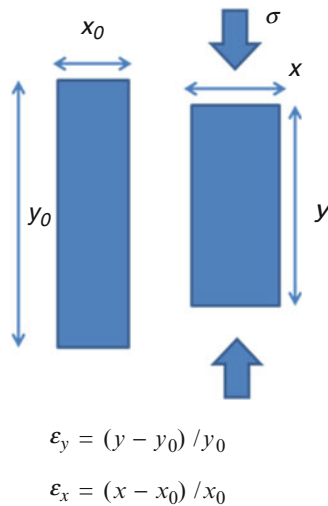


FIGURE 5.6. In a simple compression test the Young's modulus is $E = \sigma/\epsilon_y$ and the Poisson's ratio is $\nu = \epsilon_x/\epsilon_y$.

et al. 2002)! If you compress a fluid filled material quickly it will shorten and widen in the same manner as does a simple solid. However, if you hold the compressive displacement for a long time the fluid can percolate out of the matrix. As a result of fluid percolation, the stress to hold the compression decreases, the width of the specimen decreases and the volume of the matrix decreases with time—eventually reaching a steady-state value that depends only on the properties of the drained solid matrix. In many of the simpler theories for fluid filled materials it is the drained modulus and drained Poisson’s ratio of the solid matrix that are needed to predict the mechanical response of the matrix to loading. Fluid effects are governed by the permeability of the solid matrix which determines how the solid and fluid interact. A little curve ball in the use of the drained moduli for the solid in a fluid filled material is that many theories (in particular the KLM or poroelastic theory discussed below) often use a confined mechanical test that prevents the widening of the specimen to measure mechanical properties. The modulus calculated from a confined test is called the aggregate modulus (H_A). The aggregate modulus has a defined relationship with the drained Young’s modulus and drained Poisson’s ratio of the matrix as:

$$H_A = E(1 - \nu)/[(1 + \nu)(1 - 2\nu)]$$

Whether one measures the aggregate modulus directly using a confined compression test or calculates it from the drained moduli of the matrix is all the same mathematically.

5.2.5 Tensile Tests

Woo and co-workers (1976) pioneered the tensile testing of cartilage. They sliced articular cartilage into miniature “dumbbells” having a thickness of 0.25 mm and clamped them into a tensile test apparatus. Strain was measured by videotaping the specimen as it was pulled apart and measuring, in each video frame, the distance between two lines drawn across the undeformed specimen (see Fig. 4.11 for further details on this approach).

When cartilage is loaded, the way the force is shared between the fluid and solid phases is highly dependent on the nature of the loading. In tension, it is primarily the solid phase of the cartilage that carries the load; in particular, it is the collagen fibers. Figure 4.1 shows a typical tensile load-deformation relationship for a collagenous soft tissue loaded slowly. As explored in Chap. 4, these tissues exhibit a “toe” region with an initially low modulus, followed by a region of increasing stiffness. Fibers in the toe region straighten and become stretched in the linear region; in the case of cartilage, it is probably the collagen fibrils more than the proteoglycan molecules that are behaving in this way.

Cartilage produces the same sort of curve seen in Fig. 4.1 except that its upper portion remains nonlinear because the collagen fibers are not brought into full alignment. Its stress-strain relationship can therefore be described by an equation of the form

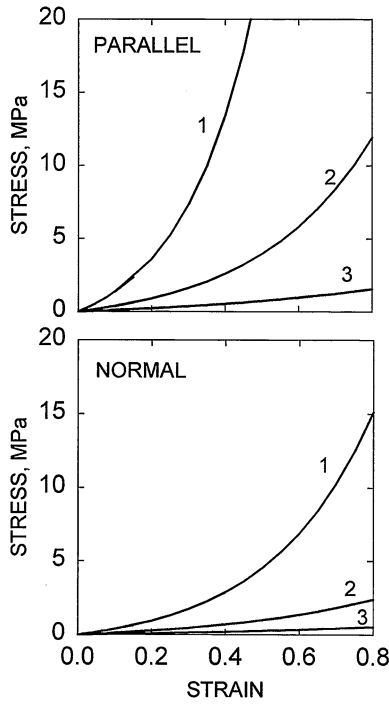


FIGURE 5.7. Stress-strain curves for bovine cartilage specimens as functions of orientation and depth. Loading rate approximately 1.2 % per second (Woo et al. 1976).

$$\sigma = A(e^{B\epsilon} - 1) \quad (5.3)$$

where σ is stress, ϵ is strain, and A and B are coefficients. To find the modulus at any point on the curve, one merely differentiates to obtain

$$E = d\sigma/d\epsilon = B(\sigma + A) \quad (5.4)$$

Figure 5.7 shows the shapes of such stress-strain curves for bovine articular cartilage as a function of depth below the surface and orientation relative to the split line direction (recall Fig. 2.20). Table 5.2 shows the elastic modulus at a stress of 2.5 MPa for each position and orientation. Note that the tensile modulus is greater parallel to the split lines. Near the articular surface fluid pressure caused by loading and the tensile stresses in the matrix at the edge of a loaded region have components that are perpendicular to the load (parallel to the surface) and tend to tear the surface of the tissue. As a result, the tangential orientation of the collagen at the tissue surface is a useful direction for resisting the tensile stress. Ultimate tensile stresses for articular cartilage range from 9 to 18 MPa and ultimate tensile strains from 60 % to 120 % (Williamson et al. 2003).

TABLE 5.2. Data for bovine cartilage

Direction	Slice # (relative depth)	Tensile modulus (MPa)
Parallel to split lines	1	23.0
2	11.2	
3	5.0	
Normal to split lines	1	2.2
2	6.6	
3	3.6	

Loading rate approximately 1.2 % per second

From Woo et al. (1976)

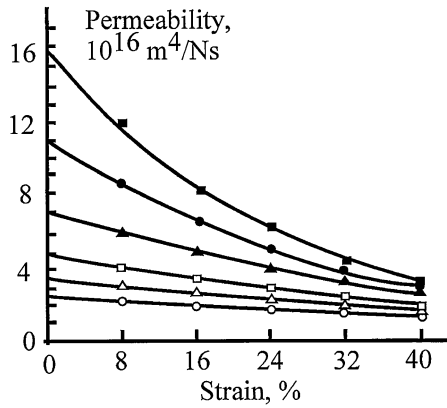


FIGURE 5.8. Permeability vs. compressive strain for cartilage. The family of curves shows how permeability decreases as increasing pressure is applied. From top to bottom, the pressures were 0.07, 0.17, 0.34, 0.69, 1.03, and 1.72 MPa. Previously published as Fig. 4 in Mow, V.C., Kuei, S.C., Lai, W.M., and Armstrong, C.G. Biphasic creep and stress relaxation of articular cartilage in compression. Theory and experiment. *Journal of Biomechanical Engineering* 1980; 102:73–84 (Mow et al. 1980).

5.2.6 Poroelastic Theory of Articular Cartilage

When a cartilage specimen is loaded in tension or compression, drops of fluid can be seen to exude and collect on its surfaces (Fig. 5.8). This is a clear sign that cartilage liquid flows under the influence of pressure gradients produced by mechanical deformation of the solid phase. Movement of the fluid phase produces some of the viscoelastic behavior of the tissue, especially for slow loading. Additionally, the permeability of the tissue changes when the solid phase is deformed. Mansour and Mow (1976) studied this behavior in 200- μm -thick specimens of cartilage clamped between two porous metal

platens. Through these porous platens a pressure gradient was imposed across the specimen: the applied pressure, P_A . The resulting flow was measured and the permeability coefficient k calculated. By compressing the specimen to different strain levels (ϵ), the relationship between permeability and tissue strain could be studied. Typical results show that an interaction occurs between the deformations of the fluid and solid phases when cartilage is loaded. The curves can be fit to the equation

$$k = Ae^{\alpha\epsilon} \quad (5.5)$$

where A and α both decrease as the applied pressure increases. To understand this interplay, Mow and co-workers developed what has become known as the KLM biphasic theory of cartilage mechanics (Mow et al. 1980). This theory assumes that cartilage consists of a soft, porous, elastic solid, the pores of which communicate and are filled with water. The materials that compose both phases are assumed to be incompressible and the fluid to be inviscid. The details of this theory are beyond the scope of this book, but three examples are presented to show where it can and cannot describe the compressive viscoelastic behavior of cartilage.

5.2.6.1 Example 1, Confined Compression Creep Test

Mow et al. (1980) conducted slow creep tests of cylindrical specimens of cartilage using the scheme shown in Fig. 5.9. A cylindrical specimen of cross-sectional area A is cored from a joint surface and placed in a cylindrical metal well with solid walls and bottom. A porous metal loading platen is placed on top of the specimen. A constant load, F , is applied at time $t=0$ and the specimen deformation is recorded until an equilibrium deformation is reached. In this situation, creep is controlled by the exudation of fluid through the porous loading platen against the specimen's top surface. The movement of fluid to this interface is modeled to understand exudation. In the biphasic theory, movement of the fluid in the tissue is determined by the permeability, k , and the fluid pressure gradients of the specimen. The

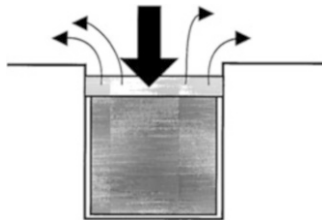


FIGURE 5.9. Schematic diagram of the experimental setup used for creep compression tests of cartilage. Fluid is allowed to flow through the porous loading platen.

equilibrium deformation is determined by the equilibrium, or aggregate, modulus of the specimen, H_A , and the applied loading. If the initial height of the specimen is h , and its displacement under stress $\sigma = F/A$ is $u(t)$, the biphasic theory solution for the ratio of the deformation to the initial specimen height as a function of time is

$$u/h = (\sigma/H_A) \left\{ 1 - 2 \sum_{n=0}^{\infty} \left[\exp\left(-\Lambda_n H_A k t / h^2\right) / \Lambda_n \right] \right\} \quad (5.6)$$

where t is time and $\Lambda_n = (n + 1/2)^2 \pi^2$ (Woo et al. 1987). Notice that as $t \rightarrow \infty$, this equation simplifies to $\sigma = H_A (u/h)$. Figure 5.10 and Table 5.3 show the results of fitting this equation to slow creep data for articular cartilage and

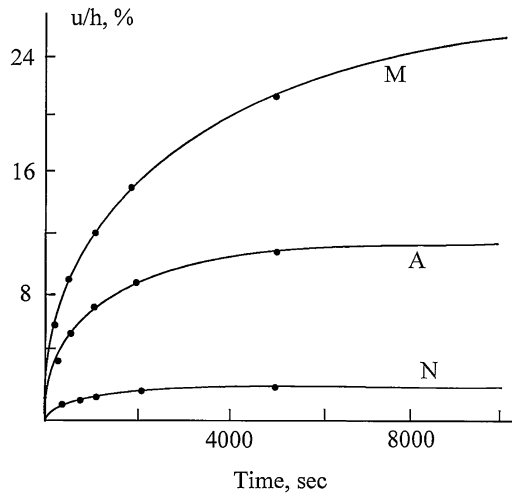


FIGURE 5.10. Strain (u/h) vs. time for articular cartilage (A), meniscus (M), and nasal cartilage (N) for slow loading. Previously published as Fig. 4.17b in Woo, S.L.-Y., Mow, V.C., and Lai, W.M. Biomechanical properties of articular cartilage. In *Handbook of Bioengineering* (edited by R. Skalak and S. Chien), 4.1–4.44. McGraw-Hill, New York; 1987.

TABLE 5.3. Compressive properties of three types of bovine cartilage

Tissue	Aggregate modulus (HA) (MPa)	Permeability (k), 10–15 m ⁴ /N s	H ₂ O by weight (%)
Nasal cartilage	5.64	0.49	75.6
Articular cartilage	0.85	4.67	80.9
Meniscus	0.41	0.81	73.9

From Woo et al. (1987)

two other kinds of cartilaginous tissue: nasal cartilage and knee meniscus. The distinct differences in the properties of these tissues are striking. Notice the inverse relationship between the aggregate modulus and the equilibrium deformation. Also, the cartilage with the greatest water content has the greatest permeability.

5.2.6.2 Example 2, Confined Compression Relaxation Test

In this experiment, the apparatus is the same as that depicted in Fig. 5.9 except that the instantaneous load application is replaced by a ramp displacement like that shown in Fig. 5.11. After a specified maximum displacement is reached, it is maintained while the reactive stress produced in the specimen is monitored. As Fig. 5.11 shows, there is a nonlinear increase in stress until maximum displacement is reached, then the stress “relaxes” to an equilibrium value. Fluid flows from the specimen early in the test (e.g., at A and B), but later the strain redistribution is accompanied entirely by internal fluid movements (e.g., at C and D).

The biphasic theory predicts the stress during both the compression and the relaxation phases of the experiment. The equation for the latter is

$$\sigma = -(H_A V_0 / h) \left\{ t_0 - 2 \sum_{n=1}^{\infty} [K_n^{-1} \exp(-K_n t) - 1] \right\} \quad (5.7)$$

where V_0 is the displacement rate, t_0 is the time at which the displacement is stopped (B in Fig. 5.11), and $K = n^2 \pi^2 k H_A / h^2$ (Woo et al. 1987). The other symbols retain their former definitions. Note again that as $t \rightarrow \infty$,

$$\sigma = H_A V_0 t_0 / h = H_A \epsilon \quad (5.8)$$

because $V_0 t_0 / h$ is the strain at the end of the slow loading period, ϵ . These equations fit experimental relaxation data quite well when t_0 is large enough

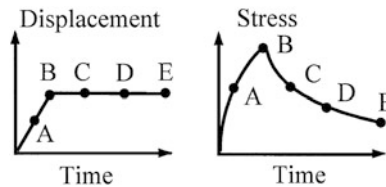


FIGURE 5.11. (Left) Displacement applied to cartilage specimen in compression relaxation test. (Right) Stress-time curve for the test. Letters A–E show equivalent positions on the two curves. Previously published as Fig. 6 in Mow, V.C., Kuei, S.C., Lai, W.M., and Armstrong, C.G. Biphasic creep and stress relaxation of articular cartilage in compression. Theory and experiment. *Journal of Biomechanical Engineering* 1980; 102:73–84 (Mow et al. 1980).

(on the order of a hundred seconds), and the values of H_A and k that achieve these fits agree well with those obtained from creep experiments. There are advantages to both methods of obtaining the aggregate modulus. The relaxation test takes less time to reach equilibrium, but it requires an apparatus to achieve a displacement curve like that shown in Fig. 5.11.

In the course of testing various kinds of specimens, it has been learned that the aggregate modulus of cartilage is inversely proportional to water content but increases in proportion to hyaluronic acid content. For example, diminishing the water content from 85 % to 75 % increases the modulus from about 0.3 to 1.0 MPa, and a 30 % increase in hyaluronic acid content can double the modulus. There is also a reduction in modulus with age, but this is not a predictive relationship because the individual variability at every age is very great.

5.2.6.3 Example 3, Unconfined Compression Relaxation Test

For rapid unconfined compression of a cylinder of cartilage modeled as a linear poroelastic cylinder of radius r filled with an inviscid fluid, the poroelastic solution for the stress is

$$\sigma(t) = \sigma_s \left[1 + \sum_{n=1}^{\infty} A_n(\nu) \exp\left(-\frac{\alpha_n H_a k t}{r^2}\right) \right] \quad (5.9)$$

where $\sigma_s (=E_s \varepsilon_s)$ is the stress after complete relaxation of the tissue, and the coefficients of the series are $A_n(\nu) = ((1-\nu)(1-2\nu)/(1+\nu)) (1/[(1-\nu)^2 \alpha_n^2 - (1-2\nu)])$, ν = Poisson's Ratio (between 0 and 0.5), α_n = characteristic equation roots, $H_a = E_s((1-\nu)/(1+\nu)(1-2\nu))$, $E_s(\varepsilon_s)$ = equilibrium Young's modulus (strain) of the material after complete relaxation, and k = permeability. The characteristic equation is $\vartheta_1(\xi) - ((1-\nu) \xi \vartheta_0(\xi)/(1-2\nu)) = 0$, where $\vartheta_0(\xi)$ and $\vartheta_1(\xi)$ are Bessel functions and α_n are the roots of the equation. This relaxation kernel is well known as the result of solving the poroelastic constitutive model for unconfined compression. As it turns out, 1.5 is the maximum size of the formula in the square brackets of Eq. 5.9 at time $t = 0$. That the *maximum* stress ($\sigma(0)$) can only be 1.5 times the equilibrium stress (σ_s) means that one should never observe peak stresses larger than this if the KLM biphasic theory holds true; however, it was observed many years ago (Armstrong et al. 1984) that the peak stresses during unconfined compression of cartilage far exceed the limit of 1.5 times the equilibrium stress, showing that the KLM biphasic theory is unable to accurately model the behavior of cartilage when loaded at rapid physiological rates.

5.2.7 *Other Models for Cartilage Mechanics*

5.2.7.1 **Nonlinearly Elastic Matrix**

Using a biphasic model with a matrix that has a phenomenological (i.e., empirical) mechanical stiffness that is high in tension and low in compression (Soltz and Ateshian 2000) can solve some of the limitations of the KLM model (i.e., its inability to predict the initial stress of an unconfined compression experiment). However, the use of this special nonlinearly elastic matrix still puts all of the tissue viscoelasticity into the fluid flow. Even though it will now be possible to reach a higher apparent stress during load, the relaxation spectrum is still too slow. The nonlinear material behavior can be modeled mechanistically (e.g. as being due to the fact that you can pull but not push on a fiber) or be simply introduced as a rule. In research a mechanistic model helps more with understanding how a tissue functions and, therefore, the former approach is preferred.

5.2.7.2 **Viscoelastic Matrix**

The key to a suitable model of early rapid stress relaxation is to use a rapidly relaxing matrix that is viscoelastic (Huang et al. 2001). The viscoelastic matrix is often introduced phenomenologically (i.e., the source of viscoelasticity is not explained) and can be combined with bi-, tri-, and other phasic fluid flow models. When a phenomenological viscoelastic material is used with a “phasic” theory the mechanisms of fluid flow and ionic interactions are approximately mechanistic, but the viscoelasticity remains unexplained. There have been small steps in the direction of mechanistic modeling of cartilage matrix viscoelasticity using polymer dynamics models (Fyhrie and Barone 2003; June et al. 2011). Polymer dynamics is a very complex field in itself (Boyd and Smith 2007) but in theory it can account for the rapid stress relaxation caused by biopolymer:biopolymer interactions in the tissue. The approach is promising and has been shown to be consistent with observations of cartilage stress relaxation in unconfined compression.

5.2.8 *Modeling Tissue Charge*

In addition to the mechanical properties of the collagen matrix and the effects of fluid flow, swelling effects due to the high density of fixed negative charges within the aggrecan molecule have been mathematically modeled. The electrostatic swelling pressure of the fixed charges can be accounted for by including the continuum Donnan equilibrium model to represent the effects of fixed charges (Lai et al. 1991). In the Donnan model, the microscopic features of the location and “grain” of the molecules are absent and only the averaged responses of the tissue are accounted for. A more granular approach is to use a micro-structural, molecular level solution of the electrostatic Poisson-Boltzmann equation in a unit cell of a charged glycosaminoglycan molecule and its surrounding atmosphere of mobile atoms to represent the

role of charge in the tissue (Buschmann and Grodzinsky 1995). The micro-modeling approach accounts for the space-varying electrical potential and electrical field among the constituents of cartilage. In a head-to-head comparison, the microstructural model accurately predicted observed experimental static swelling pressure, where the Donnan approach predicted swelling pressure several times larger than was observed experimentally (Buschmann and Grodzinsky 1995). The conclusion was that the elastostatic approach more accurately reflected the effects of GAG charge in producing mechanical swelling pressure compared to the Donnan model—at least for the static loading case.

For dynamic loading, the Donnan equation is the primary approach to modeling the effects of charge density on mechanical properties. Considering that the Donnan approach is not ideal for the static case, it would seem reasonable that a more mechanistic (i.e., non-continuum) approach to modeling dynamic behavior might be appropriate. Unfortunately, mechanistic dynamic micro-models accounting for molecular charge do not seem to exist. It is possible that polymer dynamics models accounting for free charge on the polymer molecules could be used as a starting point to model cartilage dynamics more mechanistically (Boroudjerdi et al. 2005).

5.2.9 *Inhomogeneity of Modulus*

Schinagl and co-workers (1996) loaded excised specimens of bovine articular cartilage in compression and then measured the displacements within the tissue as a function of depth below the articular surface. They allowed 40 min to elapse after applying the load to allow the displacement field within the cartilage to reach an equilibrium state. Displacements were then measured by using a microscope to record the positions of chondrocytes before and after application of the load. (The DNA in the cells was marked with a fluorescent dye to make them easier to locate.) The results showed that displacements were greater near the articular surface and diminished exponentially through the thickness of the cartilage. If all of the cartilage was fully relaxed after 40 min, the difference in strains by depth means that the modulus in the upper layers was lower than that in the depths of the cartilage. Alternatively, if the surface of the tissue relaxed viscoelastically faster than the deeper parts and 40 min wasn't enough time to reach equilibrium the result could mean that the tissue deformed more quickly at the top than the bottom. Schinagl et al. (1996) found experimentally that the compressive strain at 40 min of loading was essentially the same as that for 180 min of loading, so they rejected this alternate hypothesis and concluded that the equilibrium modulus was actually different by depth. They proposed that the layered moduli are functional, stating: "Upon loading, the surface layers of the articular cartilage deform, reducing pore space [decreasing permeability] and [slowing] the flow of fluid out of the tissue... maintaining hydrostatic

pressurization. Deformation of the relatively soft surface layer may also lead to improved congruence between opposing cartilage covered bones. This may have the effect of not only reducing stress concentrations in the joint, but also maintaining a fluid film between cartilage layers that lubricates loaded joints and reduces wear.”

Schinagl et al.’s (1996) proposal was interesting and possibly correct, but their data were for mechanical equilibrium and the proposal was for dynamic loading of cartilage at a physiological rate, so the connection between the proposal and data was not extremely strong. Silverberg et al. (2013) have demonstrated conclusively that the dynamic shear modulus of bovine cartilage (1 Hz surface shear) has depth dependence with the surface more compliant than the deeper tissue and that regularly loaded regions are more compliant at the surface than rarely loaded regions of the tissue. This result is consistent with and expands the original proposal by Schinagl et al. (1996) and suggests that the depth dependence of dynamic cartilage moduli is important to the maintenance of cartilage lubrication and pressurization during physiological loading.

5.2.10 Effect of Strain Rate on Tensile Tests

Figure 5.12 shows a graph of the effects of varying strain rate on the stress-strain relationship of bovine articular cartilage tested in tension (Woo et al. 1987). As with bone, tendon and ligament, the stiffness of cartilage increases as it is deformed more quickly. Notice that the curve for the fastest strain rate has an inflection point, but the others do not. This phenomenon is thought to be caused by resistance to fluid flow. At slow strain rates, there is time for fluid to flow without offering much resistance to the applied deformation. At higher strain rates, fluid pressure resists the loading early in the test. This pressure also expresses fluid from the tissue, reducing its participation in the production of stress and allowing the slope of the stress-strain curve to decline (i.e., to inflect). The biphasic theory produced the following equation to describe this kind of tensile test:

$$\sigma = E_s \varepsilon + \left(h^2 d\varepsilon/dt \right) / (H_A k) \left[A_0 - \sum \left(A_n \exp \left(-4\alpha_n^2 H_A k t / h^2 \right) / \alpha_n^2 \right) \right] \quad (5.10)$$

where ε is strain and $d\varepsilon/dt$ is strain rate, E_s is the elastic modulus of the solid matrix, and A_n and α_n are constants. The summation and other symbols are as in Eqs. 5.6 and 5.7. This result describes only the low strain portion of Fig. 5.14 (i.e., $\varepsilon < 10\%$) because the biphasic theory assumes infinitesimal strains and linearity. At higher strains the collagen fibers have been stretched to the point that their participation in resisting the applied deformation is increasingly important. It is postulated that the three curves have similar slopes at their right ends because by then the stretched-out collagen fibers are

carrying virtually the entire load, and fluid effects are insignificant. To date tensile results such as these have not been examined as polymer mechanics or electrodynamics problems. It is unlikely, however, that the viscoelasticity of cartilage in tension results solely from fluid flow.

5.2.11 Effect of Strain Rate on Compression Tests

The tensile elastic modulus of the specimens in Fig. 5.12 is in the neighborhood of 1–10 MPa, depending on strain magnitude and rate. This value is similar to the aggregate modulus for compression of articular cartilage at low strain rates. There is evidence that the compressive modulus is at least an order of magnitude higher when dynamic loading is involved. Figure 5.13 shows how this modulus (measured at 10 % strain) increases as a function of strain rate. Even higher moduli have been recorded for the dynamic modulus of cartilage when the dynamic loading is superimposed on creep strain. (Imagine the cartilage in your knees creeping into a compressed state while you stand waiting for a traffic light to change, then being dynamically loaded as you step off the curb.) Under these conditions, the cartilage behaves as a linear elastic material with a modulus of 100–500 MPa, depending on the amount of creep strain (Higginson and Snaith 1979).

In summary, the mechanical properties of cartilage are partly determined by fluid flow through the collagen–proteoglycan matrix but models employing fluid flow through a porous elastic solid only work at very slow loading rates and fail to predict behavior at higher, more physiological rates of loading. Thus, tissue permeability is an important factor for slower loading rates but other less understood contributors to viscoelasticity, e.g. biopolymer interactions, also play an important role in the material behavior of cartilage. The relevance of viscoelasticity and fluid flow becomes

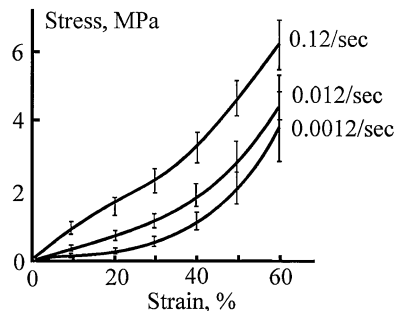


FIGURE 5.12. Stress-strain curves for cartilage at three different strain rates. Previously published as Fig. 4.24b in Woo, S.L.-Y., Mow, V.C., and Lai, W.M. Biomechanical properties of articular cartilage. In Handbook of Bioengineering (edited by R. Skalak and S. Chien), 4.1–4.44. McGraw-Hill, New York; 1987 (Woo et al. 1987).

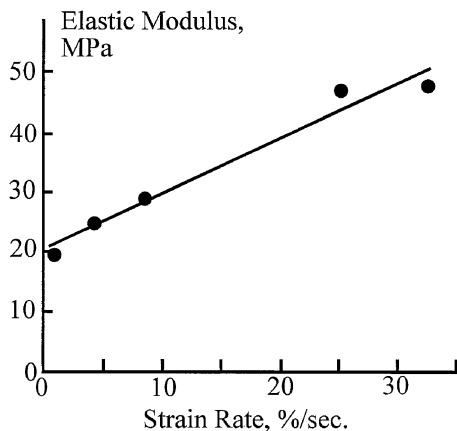


FIGURE 5.13. Compressive elastic modulus vs. strain rate for bovine articular cartilage (Radin et al. 1970).

clearer in the next section. Unlike most engineered bearings, synovial joints do not involve lubricants acting on rigid, inert surfaces. Joint lubricants are composed of water and hydrated polymers that are small enough to flow into and out of the cartilaginous bearing surfaces and larger molecules that are sieved and bound to joint surfaces. As a further complication, the bearing surfaces themselves deform significantly to accommodate both load and lubrication.

5.3 Lubrication of Joints

Synovial joints have extremely low friction coefficients, about equal to an ice skate on ice. Just how this happens is not fully understood, but much can be deduced on the basis of research on joints themselves and on engineering lubrication theory, which has a long and interesting history.

5.3.1 Friction

Friction is produced between sliding surfaces because they cannot be perfectly smooth, but have projections that interlock (Fig. 5.14). Frictional forces arise from interdigitation of the asperities on the two surfaces. The compressive force between the sliding surfaces must be overcome to enable the interdigitating asperities to slide up and over one another. In addition, adhesive forces may exist between the two surfaces.

Thus, to slide more easily past one another, the two surfaces must be separated slightly against the action of the compressive force. The primary

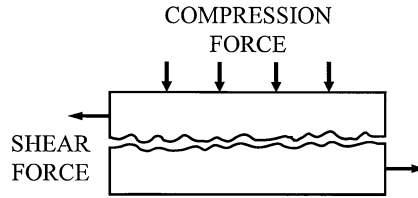


FIGURE 5.14. Representation of two surfaces sliding on one another. The required shear force must exceed the frictional force resulting from interactions between the asperities of the two surfaces.

purpose of lubrication is to reduce interlocking and adhesion between the asperities on the two surfaces. The surfaces of articular cartilage seem to have ridges not unlike those in the figure, at a microscopic level. The roughness of these surfaces has been observed to increase with aging and joint disease. However, drying of the tissue during preparation for microscopy may exaggerate such features, and there is currently uncertainty about the magnitude of the surface ridges on cartilage.

The two basic laws of friction are known as Amontons' laws²

- The friction force is directly proportional to the applied load.
- The friction force is independent of the apparent area of contact.

The friction force is defined as the shear force required to pull one surface across another. The friction coefficient, μ , is just the ratio of the friction force, F_F , to the applied load, or normal force, N , which pushes the surfaces together:

$$F_F = \mu N \quad (5.11)$$

Static friction coefficients usually are greater than sliding friction coefficients. Typically, friction coefficients between unlubricated surfaces are about 0.3 and those in lubricated machinery bearings are 0.05 or less (Table 5.4). Friction coefficients in synovial joints are among the smallest ever measured, about 0.001–0.005.

The reason that the friction force is independent of the bearing surface area is that when the area increases, the number of barriers to sliding increases, but the compressive force across each one decreases. Therefore, making an articular surface larger reduces the continuum level or macroscopic stresses in the tissues of the joint, but in principle it does not affect the frictional force.

² The French scientist Guillaume Amontons presented his classic paper on this subject to the Academie Royale in December of 1699.

TABLE 5.4. Typical friction coefficients

Materials and conditions	μ
Rubber on concrete, wet or dry, static	1.0
Brake material on cast iron, clean and dry, static	0.4
Brake material on cast iron, lubricated with mineral oil, static	0.1
Graphite on steel, static	0.1
Hickory on dry snow, waxed, 4 m/s, -3°C , dynamic	0.18
Hickory on dry snow, unwaxed, 4 m/s, -3°C , dynamic	0.08
Ice on ice, 4 m/s, 0°C , dynamic	0.02
Articular cartilage in human joints, dynamic:	
Human knee; Charnley (1960)	0.005–0.023
Porcine shoulder; McCutchen (1962)	0.02–0.35
Canine ankle; Linn (1967, 1968)	0.005–0.01
Unsworth et al. (1975)	0.01–0.04
Malcom (1976)	0.002–0.03

Values from CRC Handbook of Chemistry and Physics (1984) and Mow and Soslowsky (1991)

5.3.2 Wear

Wear occurs in a bearing because projections become broken by crashing into one another or tear off because of adhesion, or because foreign bodies create additional “mountains” that temporarily lodge on one surface and gouge the opposite surface. These phenomena occur in synovial joints as well, causing the cartilage surfaces to wear away over time. Foreign bodies play little part in wear of normal joints, but can play a major role in prosthetic joints, where wear particles become foreign bodies and accelerate the wear problem. They can also migrate to the periphery of the joint, contact the synovial membrane, and cause an inflammatory response that induces chondrocytes to produce catabolic enzymes that digest the cartilage. This response may also cause bone resorption and pain. Reduced friction is associated with reduced wear (Table 5.4).

5.3.3 Types of Lubrication

There are several different mechanisms by which lubrication can be achieved. Each of the following mechanisms is thought to play a role in the lubrication of synovial joints.

5.3.3.1 Boundary Lubrication

“All things and everything whatsoever however thin it be which is interposed in the middle between objects that rub together lighten the difficulty of this friction.”

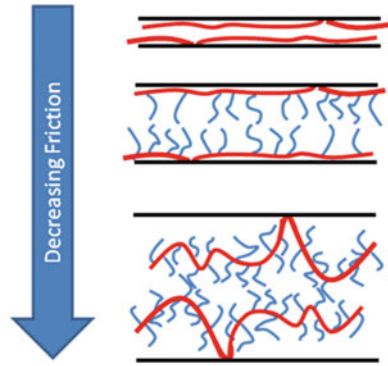


FIGURE 5.15. Polymer lubricants (red and blue) between cartilage surfaces (black lines) greatly reduce friction by trapping a water film (water molecules not drawn).

So said Leonardo da Vinci, and while this translation may not win a prize for punctuation, it is perhaps as good a description as any of the idea behind boundary lubrication. Placing a “boundary” layer (i.e., an extremely thin layer, perhaps one molecule thick) of almost any liquid-like material between the bearing surfaces holds the surface projections slightly apart, reduces adhesion, and keeps the projections from interlocking (Fig. 5.15).

5.3.3.2 Squeeze Film Lubrication

Another means to protect the joint surfaces is to separate them by creation of a pressurized layer of fluid. One way to achieve this is as follows. When you run and your leg swings forward, its knee joint is minimally loaded, allowing the cartilage surfaces to separate and minimizing the friction. When this leg enters its stance phase, body weight pushes the femoral condyles down through the synovial fluid toward the tibial side of the joint. When contact is made, frictional forces greatly increase. The viscosity of the fluid resists the descending femur and increases the amount of time required for the two sides of the joint to meet. In the meantime, friction remains low. The bearing surface separation produced by such load-and-unload cycles is called squeeze film lubrication.

A slightly different but related mechanism for pressurization in joints that improves lubrication is the tough seal formed by the collagenous labrum that encloses articulating surfaces (Haemer et al. 2012). The labrum prevents fluid from being pushed out when the joint is loaded. If the labral seal is tight, then the joint fluid will be pressurized during loading and will reduce the friction in the joint. Squeeze film lubrication doesn’t depend on the anatomy of the whole joint but on having the lubricant trapped between the surfaces, so it isn’t affected if the labrum and/or meniscus are damaged. It is not known at this time whether lack of meniscal sealing contributes to the high incidence of

osteoarthritis of the knee after damage to the menisci but it is reasonable to speculate that damage to the joint seal could play a role in increased friction and later development of OA.

5.3.3.3 Hydrodynamic Lubrication

A way to achieve prolonged fluid pressure to hold up body weight and separate the joint surfaces is to use parallel rather than perpendicular relative motion between them, that is, sliding motion of one surface across the other (Fig. 5.16). The fluid next to each surface sticks to it. As one surface moves across the other, friction between adjacent fluid molecules drags the intervening fluid into the space between them, creating hydrodynamic pressure that forces them very slightly, but completely, apart. The friction force is essentially only that between the fluid molecules. For this kind of lubrication to work, the bearing or joint must be moving fairly rapidly. Thus, it can function in synovial joints only during flexion or extension motions.

Hydrodynamic lubrication was discovered by Beauchamp Towers when he was asked in 1882 by a royal commission in England to study the problem of friction in machinery (Dowson 1979). Soon after, in 1886, Osbourne Reynolds published a theory explaining how hydrodynamic lubrication worked. While the application of this research was to be in the field of conventional engineering, Reynolds ended his classic paper with the speculation that hydrodynamic lubrication was important in biologic as well as engineering bearings (Reynolds 1886). Reynolds squarely identified the issue in joint lubrication that remains uncertain today: how can a reciprocating bearing maintain a state of hydrodynamic lubrication?

The question as to whether this action can be continuous or not turns on whether the action tends to preserve the matter between the surfaces at the points of pressure as in the apparently unique case of the revolving journal, or tends to sweep it to one side as is the result of all backwards and forward rubbing with continuous pressure.

The fact that a little grease will enable almost any surfaces to slide for a time has tended doubtless to obscure the action of the revolving journal to



FIGURE 5.16. Hydrodynamic lubrication results in complete separation of the surfaces.

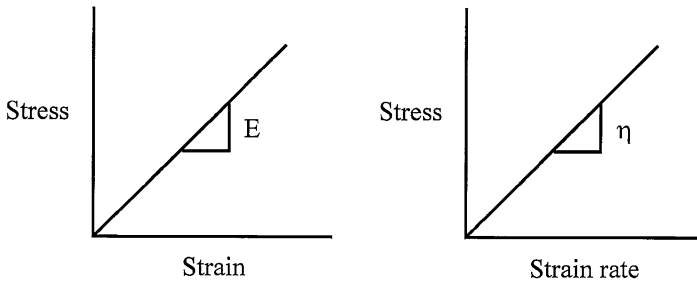


FIGURE 5.17. The correspondence between elastic modulus in solids and viscosity in fluids.

maintain the oil between the surfaces at the point of pressure. And yet, although only now understood, it is this action that has alone rendered our machines and even our carriages possible. The only other self-acting system of lubrication is that of reciprocating joints with alternate pressure on the separation . . . of the surfaces. This plays an important part in certain machines, as in the steam engine, and is as fundamental to animal mechanics as the lubrication action of the journal is to mechanical contrivances.

It is clear that Reynolds not only foresaw the possibility that hydrodynamic lubrication occurs in synovial joints, but also that there would be difficulty in explaining how the system could work if it involved “rubbing with continuous pressure.”

To gain a more tangible feel for the criteria that must be met if a joint is to be hydrodynamically lubricated, we shall derive Reynolds’ equation and use it in an example problem. The first step in this process is to recall the definition of viscosity. In an elastic solid, stress is linearly proportional to strain and the slope of the stress-strain curve is the elastic modulus. In a Newtonian fluid, stress is linearly proportional to strain rate, and the slope of the line is called viscosity (Fig. 5.17). If the stress-strain rate relationship is nonlinear, the fluid is non-Newtonian. If the slope of the curve decreases with increasing strain rate, the fluid is said to be thixotropic. It will be seen that synovial fluid is thixotropic, and that this is advantageous for joint lubrication.

We consider a simplified joint such as that shown in Fig. 5.16. We assume that there is a flow of fluid between the joint surfaces and that:

- The joint surfaces are impermeable to the fluid.
- The fluid is incompressible and Newtonian.
- The fluid’s viscosity is homogeneous and isotropic.
- The pressure gradient in the vertical direction is much less than that in the flow direction.

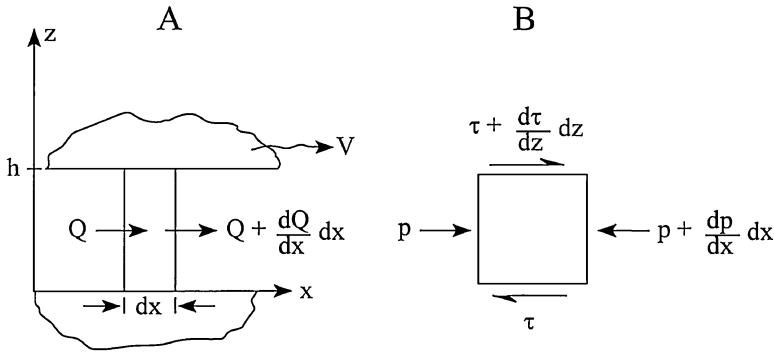


FIGURE 5.18. Diagrams for derivation of Reynolds' equation. (a) A fixed column of fluid with thickness dx defined within the space between the joint surfaces. The *upper surface* is moving with velocity V relative to the *lower surface*. (b) Free-body diagram of a small cube of fluid within the column in (a).

First, we obtain a continuity of flow condition. Consider the column of fluid of width dx sketched in Fig. 5.18a. Let Q be the flow rate (per unit width of the joint space) into the left side of the column, and $Q + dQ/dx$ be the flow rate out the right side. Because we do not want any fluid sources or sinks in the column, we require that the flow in equals the flow out, and this reduces to

$$dQ/dx = 0 \quad (5.12)$$

Now consider the balance of forces on a cube within the column (Fig. 5.18b). Assuming the flow is associated with a pressure gradient in the x -direction, the force on the left side is $p \, dydz$ and that on the right side is $(p + dp/dx) \, dydz$. There are also shear stresses on the top and bottom surfaces of the cube because one side of the joint is moving relative to the other. If the shear stress is τ , then the shear force on the bottom (acting to the left) is $\tau \, dx dy$ and that on the top (acting to the right) is $(\tau + d\tau/dz) \, dx dy$. Equating the forces in each direction gives

$$p \, dydz + (\tau + d\tau/dz) \, dx dy = (p + dp/dx) \, dydz + \tau \, dx dy \quad (5.13)$$

which reduces to

$$dp/dx = d\tau/dz \quad (5.14)$$

If the fluid is Newtonian, then the shear stress is related to the velocity gradient in the z -direction by the viscosity, η , so that

$$dp/dx = d/dz(\eta dv/dz) \quad (5.15)$$

where v is velocity. Because the pressure, p , is not a function of z , we may integrate this equation to obtain

$$\eta dv/dz = z dp/dx + C_1 \quad (5.16)$$

where C_1 is the constant of integration. Assuming, as we have, that η is not a function of z either, we integrate again to get

$$\eta v = z^2 dp/dx + C_1 z + C_2 \quad (5.17)$$

where C_2 is a second integration constant. To find C_1 and C_2 , we use the conditions $v=0$ at $z=0$ and $v=V$ at $z=h$, where $h(x)$ is the distance between the joint surfaces. This gives

$$C_1 = \eta V/h - (h/2) dp/dx \text{ and } C_2 = 0$$

Substituting these into Eq. 5.17, one has

$$v = [(z^2 - zh)/2\eta] dp/dx + Vz/h \quad (5.18)$$

The gradient of this velocity in the z -direction is

$$dv/dz = (z - h/2)/\eta dp/dx + V/h \quad (5.19)$$

Integrating the flow velocity over the joint separation, h , gives the flow rate Q in the x -direction per unit width of the joint in the y -direction:

$$Q = -\left(h^3/12\eta\right) dp/dx + Vh/2 \quad (5.20)$$

Substituting this into the continuity equation, Eq. 5.12, we obtain

$$d/dx \left[\left(h^3/12\eta\right) dp/dx \right] = (V/2) dh/dx \quad (5.21)$$

which can be integrated with respect to x to yield

$$h^3 dp/dx = 6V\eta h + C_3 \quad (5.22)$$

where C_3 is the integration constant. To find a value for this constant, we assume that there is some point along x where the pressure separating the two surfaces is a maximum. At that value of x , $dp/dx = 0$. Letting $h = h^*$ there, we find that $C_3 = -6V\eta h^*$. Making this substitution, we finally obtain the one-dimensional version of Reynolds' equation:

$$dp/dx = 6V\eta(h - h^*)/h^3 \quad (5.23)$$

In this equation, V is the relative velocity of the two bearing surfaces, h is the distance between them, and h^* is the value of h where p has its maximum value. Equation 5.23 captures the fact that hydrodynamic lubrication depends on the product of three factors: the viscosity of the lubricant, η , the velocity of one joint surface relative to the other, V , and variations in the lubricant film thickness, h .

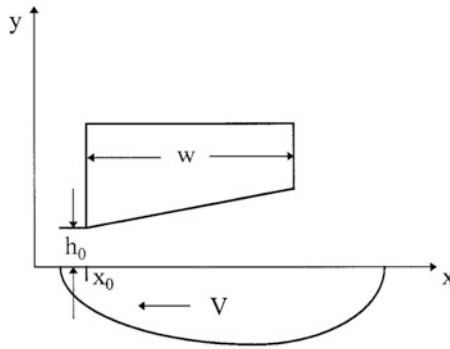


FIGURE 5.19. A flat lower surface moves to the left with velocity V relative to a fixed, canted upper surface having width w and slope s .

Example Problem To illustrate the use of Reynolds' equation and to test the feasibility of hydrodynamic lubrication between articular surfaces, a simple problem will be solved. Suppose the two bearing surfaces of a joint may be approximated by planes, the upper one of which is fixed in space at a slight angle to the other (Fig. 5.19). An x - y coordinate system is established as shown. The lower surface slides to the left with velocity $-V$. The value of x at the left end of the upper surface is x_0 and that at the right end is $x_0 + w$. The slope of the upper surface is s and the minimal distance between the two surfaces is h_0 . Both s and h_0 are very small. In general, the distance h between the surfaces is given by $h = sx$; therefore, $dx = dh/s$. Making this substitution, Reynolds' equation becomes

$$dp = -6V\eta s^{-1} \left(h - h^*/h^3 \right) dh \quad (5.24)$$

Integrating, one has

$$dp = -6V\eta s^{-1} \left(-(1/h) + (h^*/2h^2) + C_4 \right) \quad (5.25)$$

where C_4 is another integration constant. C_4 and h^* may be found using the following boundary conditions. To the left and right of the region between the two surfaces, the pressure is assumed to be zero. Thus,

$$\text{At } x = x_0, \quad h = h_0, \text{ and } P = 0,$$

$$\text{At } x = x_0 + w, \quad h = h_1, \text{ and } P = 0,$$

Making these substitutions and solving the resulting pair of equations yields

$$h^* = 2h_0h_1/(h_0 + h_1) \quad \text{and} \quad C_4 = 1/(h_0 + h_1)$$

When these equations are substituted into Eq. 5.25 and the leading minus sign is taken inside the square brackets, one has

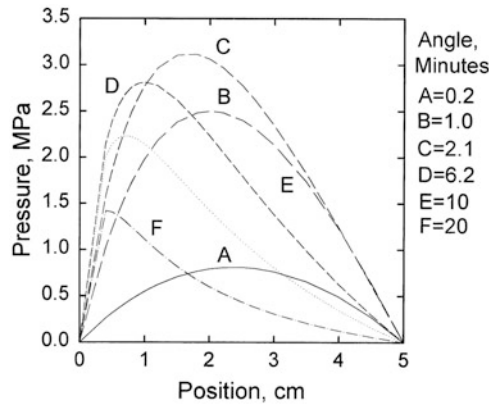


FIGURE 5.20. Effect of cant angle on pressure distribution for hydrodynamic lubrication of the system in Fig. 5.19.

$$p = 6\eta V[(1/h) - h_0 h_1 / (h_0 + h_1) b^2 - 1 / (h_0 + h_1)] / s \quad (5.26)$$

By substituting $h = sx$ into this equation and doing some algebraic manipulation, it can be put in the form

$$p = 6\eta V(sx - h_0)(h_1 - sx) / [(h_0 + h_1)s^3 x^2] \quad (5.27)$$

This equation clearly shows that the pressure between the bearing surfaces is a parabolic function of x with zeros at each end (where $sx = h_0$ and h_1 , respectively).

Figure 5.20 shows the pressure p as a function of position between the two bearing surfaces and the cant angle between them for the following parameters:

$$V = 0.25 \text{ m/s} \quad \eta = 0.05 \text{ Pa} \cdot \text{s} \quad h_0 = 5 \mu\text{m} \quad w = 5 \text{ cm}$$

Note first that the pressures are the same order of magnitude as the contact pressures between joint surfaces, suggesting that hydrodynamic lubrication may indeed be able to separate the surfaces under physiologic load conditions. Note also the extremely small cant angle that is required for hydrodynamic lubrication to be effective. Separation pressures drop off rapidly if the angle is less than a half minute or more than half a degree. This kind of behavior is largely independent of the bearing surface geometry.

The integral of p across the joint (from left to right in Figs. 5.19 and 5.20) must equal the joint force if the separation between the two joint surfaces is to be maintained. Equation 5.22 shows that this integral (the joint force) must equal the product of η , V , and a geometric function of the interbearing fluid space. Another way of saying this is that a load-normalized “visco-velocity factor,”

$$\frac{\text{viscosity} \times \text{velocity}}{\text{joint force}}$$

should be constant for a given joint geometry.

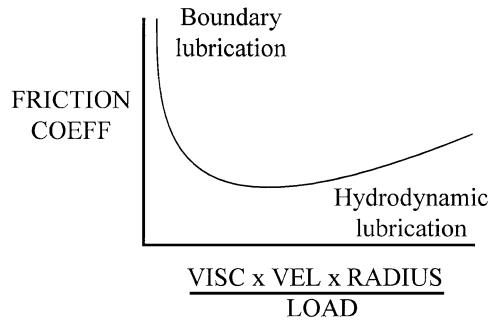


FIGURE 5.21. Schematic diagram of the effect of the viscovelocity factor on the measured friction coefficient.

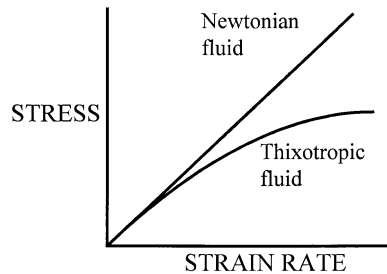


FIGURE 5.22. In a thixotropic fluid, the slope of the stress-strain rate curve decreases as the strain rate increases.

Plotting the friction coefficient (i.e., the ratio of frictional force to joint force) vs. this ratio for bearings under various conditions gives Fig. 5.21. At low velocities (or viscosities, or high loads), friction is relatively high because lubrication is by boundary or “thin-film” lubrication. As the velocity increases, hydrodynamic lubrication begins and friction falls. Then, however, as velocity continues to increase, friction begins to rise again because of “internal friction” in the lubricant. As the strain rate in a Newtonian lubricating fluid increases, the shear stress increases as well, increasing the frictional force. Thus, such lubricants limit the remarkable advantage obtained from the hydrodynamic lubrication mechanism.

To prevent the deterioration of lubrication efficiency at high velocities, one may take advantage of the fact that some liquids containing long, chainlike molecules have a peculiar kind of viscosity that decreases as the rate of shearing increases (Fig. 5.22). If such a thixotropic liquid is used as the lubricant, the friction force does not rise so fast at high sliding speeds. It turns out that healthy synovial fluid is a very thixotropic lubricant. Furthermore, measurements have shown that the degree of thixotropy is decreased with age and by various joint diseases (Fig. 5.23). This result suggests that the efficiency of lubrication, and thus the avoidance of wear, declines with age and the appearance of disease. Whether the diminishment of thixotropy is a primary cause of joint degeneration is unknown, but it would seem to be a contributing factor.

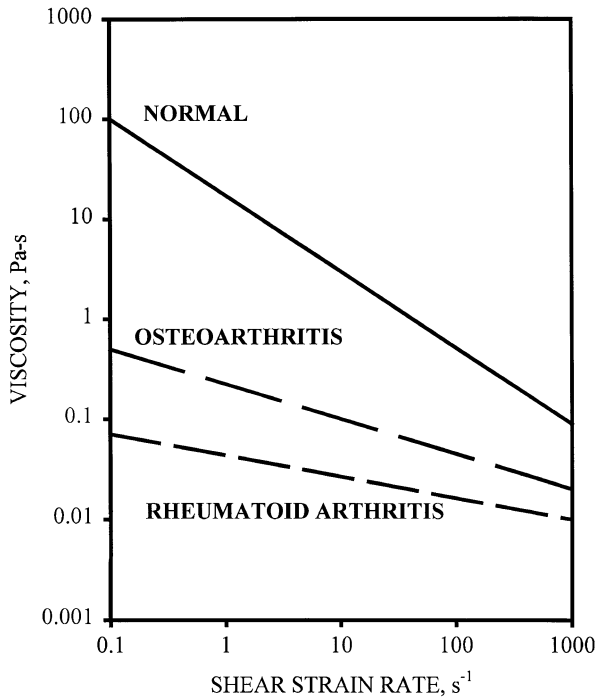


FIGURE 5.23. Synovial fluid viscosity vs. strain rate for representative normal, osteoarthritic, and rheumatoid arthritis joints. Water would be a horizontal line at 1.8×10^{-3} Pa s (Wright and Dowson 1976).

5.3.3.4 Elastohydrodynamic Lubrication

This term refers to situations in which the bearing surfaces are compliant enough (i.e., have a low elastic modulus) so that they deform significantly under the pressures produced in the lubricating fluid as it passes over them. It seems clear that this applies to articular cartilage, but the difficulty in mathematically modeling the situation has prevented a credible demonstration of elastohydrodynamic effects in synovial joints.

5.3.3.5 Weeping Lubrication

The source of the water that fills the pores of cartilage is the synovial fluid in the adjacent joint space. However, the hyaluronic acid molecules are too large to pass into the pore spaces. Therefore, the fluid that is contained in the pores of cartilage is actually a filtrate of synovial fluid. Lewis and McCutchen (1959) postulated that, just as this fluid is expressed from a cartilage specimen during a compression test, it is also exuded into the load-bearing region when two joint surfaces are pressed together during joint flexion. This flow of fluid under pressure was hypothesized to contribute to separation of the joint

surfaces. As in the case of elastohydrodynamic lubrication, it has been difficult to demonstrate convincingly that such an effect really occurs.

5.3.3.6 Boosted Lubrication

This theory of joint lubrication was proposed independently by Maroudas (Maroudas 1967; Walker et al. 1968). The idea is quite simple. The large hyaluronic acid molecules of the synovial fluid that get left behind when synovial fluid is transported into the cartilage is hypothesized to form a gel-like layer on the articular surface. This “concentrated lubricant” is postulated to play a major role in the lubrication process. The problem is to understand where and when on the joint surface this happens. Weeping lubrication claims that fluid flows out of the loaded surface, not in. - Squeeze-film lubrication has the synovial fluid being squeezed out to each side of the contact region. The proponents of boosted lubrication claim, on the other hand, that because of the microscopic ridges and grooves on the surface, when the two cartilage surfaces are pressed together, pools of synovial fluid are trapped and the water is forced into rather than out of the cartilage. These conflicting ideas have generated intense debate.

5.3.4 Synovial Joint Lubrication

Before trying to draw some conclusions about which of these lubrication mechanisms is most important in synovial joint lubrication, it is useful to first explore some of the early research in this area.

5.3.4.1 Pendulum Experiments

After Reynolds’ comments, there seems to be no discussion of hydrodynamic lubrication in joints again until MacConaill proposed it in 1932. A few years later, simple experiments were done in which joints were set to swinging by attaching one limb segment to a pendulum, and the rate of decay of the swinging motion was studied (Jones 1934, 1936). The fact that the amplitude of the motion decayed nonlinearly led to the conclusion that viscous damping was at play, providing support for the theory that a layer of fluid separated the joint surfaces, and hydrodynamic lubrication existed in the joint. Many years later, Charnley (1960) pointed out that the nonlinear decay could be for another reason: that the soft tissues about the joint resist large motions more than small motions. Soon, Barnett and Cobbold (1962) did an experiment that showed that this indeed happens (Fig. 5.24). A dog’s ankle joint was rigged as the fulcrum of a pendulum, showing that the resistance to motion became less, and the decay in motion more linear, when the adjacent soft tissues were removed. However, to properly interpret this experiment, it needs to be understood that the joint was rigged “upside down” so that the pendulum weight compressed the joint surfaces together continuously. Thus,

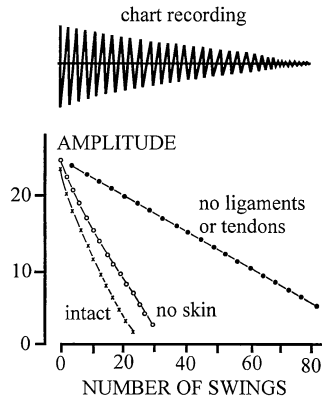


FIGURE 5.24. Experimental results from an early pendulum test of synovial joint lubrication. The *upper trace* shows oscillation decaying over time. The *lower graph* shows the nonlinearity of the decay was a function of the adjacent soft tissues. Figure is a composite of Figs. 4 and 5, previously published in Barnett, C.H., and Cobbold, A.F. Lubrication within living joints. *Journal of Bone and Joint Surgery* 1962; 44B:662–674 (Barnett and Cobbold 1962).

this experiment did not examine the possibility of short-term hydrodynamic lubrication following an unloaded state.

Another pendulum experiment suggests that hydrodynamic lubrication does occur in joints after all if the joint load is not too great. For boundary lubrication, the friction coefficient, μ , is independent of the joint contact force, N . When the joint surfaces are separated by a layer of fluid, as in hydrodynamic lubrication, the friction coefficient will decrease as the joint load increases because of the relationship $\mu = F/N$ and the fact that, in this case, the frictional force, F , is largely independent of the normal force, N . Figure 5.25 shows how the friction coefficient for the distal interphalangeal joint of a human finger decreases with increasing force across the joint (expressed as mass), strongly suggesting that the lubrication mechanism is not simply boundary lubrication. However, the plateauing of the curve when the force exceeds 900 g indicates that sufficient joint force can reduce the fluid layer between the articular surfaces to a thin film (i.e., boundary lubrication).

5.3.4.2 How Much Separation Is Needed?

Dintenfass (1963a, b) argued that elastohydrodynamic lubrication was probably occurring because the required pressures would deform cartilage significantly relative to the fluid film thickness. The minimal fluid layer necessary to produce hydrodynamic lubrication should be greater than the surface roughness of cartilage, which has been measured by several investigators at 1–6 μm (Dintenfass 1963b). Several analyses of elastohydrodynamic

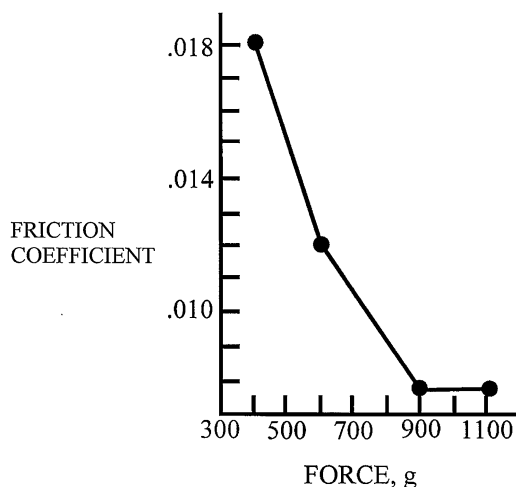


FIGURE 5.25. Friction coefficient vs. force across the joint for a human finger joint. Taken from Fig. 15, previously published in Barnett, C.H., and Cobbold, A.F. Lubrication within living joints. *Journal of Bone and Joint Surgery* 1962; 44B:662–674 (Barnett and Cobbold 1962).

lubrication as a mechanism for joint lubrication have yielded fluid film thickness of less than 1 μm , however (Mow and Soslowsky 1991). To circumvent this objection, Dowson and Jin (1986) argued that the physiologic pressures produced in joints would flatten the unevenness of the cartilage surface, permitting “microelastohydrodynamic lubrication” to work, but this idea has not been widely accepted. More recently, attempts to refine the measurement of surface roughness of cartilage have suggested that the *in vivo* roughness is much less than the measured on dried specimens, and that the measurement probe may “cut through” a superficial layer of smooth material and register instead the subsurface structure (e.g., proteoglycan and collagen fiber bundles). Thus, just plain hydrodynamic lubrication may be an adequate theory after all.

5.3.4.3 What Are the Lubricating Ingredients?

The principal argument against hydrodynamic lubrication in joints, and in favor of boundary lubrication, is the obvious fact that these bearing surfaces are not in continuous motion. Indeed, they are stationary much of the time and, it could be argued, often move slowest when carrying the highest loads. Experiments conducted so that only boundary lubrication could operate show remarkably low friction coefficients for cartilage on cartilage, supporting such a hypothesis. It has been suggested that the boundary lubricant could be (1) lipids in the cartilage (because fat solvents increase the friction coefficient), (2) a hyaluronan gel produced by synovial fibroblasts

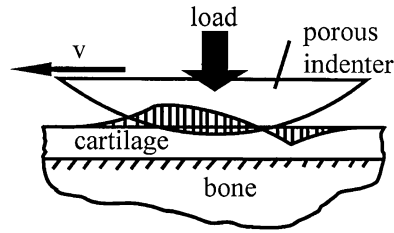


FIGURE 5.26. Flow out of the cartilage in front of an advancing porous indenter (at *left*) and back into the cartilage behind it (at *right*). Previously published as Fig. 13 in Van C. Mow, V.C., Lai M.W. *Recent Developments in Synovial Joint Biomechanics*. SIAM Review 1980; 22:3, 275–317 (Mow and Lai 1980). Copyright ©1980 Society for Industrial and Applied Mathematics. Reprinted with permission. All rights reserved.

and precipitated from synovial fluid, or (3) a lubricating glycoprotein called lubricin (Little et al. 1969; Swann et al. 1981) that is found on both synovial lining and the cartilage surface (Jay et al. 2000). Later, a phospholipid called dipalmitoyl phosphatidylcholine (DPPC) was put forth as the active boundary lubricant in synovial joints (Hills 1989).

5.3.4.4 Back to the Poroelastic Theory

Some of the many questions about the role played in lubrication by the movement of fluid into and out of the cartilage have been clarified by extension of the biphasic theory to the problem of lubrication, as opposed to the mechanical properties of cartilage (Mow and Soslowsky 1991). Analysis indicates that when a parabolically shaped porous indenter is slid across a layer of cartilage at physiologic speeds, fluid is expressed from the cartilage into the leading-edge region and fluid is imbibed back into the cartilage near the trailing edge (Fig. 5.26). Fluid “weeping” from the cartilage begins well in advance of the arrival of the load, and as the load increases, so does the flow. Then, elastic recoil of the cartilage allows fluid to reenter the tissue when the load is past, presumably leaving a layer of hyaluronic acid gel on the surface. Thus, the theory predicts that both the weeping and boosted lubrication theories have a role to play, as do boundary and hydrodynamic lubrication. It must be understood, however, that the values of many of the variables in the theoretical model are poorly defined, and much remains to be done in taking full advantage of the potential of the model to elucidate the etiology of various joint diseases.

5.3.5 Diversity of Joint Architecture

In concluding this chapter, we should take care to point out that there is great variety in the anatomy of synovial joints (Adams 1993). The implications of this with respect to lubrication are unclear, but it certainly raises the prospect

that different combinations of the proposed lubrication mechanisms may predominate in different kinds of synovial joints. To make this point more tangible, we briefly describe four human synovial joints having very different anatomical conformations. Often, there seems to be a trade-off between the need for range of motion and the need for stability under load.

5.3.5.1 Temporomandibular Joint

Motions of the mandible during chewing and speaking are extremely complex, with up-down, side-to-side, and in-out motions that must be regulated with great precision. Moreover, the two temporomandibular joints are linked by the mandible and the skull, so that motion in one joint produces motion in the other joint. It is clear that a highly constrained joint would be inadequate for these needs. The articular surface of the mandible does not directly bear on the articular surface of the temporal bone; instead elastic fibrous articular disks (menisci), 2- to 3-mm thick, lie between them. Rolling and sliding of the mandible on the disk, and the disk on the temporal bone, allow considerable freedom of motion, and it is likely that the presence of the disk is the evolutionary solution to reduce shear forces on the cartilage surface in a joint that slides. Stability is provided primarily by muscles and to a lesser extent by ligamentous and bony restraints.

5.3.5.2 Trapeziometacarpal Joint of the Thumb

The base of the thumb articulates with two interlocking saddle-shaped joint surfaces; that is, each surface is concave in one direction and convex in the orthogonal direction. The two surfaces are not precisely congruous. This geometry allows considerable freedom of motion when the thumb is unloaded. When the thumb is loaded (e.g., in a pinch grip) the two saddles push together and the joint geometry contributes to stability. Relative motion of the two joint surfaces is entirely by sliding because a rolling motion is not possible with this geometry. Because friction is clearly increased in sliding contact, this restriction has lubrication implications.

5.3.5.3 Hip Joint

The ball-and-socket joint of the hip consists of the spherical femoral head tethered to the socket shaped acetabulum of the pelvis by a short ligament, seemingly a situation in which joint contact is continuous. However, the convex part is actually larger in diameter than its concave articulation, so that at low loads joint contact is confined to the periphery. As loads increase, the acetabulum undergoes elastic deformation and the two surfaces become congruent. When the joint is again unloaded, the residual force in the acetabulum pops the femoral head partially back out of the joint. This was nicely demonstrated in a cadaver study by Bay and colleagues (1997) who measured substantial periacetabular strains as a function of hip loading. During gait, the femoral head is not only rotating within the acetabulum, but is moving in and out of the acetabulum with each step. This movement may support

squeeze-film lubrication in this large, heavily loaded joint. The hip joint also has a strong labrum that seals the joint when it is healthy. In fact, to remove a healthy femoral head from its acetabulum in the anatomy laboratory it is necessary to “break the seal” of the labrum to allow air to enter the joint and permit separation of the femoral and acetabular joint faces.

5.3.5.4 Tibiofemoral Joint at the Knee

The distal femur terminates in two rounded condyles that articulate with the much flatter surface of the proximal tibia. This arrangement allows for a combination of rolling and sliding as the joint moves, but contributes little to joint stability, which depends on dynamic muscle action and on ligamentous restraints. These stability mechanisms are often inadequate, and soft tissue injuries about the knee are common. The menisci are interposed between the joint surfaces presumably to help distribute forces across the joint, and may help to provide the extraordinarily narrow fluid spaces necessary for hydrodynamic lubrication. The menisci also have a stabilizing and sealing function that is noticeable in the anatomy laboratory. This is particularly true on the medial side of the joint where the fibers of the joint capsule interdigitate with the outer rim of the meniscus when separating the femur from the tibia, the seal between the distal femur and the menisci must be broken to allow air to enter the space if the joint is to be easily separated.

These few examples give a glimpse of the diversity of joint architecture. Obviously, caution is required in applying Reynolds’ equation, or any other theory, to synovial joints. It seems likely that several different lubrication mechanisms function to varying degrees in different joints.

5.4 Summary and Further Reading

We have described the mechanics of cartilage and joint lubrication in synovial joints. Arthritic diseases involve interactions between mechanical damage and inflammatory responses within a joint and the adjacent bone. The mechanics of cartilage are dominated by its high water content, high negative charge density and the matrix of proteoglycan and collagen molecules, which retain the water during compression and resist tensile stresses, respectively. Under static loading, electrostatic repulsion between the fixed charges explains most of the modulus of the tissue. At very slow loading rates tissue viscoelasticity is explained by fluid flow and permeability is an important variable. The biphasic theory has been extensively used for interpreting the results of compression tests and predicting the mechanical behavior of cartilage, but it is inaccurate for other than very slow loading rates. For more rapid loading, the viscoelasticity of the matrix is an essential part of the mechanical response. This has been modeled using phenomenological viscoelasticity or polymer mechanics with some success. The polymer mechanics approach has the advantage of being a mechanistic theory, but it is presently

less developed than the phenomenological approach. The lubrication of synovial joints is similarly poorly understood, but is thought to involve both boundary and hydrodynamic lubrication mechanisms. The flow of synovial fluid in and out of the cartilage on the joint surfaces seems to be important as well.

For additional information on this subject, there are several books: *Tribology of Natural and Artificial Joints*, by Dumbleton (1981), *Biomechanics of Diarthroidal Joints*, edited by Mow et al. (1990), and *Mechanics of Human Joints* edited by Wright and Radin (1993). There are also several reviews by Mow and his collaborators (Mow et al. 1986, 1991; Mow and Soslowsky 1991; Mow and Mak 1987), and by Woo et al. (1987). An early paper on the biphasic theory is that by Mow et al. (1980). More recent papers on the “triphasic” extension of the theory (the third element being ions) are by Lai et al. (1991) and Gu et al. (1993). A clearly written engineering textbook on lubrication theory is *Basic Lubrication Theory* by Cameron (1981), and *History of Tribology* by Dowson (1979) provides fascinating context for this chapter. The second edition of the book *Basic Orthopaedic Biomechanics*, edited by Mow and Hayes (1997), devotes four excellent chapters to synovial joints. For a general overview of cartilage and its relationship to bone: *Bones and Cartilage: Developmental and Skeletal Developmental Biology*, by Hall (2005) is a very interesting overview that takes an evolutionary standpoint. A general overview of tissue mechanics that includes cartilage is: *Mechanics of Biological Tissue*, Editors: GA Holzapfel and RW Ogden, Springer-Verlag Berlin Heidelberg, 2006 (Holzapfel and Ogden 2006). This book is interesting for including modeling of tissue growth integrated with the mechanical property development.

5.5 Exercises

- 5.1 Determine the relative amounts of strain energy absorption in the cartilage and bone of the proximal half of the tibia when it is loaded uniformly over the articular surface by an arbitrary compressive force. Simplify the situation by ignoring the fibula and assuming that the tibia consists of three parts: a cylindrical diaphyseal segment of cortical bone, a metaphyseal segment of cancellous bone in the form of a truncated cone, and a disk-shaped cartilage layer (Fig. 5.26). The dimensions shown in the sketch are $a = 10$ mm, $b = 30$ mm, $c = 50$ mm, $e = 4$ mm, $f = 70$ mm, and $g = 130$ mm. Assume the diaphysis is a hollow circular cylinder having inner and outer diameters as shown, and the cartilage and cancellous regions are uniformly filled with tissue. Assume also that stresses are uniformly distributed throughout each individual segment of the model and each tissue is linearly elastic with the following elastic moduli:

Cortical bone	20,000 MPa
Cancellous bone	200 MPa
Cartilage	20 MPa (Radin et al. 1970) or 200 MPa (Higginson and Snaith 1979)

Calculate the total strain energy in each segment of the model as a percent of the total energy in the model. What is this percentage for the cartilage using the two different dynamic modulus values?

- 5.2 Calculate the average stress and strain values in the cartilage and cortical bone in Exercise 5.1 if the uniform compressive load is the body weight of a 70-kg person (i.e., approximately 700 N).
- 5.3 In Fig. 5.23, viscosity is plotted as a function of strain rate. Show that when a fluid is sheared between two surfaces, one of which is moving with velocity V relative to the other, the strain rate is equal to the velocity gradient across the fluid. Then calculate the relative velocity between the joint surfaces associated with a mean strain rate of 1000 s^{-1} if they are separated by a distance of $1 \mu\text{m}$ (10^{-6} m). Compare this to an estimate of the relative velocity in the knee joint.
- 5.4 Suppose a hydrodynamically lubricated joint consists of a moving horizontal surface on one side and a stationary, exponentially curved surface on the other, as depicted schematically in Fig. 5.27. The joint is $W = 0.04 \text{ m}$ wide and the equation for the upper surface is

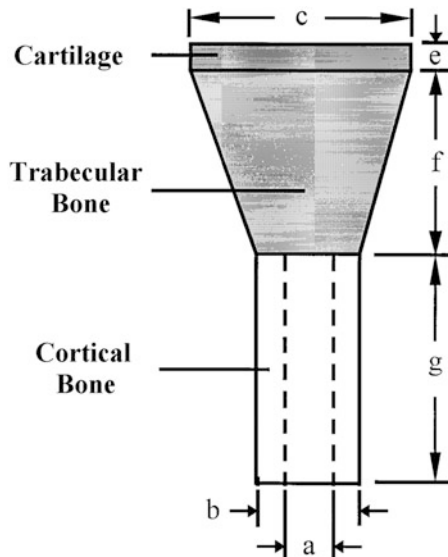


FIGURE 5.27. Diagram for exponentially curved joint surface problem (not to scale).

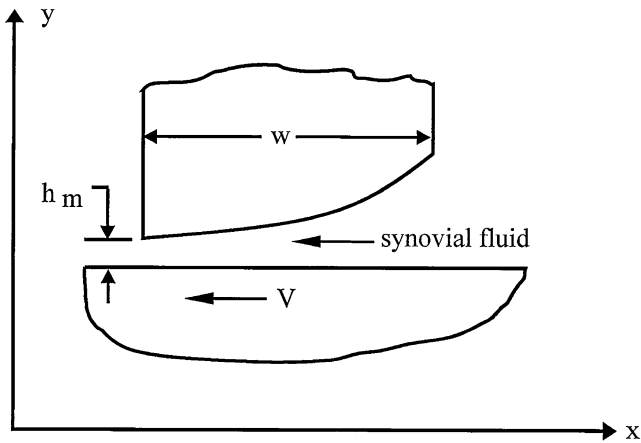


FIGURE 5.28. Model for the proximal half on the human tibia.

$$y = h_m e^{kx} \quad (7.27)$$

where k is a constant and the minimum distance between the joint surfaces is $h_m = 5 \mu\text{m}$. The viscosity of the synovial fluid is 0.05 Pa s . The relative velocity between the joint surfaces is 0.25 m/s . Using Reynolds' equation; obtain an equation for the fluid pressure, p , as a function of position between the two surfaces (Fig. 5.28).

- 5.5 In Exercise 5.4, find the value of k that maximizes the hydrodynamic pressure. Plot the corresponding surface.
- 5.6 In Exercise 5.4, calculate the force exerted on the joint surfaces by p , assuming that the depth of the joint surfaces (into the page in Fig. 5.28) is 0.04 m . How does this force compare to a body weight of 700 N ?
- 5.7 In the experiments that oscillated joints as pendulums, it was necessary to find a relationship between the decay of the oscillation and the friction coefficient in the joint. Solve this problem; that is, show that the friction coefficient is approximately

$$\mu = (a - b)/4R$$

where $a - b$ is the change in amplitude during one swing and R is the distance from the center of rotation of the joint to the articular surface. Other features of the problem are shown in Fig. 5.29. (Hint: use energy methods. If you really get stuck, look in Barnett and Cobbold (1962)!)

- 5.8 Using a flat-ended cylindrical indenter of radius $R = 2.5 \text{ mm}$, Obeid and co-workers (1994) measured the mechanical properties of human

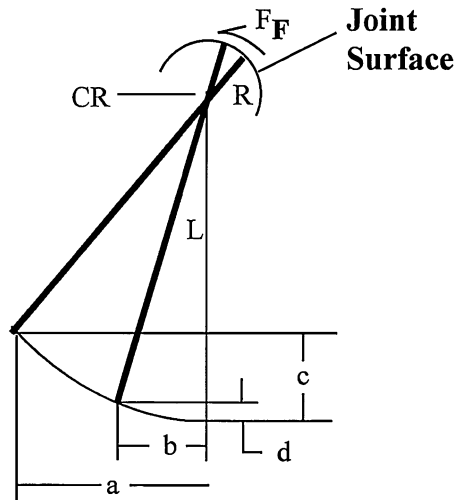


FIGURE 5.29. Diagram for pendulum experiment. F_F is the friction force against the articular joint surface; CR is the center of rotation. Previously published as Fig. 17 in Barnett, C.H., and Cobbold, A.F. Lubrication within living joints. *Journal of Bone and Joint Surgery* 1962; 44B:662–674 (Barnett and Cobbold 1962).

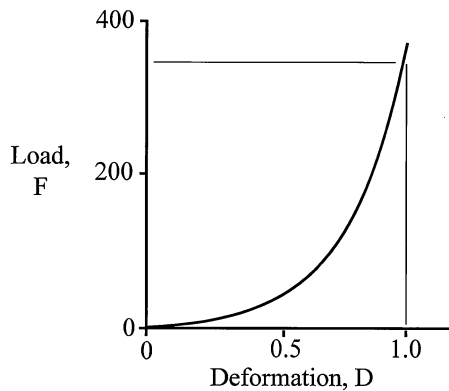


FIGURE 5.30. Load vs. deformation curve for indentation test of human articular cartilage. Fig. 2 in Obeid, E.M.H., Adams, M.A., and Newman, J.H. Mechanical properties of articular cartilage in knees with unicompartmental osteoarthritis. *Journal of Bone and Joint Surgery* 1994; 76B:315–319 (Obeid et al. 1994). With kind permission of Springer Science + Business Media.

articular cartilage from the tibial plateau in the knee joint. A typical load-deformation curve was as shown in Fig. 5.30 when the load was applied in about 0.5 s. The theoretical relationship between load, F , and the maximum deformation, D , produced by such an indenter (Lewis and McCutchen 1959) is

$$F = 2RED/(1 - \nu^2)(7.28)$$

where E is the elastic modulus and ν is Poisson's ratio (Figs. 5.27 and 5.29).

Calculate the average (secant) modulus for the specimen shown in the figure as it was loaded to 350 N and 1 mm deformation. Assume that Poisson's ratio is 0.4. How does your result compare with typical values of cartilage's aggregate modulus? Explain the difference. How would the behavior shown in Fig. 5.30 change if the same deformation were achieved in 0.1 s, as when a person jumps?

References

- Adams LM. The anatomy of joints related to function. In: Wright V, Radin EL, editors. *Mechanics of the human joints*. New York, NY: Dekker; 1993. p. 27–81.
- Armstrong CG, Lai WM, Mow VC. *J Biomech Eng*. 1984;106:165–73.
- Barnett CH, Cobbold AF. Lubrication within living joints. *J Bone Joint Surg*. 1962;44B:662–74.
- Bay BK, Hamel AJ, Olson SA, Sharkey NA. Statically equivalent load and support conditions produce different hip joint contact pressures and periacetabular strains. *J Biomech*. 1997;30(2):193–6.
- Benninghoff A. Form und bau der gelenkknorpel in ihren beziehungen zur funktion. Erste Mitteilung *Zeitschrift für Anatomie und Entwicklungsgeschichte*. 1925;76:43.
- Boroudjerdi H, Kim Y-W, Naji A, Netz RR, Schlagberger X, Serr A. Statics and dynamics of strongly charged soft matter. *Phy Rep*. 2005;416:129–99.
- Boussinesq J. *Applications des potentiels a l'étude de l'équilibre et du mouvement des solides elastique*. Paris: Gauthier-Villars; 1885.
- Boyd R, Smith G. *Polymer dynamics and relaxation*. Cambridge: Cambridge University Press; 2007.
- Brand RA. Joint contact stress: a reasonable surrogate for biological processes? *Iowa Orthop J*. 2005;25:82–94.
- Buckwalter JA, Mow VC. Cartilage repair in osteoarthritis. In: Moskowitz RW, Howell DS, Goldberg VM, et al., editors. *Osteoarthritis: diagnosis and medical/surgical management*. Philadelphia, PA: Saunders; 1992. p. 71–107.
- Buschmann MD, Grodzinsky AJ. A molecular model of proteoglycan-associated electrostatic forces in cartilage mechanics. *Trans ASME J Biomech Eng*. 1995;117(2):179–92.
- Cameron A. *Basic lubrication theory*. New York, NY: Halsted; 1981.
- Charney J. The lubrication of animal joints in relation to surgical reconstruction by arthroplasty. *Ann Rheum Dis* 1960; 19:10–19
- CRC Handbook of Chemistry and Physics. Boca Raton, FL: CRC Press; 1984.
- Dintenfuss D. Lubrication in synovial joints: a theoretical analysis. *J Bone Joint Surg*. 1963a;45A:1241–56.
- Dintenfuss D. Lubrication in synovial joints. *Nature*. 1963b;197:496–7.
- Dowson D. *History of tribology*. London: Longman; 1979.

- Dowson D, Jin ZM. Micro-elastohydrodynamic lubrication of synovial joints. *Eng Med*. 1986;15:63–5.
- Dumbleton JH. *Tribology of natural and artificial joints*. Amsterdam: Elsevier; 1981.
- Fyhrie DP, Barone JR. Polymer dynamics as a mechanistic model for the flow-independent viscoelasticity of cartilage. *J Biomech Eng*. 2003;125(5):578–84.
- Gu WY, Lai WM, Mow VC. Transport of fluid and ions through a porous-permeable charged-hydrated tissue, and streaming potential data on normal bovine articular cartilage. *J Biomech*. 1993;26:709–23.
- Haemer JM, Carter DR, Giori NJ. The low permeability of healthy meniscus and labrum limit articular cartilage consolidation and maintain fluid load support in the knee and hip. *J Biomech*. 2012;45(8):1450–6.
- Hall BK. *Bones and cartilage: developmental and skeletal developmental biology*. San Diego: Elsevier; 2005.
- Hayes WC, Keer LM, Herrmann G, Mockros LF. A mathematical analysis for indentation tests of articular cartilage. *J Biomech*. 1972;5:541–51.
- Higginson GR, Snaith JE. The mechanical stiffness of articular cartilage in confined oscillating compression. *Eng Med*. 1979;8:11–4.
- Hills BA. Oligolamellar lubrication of joints by surface active phospholipid. *J Rheumatol*. 1989;16:82–91.
- Holzhappel GA, Ogden RW, editors. *Mechanics of biological tissue*. Berlin: Springer; 2006.
- Huang CY, Mow VC, Ateshian GA. The role of flow-independent viscoelasticity in the biphasic tensile and compressive responses of articular cartilage. *J Biomech Eng*. 2001;123(5):410–7.
- Hunter W. Of the structure and diseases of articulating cartilages. *Philos Trans R Soc Lond*. 1743;43:514.
- Jay GD, Britt DE, Cha CJ. Lubricin is a product of megakaryocytes stimulating factor gene expression by human synovial fibroblasts. *J Rheumatol*. 2000;27:594–600.
- Jones ES. Joint lubrication. *Lancet*. 1934;228:1426–7.
- Jones ES. Joint lubrication. *Lancet*. 1936;230:1043–4.
- June RK, Mejia KL, Barone JR, Fyhrie DP. Cartilage stress-relaxation is affected by both the charge concentration and valence of solution cations. *Osteoarthritis Cartilage*. 2009;17(5):669–76.
- June RK, Neu CP, Barone JR, Fyhrie DP. Polymer mechanics as a model for short-term and flow-independent cartilage viscoelasticity. *Mater Sci Eng*. 2011;31(4):781–8.
- Korhonen RK, Laasanen MS, Toyras J, et al. Comparison of the equilibrium response of articular cartilage in unconfined compression, confined compression and indentation. *J Biomech*. 2002;35(7):903–9.
- Kuettner KE. Biochemistry of articular cartilage in health and disease. *Clin Biochem*. 1992;25:155–63.
- Lai WM, Hou JS, Mow VC. A triphasic theory for the swelling and deformation behaviors of articular cartilage. *Trans ASME J Biomech Eng*. 1991;113(3):245–58.
- Lewis PR, McCutchen CW. Experimental evidence for weeping lubrication in animal joints. *Nature (London)*. 1959;184:1285.
- Linn FC. Lubrication of animal joints: I. The arthrotrip-someter. *J Bone Joint Surg*. 1967;49A:1079–98.

- Linn FC. Lubrication of animal joints: II. The mechanism. *J Biomech.* 1968;1:193–205.
- Little T, Freeman M, Swanson A. Experience on friction in the human joint. In: Wright V, editor. *Lubrication and wear in joints*. London: Sector Publishing; 1969. p. 110–4.
- Malcom LL. An experimental investigation of the frictional and deformational responses of articular cartilage interfaces to static and dynamic loading. Doctoral dissertation, University of California at San Diego; 1976.
- Mansour JM, Mow VC. The permeability of articular cartilage under compressive strain and high pressures. *J Bone Joint Surg.* 1976;58A:509–16.
- Maroudas A. Hyaluronic acid films. In: *Proceedings, Institute of Mechanical Engineering*, London 1967;181:122–24.
- McCutchen CW. The frictional properties of animal joints. *Wear.* 1962;5:1–17.
- Mow VC, Hayes WH. *Basic orthopaedic biomechanics*. Philadelphia, PA: Lippincott-Raven; 1997.
- Mow VC, Lai WM. Recent developments in synovial joint mechanics. *SIAM Rev.* 1980;22(3):275–317.
- Mow VC, Mak AF. Lubrication of diarthroidal joints. In: Skalak R, Chien S, editors. *Handbook of bioengineering*. New York, NY: McGraw-Hill; 1987. p. 5.1–5.34.
- Mow VC, Soslowsky LJ. Friction, lubrication and wear of diarthroidal joints. In: Mow VC, Hayes WC, editors. *Basic orthopaedic biomechanics*. New York, NY: Raven; 1991. p. 245–92.
- Mow VC, Kuei SC, Lai WM, Armstrong CG. Biphasic creep and stress relaxation of articular cartilage in compression. Theory and experiment. *J Biomech Eng.* 1980;102:73–84.
- Mow VC, Kwan MK, Lai WM, Holmes MH. A finite deformation theory for nonlinearly permeable soft hydrated biological tissues. In: Schmid-Schonbein GW, Woo SL-Y, Zweifach BW, editors. *Frontiers in biomechanics*. New York, NY: Springer; 1986. p. 153–79.
- Mow VC, Ratcliffe A, Woo SL-Y. *Biomechanics of diarthroidal joints*. New York, NY: Springer; 1990.
- Mow VC, Zhu W, Ratcliffe A. Structure and function of articular cartilage and meniscus. In: Mow VC, Hayes WC, editors. *Basic orthopaedic biomechanics*. New York, NY: Raven; 1991. p. 143–98.
- Mow VC, Ratcliffe A, Chern KY, Kelly MA. Structure and function relationships of the menisci of the knee. In: Mow VC, Arnoczky SP, Jackson DW, editors. *Knee meniscus: basic and clinical foundations*. New York, NY: Raven; 1992. p. 37–57.
- Nimer E, Schneiderman R, Maroudas A. Diffusion and partition of solutes in cartilage under static load. *Biophys Chem.* 2003;106(2):125–46.
- Obeid EMH, Adams MA, Newman JH. Mechanical properties of articular cartilage in knees with unicompartamental osteoarthritis. *J Bone Joint Surg.* 1994;76B:315–9.
- Radin EL, Paul IL, Lowy M. A comparison of the dynamic force transmitting properties of subchondral bone and articular cartilage. *J Bone Joint Surg.* 1970;52A:444–56.
- Reynolds O. On the theory of lubrication and its application to Mr. Beauchamp Towers' experiments. *Philos Trans R Soc Lond.* 1886;177:157.

- Schinagl RM, Ting MK, Price JH, Sah RL. Video microscopy to quantitate the inhomogeneous equilibrium strain within articular cartilage during confined compression. *Ann Biomed Eng.* 1996;24:500–12.
- Silverberg JL, Dillavou S, Bonassar L, et al. Anatomic variation of depth-dependent mechanical properties in neonatal bovine articular cartilage. *J Orthop Res.* 2013;31:686–91.
- Skalak R, Chien S. *Handbook of bioengineering.* New York, NY: McGraw-Hill; 1987.
- Sneddon IN. The relation between load and penetration in the axisymmetric Boussinesq problem for a punch of arbitrary profile. *Int J Eng Sci.* 1965;3:47–57.
- Sokoloff L. Elasticity of aging cartilage. *Fed Proc.* 1966;25:1089–95.
- Soltz MA, Ateshian GA. A conewise linear elasticity mixture model for the analysis of tension-compression nonlinearity in articular cartilage. *Trans ASME J Biomech Eng.* 2000;122(6):576–86.
- Swann DA, Slayter HS, Silver FH. The molecular structure of lubricating glycoprotein-I, the boundary lubricant for articular cartilage. *J Biol Chem.* 1981;256:5921–5.
- Thomas TR, Sayles RS, Haslock I. Human joint performance and the roughness of articular cartilage. *J Biomech Eng.* 1980;102:50–6.
- Unsworth A. Cartilage and synovial fluid. In: Dowson D, Wright V, editors. *An introduction to the biomechanics of joints and joint replacement.* London: Mechanical Engineering; 1981. p. 107–14.
- Unsworth A, Dowson D, Wright V. The frictional behavior of human synovial joints: I. Natural joints. *J Lubr Technol.* 1975;97:360–76.
- Walker PS, Dowson D, Longfield MD, Wright V. Boosted lubrication in synovial joints by fluid entrapment and enrichment. *Ann Rheum Dis.* 1968;27:512–20.
- Williamson AK, Chen AC, Masuda K, et al. Tensile mechanical properties of bovine articular cartilage: variations with growth and relationships to collagen network components. *J Orthop Res.* 2003;21(5):872–80.
- Woo SL-Y, Akeson WH, Jemmott GF. Measurements of nonhomogeneous, directional mechanical properties of articular cartilage in tension. *J Biomech.* 1976;9:785–91.
- Woo SL-Y, Mow VC, Lai WM. Biomechanical properties of articular cartilage. In: Skalak R, Chien S, editors. *Handbook of bioengineering.* New York, NY: McGraw-Hill; 1987. p. 4.1–4.44.
- Wright V, Dowson D. Lubrication and cartilage. *J Anat.* 1976;121(Pt 1):107–18.
- Wright V, Radin EL. *Mechanics of human joints.* New York, NY: Dekker; 1993.
- Yamada H. *Strength of biological materials.* Huntington: Krieger; 1973.

6

Mechanical Adaptability of the Skeleton

Every change in the form and function of . . . bone[s] or of their function alone is followed by certain definite changes in their internal architecture, and equally definite secondary alterations in their external conformation, in accordance with mathematical laws. J. Wolff as quoted by Keith (1918).

Abstract Bone is exquisitely sensitive to its mechanical environment. The mechanisms by which this occurs from the biochemical to the organ level are explored in this chapter. Various models used to explain mechanical adaptation are also described. Cartilage is less adaptive, but several models have been proposed to explain how cartilage develops and adapts to loads.

Keywords Wolff's law • Mechanotransduction • Strain • Mechanostat • Habituation • Fluid flow • Damage repair • Chondral modeling

6.1 Introduction

The idea that bone structure is somehow controlled locally to suit its current mechanical function has become a central tenet of orthopaedic medicine and science. The goal of this chapter is to come to grips with this idea. As a scientific hypothesis, the putative ability of bone not only to repair itself, but to adapt to variations in imposed stresses, which we now know as Wolff's Law, is an extremely appealing idea to engineer and biologist alike. However, developing this idea into a scientific theory supported by appropriate experimental results has proven to be a hefty scientific problem. In the following pages we take up this problem, trying to organize into a coherent theory the principal hypotheses and more recent experiments found in the literature on this subject. First, however, it is helpful to review the origins of Wolff's law and the concepts that it represents.

6.2 The Historical Context

What is today called Wolff's law incorporates several different concepts, most of which Wolff himself said nothing about. Following the lead of Roesler (1981, 1987), we focus on three key concepts that arose in the nineteenth century:

1. Optimization of strength with respect to weight
2. Alignment of trabeculae with principal stress directions
3. Self-regulation of bone structure by cells responding to a mechanical stimulus

The idea that bone senses and adapts to its mechanical environment is an old one. Julius Wolff wrote about the link between mechanics and architecture in the late 1800s and is routinely given credit for this idea, but he was certainly not the first to make the observation. Galileo (1638) pointed out the mechanical implications of the shapes of bones, as did other early writers (Monro 1976). Bell (1827) observed that cancellous bone structure had "reference to the forces acting on the bone." Both he and Bourguery (Bourguery 1832) apparently recognized that trabecular bone architecture is influenced by mechanical forces, and thought that it maximized strength relative to the amount of material used. This relationship is the first important concept that we want to recognize: optimization of bone strength (or stiffness) with respect to its weight. This is vitally important because the density of bone tissue (about 2 g/cm^3) is twice that of other tissues in the body. The metabolic cost of carrying about extra bone is very high. It is perhaps unlikely that Bell and Bourguery ever considered this; they may simply have been struck by the obviously exquisite structure of cancellous bone, which looks so efficient!

Ward (1838) compared the trabecular arrangement in the neck of the femur to a bracket designed to hold a street lamp to its post (Fig. 6.1). We now apply the name Ward's triangle to the region of sparse trabecular struts corresponding to the central opening in this "bracket." Ward's comparison is significant because it is an early recognition that bone structure is analogous to structures engineered by people and that bone does not remain where it does not need to be.

We come now to the second key concept, which arose from a legendary conversation between Karl Culmann and Hermann von Meyer. The story goes that 1 day in 1866 Prof. Culmann, a well-known engineer, visited the laboratory of Prof. Meyer, an equally well-known anatomist, and happened to see a specimen of a proximal human femur that had been sectioned longitudinally to reveal its pattern of trabecular arcades. Culmann was immediately struck by the similarity of this pattern to that of the principal stresses within a crane structure which he had analyzed (Fig. 6.2 shows sketches of these two objects). He is said to have cried out "That's my

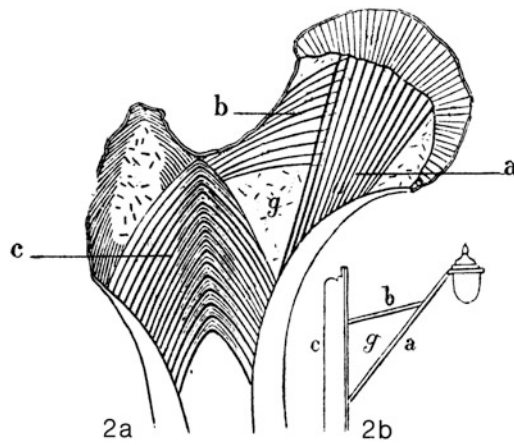


FIGURE 6.1. Sketch showing Ward's analogy between the trabecular arrangement in the neck of the femur and a streetlamp bracket (Ward 1838).

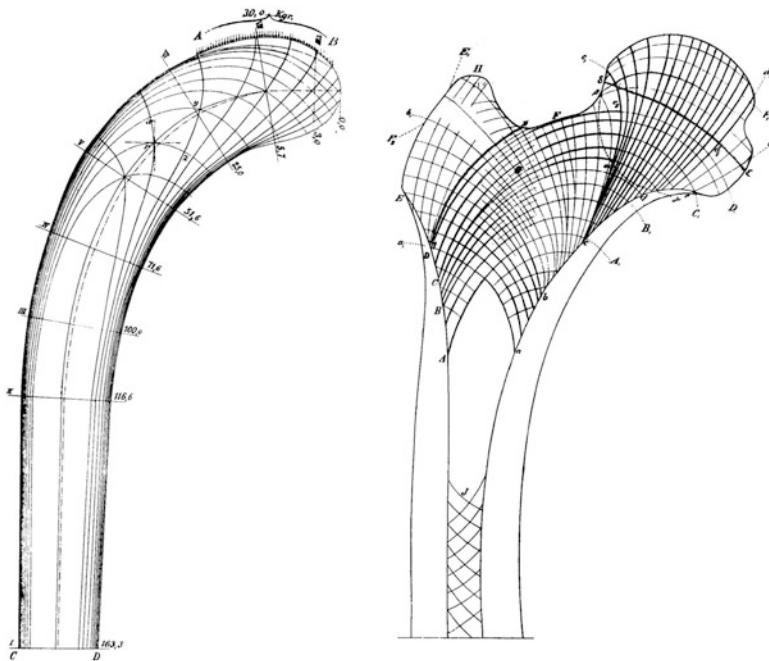


FIGURE 6.2. (Left) Principal stress trajectories in a Fairbairn crane as calculated by Culmann. (Right) The trabecular arrangement in a proximal human femur as sketched by von Meyer.

crane!” (Thompson 1992).¹ The crane was actually named after Fairbairn, the Scots engineer who designed it. Its flowing, bonelike shape is typical of the trial-and-error designs common during the early industrial revolution in Great Britain. It is interesting to note that elasticity theory and the science of strength of materials were rapidly developing during this period of time. Augustin Cauchy had, in about 1822, discovered the existence of principal stresses, but decades passed before engineers began to make practical use of the concept. Culmann was one of the first to understand and use the concept of principal stress directions; he was one of several engineers from the continent who led the way in developing a mathematical theory of mechanics, which eventually refined the experimental approaches of British and American engineers. He had apparently been applying his graphical methods for calculating principal stress directions to Fairbairn’s crane. In any event, the similarity between the trabecular arcades and the stress trajectories is indeed striking. Clearly, however, the shapes and loads are substantially different for the crane and the femur. Not only that, but the femur is a discontinuous collection of bony struts filled with marrow, while the crane is a homogeneous structure. Principal stresses do not exist for a discontinuous structure. The significance of this was overlooked for a long time (Cowin 1997). We shall return to this issue several times in this chapter.

This conversation was noted by others interested in bone, including Wolff, and soon afterward he began to publish papers on this subject (Wolff 1869, 1870). His interpretations of cancellous architecture quickly rose to dominance. He accepted at face value the implications of the Culmann–Meyer stress-trajectory hypothesis and asserted that there was a mathematically perfect correspondence between trabecular architecture and principal stress trajectories. This similitude is what Wolff’s law meant to Wolff, and it is our second key concept.

Wolff went on to try to explain the trabecular structure within many pathologic specimens using this basic idea. In 1892, he published his famous book, *Das Gesetz der Transformation der Knochen* (The Law of Bone Transformation), which summarized his earlier writings.² It must be appreciated, however, that while he spoke of mathematical rules governing trabecular architecture, he was not an engineer or mathematician, and never attempted to formulate a mathematical theory of “Wolff’s law.” Furthermore, he believed that bone was formed interstitially (i.e., cells pushed out new bone from within the calcified matrix—a physical impossibility), and had no understanding of bone modeling and remodeling as we know them.

¹ Some sources have placed this encounter at a scientific meeting, and others have said the femur was actually a metatarsal.

² This book has been translated into English by Maquet and Furlong. They translated the German word “transformation” as “remodeling,” although Wolff was speaking about growth and what we now term modeling.

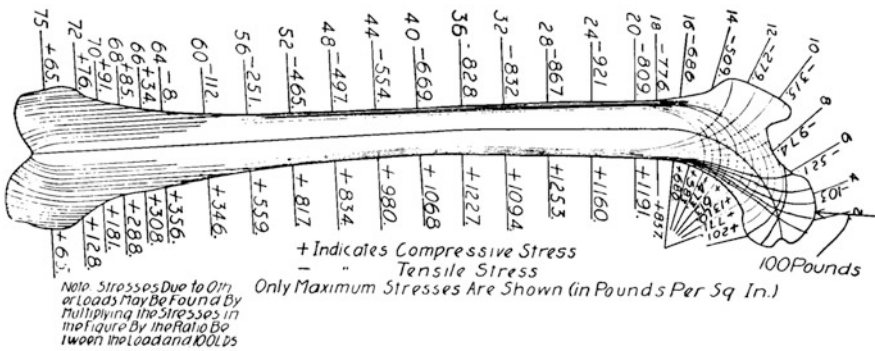


FIGURE 6.3. Koch's computation of the principal stress trajectories and values in a human femur. Note the applied load vector on the femoral head (Koch 1917).

Later, Koch (1917), an engineer as well as a physician, resolved some of the issues regarding the comparison of Culmann's stress trajectories for the Fairbairn crane with a human femur. He calculated the stress trajectories in a human femur based on a reasonably accurate load applied to the head and geometry obtained from actual anatomic measurements. In doing so, he was able to demonstrate that trabecular orientations are, after all, similar to mathematical calculations of stress trajectories for a homogeneous structure of the same shape (Fig. 6.3).

We must return now to the nineteenth century, to note the emergence of the third key concept regarding the adaptation of bone to mechanical loading: the idea that this occurs via a self-regulating biologic mechanism. While it may seem obvious to you that this must be the case, that is only because you have been educated in an era when self-regulation is accepted as an essential part of biological organisms. That was not the case in the time of Wolff. It was in his century that Claude Bernard (1813–1878) introduced the idea of homeostasis to physiology: the concept that physiologic systems self-regulate so as to keep a critical internal variable constant. However, it was many years before the full significance of this concept came to be appreciated and accepted. Thus, it was still a fundamentally new idea when, in 1881, W. Roux, influenced by Darwin and others, hypothesized that organisms possess the ability to adapt to changes in their living conditions (Note that we are not speaking of evolution here, which involves long-term, genetic changes between one generation and the next, but changes in an individual during its lifetime, even day-to-day or minute-to-minute changes. Also, we are not speaking of adjusting to internal, physiologic changes, but to changes involving interactions with the external environment). Roux specifically proposed that the ability of bone to align trabeculae with stress trajectories was accomplished by cells forming and resorbing bone according to variations in a functional stimulus, in this case, a mechanical stimulus. While his

imperfect understanding of biology and mechanics led him to make many mistakes in developing his theory, his concept of self-regulated functional adaptation governed by cell sensitivity to a functional stimulus was extremely important.

In summary then, as Roesler (1987) has pointed out, nineteenth-century researchers provided us with three key concepts regarding bone's ability to adapt to changing mechanical loads: bone structure optimizes strength with respect to the amount of material used; trabeculae line up with principal stress directions; and these things are accomplished by a self-regulating system of bone cells responding to a mechanical stimulus. Today, Wolff's law remains a rather poorly defined "law," but it more or less includes these three core principles.

6.3 Mechanical Adaptability and Evolutionary Adaptability

There has been considerable debate among engineers, biologists, and clinicians concerning the balance between genetics and mechanical adaptability in deciding the structures of bones. Some have argued that the anatomy and histology of a bone such as the human femur are determined primarily by genetic information expressed by various cells as they form the skeleton and maintain it throughout life. Others have come to believe that mechanical factors strongly influence the expression of the genetic plan carried by the cells. It is appropriate to consider the relationships between mechanical adaptability and evolutionary change as determinants of skeletal structure. This discussion is based on the ideas of Bateson (1963, 1987).

6.3.1 *Somatic Versus Evolutionary Adaptation*

The adaptation of bone to alterations in mechanical loading is an example of somatic change, that is, a change in body form or function that does not entail a genetic change and is therefore not passed on to succeeding generations. Somatic change is distinct from evolutionary change, which does involve a modification of the genome of the organism. A simple example of somatic change is obtaining a sun tan during a beach vacation; obviously, this does not alter the genetic code for skin color, and the tan is not passed on to the vacationer's children. Nevertheless, somatic change plays an important role in evolution.

Bateson pointed out that the ability to adjust to environmental changes by somatic modification of body systems is essential to survival because evolutionary change operates relatively slowly, through chance mutations. In addition, somatic adjustments allow the organism to survive when a genetic mutation, which might ultimately have survival value, disrupts one or more

body systems. Conversely, genetic adjustments to the somatic system itself allow the homeostatic range to be restored when a mutation has driven it to an extreme. Thus, evolution involves constant interaction between slow, haphazard genetic changes and rapid, predictable somatic changes. In short, somatic flexibility allows time for adaptive modifications of the genome to occur, and also helps the organism tolerate their adverse side effects.

Work in the burgeoning field of epigenetics has demonstrated that interactions between somatic and genetic adaptations are far more complex and inter-related than was once thought. We now know that it is possible for somatic changes to an organism to be passed on to subsequent generations, not through mutations in the genetic blueprint (DNA sequence) but through somatic events such as DNA methylation and histone modifications that change the expression of specific DNA sequences. The processes of epigenetics are well appreciated during growth and development as specialized cells differentiate along countless pathways, depending upon morphogens and other factors within their immediate environment. Subsequent daughter cells inherit the characteristics and phenotypes of their differentiated predecessors. This often occurs through an epigenetic process in which an enzyme called methyltransferase adds a methyl group to the carbon 5 position of cytosine (one of the four bases that comprise DNA) converting it to 5-methylcytosine. High levels of 5-methylcytosine alter local transcription patterns and cell phenotype. Going from cells to the whole organism, genetic imprinting occurs when these heavily methylated regions of the genome are passed from parent to offspring through the germline. This tends to silence expression of the allele from that parent in favor of expression of the other parent's allele. Recent work suggests that epigenetics may play an important role in transgenerational adaptation to environmental factors. Cooney and colleagues (2002) found that feeding maternal mice methyl-supplemented diets strongly affected the distribution of coat color patterns in their offspring, and correlated with methylation of the sequences controlling its phenotypic expression. A pair of multi-generational studies in humans found associations between early paternal smoking and the body mass index of their sons, and even more striking, between the food supply of paternal grandfathers and the mortality risk ratios of their grandsons (Pembrey et al. 2006). It is not beyond the realm of possibility that epigenetic mechanisms play an important role in bone adaptation to environmental or lifestyle changes (e.g. diet, physical activity) in the individual during development and maybe even in subsequent generations (Grossniklaus et al. 2013).

6.3.1.1 A Non-skeletal Example

Consider, for example, what happens when the vacationer mentioned above goes instead to visit Machu Picchu, the ancient Incan ruins in the Peruvian Andes. After flying to Lima, the tourist's next stop is the city of Cuzco with

an elevation of 3700 m (12,000 ft). Picking up her backpack and heading for her hotel, she soon finds her heart pounding and her breath coming in gasps because of the thin air. Clearly, these are responses to the trouble her cells are having in getting enough oxygen. Her increased volume and rate of breathing and pumping blood are somatic changes allowing her to cope with the environmental changes. By the time she leaves Cuzco 2 days later to continue her journey to Machu Picchu itself, carrying her backpack no longer requires her heart and respiration rates to be adjusted upward so much because her hematopoietic system has produced more red blood cells. This second level of somatic adaptability requires more time to put in place, but much less metabolic energy than pumping blood and air faster.

However, more red blood cells per unit volume of blood (called hematocrit) is not the best physiologic adjustment for living permanently at high altitude, either. “Lowlanders” who go to high altitudes may suffer severe cardiovascular problems caused by the effects of increased hematocrit on the physical properties of their blood. Some individuals can adjust to living at high altitudes, but many cannot. However, groups of people have lived at high altitudes (principally in the Andes and Himalayan mountains) for many generations and appear to be free of altitude sickness. It is possible that these populations have experienced genetic changes, through differential reproductive success, which have adapted them to high altitudes in ways that avoid the problems associated with increased hematocrit (Cruz-Coke 1978).

6.3.1.2 Somatic Change in the Skeleton

Comparing bone’s mechanical adaptation to Bateson’s altitude adaptation example, the correlate to increased heart and respiration rates is an increased remodeling rate to repair load-induced microdamage. The correlate to increased hematocrit is periosteal and endosteal modeling to decrease the stresses in the bone. These skeletal responses can have problematic sequelae, just as the respiratory ones can. The increased rate of remodeling requires metabolic energy, and it introduces remodeling space porosity, which can increase strains and the damage formation rate. The formation of new bone on periosteal and endosteal surfaces also requires metabolic resources, and the additional bone mass must be carried about each day, at a very significant energy cost.

6.3.2 Somatic Versus Evolutionary Effects

Bateson’s point was that all evolutionary change involves endless somatic and genetic adjustments, which together achieve adaptations that would be impossible if only one or the other existed. Thus, it is shortsighted to debate which caused what, exactly. It is more important to learn how somatic and genetic changes interact in the skeleton, as they do in other organ systems. This kind of knowledge should ultimately be very useful in treating many bone injuries and diseases.

Box 6.1 Teleology and Adaptation, Optimization and Evolution

An engineer who has designed a structure can say, “The purpose of this strut is to reduce bending moments on this beam,” but we cannot analogously say that the purpose of a strut of trabecular bone is to reduce the bending moment on an adjacent trabecula. The difficulty lies with the word purpose. The attribution of purpose to physiologic mechanisms is called teleology, and in this text we have tried to avoid it. We might say instead, “a consequence of this trabecular strut is that bending moments on this adjacent trabecula are reduced.” This is not just an esoteric point of semantics, it is important in understanding the nature of biological adaptation, whether it is somatic or evolutionary.

In the course of natural selection a group of organisms produces a great many genetic variants, some of which are better adapted in the sense that they will leave more progeny in the next generation. By this mechanism, species evolve, and adaptations such as bone remodeling come into being and are preserved. The process is a passive and stochastic one: the concept of purpose does not enter into it.

There are two sides to biological adaptation—the organism and its environment. When the environment is noisy and unpredictable, the adaptive process has a more difficult problem to solve. For an example see: Crummett et al. (2013), Ng et al. (2013), Jansen et al. (2011) and Eranen et al. (2009). As we saw in Chap. 1, the mechanical environment of vertebrate bones is provided to a large extent by their own muscles; much of this loading is predictable and controllable. But the outside environment also directly loads bones, and these forces may be less predictable and controllable. If a creature walks on level ground all day, and at the end of the day steps unexpectedly into a hole, which of these is mechanically most significant in driving adaptation, the 20,000 cycles of gait or the one cycle of tripping that generates twice as much skeletal stress? Lanyon (1993) has favored the idea that skeletal adaptation is primarily an error-driven process, and it is the unusual loads that are most significant, but this important question has not been subjected to much research. More recent work by Currey and colleagues (2007) tends to suggest otherwise.

Organisms that live in water inhabit a mechanical environment that is seemingly more predictable than the terrestrial one, which should allow for more metabolically efficient adaptation of their skeletons relative to land dwellers. But the buoyancy of water works against the adaptive advantage of a lighter skeleton in that a heavy skeleton balances low-density fat, resulting in an organism with an average density close to that of water, thereby facilitating aquatic locomotion. Conversely, birds (that is, flying birds, not terrestrial birds like chickens and ostriches) find exceptional adaptive value in a minimal-weight skeleton, which improves the efficiency of flight. Other factors contribute to the skeleton’s mechanical milieu as well, including pregnancy and other seasonal fluctuations in body weight, antler growth, hibernation, migration, and fighting among males. These factors must complicate the optimization of skeletal structure.

This chapter is devoted to discussing the proposition that the skeleton contains adaptive mechanisms that optimize its structure in some sense. Deterministic models (Nowlan and Prendergast 2005) suggest that bones do not evolve an optimal structure or properties from any one perspective, because there are competing interests (e.g. mechanical demand vs. mineral homeostasis vs. nutrient transfer vs. immune function), although they may approach optimal. It would therefore be a mistake to

hypothesize that evolution is directed toward maximizing efficiency of structures. Evolution is directed toward maximizing genetic representation in subsequent generations, which in turn is associated with efficiency, but only rather loosely. Indeed, the physiques of many ancient relict species contain features that are hardly paragons of efficiency. Evolution can only modify the basic plan that exists. Take for instance the now well-known example of The Panda's Thumb popularized by Gould (1980). It works quite well and is adapted to its role, but it is certainly not designed in a way an engineer would design it were she to start "from scratch." The human knee joint, and the vertebral column, are two other examples, perhaps more immediate to those using this book.

6.4 What Is the Design Goal: What Does Bone Adapt to?

It is difficult to discuss mechanical adaptation of bone unless we can define the term "adaptation." What does this really mean in the context of purely mechanical skeletal function? *For the purposes of our following discussion we propose that bone structure (its shape and mass) is regulated to find the best (the optimum) solution to minimize both fracture risk and bone mass for the loadings that the bone experiences.* The adaptation of bone is a somatic change, the limits of which are determined by the genome of the organism. We will show later that the unknown genetic potential for somatic change severely hampers the formulation of a generally applicable theory for the adaptation of bone to mechanical load.

Clearly, minimization of fracture risk is important in most bones, and minimizing bone mass allows the organism to reduce the metabolic energy required for its daily activities, increasing its ability to survive. Following the lead of Currey (1984), we speak of optimizing stiffness because bones need to support loads without bending too much, but they also need some flexibility for shock absorption and to avoid fracture. The optimal stiffness surely varies with the function of the particular bone. For example, Currey (1981) has discussed the divergent functions of three bones. The antlers³ of the male red deer require impact strength and energy-absorbing power; stiffness is deemphasized. The femoral cortex of the cow requires more stiffness. The tympanic bulla in the ear of the fin whale requires a high acoustic impedance ($= \sqrt{\rho E}$) and therefore very high stiffness. Table 6.1 shows how the elastic modulus and other attributes vary among the three bones. Although most bones would be more like the cow femur than antler or tympanic bulla in their function, the optimal stiffness obviously varies from bone to bone. It should be understood that this is only a trial hypothesis to serve as a

³ Deer antlers are made of bone. Cow horns, on the other hand, are keratinous tissue, as are fingernails.

TABLE 6.1. Material properties of three different bones

Variable	Red deer antler	Bovine femur	Whale bulla
Work-of-fracture, J/m ²	6190	1710	200
Bending strength, MPa	179	247	33
Elastic modulus, GPa	7.4	13.5	31.3
Density, g/ml	1.86	2.06	2.47
Acoustic impedance, 109 kg/m ² /s	3.71	5.27	8.79
Mineral content, % by weight	59.3	66.7	86.4

After Currey (1981)

benchmark for exploring this subject. At this point it seems clear that fracture risk and bone weight need to be minimized, but we shall see several alternative possibilities for the optimized variable in the following pages.

6.4.1 Optimization, Mechanical Adaptability, and Structure-Sensitive or Insensitive Properties

There is always a trade-off in nature; improvement in one property may be detrimental to another, and to maximize evolutionary fitness requires balance. In bone, this is reflected by the relationship between stiffness and toughness (or work to fracture), which are typically inversely related. Stiffer bones tend to absorb less energy prior to fracture (Currey 2004) and therefore optimizing stiffness does not necessarily mean that fracture risk will be reduced. Osteopetrosis is a good example. This is a condition in which the osteoclasts are not working well, so that bone is not resorbed as it should be (which also means that new bone cannot be formed by remodeling processes). The bone as a whole thus becomes more highly mineralized, and very stiff and strong. However, it has little capacity to absorb energy once it has passed the yield point, and therefore breaks rather easily.

Currey proposed that bone adapts to these competing needs by optimizing towards structure-insensitive mechanical properties (Currey et al. 2007). Structure-insensitive properties are those that are not greatly influenced by small flaws or inhomogeneities in the structure (Ashby 2004). These properties tend to be less variable across bones or species, than those that are structure-sensitive. For bone, structure-insensitive properties are those that determine properties in the pre-yield region of the stress-strain curve, such as elastic modulus and bending strength. Structure-sensitive mechanical properties are more greatly affected by structural inhomogeneities, and include toughness and work to fracture that will determine the energy expended in the post-yield region, as well as impact strength.

Evolutionarily, Currey argues, bone adapts to the more predictable, i.e. less variable, properties. Thus, stiffness would be deterministic, even at

the expense of greater uncertainty in toughness. Mineral content and collagen cross-linking determine stiffness and toughness, albeit in an inversely proportional way. In the trade-off between stiffness and toughness, the fittest individuals would be those with the greatest certainty in pre-yield properties, even if this created less predictability in post-yield properties. From a structure-function standpoint, this makes a great deal of sense given that we spend most of our time loading our bones in the pre-yield range. It also makes sense in a self-repairing structure like bone, in which damage caused by unpredictable loads can eventually be repaired so that the animal can survive to reproduce another day.

6.5 How Does Bone Adapt: What Are the Signals and How Are They Transduced?

Mechanotransduction is a word used frequently by researchers in the field, but not often defined. It is a general term that can have a variety of meanings, depending on the context in which it is used. Mechanotransduction should be considered as a series of inter-related mechanical and biological events, rather than a single process. It was defined most clearly and thoroughly by Duncan and Turner (1995), who divided the sequence of events from the initial mechanical signal to the eventual biological effect into four stages: Mechanocoupling, Biochemical Coupling, Signal Transmission, and Effector Cell Response.

Mechanocoupling refers to the mechanical stimulus that initiates the events leading to a particular cellular response. We treat the potential “actors” in this play extensively in this chapter, concluding that strain may or may not have a direct impact on cellular response, but it certainly is the driving force behind fluid movements in bone that cause fluid shear stresses on osteocyte and canalicular membranes which in turn produce (although this is less likely to be significant) electrical potentials (streaming potentials) to which bone cells can respond. Fluid motion also modifies the diffusional mixing of extracellular signaling molecules which can affect cell behavior completely independently of fluid shear stress. Fluid mixing of the molecules is important because the concentration and the rate of presentation of signaling molecules changes cell response in complex and incompletely understood ways.

Once the mechanical stimulus has been perceived by the cell, it must transduce that mechanical signal into an intracellular biochemical response in the sensor cell (the osteocyte in bone) that has the potential to either transmit the signal to other cells or to initiate the recruitment and eventual maturation process that causes bone formation or resorption. This biochemical coupling may take several forms, and include physical changes to the cell (e.g. integrin—cytoskeleton-nuclear matrix deformations (Wang et al. 1993;

Chen and Ingber 1999)), alterations in G-protein pathways, phosphorylation of proteins or regulation of mineral channels (e.g. calcium, phosphorus, or chlorine channels are common), or combinations of physical and biological regulation (e.g. stretch activated protein channels).

Once sensor cells, now commonly assumed to be osteocytes and bone lining cells (pre-osteoblasts), detect the signal and begin to respond, they must transmit this information to other cells that will actually do the work of making more bone, or removing it. Signal transmission may involve a host of autocrine and paracrine factors, but the messages themselves must be transmitted through a dendritic network that is much like the neural network in the brain (Turner et al. 2002). This transmission seems to require the help of connexins, in particular connexin-43, but probably also others.

Finally, effector cells that actually produce the response must mature and go to work. In bone, these are the osteoclasts and osteoblasts. How they are controlled is the subject of other chapters in this book (see Chaps. 2 and 3). It is worth emphasizing, though, that they do not work in a vacuum generated only by the information provided by the sensor cells, but are regulated during their activities by hormones, other metabolic products, by noncollagenous matrix proteins and by communication among themselves.

In this chapter, we will concentrate mostly on the first step in the mechanotransduction process, mechanocoupling, and to a lesser degree on signal transmission, as the other steps are largely complex biological processes, outside the scope of this text. If cells adapt bone to mechanical loads by responding to a mechanical stimulus, it is obviously important to know what that stimulus is and to be able to measure it. One attractive suggestion, supported by many experiments, is that bone adapts to strain. Cells attached to bone surfaces may “feel” the displacements of the surface and respond to them. Strain is also an appealing variable from the point of view of minimizing fracture risk. We will see in Chap. 7 that yield strains are reasonably constant for bone. Therefore, keeping strains near a point well below the yield strain should maintain a margin of safety against monotonic, creep, and fatigue failure. Finally, if stiffness is to be optimized in some sense, the fact that strain is a measure of bone stiffness under a given load (e.g., body weight) may be significant. These features and others make it easy to see why strain is targeted as the prime functional stimulus, and why measurement of bone strains *in vivo* has been such an important component of the experimental work aimed at bone adaptation.

6.5.1 *Measurement of Strain in Living Animals*

Until the advent of *in vivo* strain gauging techniques, the adaptation of bone to mechanical loading could not be effectively analyzed because bone’s “mechanical environment” could not be quantitatively assessed at a level commensurate with cellular activity. Lanyon and Smith (1969, 1970)

reported the first measurement of in vivo bone strains in animals when they implanted uniaxial strain gauges on the medial tibial cortex in sheep. They found mean longitudinal surface strains of 260 microstrain during walking and a 16 % increase when the animals were trotting.

Subsequent to this, Lanyon (1971, 1972) applied strain gauges to the thoracic and lumbar vertebrae in sheep. He reported average peak surface strain ranges in thoracic vertebrae during walking of 142 microstrain ($\mu\epsilon$), increasing 25 % to 187 $\mu\epsilon$ during trotting. Cochran (1972, 1974) applied strain gauges to the tibia and ulna in dogs, measuring strains in the tibia between 103 and 665 $\mu\epsilon$, depending on the load and the anatomical position of the strain gage. The mode of the strain, compressive or tensile, was not specified. Shortly after this, Barnes and Pinder (1974) and Turner et al. (1975) added the radius, tibia, metacarpus, and metatarsus of horses to this list. Lanyon et al. (1975) demonstrated that strain magnitudes in the human tibia during locomotion were similar to those in bones of other animals.

6.5.2 *Principal Stresses, Strains and Trabecular Orientations*

About this time, Lanyon began to use rosette strain gauges,⁴ permitting the measurement of principal strain directions. Applying these gauges to the calcaneus of a sheep, Lanyon (1973, 1974) demonstrated that the principal orientations of trabeculae coincided with the principal compressive and principal tensile strain directions. Later, Hayes and Snyder (1981) conducted an intensive study of the human patella which concluded that its trabecular orientations are highly correlated with principal stress directions, as determined by finite element analysis. Although this study did not examine in vivo strains on the patella, both studies suggest that the proposed “trajectory criterion” of Culmann, Meyer, and Wolff is essentially correct and trabeculae do align themselves with the continuum level principal stress (or strain) directions (i.e., the principal directions in a homogenous model structure).

There have been several efforts to understand mathematically the trajectory criterion for cancellous bone by combining properties of the stress (or strain) with properties of cancellous bone structure and/or cancellous bone mechanical properties. Cortical bone (Dempster and Liddicoat 1952) and cancellous bone (Harrigan and Mann 1984) were experimentally demonstrated to be mechanically anisotropic and probably to be essentially orthotropic—having three orthogonal principal material axes similar to the three principal axes of stress. If cancellous bone is an orthotropic material, then if one aligns the largest principal modulus (stiffness) with the largest magnitude of principal stress and so on down to the smallest principal

⁴ A group of three gauges oriented at 45° to one another.

modulus and stress, the elastic energy put into the material by an arbitrary stress is minimized (Fyhrie and Carter 1986). This suggests that trabecular orientation is a result of elastic energy minimization through alignment of the material properties with the applied stress (There are some constraints on the size of the shear moduli in that theory). It was also demonstrated experimentally that the orientations of the surfaces of the trabeculae could be modeled fairly accurately as a second rank tensor called the mean intercept length (MIL) and that the orientations of the principal axes of the MIL tensor matched the orientation of the orthotropic material properties of the bone (Cowin 1985). As a result, there are three mathematical objects that one can consider aligned in cancellous bone in order to try and explain the trajectory criterion: (1) the principal directions of the orthotropic material properties, (2) the principal directions of the mean intercept tensor and (3) the principal stress (or strain) directions. These alignments can be condensed into the rule that trabecular orientation (i.e., orientation of the mean intercept ellipsoid) matches the orientation of the principal stresses (Cowin 1986) and, with additional arguments, that their orientation is driven by the goal (an optimization) to make the principal apparent axial strains all the same (Turner 1992).

All of these ideas for mathematically explaining the orientation of cancellous bone with loading result in solely orthogonal networks of trabeculae because all of the principal axes that are being aligned are orthogonal. Unfortunately, trabeculae don't always meet orthogonally—for example, the two main portions of the trabecular network in the proximal femur meet at an angle different from ninety degrees. This was recognized early (von Meyer 1867) but was improperly rejected by arguing that early work on the theory of elasticity predicted orthogonal stresses (Wolff 1892) [For the philosophers among our readers, this was a triumph of the Platonic Ideal (mathematical theory) over Aristotelian Materialism (anatomical data)]. Non-orthogonality of trabeculae was mathematically explained by a superposition of multiple loading cases at both the average (Carter et al. 1989) and trabecular level (Mullender et al. 1994), but the underlying biological mechanisms for the sensation and correction of trabecular orientation to loading are still not known.

6.5.2.1 Peak Strains Are Similar Across Species

Subsequently, Rubin and Lanyon (1982) surveyed their own work and that of others and published the data shown in Table 6.2. These data show that when an adult animal's musculoskeletal system is functioning at a high physiologic level (e.g., a dog is running or a bird is flying), peak periosteal strains usually lie between 2000 and 3000 $\mu\epsilon$. The fact that measured strains were nearly equivalent on different bones and in different animals supported the hypothesis that the skeletons of all animals are adapted to control strain.

There are at least two notable exceptions to the data in Table 6.2. First, strains on the human tibia at a site where stress fractures occur have not been observed to exceed 2000 $\mu\epsilon$, even during strenuous activity by Special Forces

TABLE 6.2. Peak functional strains in various animal bones

Bone	Activity	Peak strain	Reference
Horse radius	Trotting	−2800	Rubin and Lanyon (1982)
Horse tibia	Galloping	−3200	Rubin and Lanyon (1982)
Horse metacarpal	Acceleration	−3000	Biewener et al. (1983)
Dog radius	Trotting	−2400	Rubin and Lanyon (1982)
Dog tibia	Galloping	−2100	Rubin and Lanyon (1982)
Goose humerus	Flying	−2800	Rubin and Lanyon (1982)
Cockerel ulna	Flapping	−2100	Rubin and Lanyon (1982)
Sheep femur	Trotting	−2200	Rubin and Lanyon (1982)
Sheep humerus	Trotting	−2200	Rubin and Lanyon (1982)
Sheep radius	Galloping	−2300	O’Conner et al. (1982)
Sheep tibia	Trotting	−2100	Lanyon et al. (1982)
Pig radius	Trotting	−2400	Goodship et al. (1979)
Fish hypural	Swimming	−3000	Rubin and Lanyon (1982)
Macaca mandible	Biting	−2200	Hylander (1979)
Turkey tibia	Running	−2350	Rubin and Lanyon (1984)

After Rubin and Lanyon (1982)

soldiers (Burr et al. 1996). More recent experiments, however, show that strains in excess of 3000 $\mu\epsilon$ can be generated during jumping activities, which may explain why “shin splints” are a relatively common injury among basketball players. See Yang et al. for a comprehensive review of in vivo bone strain measurements in people (Yang et al. 2011). Second, thoroughbred racehorses generate 5000–6000 $\mu\epsilon$ on the surface of their metacarpal (cannon) bone when running at racing speeds. However, these horses are special animals, bred and trained to run exceptionally fast while they are still juveniles. Strains of this magnitude probably only occur in young animals (or people) in which bone turnover is rapid and many osteons are still mineralizing.

6.5.3 Compressive Strains Are Larger than Tensile Strains

Because compressive strains in a long bone are usually superimposed on bending strains, the neutral axis is shifted and the peak compressive strains are usually greater than the peak tensile strains (Fig. 6.4).⁵

⁵ This is not universally true, however. An example would be in the palatamaxillary segment of some lizards, where tensile strains exceed compressive strains (Smith and Hylander 1985).

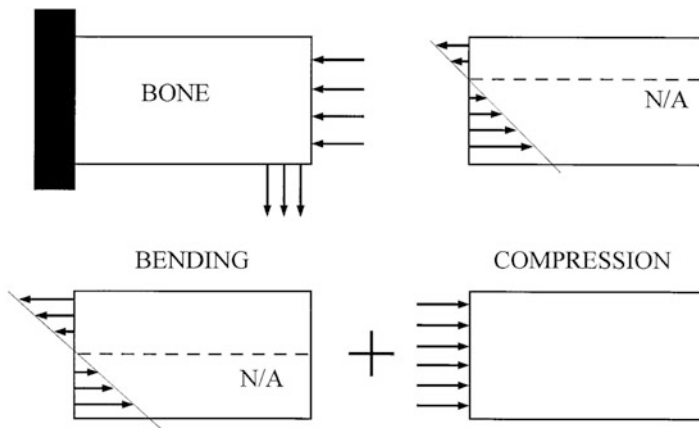


FIGURE 6.4. Most long bones are loaded in combined bending and compression, producing an asymmetric stress distribution in which compressive stresses are greater than tensile stresses.

Lanyon et al. (1982) demonstrated that the ratio of (the absolute values of) peak compressive strain to peak tensile strain on the sheep radius varies between 1.45 and 2.34, and tensile strain was never greater than compressive strain during normal locomotion. In fact, tensile strain was always less than, and compressive strain was always greater than, $1000 \mu\epsilon$. Even after osteotomy of the adjacent ulna, the tensile strain on the radius was less than $1000 \mu\epsilon$. High fidelity cadaver simulations of human gait suggest that this phenomenon also holds true in the weight bearing bones of humans. Such experiments enable comprehensive but otherwise prohibitively invasive assessments of bone strain through placement of rosette gauges at multiple sites on a single bone (Goodwin and Sharkey 2002; Peterman et al. 2001). Results from studies focused on the tibia under loading conditions representative of normal walking nicely illustrate the greater compressive strains experienced in the posterior aspect of the tibial diaphysis relative to the tensile strains present on its anterior aspect (Fig. 6.5).

6.5.3.1 What Aspects of Strain Regulate Adaptation?

The pioneering studies by Lanyon and others to measure strains on the surface of bones during activity provide some guidelines about which features in the strain environment may be critical to the adaptive response. The experiments of Lanyon and Rubin in particular showing that strain magnitude seems constrained within relatively strict boundaries are convincing that the magnitude of strain is a component of the signal. The uniformity of peak strain across species, and the correspondence between trabecular orientation and principal strain directions—are strong evidence that strain magnitude is one regulator of skeletal adaptation to the mechanical environment.

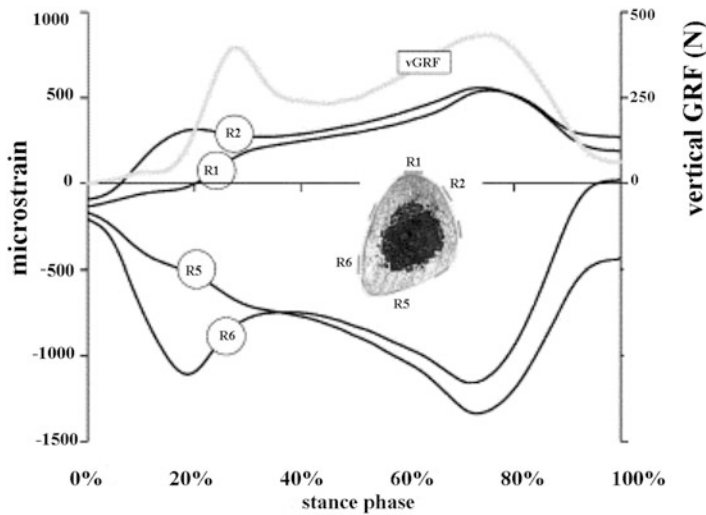


FIGURE 6.5. Axial strain profiles at two anterior (R1 and R2) and two posterior (R5 and R6) sites of the tibial shaft, illustrating compressive loading superimposed with bending. The relationship between vertical ground reaction force (vGRF) and tibial strains during stance is also readily apparent. The left ordinate axis has units of microstrain ($\mu\epsilon$) and the right ordinate axis corresponds to the dashed vGRF profile, with units of Newtons (Peterman et al. 2001).

The *principle of uniformity of peak strain* implies a threshold-driven process. To further identify the relevant magnitudes of strain that could elicit a response, Turner et al. (1991) developed a rodent model in which bone formation rate (BFR) in the tibia could be measured in response to application of different magnitudes of strain generated in four point bending. Turner et al. (1994) loaded a series of rat tibiae at strains from 400 to 2000 $\mu\epsilon$, and found no difference in BFR in the loaded tibia when compared to the contralateral nonloaded tibia when strains were less than 1050 $\mu\epsilon$. When strains were greater than this, bone formation increased monotonically by sevenfold as strain increased to 2000 $\mu\epsilon$. The idea that bone formation was linearly related to strain magnitude was not a new one, but the experimental demonstration that there was a threshold, and about where that threshold lay, was a new observation.

Nine years earlier, Rubin and Lanyon (1985) performed a similar experiment in which they applied loads to avian forelimbs from zero (unloaded, and in disuse) to 4000 $\mu\epsilon$ and noted a linear increase in cortical bone area related to increasing strain. They did not identify a threshold for response in their earlier experiment, but presented a linear response over the entire range of loads. In more recent work, this group has continued to assert this linearity of response over the range of strains and activities, from disuse, to very

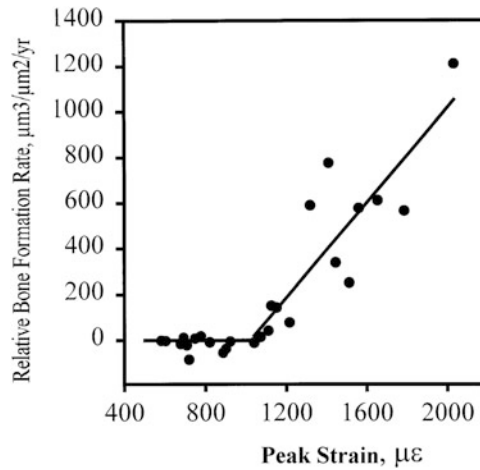


FIGURE 6.6. Apposition rate vs. applied bending load for endosteal surface of rat tibia (Turner et al. 1994a).

rigorous loads involving high strains with adaptation by woven bone (Sugiyama et al. 2012). However, when their data are viewed more closely, it becomes clear that there is a threshold for response, and that this threshold occurs slightly above 1000 $\mu\epsilon$, just as in the Turner experiment using rats (Fig. 6.6). Although presented as loads rather than strains, their more recent data also suggest a threshold driven process because mean values for cortical and cancellous bone areas and volumes do not change substantially when peak loads (in mice) are less than 6 N. These studies appear to validate the idea of a threshold for adaptive response, and imply that this threshold may be similar across species.

The concept of a regulatory threshold is consistent with the view that the osteocyte-canalicular system performs as a neural (Turner et al. 2002) or “small world” (Gross et al. 2004) network that connects osteocytes with osteoblasts and lining cells to enhance mechanical signals. As Gross points out, such networks exhibit threshold-related behavior, and also have the capacity, once the threshold has been reached, to synchronize the signal, which effectively can amplify it. This does not necessarily mean, however, that the strain threshold is exactly the same at all locations on bones. In fact, the idea of a “neural network” involving osteocytes suggests the opposite: the capacity for learning and memory, as in a true neuronal network. Using the rat ulnar loading model, Hsieh et al. (2001) showed that bone formation was initiated at strain thresholds that varied from about 1300 to 3000 $\mu\epsilon$ along the length of the ulna. These thresholds correlated to regions of the ulna that were normally under higher strain, suggesting that bone tissue habitually receiving higher mechanical strains can become less responsive to loading, i.e. the system “remembers” its local loading history and responds

accordingly. In other words, cells accommodate (or habituate) to their routine loading history and probably their hormonal environment, and only respond when the perceived stimulus is outside what they normally experience. The biological mechanism for such an accommodation has not been determined, but may reside in the osteocyte network (Turner et al. 2002).

Box 6.2 Cellular and Acellular Bone

Not all bone contains an osteocyte network. It has been known since the middle part of the nineteenth century that bone in some fish is acellular, although cells may be present during skeletal development. It may be, however, that fish have limited need for cellular signaling to repair bone damage. Being in an aquatic environment certainly lowers gravitational forces associated with muscle contraction. Nevertheless, muscle contraction requires calcium and so fish must use other mechanisms for systemic calcium regulation, and indeed they do. During development, humans develop a structure called the ultimobranchial body, which is eventually incorporated into the thyroid gland. Cells from this structure give rise to C cells in the thyroid gland that produce calcitonin, which helps to regulate plasma calcium levels. Fish also have an ultimobranchial body that produces calcitonin. When calcitonin is administered to fish with osteocytes, plasma calcium levels decrease. However, when calcitonin is administered to fish without osteocytes, plasma calcium does not change. To see reviews of this work, refer to Parenti (1986). This suggests that osteocytes are important for the systemic regulation of calcium in the blood, which they must perform by signaling other cells to resorb bone to release calcium stored there or stop resorbing when circulating Ca^{++} levels become too high. When osteocytes are not present, those signals are not generated and plasma calcium is invariant. Moss (1962) performed an interesting experiment to demonstrate differences in fracture repair between fish with cellular and those with acellular bone. Moss fractured the jaws of both kinds of fish, and observed the repair process under normal conditions in an aquatic environment with normal levels of calcium. In fish both with and without cells, fracture healing occurred via endochondral ossification by the formation of a large cartilage callus that subsequently calcified. When a similar experiment was performed in water that had no calcium salts in it, the fish with cellular bone healed and mineralized normally, but those with acellular bone failed to calcify the cartilage callus. Again, this suggests that fish without cellular bone are not able to internally regulate their calcium balance as do fish with osteocytes. These studies demonstrate the capacity of osteocytes to signal other cells, and the critical role played by osteocytes in metabolic regulation, but do not speak to the exclusivity of osteocytes in transducing mechanical signals. In this regard, there is preliminary evidence that acellular fish bone may have the capacity to form new bone in response to experimentally imposed mechanical loading, suggesting other, perhaps more primordial, mechanisms of mechanotransduction that function independent of osteocytes (Shahar et al. 2013). Regardless of eventual conclusions, cellular and acellular fish bone provides a natural model to further investigate the regulatory pathways responsible for bone growth and maintenance.

Thus, bone adaptation can be considered a location-specific and error-driven process. This has formed the basis for a number of computational models that attempt to predict bone's modeling and remodeling behavior, including those devised by Cowin (Cowin and Hegedus 1976; Cowin and Nachlinger 1978), Hart and Davy (Hart and Davy 1989), Carter and Beaupre (Beaupre et al. 1990a, b), and Frost (Frost 1987b). For this to be true, of course, the opposite must also be demonstrated: loss of a mechanical signal should result in bone loss.

6.5.4 Canine Disuse Experiments

Uthoff and Jaworski (1978) and Jaworski and Uthoff (1986) studied the effects of disuse in growing dogs by casting one of their forelimbs for 40 weeks. The metacarpal diaphyseal bone loss in growing dogs was characterized primarily by reduced formation and increased resorption on the periosteal surface, indicating that the normal periosteal bone formation accompanying growth was inhibited or converted to resorption. Relative to the contralateral control bone, bone mass decreased rapidly at first, then more slowly. After 40 weeks of immobilization, it was asymptotically approaching about half the bone mass of the contralateral control bone. When the cast was removed after 32 weeks of immobilization, about 65–70 % of the deficit in bone mass was recovered in 28 weeks. In the diaphysis, this occurred both through endosteal apposition and through additions of woven bone to the periosteal surface. When the experiment was repeated in older dogs, bone loss occurred primarily at the endosteal surface and through increased intra-cortical porosity (Jaworski et al. 1980). There was no effect on periosteal diameter. Older dogs were less able to recover lost bone following remobilization, recovering only 40 % of the lost diaphyseal bone. In older dogs, remobilization stopped the endosteal bone loss and caused apposition of small amounts of lamellar bone to the periosteal surface.

6.5.5 The Mechanostat and the Equilibrium Strain Range

Frost (1987a, 1986) attempted to put these ideas about how bone adapts together into a conceptual model, integrating the concepts to define not only when a response would occur, but the nature of the response (resorption or formation) and the location of the response on one of the four bone envelopes. He likened the modulation of bone adaptation to strain to a home thermostat, and therefore termed his conceptual model “*The Mechanostat*.”

The mechanostat theory begins with the concept that there is a *minimum effective* strain which must be exceeded to excite an adaptive response

(Fig. 6.6). However, instead of simply postulating that strains below a certain threshold will evoke no response, Frost suggests that there is an *equilibrium range* of strain values which will evoke no response. Strains above this range evoke a positive adaptive response (increased bone), and strains below this range cause a negative adaptive response (loss of bone). Others have suggested similar physiologic ranges, sometimes called “*dead zones*” or “*lazy zones*.” In this framework, the minimum effective strain becomes the upper limit of the equilibrium strain range, or the strain beyond which an augmenting adaptive response is initiated. There would also be a lower strain limit defining the strain below which bone loss would occur. The thresholds separating these zones or ranges are termed “setpoints.” These setpoints are determined by many factors in the hormonal and metabolic environment, and so cannot be viewed as fixed at a particular strain level. At very high strains, there is another setpoint which defines the strain range within which damage is severe enough to elicit a repair response. This can be either a remodeling response to repair internal damage to the bone matrix—microdamage—or rapid deposition of woven bone on external surfaces, such as is seen in a stress fracture.

Frost’s mechanostat theory asserts that modeling and remodeling have opposite responses to loading that pushes strains above or below the equilibrium range:

Increased loading. modeling is increased and remodeling is inhibited.

Decreased loading. modeling is inhibited and remodeling is increased.

The logic behind this algorithm is as follows. While the inhibition of remodeling by high strain may seem “backwards” because of the long-held concept that increased loads or strain stimulate “remodeling” to increase bone strength, it must be remembered that we are using the strict definition of remodeling, which does not include modeling. Now, think first of cortical bone. Osteonal remodeling increases porosity as a result of the remodeling space and the new Haversian canals created. It is therefore desirable that osteonal remodeling should be depressed by overload and stimulated by disuse. However, in those cases in which loads are so high that bone is damaged, it is desirable not to reduce remodeling, but to stimulate a repair reaction. We now know that the mechanostat theory is not entirely correct because remodeling is indeed activated at very high loads, when bone is damaged and dying osteocytes produce signals that recruit osteoclasts to remove the damage. It is also reasonable that mechanical overload should activate modeling to adapt the bone to future, similarly high, loads and, as we have seen, this has been observed experimentally in many different models.

It is important to understand that an identical mechanical strain can elicit a different response if the setpoint for response is altered. For example, if the setpoint to activate remodeling were raised, the bone could perceive that normal physical activity was decreased when it was not, activating an inappropriate remodeling response causing loss of bone. This may be in part what occurs in postmenopausal osteoporosis. With the loss of the estrogen

stimulus, the lower setpoint for remodeling is higher, causing a loss of bone from endocortical, trabecular and intracortical surfaces even in the presence of normal activity and normal strains. Parathyroid hormone (PTH) is well-known to increase osteoblast sensitivity to a mechanical load (Ryder and Duncan 2001; Li et al. 2003; Kim et al. 2003). When given intermittently, as in therapies for osteoporosis, it elicits bone formation directly on endocortical and trabecular surfaces, i.e. a modeling response. Thus, the upper threshold to stimulate modeling is lowered, making the bone “see” usual strains from normal activity as an error signal.

6.5.6 *Mechanical Compensation and the Roles of Bone Modeling and Remodeling*

The competing architectural demands of remodeling and modeling show that mechanical adaptation occurs at both the structural level and at the tissue level, and these processes must be somehow coordinated. Shape and size changes are effected on external (i.e., periosteal and endosteal) bone surfaces, where there is more modeling than remodeling during growth, and less capacity for change in adults (Frost 1986). On the other hand, changes in material properties are effected on internal (i.e., Haversian and trabecular) bone surfaces. Internal bone surfaces are governed primarily by remodeling throughout life (Frost 1986).

The gross anatomy (i.e., the shape and size) of a bone is in large part determined during growth. Modeling serves to sculpt the bone to suit both the genetic plan and the demands of current mechanical usage. In adults, modeling is reduced in rate and extent, so that alterations of bone geometry require much more time (or a mechanical stimulus large enough to produce woven bone). This restriction is important with respect to joints as well as bones. An adult who greatly increases the loads on his skeleton is, in principle, more likely to acquire osteoarthritis than a juvenile because his bones have insufficient capacity to alter their metaphyseal size to reduce joint stresses (Frost 1994).

The mechanical properties of the bone material are matched to mechanical usage by remodeling. Unlike modeling, remodeling rates can be substantially increased in adults by various events, including changes in mechanical loading. But the purpose of remodeling is to maintain bone and to repair material damage, not to change gross architecture; large architectural adaptations to mechanical loading can only be effected by modeling processes. This limitation of remodeling is significant because it is practically impossible to substantially increase the strength or stiffness of cortical bone by remodeling, there being no large voids to be filled. Therefore, in adults it is much easier to reduce cortical bone strength than to increase it.

Because of these competing goals of modeling and remodeling, in both health and disease there are compensatory modifications of structure and

material that maintain or adapt the bone to its mechanical environment. For example, postmenopausal osteoporosis involves loss of cancellous bone mass in the spine, which reduces the failure stress of the vertebral bodies. However, in men there is a concurrent increase in vertebral diameter that prevents the structural strength from being reduced to the same extent as the material strength (Mosekilde and Mosekilde 1990). A similar phenomenon occurs in long bone diaphyses (Martin and Atkinson 1977; Martin et al. 1980; Ruff and Hayes 1988; Ahlborg et al. 2003) and the neck of the femur (Beck et al. 1993), although others (Poole et al. 2010) have shown a more complex re-distribution of bone with age in the femoral neck. This form of compensation does not occur as effectively in women, for reasons that are not entirely clear. In part the difference between men and women may be related to both size differences and material property differences. Cancellous or intra-cortical bone loss in smaller bones requires greater compensatory periosteal bone formation to maintain a given stiffness than it does in larger bones. Moreover, reduced total bone stiffness during aging, which can be caused by the increased rate of remodeling at the menopause in women, requires a much greater periosteal response to maintain stiffness than does a similar percentage loss of bone mass (Jepsen and Andarawis-Puri 2012). For example, a 2 % reduction in bone apparent modulus might require twice as much apposition on the periosteal surface to maintain stiffness, as does a 2 % reduction in bone mass.

These adaptive functions between structure and material occur even during serious pathologic episodes, resulting in the strongest structure that circumstances allow. For example, human femurs drastically alter their cross-sectional geometry (by modeling of the periosteal and endosteal surfaces) to compensate for the poor material quality and large bending stresses arising from curvatures produced by rickets. Another example is shown in Fig. 6.7, which shows silhouettes of young (left) and old (right) femoral



FIGURE 6.7. Midshaft cross sections of femur from a woman in her thirties (*left*) and from a woman in her nineties (*right*).

midshaft cross sections. The neutral axis for bending during gait is approximately from the top (anterior) to the bottom (posterior) of the sections. Note that the increased porosity associated with aging is more apparent in endosteal regions and quite striking near the neutral axis, where the resulting reduction in the bone's failure stress has minimal effect on torsional and bending strength.

6.5.6.1 Cellular Habituation: Do Cells Get Tired?

The idea that bone adaptation is threshold-related (strain magnitude) and error-driven (strain magnitude and distribution) led to the notion that as long as bone is loaded above its threshold, or in an unusual direction, cells will continue to respond until bone adaptation (which after all, takes some time to achieve) lowers the strain magnitude to below some response "setpoint" (in Harold Frost's terminology). We know that bone responds to a very few number of cycles if the strain magnitude is high enough, and may not respond at all to many cycles of loading if the strain is insufficient to achieve a certain threshold. But what if bone is repeatedly loaded above the strain threshold, will that threshold eventually be insufficient to cause a cellular response? In other words, do cells themselves "adapt" to strain magnitude and become de-sensitized to it over a period of time?

Experiments suggest that cells do become less sensitive, and that the timeline and/or number of loading cycles in which this occurs can be quite short. The idea that cells may become less responsive was embedded in the early experiments of Rubin and Lanyon (1984), but because those experiments were focused on the initiation of bone adaptation, the process of habituation was not recognized from these experiments.

The Rubin and Lanyon experiments showed that, if the strains were sufficiently high, bone loss could be prevented with only 4 cycles/day of loading, and bone formation (in modeling mode on periosteal surfaces) would be stimulated with only 36 cycles per day. These experiments also showed, however, that bone formation was no greater when bone was loaded 100 cycles/day, or 1800 cycles per day at the same strains, clearly showing that just increasing the number of cycles applied did not stimulate the cells to work harder. Other studies following the early Rubin and Lanyon experiments, using different experimental designs and different animal models, suggested that when bone is loaded to about 1000 $\mu\epsilon$, the adaptive response saturates within about 200–400 loading cycles (Umemura et al. 1997) (Fig. 6.8). These findings led Robling and Turner (Robling et al. 2000, 2001a, b) to use the four point tibial bending model in rats to deliver identical numbers of load cycles each day (i.e., an identical energy input), but to vary the timing of load delivery. They either delivered the load as 360 cycles at one time, or divided those 360 loading cycles into 2, 4 or 6 bouts of 180, 90 or 60 cycles in each bout, separated by different periods of time over the course of a single day. One would expect that, without cell saturation, it wouldn't matter how the 360 cycles were delivered, since the amount of energy

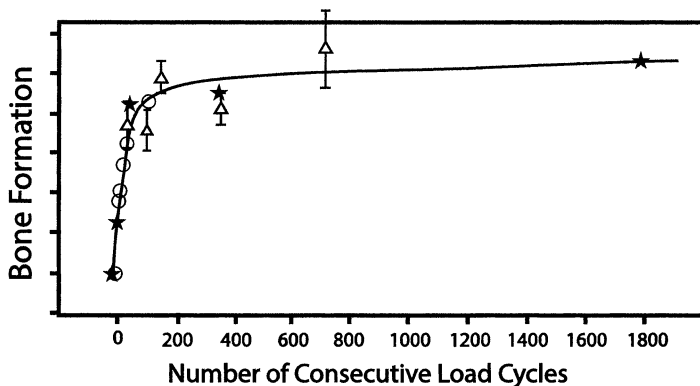


FIGURE 6.8. The relationship between bone formation and the number of load cycles per day based on experiments in roosters (★, Rubin and Lanyon 1984), rats subjected to jumping exercise (○, Umemura et al. 1997) and four point tibial bending in rats (△, Turner et al. 1994b). In combination, these data demonstrate that a small number of loading cycles creates the signals necessary for bone formation. Additional loading cycles do not increase the amount of bone formation or bone mass.

imparted to the system in each case was identical. However, these experiments showed that it did matter, and that bone formation rates (BFR/BS) were significantly increased when the load was delivered in bouts of 60 or 90 cycles each over the course of the day, compared to a group in which the 360 cycles were delivered at one time.

Box 6.3 Adaptation to Load Versus Evolutionary Adaptation

In the theory of evolution an *adaptation* (or adaptive trait) is a trait of an organism that has a current functional role. In evolutionary theory *adaptation* refers both to the trait and to the evolutionary process that led to the adaptation. For a theory of how bone responds to load the word *adaptation* means that the organism is using *phenotypic plasticity* to improve bone function. The two meanings of the word adaptation are similar, however, *evolutionary* adaptation changes the genotype and *load* adaptation is a somatic (phenotypic) adjustment that (as far as we know today) is not heritable. For bone adaptation the implication is that the bone is improved (in some sense of the word) by the somatic changes. This idea of improvement of the bone in response to loading is a *theory*, albeit one with substantial experimental support. Current experimental results do show that many of the changes in bone in response to loading should reduce the risk of fracture by, for example, increasing the cortical width in a bent bone or by reducing the curvature of an axially loaded long bone. However, this is not to say that all of the changes of bone in response to load make it better mechanically—for example, a blow on the shin can injure the periosteum and cause formation of a permanent bony lump that doesn't appear to have any mechanical function, but may have a *biological* function in repairing tissue damage. Therefore, in

general, changes of a bone in response to changed loads seem to make the bones mechanically better, but sometimes they do not. On balance, most of the responses to loading do seem to be “for the better” so it is not wrong to call the process bone *adaptation* as long as any load related changes that are mechanically irrelevant (or are mechanically maladaptive) are not forgotten or ignored.

These data convincingly demonstrate that cells habituate rather quickly to repeated loading (a relief to those of us who don’t care to run 5 miles a day to maintain our bone health). At compressive strain levels of approximately $1500\ \mu\epsilon$, cell accommodation to load occurs within the first 100–200 load cycles (because the loading cycles were delivered at 2 Hz, this represented less than 2 min. of cyclic loading). However, these experiments did not demonstrate the length of the refractory period. How long does it take cells to become re-sensitized to loads once they have become saturated? A second set of experiments was designed in which all groups were given 4 bouts of 90 cycles per bout, and the time between the bouts varied from 30 min to 8 h (Robling et al. 2001b). Bone formation rate was maximized when bouts were separated by 4–8 h (Fig. 6.9). However, re-sensitization did not occur all at once, but was a gradual process that extended beyond the 8 h inter-loading period used in these experiments. Inter-loading periods longer than 8 h were not examined, but extrapolating data from the shorter periods suggests that return to full response may require as long as 12 h (Fig. 6.9).

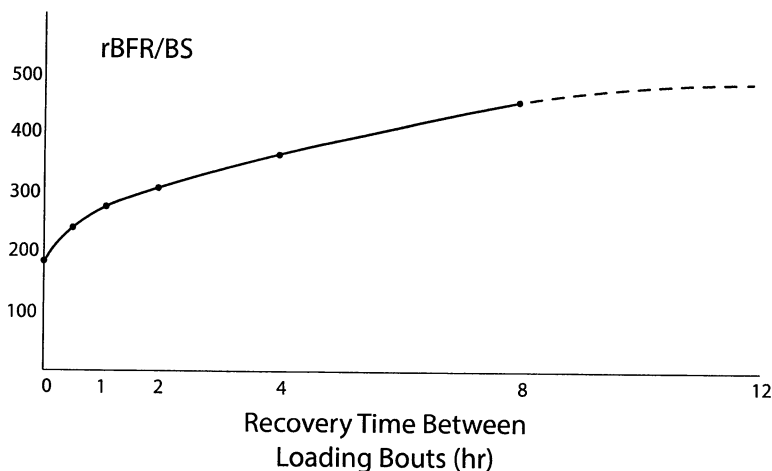


FIGURE 6.9. Bone cells have a refractory period following loading, and only gradually recover their full sensitivity to subsequent loads. Osteoblasts recover most of their responsiveness to loads within about 8 h, but extrapolation from that data indicates that full recovery may not occur for about 12 h following a loading session (Robling et al. 2001b).

Related studies have been performed using low strain, high frequency vibrations applied to mesenchymal stem cells (MSC) in culture (Sen et al. 2011). In this case, refractory periods of 1–3 h between loading bouts enhanced the response of the MSCs, more so with the longer refractory period. This effect appeared to be independent of the magnitude or the frequency of the loading stimulus, further suggesting that strain magnitude may not be the predominant stimulus for bone adaptation. Taken as whole, these results have significant implications for the design of training and exercise programs intended to enhance skeletal health. On the one hand, skeletal health can be maintained or improved by several short periods of exercise over the course of the day. It is also true that, if skeletal health is the goal, injuries such as stress fractures can be avoided by designing training programs that include modest amounts of exercise. However, it should be kept in mind that short bouts of exercise are insufficient for proper maintenance of cardiovascular fitness.

6.5.7 Optimization and the Mathematical Basis of Wolff's Law

Our assumption for this discussion is that *bone structure is regulated to find the optimum solution to minimize both fracture risk and bone mass for the loadings that the bone experiences*. A mathematical difficulty in this formulation is that we do not specify whether this is a “global” or a “local” optimum for the bone. A local optimum is a state of bone mass and geometry that is “better than any close-by.” That is, small changes in the mass or structure make the bone worse (higher fracture risk or greater bone mass) for the loading experienced. A global optimum is (not to be too specific) the best local optimum in the whole range of the possible variation of the system—as an example, the absolute minimum fracture risk for the given loads. The fact that bone could have more than one local minimum means that a real (as opposed to a mathematical) bone may not reach the “best solution” during adaptation because the changing tissue could easily fall into a local optimum where *any further changes would be stopped* because all perturbations of the tissue from the local optimum make it *worse* with-respect-to the load to which it is adapting.

Experiments performed by Robling et al. (2002) suggest that this could be happening in real bones. Using an experimental approach similar to that previously described, rat ulnae were given 360 loading cycles, either delivered at one time or in 90 cycle “bouts” four times throughout the day 3 days per week, for 16 weeks. They discovered that bones exposed to 360 cycles at one time never achieved the strength or mass of those that were loaded in shorter bouts. After 4 months, bone formation rate (BFR/BS) was threefold higher in the group which received the loads in bouts with fewer cycles. In terms of the remodeling cycle, 4 months in a rat is about equivalent to 1.5–2

years in a human, certainly sufficient time for bone to fully adapt to the mechanical stimulus. Although stiffness as a specific endpoint was not investigated, bone mineral content, whole bone areal bone mineral density, cross-sectional area and rigidity in the direction of bending at the midshaft were all 40–50 % greater in animals that received the loads in smaller doses. This is a remarkable result—an identical input of mechanical cycles resulted in two different “optimal” structures.

The explanation for there being two results was that cells become desensitized following relatively short periods of stimulation, and require a period of recovery before they can fully respond again (Robling et al. 2002). Another possible explanation is that the loading with shorter groups of cycles found a different local optimum because the bone perceived the two loadings as *different*. The bone does not have any particular long-term agenda, but only a series of short-term goals that reduce its risk of breaking. But the experiments raise a series of other important questions: If bone loading continued in short bouts, would the bone continue to respond, and for how long? Which cell populations become desensitized—the sensors (osteocytes) or the effectors (osteoblasts)? These are all exciting questions that remain to be answered by biologists and engineers.

6.5.8 Some Final Thoughts on Optimization, Adaptation and the Scientific Method

We have placed much of the previous discussion of bone’s response to load in the context of an “adaptation” or the seeking of an optimal state for the bone structure under the given loads. The idea of optimization of a mechanical structure is very seductive both for explaining the adaptation of bone to mechanical loading and, in fact, for explaining the process of evolution. It is unfortunate that it is wrong for both! The problems with this idea were beautifully presented by the celebrated evolutionary theorists Stephen Jay Gould and Richard C. Lewontin in a highly notable paper from 1979 (Gould and Lewontin 1979). In that paper they strongly questioned the idea of “natural selection as so powerful and the constraints upon it so few that direct production of adaptation through its operation becomes the primary cause of nearly all organic form, function, and behavior.” Viewed in this light, bone will adapt by finding an *acceptable* solution to a biological problem, but this solution may not be the one an engineer would devise as optimal, either in the short term or in the long-term.

The weakness of a bone adaptation theory that uses the concept of optimality is an assumption that natural selection is so strong that it produces a set of genetically defined mechanisms that result in *mechanically optimal somatic adaptations of bone* to applied load. In other words, optimization theories for bone adaptation are weak for the same reasons that optimization arguments are weak in evolutionary theory. In general, when an

optimization theory is proposed as a broad explanation for the adaptation of bone to load it generally ignores the biological constraints or, more frustratingly, acknowledges and then ignores them.

Taken to the extreme, one might think that any use of optimization principles for understanding bone changes in response to load would be completely useless. Fortunately for the engineer interested in bone, optimization isn't completely useless; instead it provides a useful *ideal* case that can be used to help build experiments. For example, in the previous section some of the concepts of optimization were used to better understand the experiments of Robling et al. (2002) who deliberately tested whether “identical loadings” (i.e., same number of cycles of the same kind) applied with different timing would produce the same biological result. The researchers formulated an experiment that could test whether the two loading scenarios were truly the same in terms of the adaptive response they elicited—with the underlying assumption (inherent to optimization) that there is a single “best” result for identical loads. Their experiment worked well because the results were *not* the same, leading to the conclusion that the loading scenarios were perceived as different by the real biological system, despite any interesting thoughts related to structural optimization based purely on mechanical input.

So we see that a theoretical construct for bone adaptation (that could be formulated as an optimization theory) was used to design a successful and very informative experiment on real bones. But what if the two outcomes had been the same? Could the experimenters then claim that bone uses an optimization scheme to find the single best solution? The answer is a resounding no because *despite* having used some ideas from optimization approaches to create the initial hypothesis it is a scientific mistake to use the results of an experiment to proclaim as truth the reasoning that led to the hypothesis on which the experiment was based. Remember that a scientific hypothesis can only be falsified and cannot be shown to be correct. This is because the result of any test of that hypothesis is only valid for the conditions under which the test is run. If the hypothesis cannot be invalidated, that does not suggest that it is universally true, nor does it suggest that we have even measured the proper outcomes.

A second less appreciated and perhaps more conceptually difficult reason is that even if a hypothesis is in fact true, the idea or reasoning leading to the stated hypothesis is totally irrelevant. To make a joke to illustrate this second point—an immigrant Indiana Leprechaun could have told Robling et al. that the bone adaptation results would be the same for the two loadings (Begorrah!). Despite magical intervention to create the hypothesis, the experiment would still be valid and falsify the hypothesis in a meaningful manner. And if the results had been same (i.e. the hypothesis was not rejected) the researchers could not have claimed with certainty that it was due to leprechauns, magic, biological optimization or any other reason. In this feeble joke, you can identify optimization theory with the leprechaun—whatever

the outcome of the experiment, the source of the original hypothesis is not relevant and the validity of optimization isn't verified. So, maybe the *hypothesis* is correct—but the source for the hypothesis (leprechaun or optimization) falls outside of the scientific method. This difficulty underlies all theories that are accepted as true natural laws. Even the extremely validated and useful Newton's laws of motion give us only a *hypothesis* that the sun will rise in the east tomorrow. Even if it rises tomorrow (and forever after for the remaining history of our species) that doesn't mean that the sun is governed by Newton instead of Ra—because *why* we think it will rise doesn't have anything to do with whether what we think is true or false. For all practical purposes, of course, theories and rules that make many correct predictions (many correct hypotheses) are eventually acceptable as scientific truths; but even the extremely useful Newton's laws of motion aren't correct—ask the shade of Dr. Einstein. For optimality based bone adaptation theories there are more exceptions to the rules than there are to Newton's—so all theories approaching bone's response to load from that direction should be used carefully as tools rather than accepted as truths.

6.5.8.1 Statics and Dynamics in Bone Adaptation

The experiments explored in this chapter have demonstrated that very few load cycles are necessary to maximize bone's adaptive response. However, they also showed that there is a lower limit of loading below which bone is lost. That means that one single loading event is not sufficient to prevent bone loss. Taken to its natural extreme, a single loading event is the same as one statically applied load, suggesting that dynamic loading is required to generate an adaptive response in bone. This, in-turn, suggests that bone adaptation is a rate-related phenomenon. We already know that bone is responsive to strain magnitudes that exceed local thresholds or setpoints. Is bone also sensitive to strain rate and if so, what rate is optimal, what are the thresholds (if any), and what is the relationship between strain rate and strain magnitude?

It has been known for 50 years or more that bone does not respond to a static load unless, as in the case of orthodontia, a dynamic load is imposed on top of the static load. Static loads, even those that are quite high, do not elicit any adaptive response. Early experiments by Hèrt in 1969 first suggested that dynamic loads were necessary for bone to respond. Later, this was underscored by Lanyon and Rubin (1984), who applied static loads to the avian ulna for 2 months and observed bone loss endocortically and increased intracortical porosity, changes identical to those found in immobilized animals. Dynamic loads at equivalent peak strain increased net bone formation periosteally, and prevented loss both endocortically and within the cortex. Mosley and Lanyon (1998) followed this work by investigating low, moderate and high strain rates using the rat ulnar axial loading model. In this work, they found that low strain rate loading suppressed the normal periosteal bone formation response, whereas high strain rate loading significantly increased

the amount of periosteal new bone. These were growing animals, however, and other studies have shown that low to moderate impact loads can maintain skeletal mass and reduce the amount of loss during aging

Lanyon and Rubin (1984) observed that when no loading was applied over an 8-week period, the ulna suffered endosteal resorption and increased osteonal remodeling. Compressive static loading to 2000 $\mu\epsilon$ for the same period of time led to a similar result, but the same strains applied dynamically at 1 Hz produced a 24 % increase in cortical area, primarily from woven bone formation on the periosteal surface. Subsequently, Rubin and Lanyon (1985) did a dose-response experiment for 1 Hz loading in a bending mode such that peak compressive strains were about twice the magnitude of the tensile strains. When 100 cycles/day were applied at peak strains from 0 to 4000 $\mu\epsilon$, there was a linear relationship between cortical area after 8 weeks and peak strain magnitude (% change in area = $-15.3 + 0.015 \times$ (peak microstrain), with $R^2 = 0.69$). Setting the left side of this equation equal to zero indicates that an equilibrium loading for maintenance of bone mass (zero change in area) would be 100 cycles/day at about 1000 $\mu\epsilon$. This was an important benchmark. These investigators did another frequently cited experiment (Rubin and Lanyon 1984) in which 2050 $\mu\epsilon$ was applied at 0.5 Hz in daily loading bouts of 4, 36, 360, and 1800 cycles. It turned out that only 4 cycles/day were sufficient to maintain bone mass at this strain magnitude. After 4 weeks of loading, each of the three higher numbers of cycles produced the same increase in bone mass, about 40 ± 10 %, indicating that the woven bone response was an all-or-nothing phenomenon.

Subsequently, an experiment performed by Robling (2001c) using the rat axial ulnar loading model further demonstrated this concept. Bone formation on the periosteal surface in rats loaded statically at 8.5 and 17 N showed the same amount of bone formation, and was not different than formation on the non-loaded contralateral limb. Even at double the load, a static application did not generate a response; however, when a dynamic load was applied at 17 N for an equivalent length of time, there was a large proliferation of bone on the periosteal surface. Thus, loads of equivalent magnitudes have very different effects, depending on whether they are applied statically or cyclically.

6.5.9 How Frequently Should Dynamic Loads Be Applied: The Role of Strain Rate

Dynamic load is required for bone's mechanical adaptation and the rate of load application can be fast or slow for the same peak load. How fast the load is applied could be an important signal to bone or bone cells, since we already know that a static load (a rate of "zero") doesn't stimulate them. Exercises that involve loads applied at higher rates, such as running, jumping and so forth, are more osteogenic than other forms of exercise in many experiments. Rough estimates have suggested that 68–81 % of the variance in bone surface

adaptation may be accounted for by strain rate (O'Connor et al. 1982). A recent study (Deere et al. 2012) reported that in adolescent boys only high impact (>4.2 g) loads were related to significant increases in femoral neck bone mineral density and that accelerations lower than that had no effect. High impact loads were associated both with increased periosteal bone formation (measured by hip structure analysis) and reduced endosteal width. Studies prior to 2012 similarly discovered that jumping exercises in young boys and girls had a positive effect on hip structural properties, including strength (Fuchs et al. 2001; Kato et al. 2006).

Experiments using the rat ulnar axial loading model have demonstrated the importance of strain rate to bone response. When loaded to a high strain, at strain rates corresponding to walking, (0.01/s), running (0.03/s), or jumping (0.05/s and higher), the response is 54 % higher at the strain rate for jumping than at the strain rate for running, and even greater compared to the rate generated during walking (Mosley and Lanyon 1998). Ulnar loading at frequencies that are 5–10 Hz are associated with higher periosteal bone formation rates than those delivered at equivalent strains but only 1 Hz (Hsieh et al. 2001). Thus, given an equivalent strain, bone will respond more strongly at higher rates of load application.

How can this be reconciled with the idea that bone formation initiates in response to strain magnitude? This is because strain rate depends on strain magnitude (ϵ) and loading frequency (f). For a sinusoidal strain ($\epsilon(t) = \epsilon_{\max} \sin(2\pi ft)$):

$$\text{Strain rate} = 2\pi f \epsilon_{\max} \cos(2\pi ft) \quad (6.1)$$

so that as maximum strain and frequency increase so does maximum strain rate. During our normal gait cycle, strain is (approximately) applied in a sine wave, and any increase in frequency, e.g., when we are running, will be translated into an increased strain rate (Fig. 6.10). Experimentally, rate and

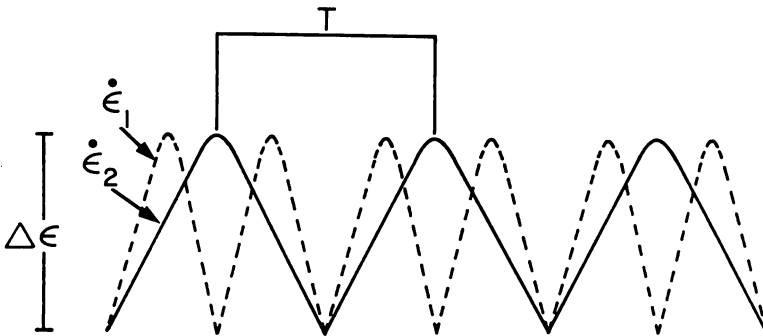


FIGURE 6.10. Two strain-time curves showing that as frequency increases, strain rate increases when the peak strain is the same. Frequency = $1/T$ where T is the interval between peak loads. $\Delta\epsilon$ = peak strain magnitude. The $\dot{\epsilon}$ are differing strain rates for the low and high frequency.

frequency can be dissociated by utilizing a triangular waveform for loading and varying the period between cycles to alter either strain rate or frequency independently. Using this concept, Turner et al. (1994a) showed that given the same peak strain and strain rate, frequencies above 0.2 Hz will generate an adaptive response, whereas those below that will not. Given our natural walking cycle of about 1 Hz, this is encouraging and “makes sense” as a story about the evolution of mammalian physiology. In a second experiment, Turner and colleagues (Turner et al. 1995a) loaded rat tibiae in bending to the same maximum load (54 N). The load was separated into a *constant* component plus a 2 Hz sinusoidal component of varying magnitude (0, 18, 36, or 54 N) so that the constant plus peak sinusoidal load was always 54 N. In this experiment, cycling over a strain range of 0 is equivalent to static loading (loading to 54 N and holding it there), and one would not expect any response. This is what occurred. But over the range of constant-plus-a-sinusoidal-component of 18–54 N, there was a dose-response effect with increasing bone formation rate as strain rate (peak sinusoidal component) increased.

Taking these two experiments, one in which frequency was varied, and one in which strain rate was varied, Turner (1998) calculated the total *strain stimulus* (S_σ) as

$$S_\sigma = k_i \sum_{i=1}^n \epsilon_{\max_i} f_i \quad (6.2)$$

where “i” denotes the number of loading types, k_i is an empirical constant, $\epsilon_{\max_i}$ is the peak strain achieved and f_i is the frequency of load application. This equation looks very much like Eq. (6.1) for strain rate, indicating that over a series of cycles, the adaptive response (S_σ) will be proportional to strain and frequency. When the strain stimulus is calculated for the two experiments in which either frequency or cyclic strain magnitude are varied, the relationship between strain stimulus and bone formation rate is almost identical. This demonstrates that it is a combination of strain magnitude and frequency—strain rate—that drives the adaptive response.

6.5.10 Effects of High Frequency Loading: Vibration

Since strain rate is an important factor in skeletal adaptation, skeletal mass and strength could be governed by the animal’s most dynamic activities even if those activities are relatively infrequent. The relationship between strain and frequency suggests that using high loading frequencies could minimize the actual strains that are required to elicit a bone adaptive response. This is the theory behind recent attempts to use high frequency vibrations, which are applied at extremely low strain magnitudes, to increase bone formation or to reverse the effects of bone loss.

Several variations of vibrational “exercise” have been explored (and marketed). In one, high frequency sinusoidal vibrations that generate higher lateral than vertical accelerations are used. Another form uses vertical displacement to generate higher strains along the longitudinal axis than transversely (Rittweger et al. 2001; Gusi et al. 2006). A further division can be made between those devices that deliver high frequency and large accelerations (>1 g) and those that deliver high frequencies at lower intensity (<1 g) (Wysocki et al. 2011). Although these devices have been associated with strong tonic muscle contraction (contractions up to 82 % of maximal muscle contraction—enough to cause tomorrow’s muscle pain!), and may affect biochemical indices of bone resorption, they have not been shown consistently to have any significant effect on bone density (Elmentaser et al. 2012; Iwamoto et al. 2012).

6.5.11 When Theory Meets Reality

Have you ever heard someone say: “Well, at least in theory this should work”? When an experiment doesn’t conform to theory, either the theory is wrong, or something has been overlooked in the experiment. According to our theory about the importance of strain rate, low strain-high frequency loading should be effective at building bone if the strain rates created are sufficiently high. So far, however, this just doesn’t seem to be correct. Let’s explore the reasons why this promising theory has predicted an approach that has not been effective.

One possibility is that there really is a strain magnitude threshold for response of cells, and that very low strains are just not sufficient to cross that threshold, no matter how many of them there are. In animal models, the high frequency (15–30 Hz) strain signals are more than 20 times smaller in magnitude than low frequency (<15 Hz) strain components. Cell culture experiments have demonstrated repeatedly that connective tissue cells (e.g. osteoblasts and fibroblasts) are highly sensitive to frequencies in the 15–30 Hz range. We discussed earlier in the chapter that there appears to be a threshold for the initiation of bone formation at about 1000 $\mu\epsilon$, but some investigators have shown that strains of only 250 $\mu\epsilon$, when delivered at 15 Hz, are sufficient to induce bone formation (McLeod et al. 1990). However, strain gage experiments using dogs walking on a treadmill (Turner et al. 1995b) show that strains within the 15–20 Hz frequency range are <50 $\mu\epsilon$, and even lower at slightly higher frequencies. There is no convincing evidence that strains this low will cause net bone formation. It may simply be that there is insufficient energy in that signal to create a cellular signal *in vivo*.

A second possibility is that strain and strain rate are not a direct stimulus to the sensor cells, and that downstream events are affected differently by very high frequency loading. The idea that strain and strain rate are a direct stimulus to cells implies that the effector cells (the ones doing the work,

osteoblasts and osteoclasts) sitting on bone surfaces sense deformation when the bone is deformed. That is, there is a direct link between the mechanical stimulus and the cell that produces the response, but there is good evidence now that *osteocytes* are the mechanical sensor cells in mammalian bone (Bonewald 2011), and that they are the primary cells controlling the induction of bone remodeling. Osteocyte death stimulates bone resorption in cases of estrogen deficiency or disuse, and is directly associated spatially and temporally with bone loss (Tomkinson et al. 1997; Noble et al. 1997; Divieti 2009). Osteocytes and osteoblasts produce osteoprotegerin (OPG) and RANKL (a ligand for a receptor on osteoclasts) in response to bone damage, eliciting osteoclast recruitment and activation, and subsequently bone resorption to repair damage (Noble et al. 2002; Cardoso et al. 2009; Kennedy et al. 2012). In mice with RANKL removed only in osteocytes, bone loss does not occur in response to unloading (Xiong et al. 2011). Moreover, mice in which osteocytes cannot produce RANKL become osteopetrotic (Nakashima et al. 2011).

On the other side of the bone remodeling coin, the formation of new bone on periosteal surfaces following mechanical loading is associated specifically with the downregulation of sclerostin production by osteocytes (Robling et al. 2008). Transgenic mice engineered to maintain high levels of sclerostin during mechanical loading suppressed new bone formation by 70–85 % during mechanical loading (Tu et al. 2012). Moreover, two osteocyte specific proteins related to matrix mineralization, DMP1 and MEPE, are upregulated after mechanical loading (Harris et al. 2007). Recent evidence has shown that variations in new bone formation within a single bone (using the mouse tibia) are more closely related to the absence of sclerostin production by osteocytes than peak strains (Moustafa et al. 2012). All of these lines of evidence suggest that osteocytes sense mechanical loading in some manner. However, osteocytes sit within a space within bone, and although they are tethered to the walls of their lacunae by proteins (the glycocalyx) and could potentially “feel” deformation, it is likely that they are protected to some degree from significant amounts of deformation. How then, can they sense a mechanical stimulus? Or perhaps the better question is: what mechanical stimuli are they sensing?

6.5.12 *Is Strain a Direct Stimulus to Bone Cells? Do Strain Gradients and Distributions Regulate Regional Bone Cell Activity?*

Changes in surface curvature induced by bending create stress gradients perpendicular to the surface. Increased surface convexity is associated with tangential strain that become more tensile as one moves toward the surface (Fig. 6.11). Decreased surface convexity (or increased surface concavity) is associated with strain that become less tensile nearer the surface. Because

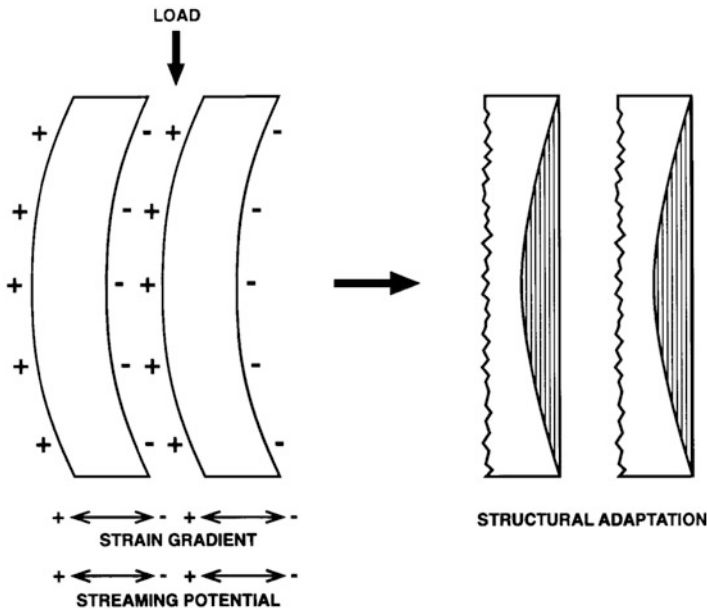


FIGURE 6.11. Bending is associated with stress (and strain) gradients. If the polarity of these gradients is as defined in the text, then they would have + and – signs as shown in the longitudinal section at *left*. The streaming potentials produced by these stress gradients would have similar polarity. If *positive* and *negative* signals produce resorption and formation, respectively, the bone would straighten as shown in the *right* diagram. Computational analyses are consistent with the prediction that bone formation will be enhanced in regions where curvature-induced bending moments are high. Bones will not completely straighten, however, perhaps because some curvature is necessary to control stresses generated by limb acceleration (Dodge et al. 2012).

there is not a perfect correspondence, however, between experimentally measured strain magnitude and specific locations of bone formation, and also because that model would lead to results that were inconsistent with observations of bone adaptation, Frost suggested that it is these strain gradients, rather than the magnitude of the strain itself, which regulates the location and magnitude of bone adaptive response. This idea has been tested experimentally using several different avian models. Although strains, by necessity given the limitations of current technology, are typically measured on external surfaces of bone, loads create distributions of strain within the bone that result in a strain gradient, which is affected both by the properties of the bone and by its external shape or curvature. This idea was tested by Rubin and Lanyon (1984, 1985) and Rubin et al. (1996) in a set of classic experiments using turkeys and roosters in which they altered the direction of application of a load in a limb otherwise deprived of any load bearing. They

discovered that when bone was loaded in directions in which the bone was not loaded during normal movements (in this case 90° to that for wing flapping), periosteal or endosteal woven bone could be produced at loads below those required to cause bone formation under normal conditions.

Thus, when the strain distribution is usual, the threshold to initiate an adaptive response may be higher than it is for strain distributions produced by unusual or less frequent loading conditions. The implication of this is that bone is adapted to a range of strain situations, and that adaptive responses are more easily generated by a variety of different loading conditions than by increased repetitions of the normal pattern of loading. This may be one reason that exercise physiologists find it so challenging to identify a significant bone response to exercise in adult people who may be already relatively well adapted to the activities they are asked to perform. The other side of the coin, however, is a cautionary tale—attempting to understand bone's adaptive response to loads that are applied in non-physiological directions in an animal model may provide solutions to problems that don't exist!

An elegant demonstration of the complexity of strain distribution to bone adaptation (Gross et al. 1997) used the turkey ulnar osteotomy model to isolate the radius, which then was subjected to a 4 week experimental loading protocol after which periosteal bone formation was measured in 24 different sectors around the circumference of the radius. Interestingly, there was bone formation in every sector, and the amount of periosteal new bone was not well predicted by the levels of strain. New bone in every sector would not be predicted under a model of adaptation that only considered strain magnitude; one would reasonably expect that bone formation close to the neutral axis would be absent or negligible. However, about 60 % of the new bone formation was predicted by a combination of circumferential, radial and longitudinal strain gradients. A subsequent study subjected intact roosters that had strain gages attached to their tarso-metatarsal bones to high speed running for up to a year. Again, it was discovered that peak strains and even peak strain rates were poorly correlated with the amount and location of new bone formation, but that circumferential strain gradients predicted 63 % of the response.

These experiments address one question that had perplexed scientists for years. That is, if bone only adapts to strain magnitude, then why doesn't the bone along the predominant neutral axis simply disappear over time? One answer to this, of course, is that we don't use our limbs always in the same ways, and so a specific place in a bone may not always see very high or very low strains. Another explanation is that bone is not optimally adapted (remember the discussion earlier in this chapter) but only finds biologically acceptable solutions. However, understanding the more complex aspects of strain distribution and gradients in bone, and how it changes with different activities, may provide a more satisfying explanation to this mystery.

The importance of strain gradients suggests that it is less important whether a region of bone is under compressive or tensile strains, but whether

that region is becoming more or less compressive or tensile. Consider the flow of fluids in a bone as it is bent. About 23 % of the volume of bone is composed of fluid, some of it bound to collagen, but much of it free to flow through the canaliculi and lacunae. In general, fluids flow down the fluid pressure gradient from more compressed regions to less compressed regions. Weinbaum et al. (1994) developed a mathematical theory predicting that physiological stress gradients will produce fluid flows in the annular spaces between the osteocyte processes and the canalicular walls. As these fluids flow over the membranes of these processes and interact with the glycocalyx, the resulting fluid shear stresses on the cell wall were predicted to be in the range that causes the release of intra-cellular calcium ions in osteoblasts and endothelial cells. Similar calculations show that stresses of similarly high magnitude are not created by fluid motion on surfaces adjacent to larger spaces, such as Haversian canals or on marrow cavity surfaces where lining cells (pre-osteoblasts) reside. Partitioning the mechanically relevant signaling to the osteocyte lacunar space suggests that the effect of mechanical loading is not a direct one to the lining cells or osteoblasts, but instead is communicated through the osteocyte network to these cells to activate modeling and remodeling.

6.5.13 *Fluid Flow*

The idea that fluid flow through canalicular channels is the direct stimulus for cellular activity is consistent with the observation that bone is strain rate sensitive, because the rate of flow is directly proportional to the strain rate. The more rapidly bone is loaded the faster fluid is forced through these channels. In the late 1990s, several studies were able to demonstrate that osteoblasts seemed to respond to the flow of fluid over them, independent of the amount of membrane deformation (Owan et al. 1997; You et al. 2000). It had always been a conundrum how osteocytes in particular, sitting within a mineralized matrix, could sense physical deformations. These experiments suggested that the signal for cellular response was not the deformation itself, but the deformation-induced fluid flow through channels, which could create either a membrane shear stress or chemical transport that would stimulate cells to react. This was predicted earlier by the important modeling studies of Weinbaum and Cowin (Weinbaum et al. 1994; Wang et al. 2000), which also predicted that fluid velocities within the canaliculi are proportional to loading rate. More recent in vivo work showed that high frequency oscillatory axial displacement (45 Hz) applied to mouse lower limbs in the absence of weight bearing led to significant increases in trabecular bone volume (17 %) and stiffness (38 %) without bone deformation (Garman et al. 2007; Ozcivici et al. 2007, 2010). Although this experiment demonstrates rather convincingly that bone matrix deformation is not required for cellular response, it does not preclude the possibility of stimulatory cell signaling from tissues

adjacent to the periosteum (muscle and fascia) and endosteum (marrow) rather than from osteocytes exclusively.

It is now generally accepted that both osteoblasts and osteocytes are stimulated by fluid flow (Riddle and Donahue 2009), and that the specific response is in part dictated by whether the flow is continuous, pulsatile, or oscillatory. Flows in the range of 0–20 dynes/cm² have been shown to cause osteoblast proliferation, and to upregulate the cellular production of a range of factors important for the regulation of bone formation, including nitric oxide, cyclo-oxygenases and the prostaglandins (PGE2) that are formed downstream of them, and osteopontin. It also stimulates cellular processes resulting in release of ATP, and the stimulation of purinergic cell-surface receptors (e.g., P2X7).

Fluid motion following bone loading has now also been demonstrated in a more dynamic manner using fluorescein tracers and photobleaching (FRAP) techniques (Wang et al. 2005; Price et al. 2011) (Table 6.3). Although prior studies examined tracer perfusion (usually procion red) through the lacunar-canalicular system with loading (Knothe-Tate et al. 1998, 2000), they were only able to examine the extent of perfusion in a static manner at defined time points, rather than true dynamic flow. More recent dynamic studies demonstrated that physiologic strain magnitudes (400 µε) increase diffusion of fluorescein tracer by 31 % (compared to non-loaded bones), and produce a flow rate of 60 µm/s (Price et al. 2011). This was calculated to impart a fluid shear stress on the membrane of the osteocyte or its processes of about 50 dyn/cm², more than double the shear stress that is typically used in *in vitro* laminar flow studies of cellular response. Of course, this calculation is highly dependent on the morphology of the lacunar-canalicular system, but these were experimentally measured and validated using a model of transport kinetics that was not dependent on bone’s poroelastic material properties (Table 6.3). However, this approach allows one to estimate how variations in canalicular connectivity would affect transport, membrane shear, and ultimately the capacity of cells to sense a mechanical signal. The model predicts that more canaliculi, or presumably greater connectivity, would linearly increase transport rate, whereas a decline in canalicular length would result

TABLE 6.3. Experimentally measured lacunar-canalicular dimensions and average diffusivity

Ca.Le (µm)	Lc.Rd (µm)	Ca.Ar (µm ²)	Lc.Ar (µm ²)	Diffusivity (10–10 m ² /s)	Loaded diffusivity (10–10 m ² /s)
30 ^a	3.9–8.7	0.315	16.3	2.72	3.56
26 ^b	2.3–11.6	NM	91.3	3.3	NM

Lacunar areas from Price et al. appear to be low, based on the measured radius

NM not measured

^aPrice et al. (2011)

^bWang et al. (2005).

in a nonlinear decrease in transport, and presumably impaired mechanical signaling. This is an interesting result given experimental observations that osteocyte apoptosis occurs in disuse (Tomkinson et al. 1997), estrogen deficiency (Noble et al. 1997), or microdamage formation (Mori et al. 1997; Cardoso et al. 2009; Kennedy et al. 2012), the last of which significantly reduces the canalicular network, possibly preventing adaptation, or in some cases repair.

The idea that fluid motion provides the mechanocoupling signal that will lead to the cell's biochemical response can explain the observation that very high frequency loading is ineffective at generating a strong cellular response. We know that the voltage and calcium gated channels that control trafficking of molecules into and out of cells are themselves frequency dependent. We also know that there is a marked decline in the rate of fluid flow within the canalicular system when the frequency of loading is increased from 1 to 30 Hz (Gross et al. 2004; Srinivasan et al. 2002; LaMothe and Zernicke 2004). If strain-induced flow of fluid through bone is the local mechanical stimulus for the cellular response, then high frequency, low magnitude loading could compromise mechanical transduction simply because there is insufficient time for the fluid movement, at least to the degree required for osteocyte stimulation. This, then, may be the solution to our question about why low magnitude—very high rate or high frequency vibration, does not generate massive responses. Fluid motion above a certain magnitude is necessary in addition to the number of cycles in order to elicit a full cellular response.

The nature of the osteocyte (sensor cell) and osteoblast and osteoclast (effector cells) response, however, is likely dependent on the nature of the flow itself. Physiologically, bone cells are probably not exposed to continuous or unidirectional pulsatile flow, but rather to flows that are oscillatory. Cell response to oscillatory flow is heavily modulated by chemotransport, unlike responses to continuous or unidirectional pulsatile flow regimes (Haut Donahue et al. 2003; Riddle et al. 2008; Bakker et al. 2001). Oscillatory flow also is a much weaker stimulus to cell response, and requires a longer period of exposure. This is somewhat inconsistent with the rapid saturation of response that is observed from in vivo studies (Robling et al. 2000).

6.5.14 Summary

The experiments reviewed in this chapter thus far suggest that the mechanical adaptation of bone is governed by a set of controlling mechanisms that can be summarized as follows:

- Trabeculae seem to be aligned with the principal strain directions calculated from a continuous homogenous model.
- Disuse of a growing long bone diaphysis inhibits periosteal bone formation; disuse of a mature diaphysis results in increased endosteal resorption and intracortical remodeling.

- Mild overloading of a growing long bone diaphysis inhibits endosteal resorption, and can stimulate periosteal apposition, producing a thicker cortex.
- A set of local strain thresholds or setpoints define the bone's adaptive response to mechanical stimuli. The upper threshold for an anabolic response is probably above 1000 $\mu\epsilon$ in typical long bones, and in general this will keep peak periosteal bone strains during even vigorous activities below about 2000–3000 $\mu\epsilon$. This appears to be conserved for the long bones across species.
- Bone adaptation is governed by a combination of strain magnitude and frequency, both of which modulate strain rate. Magnitude and frequency can be varied within limits to create a predictable response.
- Strain is probably an indirect stimulus for bone adaptation; current evidence points to strain driven motion of fluid within the canalicular network as the most likely transduction signal. Sensor cells (osteocytes) can detect this flow (by as yet unknown mechanisms), and create biochemical signals that govern the response of effector cells (osteoblasts and osteoclasts).
- Fluid can only flow over a significant length of time if bone is loaded cyclically. This goes along with the observation that bone only responds to loads that are dynamically applied, and is very responsive even with small numbers of loading cycles. Cell sensitivity saturates quickly, and takes hours to recover. Oscillatory fluid flow regulates cell response by chemotransport mechanisms, whereas continuous fluid flow regulates cell response via membrane shear stress.
- High frequency vibration does not produce as vigorous an anabolic response as theory might suggest because the flow magnitude is small.
- Experiments investigating bone adaptation emphasize diaphyseal cortical bone responses, as opposed to trabecular bone architecture, which was the historical focus. A major reason for this is that, experimentally, it is very difficult to control the mechanical environment to which individual trabeculae are subjected. The exigencies of experimental science—getting statistically significant results in a few weeks or months—lead investigators to conduct experiments that may not be very representative of the sublime, gradual, day-in, day-out way in which mechanical adaptation probably works in healthy individuals. In human beings, most cortical bone adaptations are likely to involve more modest load increases. See Bertram and Swartz (1991) for more discussion of these issues.

6.6 Some Additional Theories

In this section, we shall describe several additional theories about how the skeleton adapts to mechanical loading. As was the case with the preceding section on experimental work, this section is not meant to be an exhaustive

review, but to provide some examples of the principal approaches that have been used. These theories may be organized into four categories, which attempt to explain:

1. Modeling drifts to correct for diaphyseal curvature or osteotomy of a collateral bone
2. Control of directionality in tunneling osteonal BMUs
3. Bone density distribution
4. Trabecular alignment

The first of these includes the flexural neutralization theory, which we have already, touched upon (Fig. 6.11). We present the remaining examples in sequence by category.

6.6.1 Pauwels' Stress Magnitude Theory: Control of Modeling Drifts

Pauwels (1980)⁶ gave an elegant demonstration of how simple assumptions relating surface remodeling to stress can predict the gross distortions in cross sectional geometry that occur in very curved, rachitic femurs. Rickets is a disease which, in children, interferes with the normal functioning of the growth plate. As a result, the long bones may grow with abnormal amounts of curvature that modeling drifts fail to correct. The left side of Fig. 6.12 shows three views of a femur from a man who had untreated rickets as a child. The bone has a greatly exaggerated anterior bow. This curvature increased the bending stresses at midshaft, and it was clear to Pauwels that this led in turn to the long, narrow cross-sectional geometry which is shown in the lower right portion of the figure, along with the normal cross-sectional geometry. Both the normal and rachitic cross sections are shown in their correct positions relative to the head of the femur when viewed from above. Pauwels sought a mathematical formula that would predict this adaptation. He explained in detail his method for determining the stresses in the bone, but was less clear regarding his adaptation algorithm. He began with the elliptical cross section shown in the upper middle image (rather than the normal femoral shape), and computed the surface stress, s , from a simulated hip load. He then altered the thickness of the cortex, C , at each point according to the equation

$$C = a + bs^n \quad (6.3)$$

⁶ The cited reference is a compendium of Pauwels' research papers. The actual work discussed here was published in the 1960s.

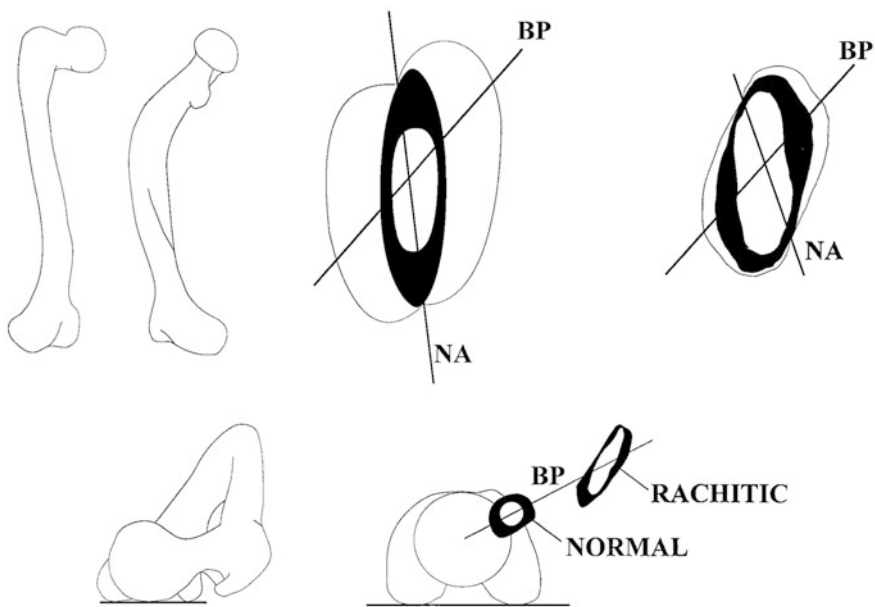


FIGURE 6.12. Effects of mechanically adaptive changes in the femur of a man who experienced untreated rickets. (*Left*) Three views of the highly curved femur. (*Lower right*) View down the shaft of a normal femur, showing the relative positions of the midshaft cross section for the rachitic bone compared to a normal bone. BP, estimated bending plane. Note the extended moment arm of the rachitic femur midsection. (*Upper middle*) Pauwels calculated the surface stresses on an elliptical cross section (black) loaded by the bending moment shown below. The distance between the curved lines and the surface of the bone section indicates the magnitude of the periosteal stresses. NA neutral axis. Using Eq. (6.1), Pauwels iteratively “modeled” the femur section, producing the shape shown at *upper right*, which is similar to the actual bone section seen below (Pauwels 1980). With kind permission of Springer Science+Business Media.

where a , b , and n were unspecified “arbitrary constants.” After this change, the stress and cortical thickness were recalculated, and the process was iterated until, presumably, a steady state was reached. The result of this adaptation is shown in the upper right cross section, which bears a striking resemblance to the actual rachitic cross section. However, it is clear from the figure that, in addition to altering the thickness of the cortex, the calculations moved it radially in and out, creating an effect which implied varying medullary resorption. Neither the rationale nor the algorithm for this aspect of the model was fully explained.

6.6.2 Adaptive Elasticity Theory: Control of Density or Modeling Drifts

These theories have been used to account for both diaphyseal modeling responses and adaptation of bone density to loading. Cowin and his co-workers developed a novel and mathematically rigorous theory of bone adaptation (Cowin and Hegedus 1976; Cowin and Nachlinger 1978; Hegedus and Cowin 1976; Cowin and Van Buskirk 1978). They viewed internal remodeling as a set of chemical reactions acting to exchange material between the solid bone matrix and the fluids in the void spaces. They employed an extension of continuum mechanics in which the chemical reactions and the bone geometry are forced to obey thermodynamic, mass, and momentum principles. Using this model, it was shown that, under the action of a uniform stress, a nonhomogeneous cylindrical bone would become homogeneous in its density (Firoozbakhsh and Cowin 1980).

Subsequently, Firoozbakhsh and Cowin (1981) showed that for uniaxial stresses and certain constraints of the remodeling coefficients, their model was equivalent to that of Pauwels for surface remodeling. Immediately thereafter, Cowin and Firoozbakhsh (Firoozbakhsh and Cowin 1981) presented a somewhat simpler model for surface remodeling that explored the implications of a generalized strain error hypothesis for mechanical adaptation. The starting point was the equation

$$M = C_{ij}(e_{ij} - e_{ij}^o) \quad (6.4)$$

where M is the net rate of addition to the bone surface (M is negative for resorption), e_{ij} is the strain tensor, and e_{ij}^o is a reference or “set point” strain tensor. This theory is much more general than most formulations of strain-governed adaptation, which typically consider only one component of strain, or some other scalar quantity. The C_{ij} are “surface remodeling rate coefficients” that can be positive or negative, depending on whether net formation or resorption occurs. Although the theory does not specify the biological mechanisms that are embodied in the C_{ij} , it allows explicit computation of adaptive results for complicated circumstances (such as realistic loading which superimposes compression, bending, and torsion).

Obviously, if this theory is to have any practical application, the values of the remodeling rate coefficients must be determined. To that end, Cowin’s group (Cowin et al. 1985) compared the predictions of their model to the results of five animal experiments (including the canine disuse and sheep experiments described earlier). The C_{ij} tensor contains coefficients for each component of strain. The authors were able to simplify matters considerably for the specific applications of the theory, so that there were just two coefficients relating surface apposition rate, M , to longitudinal stress at the surface. By adjusting these coefficients separately for each experiment, excellent

reproductions of the changes in bone cross section were reproduced. Unfortunately, widely varying values for the coefficients in each experiment were required, perhaps because the coefficients vary with species and age as well as bone surface, or because the theory uses an incorrect choice for the mechanical stimulus. A further worry is that the theory is linear in strain, so that if an experimenter reverses the loading (e.g., pulls instead of pushes on the bone) the sign of the strain error changes and reverses the direction of surface motion and/or changes whether the load causes an increase or decrease in mass.

In theoretical models similar to Cowin's, Hart and co-workers (Hart and Davy 1989; Hart et al. 1984a, b) formulated the models in terms of variables that quantify (or at least relate to) processes of cell differentiation and cell function. They assumed that some set of variables representing the bone strain history combine with genetic, hormonal, and metabolic factors to determine the regions of bone surfaces where modeling or remodeling would occur, the numbers of active osteoblasts and osteoclasts per unit area in these locations, and the average activities of individual cells. Their approach melds and extends engineering representations of mechanical influences on cells initially developed by Cowin and co-workers with formulations of cell responses initially developed by Frost (1964) and Martin (1972, 1985, 1995). Their approach was able to predict some of the results of canine osteotomy overload experiments (Burr et al. 1989a, b; Oden et al. 1995).

6.6.3 *What Controls Osteonal Tunneling Directions?*

Osteons are not precisely longitudinally oriented in a long bone diaphysis, but spiral around and through the cortex at a shallow angle (Cohen and Harris 1958). Lanyon and Bourn (1979) showed a good correspondence between this orientation and the principal strain directions. The mechanism by which such directionality is achieved has received little attention.

While the activation of osteonal BMUs may be achieved by relatively brief, isolated signals, directional control seems to require a signal that is at least intermittent. The first attempt found in the literature to explain BMU behavior was that of Martin et al. (1974). This theory asserted that the stresses in the bone surrounding the osteonal remodeling cavity play an important role in organizing and directing the cellular activities in the BMU. Like the external surfaces of a whole bone, the walls of the cavity produced by an osteonal BMU experience a variety of stresses. Most importantly, stress concentrations will occur in the vicinity of the cutting cone and reversal region. These high stresses may produce conditions that can control the cutting cone's direction of travel.

6.6.3.1 The Stresses Around an Osteonal BMU: Wolff's Law Inside Out

Figure 6.13 shows schematic diagrams of an osteonal BMU and a simple spherical cavity in a loaded bone; the bone is assumed to be loaded in simple compression, and the BMU is assumed to be aligned with the principal compressive stress direction. When a simple spherical cavity lies in a compressive stress field, p , the equatorial wall stress is increased to $3p$, and the wall stress at the pole of the cavity is $-p$ (Peterson 1974). The $-$ sign indicates that this polar stress is tensile. Thus, the surfaces of the void that are parallel to the load direction experience tangential stresses three times greater than normal, and the surfaces perpendicular to the load carry a stress of opposite polarity. In the case of a cavity having the elongated geometry of the BMU, the stress concentration factor is reduced, but the reversed polarity of the stress at the apex of the cutting cone will remain. Consequently, a BMU aligned with a compressive stress field has its osteoclasts working on a surface that carries a tensile stress and its osteoblasts working on surfaces which carry compressive stress.

Now consider the stress gradients surrounding an osteonal BMU. There is a positive gradient at the resorption surface and a negative gradient at the initial refilling surface (it is left as an exercise to show this). Thus, the cellular activities within the osteonal BMU bear the same relationships to stress gradients as do the activities of cells remodeling the external surfaces of a

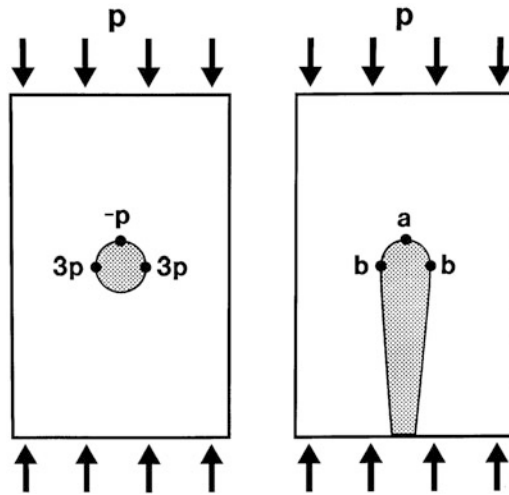


FIGURE 6.13. (Left) Stress concentration and stress of opposite polarity at apex of a spherical cavity in a uniaxial compressive stress field, p . (Right) Diagram of a BMU cavity aligned with the compressive stress direction in a long bone. The tangential stress on the resorption surface (a) would be tensile; that on the refilling surface (b) would be compressive.

curved bone shaft carrying end loads in Frost's flexural neutralization theory. Thus, it is again hypothesized that a positive strain gradient activates osteoclasts at the resorption front and a negative strain gradient activates bone formation at the beginning of the refilling region.

6.6.3.2 Steering in the Principal Compressive Strain Direction

Having seen that cell location and function may be governed by the mechanical environment of the osteonal cavity, one may now ponder how these considerations could lead to directional control of the osteon. Suppose, as depicted in Fig. 6.14, that the osteonal BMU becomes misaligned with the principal compressive stress direction. In that case, the location of the maximum positive strain gradient shifts from the apex of the cutting cone to a point associated with the principal compressive stress direction. One would then expect the most vigorous resorption to occur at this point, so that the osteon would turn in the desired direction. Thus, the principles that we have outlined provide a mechanism for automatically steering osteons in the principal compressive strain direction.

If the cortices of long bones are loaded primarily by bending and compressive end loads, and to a lesser extent by torsional loads, steering of osteons in the compressive stress direction fits nicely with the observation that osteons spiral around the shaft at a shallow angle. However, the theory encounters two difficulties in that it does not explain the tendency of BMUs to dig through the thickness of the cortex, nor does it explain the behavior of

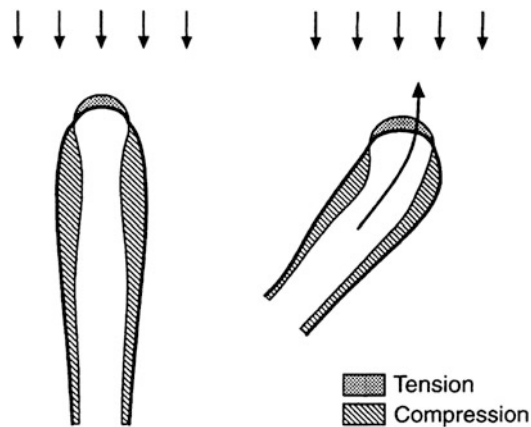


FIGURE 6.14. (Left) Surface stresses in a BMU aligned with the compressive stress direction. Thickness of hatched region inside the cavity approximates the compressive stress magnitude; thickness of the stippled region above the cavity approximates the tensile stress. (Right) Corrective steering of a BMU that has become misaligned with the compressive stress direction. The peak tensile stress has moved to the left side of the resorption surface; if resorption is enhanced there, the BMU will turn in that direction.

osteonal resorption cavities on the tensile side of bones subjected to bending. Although many of Martin's observations on the mechanical environment are relevant, based on what we now know it seems unlikely that osteoclasts and osteoblasts are independently sensing their immediate mechanical environment. Consistent with more recent evidence targeting osteocytes and fluid flow as transducers of mechanical signals, a nonlinear computational model by van Oers et al. (2008) for how local stresses (or strains) guide osteonal and hemi-osteonal activity yields interesting results that in many respects match empirical data.

6.6.4 Adaptive Finite Element Models: Control of Density

Carter et al. (1987a) and Huiskes et al. (1987) introduced the finite element approach to the development of predictive relationships between mechanical loading and trabecular bone density. For an example of this general method in a later form, see Beaupre et al. (1990a, b). We present here a simplified version of the Carter–Beaupre approach, but one that captures its essentials. This theory assumes that the mechanical stimulus is some time-averaged aspect of stress or strain. A *daily stress stimulus*, Ψ , is defined as

$$\Psi = \left(\sum n_i \sigma_i^m \right)^{1/m} \quad (6.5)$$

where n_i is the number of cycles/day of a particular kind of loading, i , which produces a continuum level “effective stress,” σ_i , at the location in question. This effective stress has assumed different forms, each a scalar quantity having units of stress, for example, $\sqrt{2E\bar{U}}$, where E is elastic modulus and $\bar{U}$ is strain energy density. The summation is over i , that is, over all the significant kinds of loads that are applied each day. Next, a finite element model is used to determine the values of σ_i at each location in the bone, each “location” corresponding to an element in the model. An error function is then defined as

$$e = \text{constant} \times (\Psi - \Psi_{AS}) \quad (6.6)$$

where Ψ_{AS} was called the “attractor state” for the stress stimulus. (Designating Ψ_{AS} as an attractor suggests that it is a limit set that the trajectories of Ψ approach over time. If Ψ is unstable and diverges, Ψ_{AS} is not an attractor.) Originally, the apparent density of each element was assumed to change according to e . In the Beaupre version of the model, the net rate of addition or removal of bone on an external or internal surface (the apposition rate, M , $\mu\text{m/day}$) is made a linear function of the error:

$$M = c_1 + c_2 e \quad (6.7)$$

where c_1 and c_2 are coefficients containing the biological mechanisms by which bone cells respond to the mechanical stimulus. For internal

remodeling, the bone formation rate (volume of bone added or subtracted per unit volume per day) can be written as

$$V_F = dB_V/dt = MkS_V \quad (6.8)$$

where B_V is the bone volume fraction, S_V is the internal surface area per unit volume (a nonlinear function of B_V —as a result, Eq. (6.8) is a nonlinear differential equation in B_V), and k is the fraction of this surface that is actively remodeling. Because apparent density, d , is the product of B_V and the density of the calcified bone matrix, Eq. (6.8) allows one to track the changes in d that occur in each element. These density changes in each element are converted to modulus changes using the Carter–Hayes correlation:

$$E = \text{constant} \times d^3 \quad (6.9)$$

[recall Eq. (7.12)]. The finite element model is solved again with the new modulus values to obtain a new stress stimulus for each element. The model is iterated to find the distribution of bone density (or external geometry), which minimizes the error function.

6.6.4.1 The Value of m

To use the model, a value must be obtained for the exponent m in the stress stimulus equation. Beaupre et al. (1990b) did this in the following way. Using data from two of Rubin and Lanyon's isolated avian ulna experiments, they plotted two combinations of applied strain magnitude and cycles per day that resulted in bone maintenance, that is, neither losing or gaining bone. These are the points on the left (4 cycles at 2000 $\mu\epsilon$ and 100 cycles at 1000 $\mu\epsilon$) in Fig. 6.15. They added to these two points a third, which represented the strain measured by Lanyon on the tibia of a human (400 $\mu\epsilon$), and assumed that this person's tibia was loaded 10,000 cycles/day and was in an equilibrium state. These three points gave a slope of 4.88; based on other studies

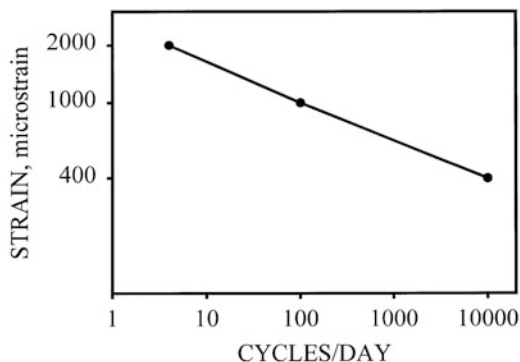


FIGURE 6.15. Applied peak strain vs. cycles/day for equilibrium of bone mass (Beaupre et al. 1990b).

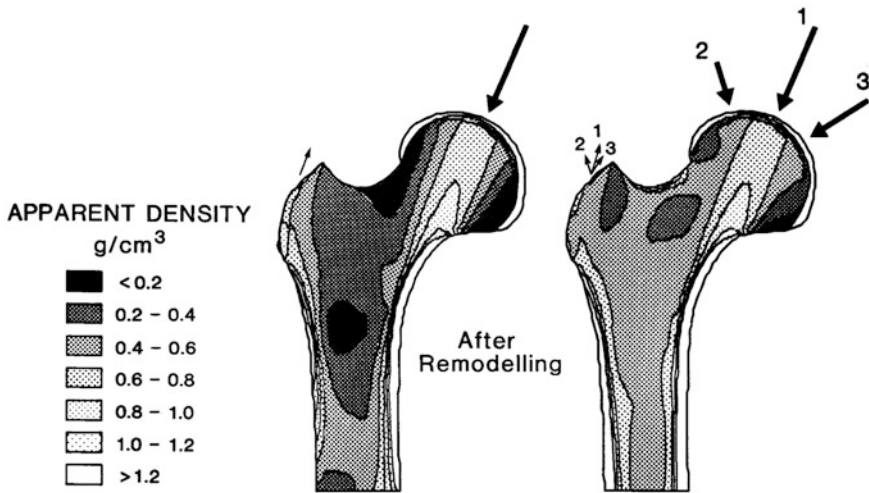


FIGURE 6.16. Results of a typical model of the kind used by Carter and co-workers. Less dense areas are indicated by *darker shading* as seen in scale at *left*. The model at *left* was loaded by a single set of hip and abductor forces (arrows); that at *right* was loaded by three different loads [i.e., different i in Eq. (6.5)], representing abduction, adduction, and normal hip positions (Carter et al. 1987a).

(Whalen et al. 1988), this was reduced to four. Figure 6.16 shows a typical pattern of bone density in the proximal femur predicted by a model like this when the loading consists of either one or three point loads applied to the femoral head and corresponding abductor loads applied to the greater trochanter. The two key similarities between the simulation result and real bones are the path of dense bone leading from the medial cortex up to the head of the femur, and the area of low density in the middle of the femoral neck known as Ward's triangle (it has an oval shape in the three load simulation).

Huiskes and co-workers (Huiskes et al. 1987) have used adaptive finite element models to study the responses of bone to the mechanical consequences of implanting a joint prosthesis. They used strain energy density as the mechanical stimulus to which the bone cells respond, and showed that such models could predict patterns of postimplantation bone density change adjacent to total hip prostheses. It has been shown that the Carter and Huiskes approaches are equivalent if the coefficients are properly chosen (Carter et al. 1989; Jacobs 1994).

6.6.4.2 Mathematical Stability Considerations

Harrigan and Hamilton (1994) analyzed the mathematical nature of models similar to those used in the original works of Huiskes' group. They found that the stability of such models depends on the relative magnitudes of the

exponents in the two principal equations. The first is for the modulus as it depends on apparent density:

$$E = k\rho^n \quad (6.10)$$

and the second is a mechanical “stimulus” that depends on the local strain energy density (SED) and the density as:

$$\Psi = SED/\rho^m \quad (6.11)$$

It turns out that $n < m$ is a necessary and sufficient condition for the simulation solution to have global stability and uniqueness. For the majority of older models $n = 2$ or 3 but $m = 0$ or 1 and the formulations did not meet the stability condition and density tended to diverge—either decreasing to zero or increasing in an unbounded manner at different points in the model. While Harrigan and Hamilton’s stability analysis did not directly apply to the Carter group’s method, density instability also occurred with that approach if it was not explicitly prevented.

At first the investigators performing the computational models did not notice the density instability because the solutions were being iterated manually and only a few steps were performed due to time constraints and computational costs (In the late 1980s a computer weaker than your smartphone could cost many thousands of dollars!). As the computer programs became automated, and more powerful computers were used, the instability became apparent and the algorithms were modified with mechanisms called “dead-zones” or “lazy-zones” that ensured that there was a range of stimulus values that caused no change in bone mass (Beaupre et al. 1990a, b). Harrigan and Hamilton (1994) noted that, “The dead zone accounts for the observations that small changes in the stress state in a bone do not seem to generate a remodeling response. (Using a dead zone) produces results which agree qualitatively with experimental results, but the use of a dead zone thwarts a unique or an optimal solution.” Although beyond the scope of this chapter, averaging the stimulus in a region (as was used to partially stabilize the self-trabeculating models discussed below) is a similar modification to using a dead zone.

It is interesting that many of the commonly used computational models for the development of bone structure have an instability that must be explicitly controlled. Such instabilities were (and are) a difficulty for computational biomechanics if the goal is predicting a smooth distribution of density. On the other hand, the instability contributes to the extremely interesting “self-trabeculating” models discussed below. The balance of stability versus instability of density in the real biological system is not known.

6.6.5 Self-Trabeculating Models: Control of Density and Trabecular Alignment

If one makes an iterative finite element model like those we have been describing, in which the density and then the modulus of each element is adjusted in proportion to its stress or strain or strain energy density (SED), and lets it run indefinitely, a very interesting thing often happens. The model may not approach an equilibrium state in which there is a smooth variation in element densities from one region to another. Instead, almost all the elements may either go to the minimum or maximum allowable density. This behavior has been termed “0–1” behavior because each elemental “square” in the model becomes either empty or full. The full elements are arranged to form a network of trabecula-like struts that carry loads from one part of the model to another.

Figure 6.17 shows one of the first such models in the orthopaedics literature, produced by Weinans et al. (1992). In this model, the mechanical stimulus was strain energy density (SED). It shows a model that has fixed elements at the bottom and carries a set of vertical point loads across the top which increase from right to left. The “1” or bone-filled elements are black and the “0” or soft tissue-filled elements are white. The bottom portion of the

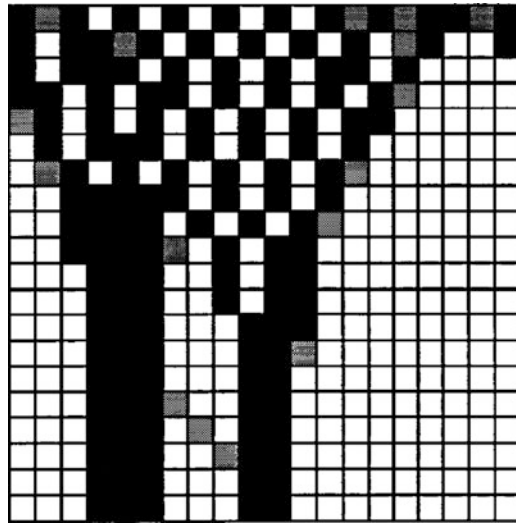


FIGURE 6.17. An example of an early self-trabeculating finite element model. A compressive load was applied to the upper surface, diminishing from *left to right*. The density of the material in each element is represented by a *gray scale*. Most elements have either maximum (*black*) or minimum (*white*) density. Two walls of cortical bone have developed below a metaphyseal region. Note the “checkerboarding” effect described in the text (Weinans et al. 1992).

model has evolved into two asymmetric cortices, and the top portion has become a network of “trabeculae.” The figure conveys the self-trabeculating concept well, although it is a relatively crude, two-dimensional model.

The 0–1 behavior, in which the model develops a trabecular architecture on its own, was very exciting to see, for if this behavior truly represented a physical phenomenon, rather than some sort of numerical processing behavior peculiar to a computer model, it could provide insight about cancellous bone structure and the mechanical adaptability hypothesis. However, there is another feature of Fig. 6.17 that is very disturbing. Notice that many of the bone-filled regions contact one another only at their corners, that is, at a single node of the finite element mesh. This effect is known as “checkerboarding,” and it signals that the finite element model is not representative of reality because there would be infinite stress at such contact points. Engineers designing optimized composite materials saw such behavior as well, and learned that this can be avoided if the modulus of each element is adjusted not on the basis of its SED (or stress or strain) alone, but based on a mean value for several neighboring elements.

Mullender et al. (1994) made a similar discovery when their subsequent models incorporated a biologically motivated feature that introduced a similar smoothing effect. They assumed that the SED per unit mass is the mechanical stimulus sensed by osteocytes and communicated through their processes. Each cell then responds according to a sum of stimuli produced at its location and received from osteocytes in other locations. They also assumed that the strength of the stimulus, SE , decays exponentially with the distance communicated. The stimulus received at a given location is of the form

$$S_E(\mathbf{x}) = \sum_i [(U_i/\rho_i - k)\exp(-d_i/D)] \quad (6.12)$$

where U_i/ρ_i is SED divided by apparent density at the location of an osteocyte labeled i , located a distance d_i away from point $\mathbf{x}$, k is an equilibrium value of SED per apparent density and D is a decay constant. The summation is over i , that is, all the surrounding osteocytes. When such a model is run, starting from a homogeneous mesh of elements having an intermediate density, the element densities again migrate to extreme values, so that 0–1 behavior persists and the model self-trabeculates, but checker-boarding does not occur (Fig. 6.18).

An even more interesting aspect of such models is that the trabeculae which develop align themselves towards the loading direction—similar to the principal directions expected for a homogeneous structure of the same shape and bearing the same loads (Fig. 6.18). Higher loads produce more and thicker trabeculae. Furthermore, if the model orients its trabeculae to loads directed one way, and then the load direction is changed, the model will realign its trabeculae towards the new load direction. This realignment occurs even though the mechanical stimulus is a scalar that contains none

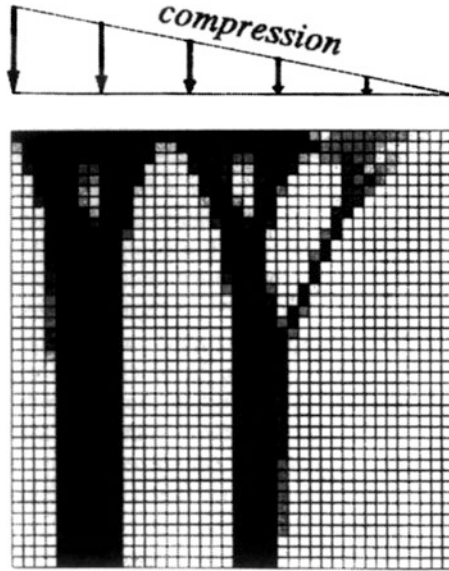


FIGURE 6.18. Self-trabeculating model similar to that of Fig. 6.17 but with a finer mesh. Checkerboarding has been eliminated by using Eq. (6.12) (Mullender et al. 1994).

of the information about the individual components of the stress tensor which an engineer would ordinarily use to compute the principal stress directions! We return to this point later.

The models we have described so far assumed that osteocytes not only sensed the mechanical stimulus, but altered the density at their location as well. Although osteocytes may make minor changes in the adjacent bone matrix, they cannot resorb or form bone to any significant degree. These processes must be carried out by osteoclasts and osteoblasts, and they must occur on bone surfaces, not in the bone midsubstance. Smith et al. (1997) modified the model of Mullender et al. (1994) to incorporate this requirement. The elements of their model represent only the trabeculae and have a constant density and elastic modulus. The model also assumes that the cells located on the surfaces of trabeculae are bone lining cells that can activate resorption or formation, depending on the sign of the error signal they receive from neighboring osteocytes and bone lining cells. Otherwise, the model is similar to previous models, with all cells sensing SED and communicating this information to nearby cells according to Eq. (6.12). If the signal received by a lining cell is positive (the nearby SED is greater than the attractor state value), formation occurs by adding a new element (representing a packet of new bone) to the surface. If the signal is negative, the element beneath the lining cells is removed from the mesh (“resorbed”), and the model’s stresses are recomputed. This model, which is more realistic from the point of view of

the actual biologic processes, continues to adjust itself to changes in load direction or magnitude just as the earlier models did.

6.6.5.1 Bone Tissue Is a 0–1 Material

The 0–1 behavior seen in some finite element models is contrary to the assumption that apparent density should vary continuously in continuum level models of a large region of bone like the proximal femur. On the other hand, 0–1 behavior is very appropriate for a finite element model that represents the structure of individual trabeculae in a much smaller region of cancellous bone. At this scale, the elements in the model should represent either bone or marrow. Although the density of bone matrix within trabeculae may vary because of variations in mineralization, it can be argued that these variations are small compared to continuum-level porosity variations. As we have seen, the mineralization of bone tissue depends primarily on the rate of remodeling and normally varies only a few percent. Thus, bone is inherently a 0–1 material. To the extent that it is not, simulation of density variations would call for tracking the mineralization of new packets of bone produced by remodeling.

6.6.5.2 A Conundrum Solved?

We noted in Sect. 6.2 that there is a fundamental problem with the essential concept of Wolff's law: that is, that trabeculae align themselves with the average principal stress directions in the loaded bone. At the continuum level of analysis, there is no problem with this because principal stress directions can be computed by overlooking the discontinuities in the cancellous bone structure. Then, the alignment of trabeculae with principal stress trajectories means that the principal axes of the fabric tensor (see Sect. 7.5) are parallel to the principal stress directions (Cowin 1986). The problem is that bone cells are incapable of understanding continuum mechanics or of overlooking the discontinuities that form the trabecular surfaces on which they may reside. Thus, the “principal stress directions for the proximal femur” are, at the cellular level, entirely fictitious (Cowin 1997).

How, then, could one expect bone cells to “know” such fictitious directions, and form trabeculae so that they line up with them? The self-trabeculating models suggest an answer to this perplexing question. Perhaps bone cells are able to do this not because they “know” what the fictitious stress directions are, but because it turns out that those directions coincide with the trabecular orientations achieved by keeping the strain energy density (or some other mechanical parameter) at a constant value. In other words, by shaping and aligning trabeculae so as to control the local mechanical environment, it may happen that the trabecular orientations in a real bone are parallel to the principal stress trajectories in a facsimile bone of the same shape and loading, but made of a continuous substance. The self-trabeculating models support this proposition, but it would be nice to have a mathematical proof that it is true.

6.6.6 *Related Engineering Models*

The more recent versions of the self-trabeculating models are three dimensional and have fine enough meshes that the individual trabeculae are several elements across, eliminating some of the concerns stemming from the crudeness of the early models. The engineering literature on optimization of structures contains similar models. The analyses of Strang and Kohn (1986) suggested that optimal structures should contain highly oriented fibers rather than a continuously varying interior, not unlike trabecular struts of bone (If you want to see examples from engineering, look up “Michell structures” or “Hencky-Prandtl nets” in the literature on structures and plasticity theory). Other papers describe finite element-based models very similar to those of Weinans and Mullender that optimize the stiffness of a structure relative to its weight. For example, the top of Fig. 6.19 shows an optimized design for a cantilever “aircraft support beam” (Reiter and Rammerstorfer 1993). It is fixed along its left end and supports a downward load at its lower right corner. There is a striking similarity between this structure and that of the inside of a vulture’s wing bone, shown in the bottom of Fig. 6.19.

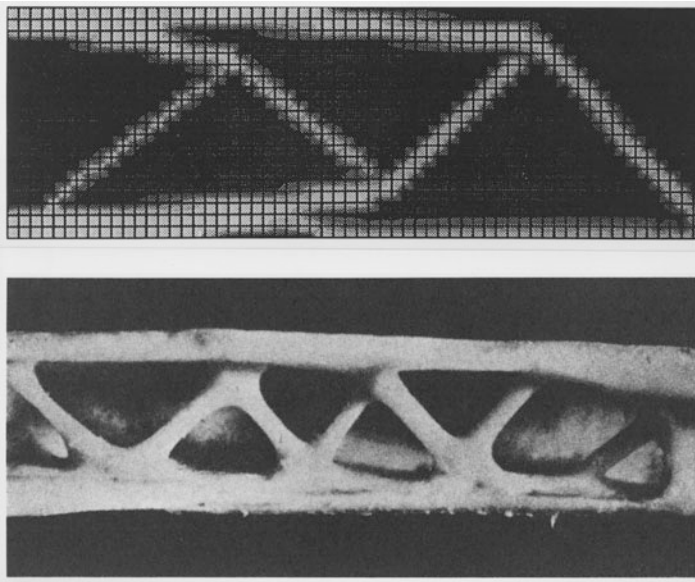


FIGURE 6.19. (Top) Optimized design of an “aircraft support beam.” (Pedersen 1993). (Bottom) Photograph of trabecular struts within a vulture’s wing (Thompson 1992).

6.6.7 *Synthesis*

While the foregoing theories vary in approach, it is possible to boil them down further than the four categories that we initially defined. The flexural neutralization and osteon alignment theories both propose that resorption and formation are determined by the stress gradient beneath the surface in question. They are essentially the same theory applied to two different structures. Both concepts are also compatible with the theory that osteocytes sense bone strain by reacting to fluid flow over their processes, because such flows would be driven by stress gradients.

Similarly, the bone density optimization and trabecular alignment theories are similar because they are based on similar mathematical principles; the latter is just formulated at a smaller scale. Also, the adaptive elasticity theories may be moved into this camp, because they also relate formation and resorption to stress or strain magnitude. Thus, we have just two kinds of theories, with the mechanical stimulus being stress (or strain) gradient in the first and some scalar measure of magnitude (e.g., strain or stress magnitudes or strain energy density) in the second. Both gradient based and magnitude based approaches are consistent with osteocytes being the mechanical sensors, and both seem to be dependent on communication across the osteocytic network, for sensing a stress gradient in one case and signal averaging in the other (to avoid “checkerboarding” in a computer model and perhaps some equivalent instability in reality). It is clear that stress magnitude does not work as the mechanical stimulus in theories of the first kind, but it is not clear that stress gradients would not work as the stimulus in theories of the second kind. Given the inhomogeneity and anisotropy of the bone tissue, stress gradients may be ubiquitous in bone, and the osteocytic network may be able to communicate information about both stress magnitude (or strain energy density) and stress gradients. Clearly, more work needs to be done to synthesize a single, general theory from the menu currently in the literature.

6.6.8 *Experimental Consistency*

We may use Frost’s mechanostat theory as a basis for comparing the results of several experiments described in the literature, some previously explored. Woo et al. (1981) exercised growing pigs and found no effect on periosteal diameter, but endosteal expansion was inhibited. This result is consistent with the mechanostat theory if the endosteal surface in young pigs is remodeling (inhibited by overuse) rather than modeling. The failure of the periosteal surface to show increased bony apposition would make sense if the rate of apposition were high already because the animals were growing. This is a reasonable supposition because pigs form plexiform bone at the

periosteal surface during growth, and the function of this kind of primary bone formation seems to be to add bone quickly.

Disuse osteopenia in the casted forelimbs of growing dogs (Uthoff and Jaworski 1978; Jaworski and Uthoff 1986) was produced both by the failure of the periosteal surface to expand and by expansion of the medullary cavity. This result is consistent with inhibition of modeling on the periosteal surface and stimulation of remodeling on the endosteal surface. However, bone loss was also seen on the periosteal surface; according to the mechanostat, this should have been produced by remodeling, but this is unclear. Remobilization of the young dogs reversed both these trends: woven bone was apposed to the periosteal surface, while the slow apposition of lamellar bone to the endosteal surface of normal young dogs was also restored. In older dogs, in which modeling is less responsive, disuse enhanced endosteal resorption and increased intracortical porosity, consistent with stimulated remodeling (Jaworski and Uthoff 1986; Jaworski et al. 1980). There was virtually no effect on periosteal diameter, consistent with the observation that this envelope almost always is restricted to modeling in the formation mode. Remobilization (Jaworski and Uthoff 1986) caused a small amount of bone to be added to both the endosteal and periosteal surfaces, but only partially restored bone mass in the diaphysis. The endosteal apposition was consistent with refilling of the remodeling space caused by an inhibition of remodeling. Much less bone was opposed periosteally than that in younger dogs, and the new bone was more highly organized, demonstrating the relative slowing of modeling processes in adults.

The isolated avian ulna model (involving adult animals) of Rubin and Lanyon (1985) also provides a good test of mechanostat predictions, and demonstrates that bone surfaces respond differently depending on whether the strain stimulus is above or below the equilibrium range. Disuse was associated with increased intracortical remodeling and endosteal resorption, as in immobilized adult dogs. Mechanical overload was associated with insignificant osteonal remodeling, a reduction in endosteal resorption, and abundant periosteal modeling producing woven bone. Similar observations concerning the effects of overload were obtained in the osteotomy studies of Lanyon and colleagues, as we have described. (It is noteworthy that increased osteonal remodeling has been observed in these experiments, but only after the woven bone formed by modeling had been present for some time and, presumably, needed remodeling to repair fatigue damage.)

Another relevant experimental model is that of Jee and co-workers (Jee et al. 1991), who unloaded one hindlimb of rats by bandaging, thus overloading the opposite limb. The increased resorption and reduced bone formation in the trabecular bone of the tibial metaphysis of the unloaded limb was consistent with increased remodeling and inhibited refilling, resulting in a 50 % reduction in bone mass and a new equilibrium state after 26 weeks. Overloading resulted in increased bone mass caused by reduced bone trabecular remodeling and increased periosteal modeling in

the formation mode. Schaffler (1990) has obtained similar results in dogs with one limb immobilized in a cast.

Thus, while some questions remain, the available experimental data support the essential features of the mechanostat theory. Reduced mechanical loading increases activation of new BMUs, increasing the number of new osteons and producing trabecular and endosteal resorption. Periosteal modeling in the formation mode is inhibited. Conversely, periosteal modeling is commonly activated by overload, which also inhibits activation of new BMUs, depressing formation of new osteons and limiting loss of bone from trabecular and endosteal surfaces. However, there must be an important limitation to the overload aspect of the mechanostat: if the overload produces damage, remodeling is activated to repair it (see Sect. 8.8).

6.6.9 Relationship of Mechanically Adaptive Responses to Other Control Factors

The response of bone to mechanical factors can be expected to depend on the balance that occurs between mechanically prompted physiologic signals and those from systemic or local nonmechanical factors. The nature of competing systemic (or global) and local signals to bone cells has been discussed by both Lanyon (Lanyon 1981, 1984; Lanyon et al. 1986) and Frost (Frost 1986). The idea is that various factors within the body compete for control of bone cells. For example, bone adjacent to marrow, particularly fatty marrow, has a tendency toward net resorption regardless of the hormonal or mechanical environment (Frost 1986; Johnson 1964). Thus, local “modulators” compete with or modify other signals, which undoubtedly accounts for some of the envelope specificity in bone remodeling (e.g., periosteal vs. endosteal).

Conversely, global factors may modify local responses, as when female distance runners lose rather than gain bone from excessive training-induced amenorrhea, a simulacrum of menopause. Exogenously supplied estrogen reverses the bone loss in these runners (Cumming 1996). Similarly, a study of male basketball players found that in a rigorous training regimen they lost bone mass unless given supplemental dietary calcium; calcium loss through perspiration resulted in a physiologic calcium demand that was met by bone resorption (Klesges et al. 1996). Skeletal health has been likened to a three-legged stool, the legs being diet, hormonal status (estrogen, vitamin D), and exercise. All three legs must be robust for the stool to stand. It is apparently also possible for local skeletal disturbances to have global effects. When marrow is destroyed within a bone, its replacement involves filling the medullary canal with woven bone, which is subsequently resorbed and replaced by new marrow. Bab et al. (1985) found increased bone formation in the mandibular condyle of rats after removal of bone marrow from the tibia. Thus, bone messengers intended to govern local responses may

“escape” to the systemic circulation and produce global remodeling responses.

Frost (1985) suggested that certain osteoporoses result from an increase in the lower limit of the physiologic strain range, so that normal mechanical loading is perceived as a disuse state and bone is lost. It is possible that estrogen affects this “set point,” so that it changes at menopause. If so, postmenopausal bone loss may occur because the normal mechanical stimulus becomes low relative to the cells’ altered reference value. Conversely, the anabolic effects of low doses of parathyroid hormone (PTH) suggest it may move the set point in the other direction (Gunness and Hock 1993). Thus, alterations in modeling and remodeling may occur because of changes in strain, changes in the physiologic strain range set points, or combinations of these factors.

6.6.10 *Mechanical Adaptability and Damage Repair*

We have hypothesized that mechanisms exist to optimize the stiffness of bone while minimizing fracture risk and bone mass. We have examined several possibilities for such mechanisms, such as the ability of osteocytes to sense fluid flow as a consequence of mechanical strain and regulate effector cells accordingly. These cells have the capability of activating modeling or remodeling to change the architecture of a bone in ways that affect its stiffness, fracture risk, and weight. However, as we discuss in Chap. 8, these cells also may activate remodeling to repair fatigue or creep damage. What is the relationship between damage repair and mechanical adaptability?

The answer has to do with the trade-off between weight and fracture risk from creep and fatigue. Let us digress again to Culmann and nineteenth-century engineering for a useful analogy. In bridges, fatigue can be an important element of failure. After some disasters, early American engineers began to build their bridges of elements large and robust enough that stresses were very low and fatigue failure would not occur. Not having the benefit of the mechanical theory being developed by Culmann and others in Europe, they overdesigned, wasting material and effort. As a young engineer, Culmann toured America and recognized this problem (Timoshenko 1953). Later, he used theory to design bridges in which stresses were low enough to avoid fatigue failure, but not unnecessarily low. Now, imagine what would happen if a way could be found to repair fatigue damage in a bridge soon after it occurs. Then the stress level could be allowed to rise, and the bridge could be made even lighter, saving further on materials and expense. That is the advantage that bones have over bridges. Damage repair by remodeling allows bones to be lighter, even though they carry stresses that would otherwise lead to fatigue failure within the organism’s lifetime.

Damage repair is not part of mechanical adaptability in the sense that it shapes elements of the skeleton to control stiffness and stresses, but it is an important part of the picture insofar as it allows higher stresses to be tolerated. The same may be said of the elements of bone tissue structure that limit and control microdamage. In fact, all these matters are components of the system that has evolved to optimize bone stiffness while minimizing fracture risk and weight. This system works similarly in the skeletons of many different animals, from hummingbirds to elephants.

6.7 Mechanical Adaptability in Cartilage

There is evidence that chondrocytes also respond to mechanical signals and that this capacity is instrumental in the development of cartilaginous structures. These mechanical effects could affect the development of bones from their cartilaginous models in the embryo and the functioning of the growth plate. We describe next a pair of theories directed at these phenomena, based on assumptions about the overall effects of stress on cartilage.

6.7.1 *The Carter–Wong Chondral Calcification Theory*

As reviewed by Pauwels (1980), several early writers proposed that the differentiation of cells to produce different kinds of connective tissue (bone, cartilage, tendon, etc.) was differentially promoted or inhibited by various kinds of mechanical stress (tension, compression, shear). This concept also included changes in chondrocytes leading to the calcification of cartilage. Carter and co-workers (Carter 1987; Carter et al. 1987b) developed this concept into a theory and explored its predictions with finite element models using methods similar to those described for bone. The basic premise of the theory is that the responses of chondrocytes to shear stresses lead to calcification of adjacent cartilage, while compressive dilatational (i.e., hydrostatic) stresses hold chondrocytes in a condition that preserves the cartilage in an uncalcified state.⁷ This hypothesis is consistent with the aqueous nature of cartilage; that is, uncalcified cartilage can support hydrostatic loads much better than shear loads. To assign scalar values to the dilatational and shear

⁷ All stress states have three independent measures of their magnitude: two of these are the hydrostatic or dilatational component, which produces a volume change but does not distort, and a shearing component, which distorts shape. The third magnitude measure doesn't have a simple mechanical interpretation and hasn't been used to develop theories of cartilage development. There are also three principal directions associated with stresses, but these also are not used in most cartilage theories.

portions of the local stress state, two variables are calculated from the three principal stresses ($\sigma_1, \sigma_2, \sigma_3$):

$$\text{Dilatational Stress} = D = (\sigma_1 + \sigma_2 + \sigma_3)/3 \quad (6.13a)$$

$$\text{Octahedral shear stress} = S = \left[(\sigma_1 - \sigma_2)^2 + (\sigma_2 - \sigma_3)^2 + (\sigma_3 - \sigma_1)^2 \right] / 3 \quad (6.13b)$$

It is proposed that the stimulus to calcify is an ossification index, I^* , given by

$$I^* \propto \sum n_i (kS_i + D_i) \quad (6.14)$$

where k is a coefficient that determines the balance between the two kinds of stress in driving calcification. As in the bone adaptation models, the subscript i denotes different kinds of daily activities or loading categories, and the summation is over all such categories. If a finite element model is made of the cartilaginous anlage of a long bone like the femur, loads are applied to its ends to simulate joint and muscle forces, the elements in the model are made to calcify and become stiffer when Eq. (6.14) reaches a threshold, and the model is iterated, interesting things happen (Carter et al. 1987b). First, a region in the center of the diaphysis calcifies; this corresponds to the primary ossification center in the developing embryonic bone. Because of shear stresses in the cartilage next to the calcified region, the calcified region spreads toward the ends of the diaphysis. As it does so, shear stresses increase in the ends of the bone, and these regions start to calcify; they correspond to the secondary calcification centers in the embryonic epiphyses. As these three ossification centers continue to grow, the regions of uncalcified cartilage shrink to four zones, corresponding to the articular cartilage on the joint surfaces and the growth plates at each end of the bone. These regions persist as cartilage because a stable mechanical situation dominated by hydrostatic stress develops in these sites. Thus, this theory holds that the development of a bone from a cartilaginous model in the embryo is not simply “genetically programmed,” but is in large part driven by the effects of mechanical stresses on the cells. This happens in much the same way that cells in the embryo are influenced by other epigenetic factors such as chemical gradients established by neighboring cells. Presumably, the growth plate calcifies and closes when the hormonal changes associated with puberty override the mechanical influences that have stabilized it.

This principle has also been used to explain the morphology of synovial joints (Carter and Wong 1988), the calcification patterns in the human sternum (Wong and Carter 1988) and the characteristic pattern of calcification in a fracture callus (Carter et al. 1988). However, mechanical loading doesn't seem to explain adequately the differences in the morphology of adult hyaline cartilage (Eckstein et al. 2006) so additional work is needed to understand the relationships between mechanical load, cartilage anatomy and mechanical properties.

6.7.2 Frost's Chondral Modeling Theory

While the Carter–Wong theory offers an explanation for the *de novo* development of the growth plate, Frost's (Frost 1990) chondral modeling theory offers mechanical reasons for changes in the orientation or shape of the growth plate as it functions. A much older theory known as the Heuter–Volkmann law holds that the physis grows faster when placed in tension and slower when normal amounts of compression are exceeded. This law is based primarily on the effects of static loads seen in clinicians' efforts to control the rate of growth of the whole growth plate. Frost's theory, on the other hand, is directed at understanding differential rates of growth within a growth plate, based on dynamic loading. It holds that higher stresses in either tension or compression accelerate growth, up to a point (Fig. 6.20). If peak compressive stress becomes too high, the growth rate will decrease (If peak tensile stress is too high, the growth plate will be torn apart). Now consider the implications of this with respect to the directionality of growth. If the dynamic joint force is not directed normal to the plane of the growth plate, but at an angle to it, bending stresses are superimposed on its normal compression. As a result, the stresses on one side of the physis are reduced relative to the other, creating a growth rate gradient across it. This gradient rotates the plane of the growth plate so as to realign it normal to the joint force. If, however, the dynamic compressive stresses increase too much on one side of the physis, growth there is slowed rather than accelerated, and the orthogonality of the growth plate and the joint load will get worse instead of better. Frost maintains that

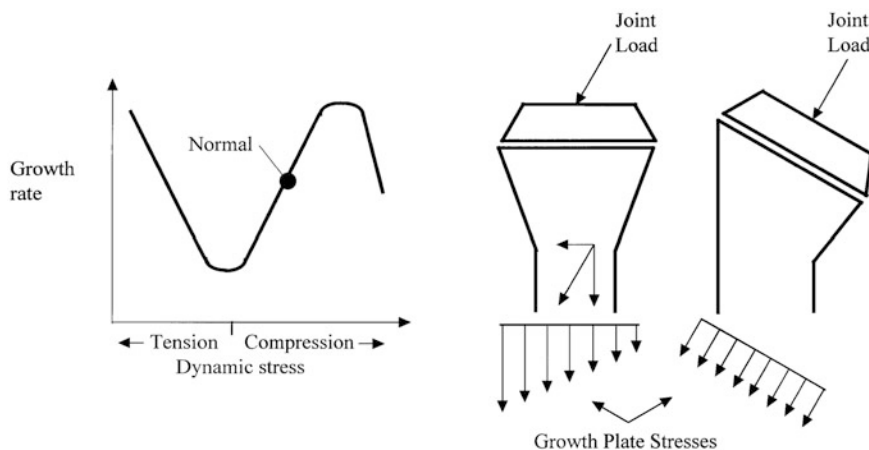


FIGURE 6.20. In the chondral modeling theory, growth rate increases with either tensile or compressive dynamic loading (*left*). Excessive compressive stresses inhibit growth, however. A misaligned joint force (*middle*) creates a stress gradient across the growth plate and a corresponding growth gradient, resulting in rotation of the growth plate and elimination of the malalignment (*right*).

this theory explains why mild problems with genu varum and valgus get better on their own and more severe ones do not. He also asserts that the curvature of long bones is established by asymmetric physal growth as well as modeling in the diaphysis.

A more complete theory, based on the actual molecular responses of chondrocytes to various stress conditions, is needed to provide a solid foundation for the Carter–Wong and chondral modeling theories, which appear to be two sides to the same coin. Together, the effects attributed to these theories may ultimately explain much of the diversity in joint morphology described in Chap. 5.

6.8 Summary and Further Reading

Wolff's law and the mechanical adaptability of bones to mechanical loading are not only of great interest to basic scientists interested in the skeleton; they are of great clinical utility as well. Orthopedists see the effects of these phenomena in their daily practices. We have reviewed the history of these concepts in some detail to gain a better understanding of how the current ideas evolved. We saw that there are three key concepts, including optimization of strength with respect to weight, alignment of trabeculae with principal stress directions, and regulation by systems of cells able to respond to a controlling stimulus. We then reviewed some of the more important experiments in this area, and saw that they have emphasized cortical rather than trabecular bone. Although disuse and overload effects are clear, it may be not be appropriate to relate the latter to normal development because woven rather than lamellar bone formation occurred.

We also reviewed some of the theories regarding mechanical adaptability of cortical and trabecular bone, and found that the putative ability of osteocytes to sense strain or strain energy, likely as a consequence of fluid movement down strain gradients, is emerging as an important element of such theories. The discovery that control of strain-energy-per-density within a loaded structure leads to 0–1 architectures with load-related trabeculae was a key development. Finally, we asserted that damage repair is an important adjunct to mechanical adaptability because it allows higher strains to be tolerated, and that damage repair and mechanical adaptability are somatic changes which extend the organism's ability to cope with environmental changes without committing to an irreversible genetic change.

Much of the historical material and interpretation at the beginning of this chapter came from two reviews by Roesler (1981, 1987); these are essential reading for anyone interested in this subject. As Roesler notes, Wolff's monograph, *Das Gesetz der Transformation der Knochen*, has often been cited but seldom read. A short time with Maquet and Furlong's translation of this book (Wolff 1892) gives one a much better concept of the insights and limitations of "Wolff's law." Several other more recent translations, of

works from both Wolff (“On the Inner Architecture of Bones and its Importance for Bone Growth” and “On the Significance of the Architecture of the Spongy Substance for the Question of Bone Growth: A preliminary publication”) and from Georg Hermann von Meyer [“The architecture of the Trabecular Bone (Tenth contribution on the mechanics of the human skeletal framework)”] with commentary, can be found in recent issues of *Clinical Orthopaedics and Related Research* (Skedros and Brand 2011; Wolff 2010, 2011). Reading these translations together with D’Arcy Thompson’s classic *On Growth and Form* (originally published in 1917) will not only enhance understanding of bone’s form and function, but also will delight those who enjoy reading science as literature—Thompson’s classic is both well written and interesting. Reading the original papers of the experiments and theories that we have described will, of course, give a much greater appreciation for the finer points of this field. Frost’s perspective remains unique in its attempt to balance biologic, mechanical, and clinical observations. An excellent and more detailed description of events associated with mechanical signaling (mechanotransduction) is given in Duncan and Turner (1995). Finally, *Bone Structure and Remodeling*, edited by Odgaard and Weinans (1995), is an excellent collection of papers by key investigators in this field, including a thought-provoking and entertaining exchange of biologic and engineering views of the subject by Currey and Huiskes.

6.9 Exercises

- 6.1. In the Lanyon–Rubin avian model for studying the effects of mechanical loading on bone, a diaphysis is isolated as shown in Fig. 6.21. Pins extend from the metal caps over the bone ends, as shown. Assume a load is applied between the ends of the pins as shown. Calculate the longitudinal strain on the upper and lower surfaces of the bone assuming it is cylindrical with the following characteristics:

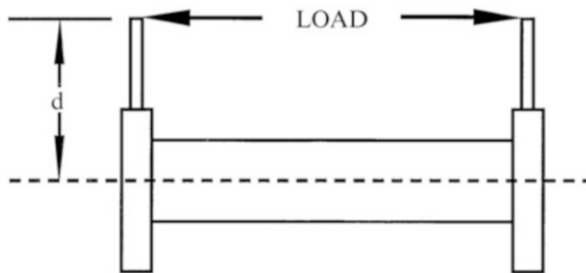


FIGURE 6.21. Isolated diaphysis loaded by projecting pins to be analyzed in Exercise 6.1.

Outer diameter = 20 mm
 Inner diameter = 10 mm
 Load = 30 N
 Moment arm (d) = 40 mm
 Elastic modulus = 20 GPa

- 6.2. Assume that a diaphyseal cross section carries a bending moment M . Let the cross section be circular with inner and outer diameters of 30 and 20 mm, respectively, and assume that the neutral axis (NA) passes through the centroid. Draw the cross section and write an equation for the stress, σ , as a function of angular position θ measured from the neutral axis. Assume that modeling adds or removes bone on the periosteal and endosteal surfaces at a rate proportional to the local longitudinal stress on the surface:

$$M_{F/R} = -K\sigma(\theta) \quad (6.15)$$

where the apposition rate $M_{F/R}$ is negative for tensile stress and positive for compressive stress and the rate constant K is the same for each kind of modeling and each surface. Sketch a line adjacent to the original periosteal surface that would represent the addition or removal of new bone at each point around the circumference after some arbitrary period of time. Do the same for the endosteal surface. (Scale your sketch so that the maximum thickness of bone added is about half the original cortical thickness.) What has happened to the bone's cross-sectional geometry? Would this change be advantageous in some way?

- 6.3. Reanalyze the bone in Exercise 6.2 assuming that $M_{F/R}$ is always formative and proportional to the **magnitude** of the longitudinal surface stress. How does the shape of this new bone distribution compare with that seen in Goodship's pig osteotomy experiment (Goodship et al. 1979)? Discuss how the shape of the periosteal new bone distribution would change if M were proportional to the fourth power of the stress.
- 6.4. Reanalyze the bone in Exercise 6.2 assuming that $M_{F/R}$ is proportional to the radial gradient of stress. Assume that a positive stress gradient produces resorption ($M_{F/R} < 0$) and a negative stress gradient produces formation ($M_{F/R} > 0$) (A positive stress gradient is defined as one in which the stress becomes more tensile or less compressive, as one approaches the surface from within the bone matrix). Let the rate constant be the same for formation and resorption, and for the periosteal and endosteal surfaces. Discuss the variations of the results in Exercises 6.2 through 6.4 and how they compare with the experimental results described in this chapter.

- 6.5. Draw a diagram to show that the stress gradients at the polar (resorption) and equatorial (reversal region) surfaces of an osteonal BMU aligned with a compressive stress direction are positive and negative, respectively, using the sign convention stated in the text and Exercise 6.4.
- 6.6. To see how 0–1 behavior develops in an adaptive model, try the following.⁸ Use the Voigt model described in Sect. 7.6 as a miniature “finite element model” containing just two elements. Let each element have the same cross-sectional area, $A = 5 \text{ mm}^2$, and let this model be loaded by a force $F = 40 \text{ N}$. The strain, ϵ , is the same in both elements. The apparent density and modulus of each element are ρ_i and E_i , respectively, where $i = 1, 2$ identifies the elements. Assume that the time rate of change of apparent density is

$$d\rho_i/dt = B[(U_i/\rho_i) - k] \quad (6.16)$$

where U_i is strain energy density ($U_i = E_i \epsilon^2/2$), B is a coefficient, and k is an equilibrium value of U_i . Assume also that the modulus-density relationship is $E_i = 2.5 \rho_i^3$ (ρ_i in g/cm^3 , E_i in GPa). Derive an expression for the strain as a function of F , A , E_1 and E_2 ; this can be used to calculate U_i . With these tools, set up a spreadsheet and iterate the model through a series of time steps to see what happens to the density of each element. Try starting with $\rho_1 = 1.9$ and $\rho_2 = 1.8 \text{ g/cm}^3$. You will find you have to restrict the densities to a reasonable range; for example, $0\text{--}2.0 \text{ g/cm}^3$. Explore the behavior of this system by varying the starting densities, Eq. (6.15), and the modulus-density equation. How can you change the latter to avoid 0–1 behavior?

- 6.7. Consider the diaphysis of an adult’s long bone subjected to a constant amount of daily torsion. Suppose that the apposition rate on the periosteal and endosteal surfaces is proportional to the difference between the average shear stress, σ_s [see Eq. (7.4)], and a set point or equilibrium value, σ_{se} . If the difference is positive, formation occurs, and if it is negative, resorption occurs. Suppose that σ_{se} is the same for both the periosteal and endosteal surfaces. What would the consequences of this be? Such a model has been discussed in detail by van der Meulen and co-workers (van der Meulen et al. 1993; van der Meulen and Carter 1995)
- 6.8. In Exercise 6.7, a stable situation could be produced by introducing a “dead zone” in place of σ_{se} , that is, a range of stress values within which the apposition rate remains zero. Sketch a graph of apposition rate (positive and negative values) vs. shear stress showing the dead

⁸ This problem is adapted from an example given in Weinans et al. (1992).

zone; label the periosteal and endosteal response zones. What geometric feature of the bone would determine the width of the dead zone? Can you think of an assumption besides a dead zone that would permit stability?

- 6.9. An archer's skeleton, leather quiver strap still around his spine, was recovered from an English ship sunk in 1545. This man's left forearm bones and the tips of the first three fingers on his right hand were obviously thickened (*Encyclopaedia Britannica* 1997, Vol. 29, p. 541). Think about the forces that produced these bony responses in the radius and ulna and in the phalanges. Considering the arm and finger responses separately, do you think the original periosteal responses involved lamellar or woven bone? Why? Should woven bone responses always be categorized as "pathologic" rather than "adaptive"?
- 6.10. Implant manufacturers attempt to design artificial joints that will function for many years, and allow the user the same strength, range of motion and flexibility that the original biological tissues had. However, they are constrained by having to work with the pre-existing structure. This limits the range of designs they can use. Suppose that you were charged with designing a human knee joint "from scratch." What would be your design goals, and how would you best achieve them?
- 6.11. Julius Wolff is routinely given credit for the idea that bone senses and adapts to its mechanical environment. This is now called Wolff's Law, although Wolff himself did not understand what mechanical signals cause bone to adapt. Discuss the features of the mechanical environment that provide signals causing bone to adapt.
- 6.12. Frost's concept of bone adaptation is codified as the "mechanostat." Although it explains some features of bone's response to its mechanical environment, some aspects of it are problematic. Design a set of experiments to test different parts of this theory using an animal model. Describe how you would separate processes associated with modeling from those associated with remodeling. Using your experiments, how would you interpret the mechanism of the adaptive response you observed based on cellular behavior?

References

- Ahlborg HG, Johnell O, Turner CH, Rannevik G, Karlsson MK. Bone loss and bone size after menopause. *N Engl J Med*. 2003;349:327–34.
- Ashby MF. *Materials selection in mechanical design*. Oxford: Butterworth Heinemann; 2004.

- Bab I, Gazit D, Massarawa A, Sela J. Removal of the tibial marrow induces increased formation of bone and cartilage in rat mandibular condyle. *Calcif Tissue Int.* 1985;37:551–5.
- Bakker AD, Soejima K, Klein-Nulemnd J, et al. The production of nitric oxide and prostaglandin E2 by primary bone cells is shear stress dependent. *J Biomech.* 2001;34:671–7.
- Barnes GRG, Pinder DN. In vivo tension and bone strain measurement and correlation. *J Biomech.* 1974;7:35–42.
- Bateson G. The role of somatic change in evolution. *Evolution.* 1963;17:529–39.
- Bateson G. Steps to an ecology of mind. Northvale, NJ: Jason Aronson; 1987.
- Beaupre GS, Orr TE, Carter DR. An approach for time-dependent bone modeling and remodeling—theoretical development. *J Orthop Res.* 1990a;8:651–61.
- Beaupre GS, Orr TE, Carter DR. An approach for time-dependent bone modeling and remodeling—application: a preliminary remodeling simulation. *J Orthop Res.* 1990b;8:662–70.
- Beck TJ, Ruff CB, Bissessur K. Age-related changes in female femoral neck geometry: implications for bone strength. *Calcif Tissue Int.* 1993;53:S41–6.
- Bell C. Animal mechanics, or proofs of design in the animal frame. Cambridge, MA: Morrill Wyman; 1827.
- Bertram JE, Swartz SM. The ‘Law of Bone Transformation’: a case of crying Wolff? *Biol Rev* 1991;66:245–73.
- Biewener AA, Thomason J, Goodship A, Lanyon LE. Bone stress in the horse forelimb during locomotion at different gaits: a comparison of two experimental methods. *J Biomech.* 1983;16:565–76.
- Bonewald LF. The amazing osteocyte. *J Bone Miner Res.* 2011;26:229–38.
- Bourguery JM. Traite Complet de l’Anatomie de l’Homme. Paris: I. Osteologie; 1832.
- Burr DB, Schaffler MB, Yang KH, Wu DD, Lukoschek M, Kandzari D, Sivaneri N, Blaha JD, Radin EL. The effects of altered strain environment on bone tissue kinetics. *Bone.* 1989a;10:215–21.
- Burr DB, Schaffler MB, Yang KH, Lukoschek M, Sivaneri N, Blaha JD, Radin EL. Skeletal change in response to altered strain environments: is woven bone a response to elevated strain? *Bone.* 1989b;10:223–33.
- Burr DB, Milgrom C, Fyhrie D, Forwood M, Nyska M, Finestone A, Hoshaw S, Saiag E, Simkin A. In vivo measurement of human tibial strains during vigorous activity. *Bone.* 1996;18:405–10.
- Cardoso L, Herman BC, Verborgt O, Laudier D, Majeska RJ, Schaffler MB. Osteocyte apoptosis controls activation of intracortical resorption in response to bone fatigue. *J Bone Miner Res.* 2009;24:597–605.
- Carter DR. Mechanical loading history and skeletal biology. *J Biomech.* 1987;20:1095–109.
- Carter DR, Wong M. Mechanical stresses and endochondral ossification in the chondroepiphysis. *J Orthop Res.* 1988;6:148–54.
- Carter DR, Fyhrie DP, Whalen RT. Trabecular bone density and loading history: regulation of connective tissue biology by mechanical energy. *J Biomech.* 1987a;20:785–94.
- Carter DR, Orr TE, Fyhrie DP, Schurman DJ. Influences of mechanical stress on prenatal and postnatal skeletal development. *Clin Orthop Relat Res.* 1987b;219:237–50.

- Carter DR, Blenman PR, Beaupre GS. Correlations between mechanical stress history and tissue differentiation in initial fracture healing. *J Orthop Res.* 1988;6:736–48.
- Carter DR, Orr TE, Fyhrie DP. Relationships between loading history and femoral cancellous bone architecture. *J Biomech.* 1989;22:231–44.
- Chen CS, Ingber DE. Tensegrity and mechanoregulation: from skeleton to cytoskeleton. *Osteoarthritis Cart.* 1999;7:81–94.
- Cochran GBV. Implantation of strain gages on bone in vivo. *J Biomech.* 1972;5:119–23.
- Cochran GBV. A method for direct recording of electromechanical data from skeletal bone in living animals. *J Biomech.* 1974;7:563–5.
- Cohen J, Harris WH. The three-dimensional anatomy of haversian systems. *J Bone Joint Surg.* 1958;40A:419–34.
- Cooney CA, Dave AA, Wolff GL. Maternal methyl supplements in mice affect epigenetic variation and DNA methylation of offspring. *J Nutr.* 2002;132(8 Suppl):2393S–400.
- Cowin SC. The relationship between the elasticity tensor and the fabric tensor. *Mech Mater.* 1985;4:137–47.
- Cowin SC. Wolff's law of trabecular architecture at remodeling equilibrium. *J Biomech Eng.* 1986;108:83–8.
- Cowin SC. The false premise of Wolff's law. *Forma.* 1997;12:247–62.
- Cowin SC, Hegedus DH. Bone remodeling. I: a theory of adaptive elasticity. *J Elast.* 1976;6:313–25.
- Cowin SC, Nachlinger RR. Bone remodeling. III: uniqueness and stability in adaptive elasticity theory. *J Elast.* 1978;8:285–95.
- Cowin SC, Van Buskirk WC. Internal bone remodeling induced by a medullary pin. *J Biomech.* 1978;11:269–75.
- Cowin SC, Hart RT, Balser JR, Kohn DH. Functional adaptation in long bones: establishing in vivo values for surface remodeling rate coefficients. *J Biomech.* 1985;18(9):665–84.
- Crummett LT, Sears BF, Lafon DC, et al. Parthenogenetic populations of the freshwater snail *Campeloma limum* occupy habitats with fewer environmental stressors than their sexual counterparts. *Freshw Biol.* 2013;58:655–63.
- Cruz-Coke R. A genetic description of high-altitude populations. In: Baker PT, editor. *The biology of high altitude peoples.* Cambridge: Cambridge University Press; 1978. p. 47–63.
- Cumming DC. Exercise-associated amenorrhea, low bone density, and estrogen replacement therapy. *Arch Intern Med.* 1996;156:2193–5.
- Currey JD. What is bone for? Property-function relationships in bone. In: Cowin SC, editor. *Mechanical properties of bone.* New York: American Society of Mechanical Engineers; 1981. p. 13–26.
- Currey JD. The mechanical adaptations of bones. Princeton: University Press; 1984.
- Currey JD. Incompatible mechanical properties in compact bone. *J Theor Biol.* 2004;231:569–80.
- Currey JD, Pitchford JW, Baxter PD. Variability of the mechanical properties of bone, and its evolutionary consequences. *J R Soc Interface.* 2007;4:127–35.
- Deere K, Sayers A, Rittweger J, Tobias JH. Habitual levels of high, but not moderate or low, impact activity are positively related to hip BMD and geometry: results from a population-based study of adolescents. *J Bone Miner Res.* 2012;27:1887–95.

- Dempster WT, Liddicoat RT. Compact bone as a non-isotropic material. *Am J Anat.* 1952;91:331–62.
- Divieti PP. Regulation of bone resorption and mineral hoemostasis by osteocytes. *IBMS BoneKEy.* 2009;6:63–70.
- Dodge T, Wanis M, Ayoub R, et al. Mechanical loading, damping and load-driven bone formation in mouse tibiae. *Bone.* 2012;51:810–8.
- Duncan RL, Turner CH. Mechanotransduction and the functional response of bone to mechanical strain. *Calcif Tissue Int.* 1995;57:344–58.
- Eckstein F, Hudelmaier M, Putz R. The effects of exercise on human articular cartilage. *J Anat.* 2006;208:491–512.
- Elementaser M, McMillan M, Smith K, Khanna S, Chantler D, Panarelli M, Ahmed SF. A comparison of the effect of two types of vibration exercise on the endocrine and musculoskeletal system. *J Musculoskelet Neuronal Interact.* 2012;12:144–54.
- Encyclopaedia Britannica. *Encyclopaedia Britannica*, Chicago; 1997.
- Eranen JK, Nilsen J, Zverev VE, et al. Mountain birch under multiple stressors— heavy metal-resistant populations co-resistant to biotic stress but maladapted to abiotic stress. *J Evol Biol.* 2009;22:840–51.
- Firoozbakhsh K, Cowin SC. Devolution of inhomogeneities in bone structure: predictions of adaptive elasticity theory. *J Biomech Eng.* 1980;102:287–93.
- Firoozbakhsh K, Cowin SC. An analytical model of Pauwel's functional adaptation mechanism for bone. *J Biomech Eng.* 1981;103:246–52.
- Frost HM. *Mathematical elements of lamellar bone remodelling.* Springfield: Thomas; 1964.
- Frost HM. The pathomechanics of osteoporoses. *Clin Orthop Relat Res.* 1985;200:198–225.
- Frost HM. *Intermediary organization of the skeleton.* Boca Raton: CRC Press; 1986.
- Frost HM. The mechanostat: a proposed pathogenic mechanism of osteoporoses and the bone mass effects of mechanical and nonmechanical agents. *Bone Miner.* 1987a;2:73–85.
- Frost HM. Bone “mass” and the “mechanostat”: a proposal. *Anat Rec.* 1987b;219:1–9.
- Frost HM. Structural adaptations to mechanical usage (SATMU): 3. The hyaline cartilage modeling problem. *Anat Rec.* 1990;226:423–32.
- Frost HM. Perspectives: a vital biomechanical model of synovial joint design. *Anat Rec.* 1994;240:1–18.
- Fuchs RK, Bauer JJ, Snow CM. Jumping improves hip and lumbar spine bone mass in prepubescent children: a randomized controlled trial. *J Bone Miner Res.* 2001;16:148–56.
- Fyhrie DP, Carter DR. A unifying principle relating stress to trabecular bone morphology. *J Orthop Res.* 1986;4:304–17.
- Galileo G. *Discorsi e dimonstrazioni matematiche*, intorno a due nuove scienze attentanti alla meccanica ed a muovimenti locali. Madison: University of Wisconsin Press; 1638.
- Garman R, Rubin C, Judex S. Small oscillatory accelerations, independent of matrix deformations, increase osteoblast activity and enhance bone morphology. *PLoS One.* 2007;2, e653.
- Goodship AE, Lanyon LE, McFie H. Functional adaption of bone to increased stress. *J Bone Joint Surg.* 1979;61A:539–46.

- Goodwin KJ, Sharkey NA. Material properties of interstitial lamellae reflect local strain environments. *J Orthop Res.* 2002;20:600–6.
- Gould SJ. *The panda's thumb.* New York: WW Norton and Co; 1980.
- Gould SJ, Lewontin RC. The Spandrels of San Marcos and the Panglossian paradigm: a critique of the adaptationist programme. *Proc R Soc Lond B.* 1979;205:581–98.
- Gross TS, Edwards JL, McLeod KJ, Rubin CT. Strain gradients correlate with sites of periosteal bone formation. *J Bone Miner Res.* 1997;12:982–8.
- Gross TS, Poliachik SL, Ausk BJ, Sanford DA, Becker BA, Srinivasan S. Why rest stimulates bone formation: a hypothesis based on complex adaptive phenomenon. *Exerc Sport Sci Rev.* 2004;32:9–13.
- Grossniklaus U, Kelly B, Ferguson-Smith AC, et al. Transgenerational epigenetic inheritance: how important is it? *Nat Rev Genet.* 2013;14:228–35.
- Gunness M, Hock JM. Anabolic effect of parathyroid hormone on cancellous and cortical bone histology. *Bone.* 1993;14:277–81.
- Gusi N, Raimundo A, Leal A. Low-frequency vibratory exercise reduces the risk of bone fracture more than walking: a randomized controlled trial. *BMC Musculoskelet Disord.* 2006;7:1–8.
- Harrigan TP, Hamilton JJ. Necessary and sufficient conditions for global stability and uniqueness in finite element simulations of adaptive bone remodeling. *Int J Solids Struct.* 1994;31:97–107.
- Harrigan TP, Mann RW. Characterization of microstructural anisotropy in orthotropic materials using a 2nd rank tensor. *J Mater Sci.* 1984;19:761–7.
- Harris SE, Gluhak-Heinrich J, Harris MA, et al. CMP1 and MEPE expression are elevated in osteocytes after mechanical loading in vivo: theoretical role in controlling mineral quality in the perilacunar matrix. *J Musculoskelet Neuronal Interact.* 2007;7:313–5.
- Hart RT, Davy DT. Theories of bone modeling and remodeling. In: Cowin SC, editor. *Bone mechanics.* Boca Raton: CRC Press; 1989. p. 253–77.
- Hart RT, Davy DT, Heiple KG. Mathematical modeling and numerical solutions for functionally dependent bone remodeling. *Calcif Tissue Int.* 1984a;36:S104–9.
- Hart RT, Davy DT, Heiple KG. A computational method for stress analysis of adaptive elastic materials with a view toward applications in strain-induced bone remodeling. *J Biomech Eng.* 1984b;106:342–50.
- Haut Donahue TL, Haut TR, Yellowley CE, et al. Mechanosensitivity of bone cells to oscillating fluid flow induced shear stress may be modulated by chemotransport. *J Biomech.* 2003;36:1363–71.
- Hayes WC, Snyder B. Toward a quantitative formulation of Wolff's law in trabecular bone. In: Cowin SC, editor. *Mechanical properties of bone.* New York: American Society of Mechanical Engineers; 1981. p. 43–68.
- Hegedus DH, Cowin SC. Bone remodeling II. Small strain adaptive elasticity. *J Elast.* 1976;6:337–52.
- Hert J, Liskova M, Landrgot B. Influence of the long-term continuous bending on the bone. *Folia Morphologica (Prague)* 1969; 17:369–399.
- Hsieh Y-F, Robling AG, Ambrosius WT, Burr DB, Turner CH. Mechanical loading of diaphyseal bone in vivo: the strain threshold for an osteogenic response varies with location. *J Bone Miner Res.* 2001;16:2291–7.
- Huiskes R, Weinans H, Grootenboer HJ, Dalstra M, Fudala B, Sloof TJ. Adaptive bone remodeling theory applied to prosthetic-design analysis. *J Biomech.* 1987;20:1135–50.

- Hylander WL. The functional significance of primate mandibular form. *J Morphol.* 1979;160:223–40.
- Iwamoto J, Sato Y, Takeda T, Matsumoto H. Whole body vibration exercise improves body balance and walking velocity in postmenopausal osteoporotic women treated with alendronate: Galileo and Alendronate Intervention Trail [sic] (GAIT). *J Musculoskelet Neuronal Interact.* 2012;12:136–43.
- Jacobs CR. Numerical simulation of bone adaptation to mechanical loading. Doctoral dissertation. Palo Alto, CA: Stanford University; 1994.
- Jansen M, Stoks R, Coors A, et al. Collateral damage: rapid exposure-induced evolution of pesticide resistance leads to increased susceptibility to parasites. *Evolution.* 2011;65:2681–91.
- Jaworski ZFG, Uthoff HK. Reversability of nontraumatic disuse osteoporosis during its active phase. *Bone.* 1986;7:431–9.
- Jaworski ZFG, Liskova-Kiar M, Uthoff HK. Effect of long-term immobilization on the pattern of bone loss in older dogs. *J Bone Joint Surg.* 1980;62B:104–10.
- Jee WSS, Li XJ, Schaffler MB. Adaptation of diaphyseal structure with aging and increased mechanical usage in the adult rat: a histomorphometrical and biomechanical study. *Anat Rec.* 1991;230:332–8.
- Jepsen KJ, Andarawis-Puri N. The amount of periosteal apposition required to maintain bone strength during aging depends on adult bone morphology and tissue-modulus degradation rate. *J Bone Miner Res.* 2012;27:1916–26.
- Johnson LC. Morphologic analysis in pathology: the kinetics of disease and general biology in bone. In: Frost HM, editor. *Bone biodynamics*. Boston: Little, Brown; 1964.
- Kato T, et al. Effect of low-repetition jump training on bone mineral density in young women. *J Appl Physiol.* 2006;100:839–43.
- Kennedy OD, Herman BC, Laudier DM, Majeska RJ, Sun HB, Schaffler MB. Activation of resorption in fatigue-loaded bone involves both apoptosis and active pro-osteoclastogenic signaling by distinct osteocyte populations. *Bone.* 2012;50:1115–22.
- Kim CH, Takai E, Zhou H, von Stecho D, Muller R, Dempster DW, Guo XE. Trabecular bone response to mechanical and parathyroid hormones stimulation: the role of mechanical microenvironment. *J Bone Miner Res.* 2003;18:2116–25.
- Klesges RC, Ward KD, Shelton ML, Applegate WB, Cantler ED, Palmieri GM, Harmon K, Davis J. Changes in bone mineral content in male athletes. Mechanisms of action and intervention effects. *J Am Med Assoc.* 1996;276:226–30.
- Knothe-Tate ML, Knothe U, Niederer P. Experimental elucidation of mechanical load-induced fluid flow and its potential role in bone metabolism and functional adaptation. *Am J Med Sci.* 1998;316:189–95.
- Knothe-Tate ML, Steck R, Forwood MR, Niederer P. In vivo demonstration of load-induced fluid flow in the rat tibia and its potential implications for processes associated with functional adaptation. *J Exp Biol.* 2000;203:2737–45.
- Koch JC. The laws of bone architecture. *Am J Anat.* 1917;21:177–293.
- LaMothe JM, Zernicke RF. Rest insertion combined with high-frequency loading enhances osteogenesis. *J Appl Physiol.* 2004;96:1788–93.
- Lanyon LE. Strain in sheep lumbar vertebrae recorded during life. *Acta Orthop Scand.* 1971;42:102–12.

- Lanyon LE. In vivo bone strain recorded from thoracic vertebrae of sheep. *J Biomech.* 1972;5:277–81.
- Lanyon LE. Analysis of surface bone in the calcaneus of sheep during normal locomotion. *J Biomech.* 1973;6:41–9.
- Lanyon LE. Experimental support for the trajectorial theory of bone structure. *J Bone Joint Surg.* 1974;56B:160–6.
- Lanyon LE. Functional strain as a determinant of bone remodeling. *Calcif Tissue Int.* 1984;36:S56–61.
- Lanyon LE. The importance of mechanical adaptation in controlling bone architecture and averting bone fracture. In: Marcelli C, Sebert JL, editors. *Architecture et Resistance Mecanique Osseuses*. Paris: Masson; 1993. p. 58–67.
- Lanyon LE. Osteoporosis and mechanically related bone modeling. In: Menczel J, Robin GC, Makin M, Steinberg R, editors. *Osteoporosis. Proceedings of the International Symposium, Jerusalem Osteoporosis Center*. Chichester: Wiley; 1981.
- Lanyon LE, Bourn S. The influence of mechanical function on the development and remodeling of the tibia. *J Bone Joint Surg.* 1979;62A:263–73.
- Lanyon LE, Rubin CT. Static vs. dynamic loads as an influence on bone remodeling. *J Biomech.* 1984;17:897–905.
- Lanyon LE, Smith RN. Measurements of bone strain in the walking animal. *Res Vet Sci.* 1969;10:93–4.
- Lanyon LE, Smith RN. Bone strain in the tibia during normal quadrupedal locomotion. *Acta Orthop Scand.* 1970;41:238–48.
- Lanyon LE, Hampson WGJ, Goodship AE, Shah JS. Bone deformation recorded in vivo from strain gauges attached to the human tibial shaft. *Acta Orthop Scand.* 1975;46:256–68.
- Lanyon LE, Goodship AE, Pye CJ, MacFie JH. Mechanically adaptive bone remodeling. *J Biomech.* 1982;15:141–54.
- Lanyon LE, Rubin CT, Baust G. Modulation of bone loss during calcium insufficiency by controlled dynamic loading. *Calcif Tissue Int.* 1986;38:209–16.
- Li J, Duncan RL, Burr DB, Gattone VH, Turner CH. Parathyroid hormone enhances mechanically induced bone formation, possibly involving L-type voltage-sensitive calcium channels. *Endocrinology.* 2003;144:1226–33.
- Martin RB. The effects of geometric feedback in the development of osteoporosis. *J Biomech.* 1972;4:447–55.
- Martin RB. The usefulness of mathematical models for bone remodeling. *Yearb Phys Anthropol.* 1985;28:227–36.
- Martin RB. A mathematical model for fatigue damage repair and stress fracture in osteonal bone. *J Orthop Res.* 1995;13:309–16.
- Martin RB, Atkinson PJ. Age and sex-related changes in the structure and strength of the human femoral shaft. *J Biomech.* 1977;10:223–31.
- Martin RB, Clark RN, Advani S. An electro-mechanical basis for osteonal mechanics. In: Brighton CT, Goldstein S, editors. *Advances in bioengineering*. New York: American Society of Mechanical Engineers; 1974.
- Martin RB, Pickett JC, Zinaich S. Studies of skeletal remodeling in aging men. *Clin Orthop Relat Res.* 1980;149:268–82.
- McLeod KJ, Bain SD, Rubin CT. Dependence of bone adaptation on the frequency of induced dynamic strains. *Trans Orthop Res Soc.* 1990;15:103.

- Monro A. The anatomy of the human bones, nerves and lacteal sac and duct. (Published in Dublin.); 1976.
- Mori S, Harruff R, Ambrosius W, Burr DB. Trabecular bone volume and microdamage accumulation in the femoral heads of women with and without femoral neck fractures. *Bone*. 1997;21:521–6.
- Mosekilde L, Mosekilde L. Sex differences in age-related changes in vertebral body size, density and biomechanical competence in normal individuals. *Bone*. 1990;11:67–73.
- Mosley JR, Lanyon LE. Strain rate as a controlling influence on adaptive modeling in response to dynamic loading of the ulna in growing male rats. *Bone*. 1998;23:313–8.
- Moss ML. Studies on the acellular bone of teleost fish. II. Response to fracture under normal and acalcemic conditions. *Acta Anat*. 1962;48:46–60.
- Moustafa A, Sugiyama T, Prasad J, Saman G, Gross TS, Lanyon LE, Price JS. Mechanical loading-related changes in osteocyte sclerostin expression in mice are more closely associated with the subsequent osteogenic response than the peak strains engendered. *Osteoporos Int*. 2012;23:1225–34.
- Mullender MG, Huiskes R, Weinans H. A physiologic approach to the simulation of bone remodeling as a self-organizational control process. *J Biomech*. 1994;27:1389–94.
- Nakashima T, Hayashi M, Fukunaga T, et al. Evidence for osteocyte regulation of bone homeostasis through RANKL expression. *Nat Med*. 2011;17:1231–4.
- Ng CYP, Choi VWY, Lam ACL, et al. The multiple stressor effect in zebrafish embryos from simultaneous exposure to ionising radiation and cadmium. *J Radiol Prot*. 2013;33:113–21.
- Noble BS, Stevens H, Loveridge N, Reeve J. Identification of apoptotic changes in osteocytes in normal and pathological human bone. *Bone*. 1997;20:273–82.
- Noble BS, Peet N, Stevens HY, et al. Mechanical loading: biphasic osteocyte survival and targeting of osteoclasts for bone destruction in rat cortical bone. *Am J Physiol Cell Physiol*. 2002;284:C934–43.
- Nowlan NC, Prendergast PJ. Evolution of mechanoregulation of bone growth will lead to non-optimal bone phenotypes. *J Theor Biol*. 2005;235:408–18.
- O'Connor JA, Lanyon LE, MacFie H. The influence of strain rate on adaptive bone remodeling. *J Biomech*. 1982;15:767–81.
- Oden ZM, Hart RT, Forwood MR, Burr DB. A priori prediction of functional adaptation in canine radii using a cell-based mechanistic model. *Trans Orthop Res Soc*. 1995;20:296.
- Odgaard A, Weinans H. Bone structure and remodeling. Singapore: World Scientific; 1995.
- Owan I, Burr DB, Turner CH, Qui J, Tu Y, Onyia JE, Duncan RL. Mechanotransduction in bone: osteoblasts are more responsive to fluid forces than mechanical strain. *Am J Physiol*. 1997;273:C810–5.
- Ozicivici E, Garman R, Judex S. High frequency oscillatory motions enhance the simulated mechanical properties of non-weight bearing trabecular bone. *J Biomech*. 2007;40:3403–11.
- Ozicivici E, Luu YK, Adler B, Qin Y-X, Rubin J, Judex S, Rubin CT. Mechanical signals as anabolic agents in bone. *Nat Rev Rheumatol*. 2010;6:50–9.
- Parenti LR. The phylogenetic significance of bone types in euteleost fishes. *Zool J Linn Soc*. 1986;87:37–51.

- Pauwels F. Biomechanics of the locomotor apparatus: contributions on the functional anatomy of the locomotor apparatus. Berlin: Springer; 1980.
- Pedersen P, editor. Optimal design with advanced materials. Amsterdam: Elsevier; 1993.
- Pembrey ME, Bygren LO, Kaati G, Edvinsson S, Northstone K, Sjöström M, Golding J. Sex-specific, male-line transgenerational responses in humans. *Eur J Hum Genet*. 2006;14(2):159–66.
- Peterman MM, Hamel AJ, Cavanagh PR, Piazza SJ, Sharkey NA. In vitro modeling of human tibial strains during exercise in micro-gravity. *J Biomech*. 2001;34(5):693–8.
- Peterson RE. Stress concentration factors. New York: Wiley; 1974.
- Poole KES, Mayhew PM, Rose CM, Brown JK, Bearcroft PJ, Loveridge N, Reeve J. Changing structure of the femoral neck across the adult female lifespan. *J Bone Miner Res*. 2010;25:482–91.
- Price C, Zhou X, Li W, Wang L. Real-time measurement of solute transport within the lacunar-canalicular system of mechanically loaded bone: direct evidence for load-induced fluid flow. *J Bone Miner Res*. 2011;26:277–85.
- Reiter TJ, Rammerstorfer FG. Simulation of natural adaptation of bone material and application in optimum composite design. In: Pedersen P, editor. Optimal design with advanced materials. Amsterdam: Elsevier; 1993. p. 25–36.
- Riddle RC, Donahue HJ. From streaming potentials to shear stress: 25 years of bone cell mechanotransduction. *J Orthop Res*. 2009;27:143–9.
- Riddle RC, Hippe KR, Donahue HJ. Chemotransport contributes to the effect of oscillatory fluid flow on human bone marrow stromal cell proliferation. *J Orthop Res*. 2008;26:918–24.
- Rittweger J, Schiesel H, Felsenberg D. Oxygen uptake during whole-body vibration exercise: comparison with squatting as a slow voluntary movement. *Eur J Appl Physiol*. 2001;86:169–73.
- Robling AG, Burr DB, Turner CH. Partitioning a daily mechanical stimulus into discrete loading bouts improves the osteogenic response to loading. *J Bone Miner Res*. 2000;15:1596–602.
- Robling AG, Burr DB, Turner CH. Skeletal loading in animals. *J Musculoskelet Neuronal Interact*. 2001a;1:249–62.
- Robling AG, Burr DB, Turner CH. Recovery periods restore mechanosensitivity to dynamically loaded bone. *J Exp Biol*. 2001b;204:3389–99.
- Robling AG, Duijvelaar KM, Gevers JV, Ohashi N, Turner CH. Modulation of appositional and longitudinal bone growth in the rat ulna by applied static and dynamic force. *Bone*. 2001c;29:105–13.
- Robling AG, Hinant FM, Burr DB, Turner CH. Improved bone structure and strength after long-term mechanical loading is greatest if loading is separated into short bouts. *J Bone Miner Res*. 2002;17:1545–54.
- Robling AG, Niziolek PJ, Baldridge LA, et al. Mechanical stimulation of bone in vivo reduces osteocyte expression of Sost/Sclerostin. *J Biol Chem*. 2008;283:5866–75.
- Roesler H. Some historical remarks on the theory of cancellous bone structure (Wolff's law). In: Cowin SC, editor. Mechanical properties of bone. New York: American Society of Mechanical Engineers; 1981. p. 27–42.
- Roesler H. The history of some fundamental concepts in bone biomechanics. *J Biomech*. 1987;20:1025–34.

- Roux W. Der zuchtende Kampf der Teile, oder die Teilauslese' im Organismus (Theorie der 'funktionellen Anpassung'). Leipzig: Wilhelm Engelmann; 1881.
- Rubin CT, Lanyon LE. Limb mechanics as a function of speed and gait: a study of functional strains in the radius and tibia of horse and dog. *J Exp Biol.* 1982;101:187–211.
- Rubin CT, Lanyon LE. Regulation of bone formation by applied dynamic loads. *J Bone Joint Surg.* 1984;66A:397–402.
- Rubin CT, Lanyon LE. Regulation of bone mass by mechanical strain magnitude. *Calcif Tissue Int.* 1985;37:411–7.
- Rubin C, Gross T, Qin Y-X, Fritton S, Guilak F, McLeod K. Differentiation of the bone-tissue remodeling response to axial and torsional loading in the turkey ulna. *J Bone Joint Surg.* 1996;78A:1523–34.
- Ruff CB, Hayes WC. Sex differences in age-related remodeling of the femur and tibia. *J Orthop Res.* 1988;6:886–96.
- Ryder KD, Duncan RL. Parathyroid hormone modulates the response of osteoblast-like cells to mechanical stimulation. *Calcif Tissue Int.* 2001;67:241–6.
- Schaffler MB. Immobilization induced bone loss: quantitative histological studies of cortical bone resorption. *Trans Orthop Res Soc.* 1990;15:187.
- Sen B, Xie Z, Case N, Styner M, Rubin CT, Rubin J. Mechanical signal influence on mesenchymal stem cell fate is enhanced by incorporation of refractory periods into the loading regimen. *J Biomech.* 2011;44:593–9.
- Shahar R, Dean MN. The enigmas of bone without osteocytes. *BoneKEY Reports* 2, Article 343; 2013.
- Skedros JG, Brand RA. Biographical Sketch: Georg Hermann von Meyer (1815–1892). *Clin Orthop Relat Res.* 2011;469:3072–6.
- Smith KK, Hylander WL. Strain gauge measurement of mesokinetic movement in the lizard *Varanus exanthematicus*. *J Exp Biol.* 1985;114:53–70.
- Smith TS, Martin RB, Hubbard M, Bay BK. Surface remodeling of trabecular bone using a tissue-level model. *J Orthop Res.* 1997;15:593–600.
- Srinivasan S, Weimer DA, Agans SC, Bain SD, Gross TS. Low-magnitude mechanical loading becomes osteogenic when rest is inserted between each load cycle. *J Bone Miner Res.* 2002;17:1613–20.
- Strang G, Kohn RV. Optimal design in elasticity and plasticity. *Int J Numer Methods Eng.* 1986;22:183–8.
- Sugiyama T, Meakin LB, Browne WJ, Galea GL, Price JS, Lanyon LE. Bones' adaptive response to mechanical loading is essentially linear between the low strains associated with disuse and the high strains associated with the lamellar/woven bone transition. *J Bone Miner Res.* 2012;27:1784–93.
- Thompson DW. On growth and form. London: Cambridge University Press (1992 republication by Dover Publications, New York.); 1942.
- Timoshenko SP. History of strength of materials. New York: McGraw-Hill; 1953.
- Tomkinson A, Reeve J, Shaw RW, Noble BS. The death of osteocytes via apoptosis accompanies estrogen withdrawal in human bone. *J Clin Endocrinol Metab.* 1997;82:3128–35.
- Tu X, Rhee Y, Condon KW, et al. Sost downregulation and local Wnt signaling are required for the osteogenic response to mechanical loading. *Bone.* 2012;50:209–17.
- Turner CH. On Wolff's law of trabecular architecture. *J Biomech.* 1992;25:1–9.

- Turner CH. Three rules for bone adaptation to mechanical stimuli. *Bone*. 1998;23:399–407.
- Turner AS, Mills EJ, Gabel AA. In vivo measurement of bone strain in horse. *Am J Vet Res*. 1975;36:1573–9.
- Turner CH, Akhter MP, Raab DM, Kimmel DB, Recker RR. A non-invasive, in vivo model for studying strain adaptive bone modeling. *Bone*. 1991;12:73–9.
- Turner CH, Forwood M, Rho J-Y, Yoshikawa T. Mechanical loading thresholds for lamellar and woven bone formation. *J Bone Miner Res*. 1994a;9:87–97.
- Turner CH, Forwood MR, Otter MW. Mechanotransduction in bone: do bone cells act as sensors of fluid flow? *FASEB J*. 1994b;8:875–8.
- Turner CH, Owan I, Takano Y. Mechanotransduction in bone: role of strain rate. *Am J Physiol*. 1995a;269:E438–42.
- Turner CH, Yoshikawa T, Forwood MR, Sun TC, Burr DB. High frequency components of bone strain in dogs measured during various activities. *J Biomech*. 1995b;28:39–44.
- Turner CH, Robling AG, Duncan RL, Burr DB. Do bone cells behave like a neuronal network? *Calcif Tissue Int*. 2002;70:435–42.
- Uthoff HK, Jaworski ZFG. Bone loss in response to long-term immobilization. *J Bone Joint Surg*. 1978;60B:420–9.
- Umemura Y, Ishiko T, Yamauchi T, Kurono M, Mashiko S. Five jumps per day increase bone mass and breaking force in rats. *J Bone Miner Res*. 1997;12:1480–5.
- van der Meulen MCH, Carter DR. Developmental mechanics determine long bone allometry. *J Theor Biol*. 1995;172:323–7.
- van der Meulen MCH, Beaupre GS, Carter DR. Mechanobiological influences in long bone cross-sectional growth. *Bone*. 1993;14:635–42.
- van Oers RFM, Ruimerman R, Tanck E, Hilbers PA, Huiskes R. A unified theory for osteonal and hemi-osteonal remodeling. *Bone*. 2008;42:250–9.
- von Meyer GH. Die Architektur der Spongiosa. *Archs Anat Physiol wiss Med*. 1867;34:615–28. Trans by Amundson PK, Skedros J, Brand R. *Clin Orthop Rel Res*. 2011; 469: 3079–84.
- Wang N, Butler JP, Ingber DE. Mechanotransduction across the cell surface and through the cytoskeleton. *Science*. 1993;260:1124–7.
- Wang L, Cowin SC, Weinbaum S, Fritton SP. Modeling tracer transport in an osteon under cyclic loading. *Ann Biomed Eng*. 2000;28:1200–9.
- Wang L, Wang Y, Han Y, Henderson SC, Majeska RJ, Weinbaum S, Schaffler MB. In situ measurement of solute transport in the bone lacunar-canalicular system. *Proc Natl Acad Sci USA*. 2005;102:11911–6.
- Ward FO. Outlines of human osteology. Published in London, England; 1838.
- Weinans H, Huiskes R, Grootenboer HJ. The behavior of adaptive bone-remodeling simulation models. *J Biomech*. 1992;25:1425–41.
- Weinbaum S, Cowin SC, Zeng Y. A model for the excitation of osteocytes by mechanical loading-induced bone fluid shear stresses. *J Biomech*. 1994;27:339–60.
- Whalen RT, Carter DR, Steele CR. Influence of physical activity on the regulation of bone density. *J Biomech*. 1988;21:825–37.
- Wolff J. Über die bedeutung der architektur der spongiosa. *Zentralblatt für die Medizinische Wissenschaft*. 1869; VI. Jahrgang: S. 223–34.
- Wolff J. Über die innere architektur der knochen und ihre bedeutung für die frage vom knochenwachstum. *Virchows Archiv für Pathologische Anatomie und Physiologie*. 1870;50:389–453.

- Wolff J. The law of bone remodeling. (Translation of Wolff's *Das Gesetz der Transformation der Knochen* by Maquet P, Furlong R.) Berlin: Springer; 1892.
- Wolff J. On the inner architecture of bones and its importance for bone growth. Trans by Heller MO, Taylor WR, Aslanidis N, Duda GN. *Clin Orthop Rel Res*. 2010; 468: 1056–65.
- Wolff J. On the significance of the architecture of the spongy substance for the question of bone growth: a preliminary publication. Trans by Amundson PK, Skedros J, Brand R. *Clin Orthop Rel Res*. 2011; 469: 3077–78.
- Wong M, Carter DR. Mechanical stress and morphogenetic endochondral ossification of the sternum. *J Bone Joint Surg*. 1988;70A(7):992–1000.
- Woo SL, Kuei SC, Amiel D, Gomez MA, Hayes WC, White FC, Akeson WH. The effect of prolonged physical training on the properties of long bone: a study of Wolff's law. *J Bone Joint Surg*. 1981;63A:780–7.
- Wysocki A, Butler M, Shamliyan T, Kane RL. Whole body vibration therapy for osteoporosis: state of the science. *Ann Intern Med*. 2011;155:680–6.
- Xiong J, Onal M, Jilka RL, et al. Matrix-embedded cells control osteoclast formation. *Nat Med*. 2011;17:1235–41.
- Yang PF, Brüggermann GP, J R. What do we currently know from in vivo bone strain measurements in humans? *J Musculoskelet Neuronal Interact*. 2011;11:8–20.
- You J, Yellowley CE, Donahue HJ, Zhang Y, Chemn Q, Jacobs CR. Substrate deformation levels associated with routine physical activity are less stimulatory to bone cells relative to loading-induced oscillatory fluid flow. *J Biomech Eng*. 2000;122:387–93.

7

Mechanical Properties of Bone

Strength of bones is dependent upon the material, the microscopic structure and the shape of the whole bone.

A.A. Rauber, 1876, as quoted in Roesler (1987)

Abstract We have now learned much about the variable structure of the material inside bones. Previously, we learned that the muscle forces acting on the skeleton are quite large relative to the forces exerted on it by the outside world. It is now time to bring these two aspects of skeletal tissue mechanics together: to learn how strong bones are relative to the loads they must bear, and how the mechanical properties of bones depend on their structure.

Keywords Stress • Strain • Strength • Fracture • Osteon • Density • Cancellous • Cortical • Composite

7.1 Introduction

We have now learned much about the variable structure of the material inside bones. Previously, we learned that the muscle forces acting on the skeleton are quite large relative to the forces exerted on it by the outside world. It is now time to bring these two aspects of skeletal tissue mechanics together: to learn how strong bones are relative to the loads they must bear, and how the mechanical properties of bones depend on their structure.

The mechanical properties of bones are governed by the same principles as those of manmade load-bearing structures. However, the ability of vertebrates to adapt their bone structure to the imposed loading, in terms of both the evolution of a genetic plan for the skeleton and the ability to acutely modify this plan by cellular activity, results in a structure that is highly complex and, when healthy, very effective. The purpose of this chapter is to explain how the various attributes of whole bones combine to determine their *monotonic* strength and stiffness. (The term monotonic refers to the behavior of a structure loaded steadily to failure, in distinction to repetitive or *fatigue* loading, which is considered in Chap. 8.)

7.2 Fundamentals of Solid Mechanics

To lay a foundation for subsequent discussion of the mechanical properties of bone, this section briefly reviews the mathematical formulation of the mechanical properties of materials and structures. Students trained in engineering may well skip this section; those with no background in this area can find other biologically oriented treatments of this material in an excellent review by Turner and Burr (1993) and books by Currey (2002), Frost (1973), Wainwright et al. (1976) and Cowin (2001).

7.2.1 *Strength and Stiffness of a Structure*

To determine the mechanical properties of a structure—or, more typically, a major part of a structure, such as a beam in a bridge or the shaft of a femur—engineers conduct a *load-deformation* test. The structure is loaded in a manner appropriate for its actual function—tension, compression, torsion, or bending—and its deformation is recorded as a function of the applied load. Figure 7.1 shows such a test for a femur shaft loaded in compression. Typically, load and deformation are linearly proportional for some time, until the *proportional limit* is reached; then, the slope of the load-deformation curve is reduced. The point on the load-deformation curve where this begins to happen is known as the *yield point* because the structure begins to “yield” or *plastically* deform. If the structure is unloaded after exceeding the yield point it will not return to its original shape, reflective of the plastic deformation that has occurred. Before the yield point the deformation is *elastic*, meaning that the structure will return to its original shape when the load is removed. If the load continues to increase, an *ultimate* load is reached; in some materials (e.g., metals), this may not be the point at which a catastrophic failure (or fracture) occurs. The point marking the latter event is known as the *failure point*. (In cortical bone, the ultimate and failure loads are usually identical.) The *strength* of the structure may be defined as the load at the yield or failure points, or as the ultimate load, depending on the circumstances. There is no absolute definition of strength, but if it is not otherwise specified, it may be taken as the load at the yield point (Figs. 7.1a and b).

The *stiffness* or *rigidity* of the structure is the load required to deform it a given amount. This is the same thing as the slope of the load-deformation curve. The slope is usually measured within the proportional range, where it is constant. *Compliance* is the reciprocal of stiffness, and is a measure of the ease of deforming the structure. The deformations at the yield and failure points are measures of the extent to which the structure may be deformed.

The *work* or *energy* required to yield or fracture the structure is another important mechanical property that is determined by measuring the area under the load-deformation curve out to the yield or failure point.

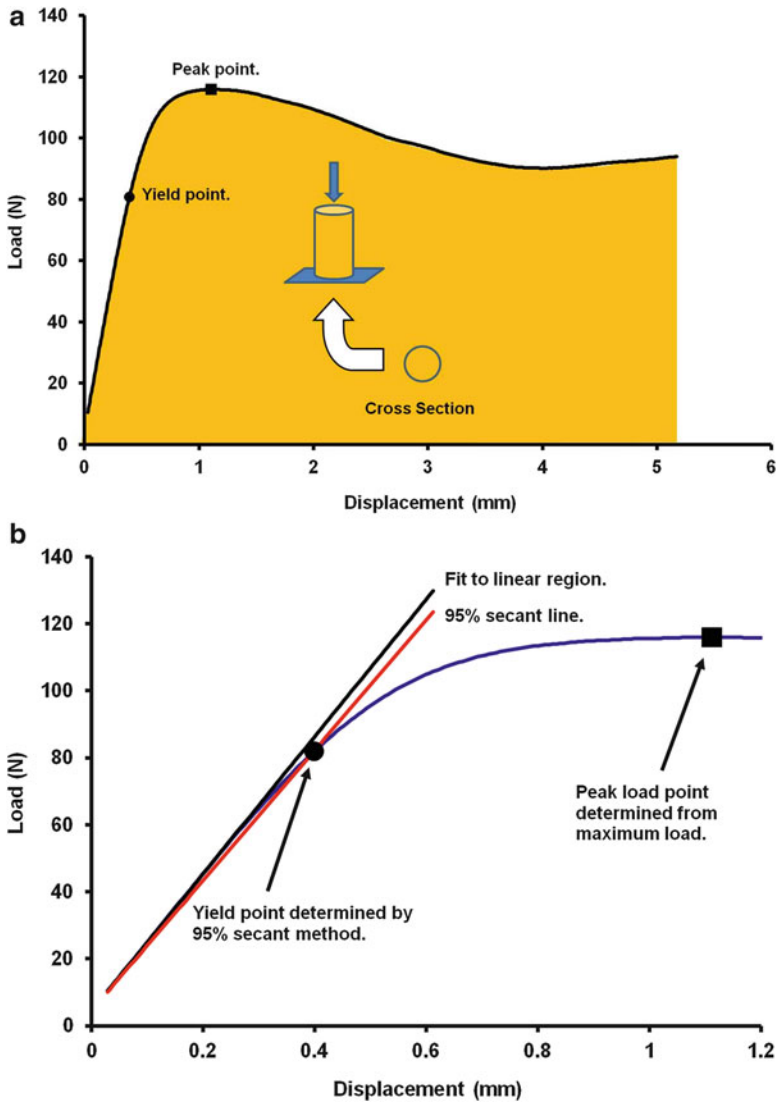


FIGURE 7.1. A compressive load-deformation test. The peak load point is the largest load carried by the specimen. The slope of the linear portion of the curve (to the yield point) is the stiffness. The area under the curve (orange) is the energy absorbed during the test (a). The yield point is determined using the 95 % secant method (ASTM Standard E-399) one of several approaches that can be used to determine yield (b).

7.2.2 Stress and Strain

The mechanical properties of a structure depend on both its geometry and the properties of the material inside. To determine the properties of the material, it is necessary to “normalize out” the geometric effects. To do this, engineers have defined *stress* and *strain*. In its simplest definition, stress is load per unit area, as shown at upper left in Fig. 7.2. The more general definition of static stress is also shown in this figure. If one imagines a small cube of material at a point inside a loaded structure, one can see that each surface of the cube has three loads acting on it: a *normal* load perpendicular to the surface, and two *shear* loads parallel to the surface. If these three loads are divided by the area of the cube’s face, they become the three stresses associated with the direction of that face. The cube has six faces, but because there must be equal forces on opposite faces for stability, it turns out that there are $3 \times 3 = 9$ *stress components* at each point in a loaded material. The first three of these are normal stress components, and the other six are shear stresses. Some of the shear stresses are equal to one another so that the cube is in *moment equilibrium*. As a result the total number of independent stress components is ultimately reduced to six: three normal and three shear. We represent these

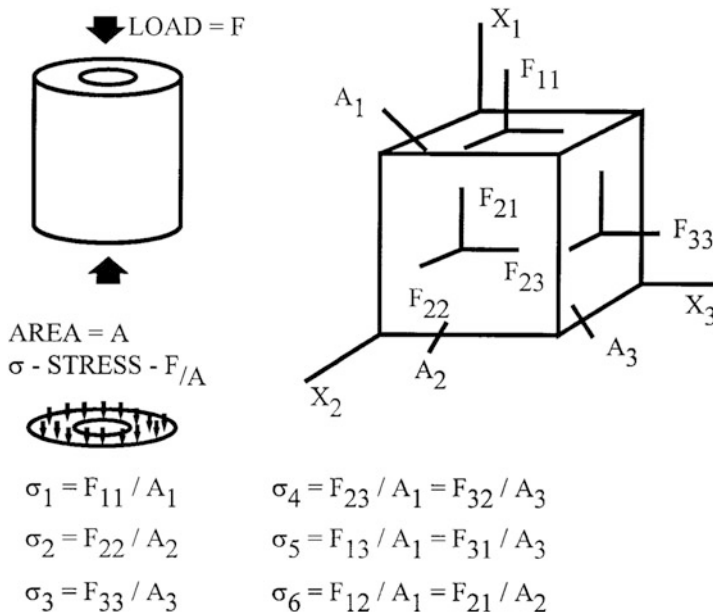


FIGURE 7.2. Simple and tensorial definitions of stress. The *cube at right* represents a free-body diagram of a small region within the loaded bone segment at *left*. F represents the normal and shear forces exerted on this region by the stresses in the surrounding bone. When divided by the areas of the cube’s faces (A_1 , A_2 , A_3), these forces become the stress tensor components at the cube’s location.

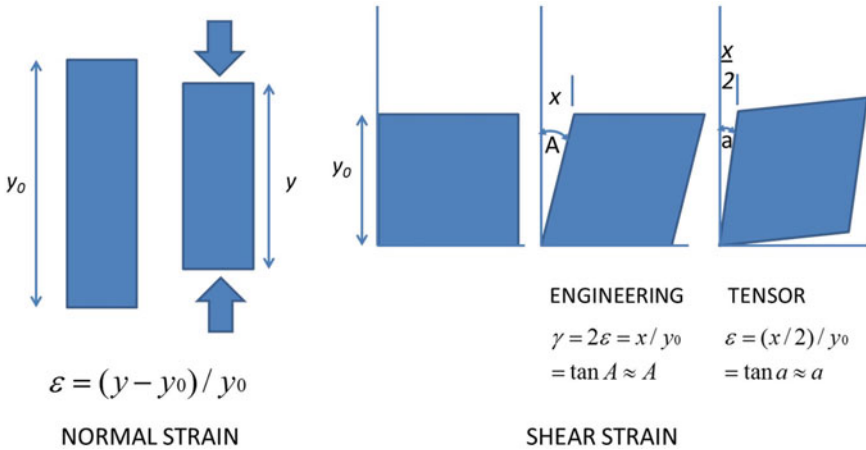


FIGURE 7.3. Normal (*left*) and shear (*right*) strains are defined. A represents the angle through which the side of the shear specimen is rotated. Strains are dimensionless; shear strains may be expressed as radians.

by the symbol σ_i where $i=1$ to 6. Stress has more components than a vector (which has only three) and is a second order tensor. Stress is commonly measured in Newtons/m² (N/m²), which is a *Pascal* (abbreviated Pa). A typical failure stress in bone would be 150×10^6 Pa which is usually abbreviated as 150 MPa (megapascals).

Strain is another second order tensor quantity that engineers define to be able to discuss the deformation of a material without regard to its structural geometry. In its simplest definition, strain is the fractional change in dimension of a loaded body; Figure 7.3 illustrates normal and shear strains. One can show that there are, in general, six strain tensor components, three normal and three shear, which describe the deformation of a cube of material at any point in a loaded structure. We represent these by the symbol ε_j , where $j=1$ to 6. Strain is a dimensionless ratio, but it is common to measure it in *microstrain*, or 10^{-6} mm/mm. For example, a strain of 0.02 mm/mm would be the same as 2 % strain or 20,000 *microstrain* (abbreviated $\mu\varepsilon$); this would be a dangerously large strain in a bone.

When a structure is loaded in tension, it becomes longer in the load direction, but it usually “waists” or becomes narrower in the transverse directions. (Of course, the opposite phenomenon occurs in a compressed specimen.) The ratio of the transverse strain component to the longitudinal strain component is known as *Poisson’s ratio*. For geometric reasons, this ratio must be less than 0.5 in ordinary materials. For incompressible materials (e.g., liquids), Poisson’s ratio is 0.5. Typical values for solids, including bone, are about 0.3. (For another biological material, cork, Poisson’s ratio is nearly 0. This makes it easier to push a cork into a wine bottle, but that

obviously has nothing to do with why the cork material grows with this unusual property.)

7.2.3 *Principal Directions*

Perhaps it is intuitively clear that the stresses associated with a cube in an arbitrarily loaded structure depend not only on the location of the cube but also on its orientation. In general each face of the cube experiences both normal and shear stresses, and the relative magnitudes of these vary if the cube is placed in different orientations. It happens, however, that it is **always** possible to find an orientation such that all the shear stresses become zero. These three mutually perpendicular directions are called the *principal stress directions*. A similar set of *principal strain directions* exists for each point in a loaded structure. In simple materials, such as steel, these two principal direction sets are identical, but in complex materials like bone they may differ. When examining stress or strain as matrix (or second order tensor) quantities, the principal values are the *eigenvalues* of the tensor and the principal directions are the *eigenvectors*. For a second order tensor, there are three eigenvalues and eigenvectors. The word *eigen* is a German adjective meaning *own* or *self* that identifies that the eigenvalues and vectors are intrinsic to the matrix.

7.2.4 *Strength and Modulus of a Material*

A load-deformation curve may be converted to a stress-strain curve by using appropriate formulas to change load to stress and deformation to strain. For example, in a compressed bone shaft, the load is divided by the cross-sectional area of the bone cortex to obtain stress, and the deformation is divided by the original length of the shaft segment to obtain strain. After this is done, all the strength terms previously defined for the structure (yield, ultimate, and failure strengths, along with structural stiffness) can be redefined for the **material** in terms of stresses (yield, ultimate and failure stress, along with elastic modulus). The area under the stress-strain curve to failure is analogous to the energy to failure of a structure and is called the modulus of toughness for the material.

Stress and strain in a solid are related to one another. In the simplest material the stress is proportional to strain. The proportional law for a simple material was discovered by Robert Hooke and is named in his honor. In its simplest form, Hooke's law relates one component of stress to its corresponding strain component (e.g., stress and strain in longitudinal tension), and it applies only to the linear portion of the stress-strain curve between the origin and the proportional limit. The slope of this line is the modulus of the **material**. For normal loading (tension or compression) this is the *elastic modulus* or *Young's modulus*. For shear loading, the modulus is

the *shear modulus*. The normal and shear moduli have the same units as stress, and values for bone are typically of the order of gigapascals (10^9 Pa), abbreviated GPa.

Box 7.1 Robert Hooke and Thomas Young: The True Theory of Elasticity or Springiness, and A Particular Explication Thereof In Several Subjects In Which It Is To Be Found . . . CEIINOSSSTTUV

Thus did Robert Hooke (1635–1703) record his law of elasticity as a Latin anagram in 1675. He did this as a result of his well-known quarrel with Isaac Newton, whom Hooke thought had stolen some of his ideas regarding planetary mechanics and other matters. This situation arose from the fact that Hooke held the position of Curator of Experiments in the Royal Society, of which Newton was a contemporaneous member. It was Hooke's responsibility to present experimental demonstrations at the society's meetings, and these experiments had often directly demonstrated aspects of what we now know as Newton's laws, well before Newton published them. Newton, often described as a reclusive and self-absorbed person, refused to acknowledge Hooke's contributions to these concepts in his *Principia* until the second edition, after Hooke's death. In 1678 Hooke presented the solution to his anagram: *ut tensio sic vis* (as is the deformation, so is the force)—and the theory of springiness that bears his name became public knowledge.

Hooke was a leading scientist of his day. His first published work was the *Micrographia* in 1665, containing his microscopic studies and introducing the term “cell” to biology. He served as Robert Boyle's assistant for several years, and invented the air pump used in the experiments leading to Boyle's law. (Some historians believe that this law also was rightfully Hooke's.) Subsequently, Hooke pursued many experiments related to respiratory physiology, involving both him and animals as experimental subjects. In the next decade Hooke turned his attention to clock-making. The impetus for this was the need to determine longitude at sea. Because pendulum clocks are inaccurate aboard ship, timekeeping based on the resonant frequency of a spring became of interest. The resulting experiments led Hooke to discover the linear relationship between applied force and extension of a spring. Interestingly, he saw this law as an extension of Boyle's law, which expresses the “springiness” of air.

Young's modulus is named for Thomas Young (1773–1829), the physician-physicist who coined the term “modulus of elasticity” and made other important contributions to mechanics. Unlike Hooke, Young had no desire to guard his scientific insights, but his inability to write clearly concealed his ideas about elasticity almost as effectively as Hooke's anagram: “The modulus of elasticity of any substance is a column of the same substance, capable of producing a pressure upon its base which is to the weight causing a certain degree of compression as the length of the substance is to the diminution of its length.”

Such unintentional obfuscation aside, Young was, like Hooke, a person of enormous breadth of interest and accomplishment. He is perhaps best known for his work in optics, having established both the wave nature of light and, by means of sometimes gruesome and painful experiments on his own eyes, the foundations of physiological optics. In addition to his practice of medicine and his work in physics,

Young helped decipher the Rosetta Stone and wrote 61 articles for the newly founded *Encyclopaedia Britannica*. Much of his scientific work was published anonymously because he was afraid that his medical practice would suffer if the public knew that he was a scientist!

7.2.5 Generalized Hooke's Law: Anisotropy

The Hooke's law relationship between the stress and strain has a matrix of constant values to inter-relate the six independent components of stress and strain. Using the vector form of the stress and strain each stress component is a linear function of the six strain components:

$$\begin{Bmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \\ \sigma_4 \\ \sigma_5 \\ \sigma_6 \end{Bmatrix} = \begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} & C_{15} & C_{16} \\ C_{21} & C_{22} & C_{23} & C_{24} & C_{25} & C_{26} \\ C_{31} & C_{32} & C_{33} & C_{34} & C_{35} & C_{36} \\ C_{41} & C_{42} & C_{43} & C_{44} & C_{45} & C_{46} \\ C_{51} & C_{52} & C_{53} & C_{54} & C_{55} & C_{56} \\ C_{61} & C_{62} & C_{63} & C_{64} & C_{65} & C_{66} \end{bmatrix} \begin{Bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \\ \varepsilon_6 \end{Bmatrix}$$

The matrix notation is written in compressed notation as

$$\sigma_i = C_{ij}\varepsilon_j \quad (7.1)$$

where both i and j assume the values 1–6 and j is understood to be summed over these values. The 36 constants of proportionality C_{ij} are defined as the elastic constants.

The vector and matrix notation is a simplification of the tensor notation for the generalized Hooke's Law where, $\sigma_{ij} = C_{ijkl}\varepsilon_{kl}$, is a summation over the stress and strain *second order tensor* components and the constant components of the *fourth order stiffness tensor* C . In matrix notation the stress and strain *matrices* are related to the stress and strain *tensor components* using Voigt notation as:

$$\begin{aligned} \{ \sigma_1 \quad \sigma_2 \quad \sigma_3 \quad \sigma_4 \quad \sigma_5 \quad \sigma_6 \} &= \{ \sigma_{11} \quad \sigma_{22} \quad \sigma_{33} \quad \sigma_{23} \quad \sigma_{13} \quad \sigma_{12} \} \\ \{ \varepsilon_1 \quad \varepsilon_2 \quad \varepsilon_3 \quad \varepsilon_4 \quad \varepsilon_5 \quad \varepsilon_6 \} &= \{ \varepsilon_{11} \quad \varepsilon_{22} \quad \varepsilon_{33} \quad 2\varepsilon_{23} \quad 2\varepsilon_{13} \quad 2\varepsilon_{12} \} \end{aligned}$$

The C_{ij} values are related to the fourth order components of the material tensor. The mapping between C_{ijkl} and C_{ij} depends on how the shear strains and stresses are ordered in the equations just above. The ordering of the shears is arbitrary (there are several ways to order them, e.g. 23-13-12, 12-23-13, 12-13-23...) and you have to make sure that all the parts of your matrix calculations match. In finite elements it is common to use the order 12-23-13 and in this book the order 23-13-12 is used.

It turns out that considerations of energy conservation require that the off diagonal terms of the stiffness matrix are equal, e.g., $C_{12} = C_{21}$, etc., so that the matrix of C_{ij} constants is symmetric. Taking into account the symmetries, the most general anisotropic material has 21 independent values, and has the form:

$$\begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} & C_{15} & C_{16} \\ & C_{22} & C_{23} & C_{24} & C_{25} & C_{26} \\ & & C_{33} & C_{34} & C_{35} & C_{36} \\ & & & C_{44} & C_{45} & C_{46} \\ & \text{Symmetric} & & & C_{55} & C_{56} \\ & & & & & C_{66} \end{bmatrix}$$

The prospect of having 21 constants to determine experimentally is a daunting and essentially impossible task for most biological materials. Fortunately, the number of independent constants needed to relate stress and strain is usually smaller than 21 for most materials because of symmetries in the material microstructure.

In *isotropic* materials (which have no “grain” but behave the same in every direction), there are only two independent elastic constants, and these are related to the elastic modulus, E , and shear modulus, G , which are sometimes known as *technical constants*. Poisson’s ratio, ν , is considered to be a third technical constant. For isotropic materials, because there are only two elastic constants, there are only two independent technical constants as well, with $E = 2G(1 + \nu)$. The stiffness matrix for an isotropic material is:

$$\frac{E}{(1 + \nu)(1 - 2\nu)} \begin{bmatrix} 1 - \nu & \nu & \nu & 0 & 0 & 0 \\ \nu & 1 - \nu & \nu & 0 & 0 & 0 \\ \nu & \nu & 1 - \nu & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 - 2\nu & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 - 2\nu & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 - 2\nu \end{bmatrix}$$

For materials like bone, which have a grain and thus are anisotropic, the number of elastic constants is between 2 and 21, and the technical constants have different values in different directions. For example, the elastic modulus in tension of a long bone’s diaphysis is greater longitudinally than it is across the shaft.

Most anisotropic materials are not completely so, but have some symmetry to their internal structure. Two common types of limited anisotropy happen to be found in bone, wood, and some other biologic materials. The mechanical properties of *orthotropic* materials are different in three perpendicular directions (e.g., the longitudinal, radial, and circumferential direction of a long bone). The orthotropic matrix of elastic constants has nine independent values:

$$\begin{bmatrix} C_{11} & C_{12} & C_{13} & 0 & 0 & 0 \\ & C_{22} & C_{23} & 0 & 0 & 0 \\ & & C_{33} & 0 & 0 & 0 \\ & & & C_{44} & 0 & 0 \\ & \text{Symmetric} & & & C_{55} & 0 \\ & & & & & C_{66} \end{bmatrix}$$

The stiffness matrix of an orthotropic material expressed in technical (engineering) constants has very complicated formulae for the constants C_{ij} . However, the orthotropic *compliance* matrix, defined by the relationship $\{\varepsilon\} = [S]\{\sigma\}$, is found as the inverse of the stiffness matrix $[S] = [C]^{-1}$ and has a simpler form. In the principal coordinates of an orthotropic material the compliance matrix $[S]$ has the particularly simple form:

$$\begin{bmatrix} \frac{1}{E_1} & -\frac{\nu_{21}}{E_2} & -\frac{\nu_{31}}{E_3} & 0 & 0 & 0 \\ -\frac{\nu_{12}}{E_1} & \frac{1}{E_2} & -\frac{\nu_{32}}{E_3} & 0 & 0 & 0 \\ -\frac{\nu_{13}}{E_1} & -\frac{\nu_{23}}{E_2} & \frac{1}{E_3} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{G_{23}} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{G_{13}} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{G_{12}} \end{bmatrix}$$

For an orthotropic material the *three* independent Young's moduli E_i , $i=1,2,3$ are for the *principal directions* of the material, there are *six* values of Poisson's ratio ν_{ij} and *three* independent shear moduli G_{ij} for the "i" face in the "j" direction. However, **the matrix is symmetric**, so that only three of the Poisson's ratios are independent—in total, therefore, there are *nine* independent material constants.

An interesting property of orthotropic materials loaded along their principal axes and of an isotropic material loaded in any direction is that there is no *shear coupling*—that is, normal strain does not cause shear stress and *vice versa*—shear stress does not cause normal strain. This property of orthotropy is related to the theory of how cancellous bone tends to "line up" with the common mechanical loads on a bone (the trajectory theory) (Murray 1936; Fyhrie and Carter 1986; Cowin 1986).

An occasionally useful special case derived from the orthotropic material is where $E_1 = E_2 = E_3 = E$, $G_{23} = G_{31} = G_{12} = G$ and all of the " ν_{ij} 's" are equal to ν . The three independent technical constants (E, G, ν) of this *cubic* material property can be useful for materials that have a shear and Young's modulus that are inconsistent with the isotropic requirement that

$E = 2G(1 + \nu)$ but for which there is not enough information to define the nine orthotropic constants.

Transversely isotropic materials behave similarly in every direction about a single axis of symmetry and have five independent constants. If the axis of symmetry for the material is in the “1” direction, then the compliance matrix (S) has the form:

$$\begin{bmatrix} \frac{1}{E_1} & -\frac{\nu_{21}}{E_2} & -\frac{\nu_{21}}{E_2} & 0 & 0 & 0 \\ -\frac{\nu_{12}}{E_1} & \frac{1}{E_2} & -\frac{\nu_{21}}{E_2} & 0 & 0 & 0 \\ -\frac{\nu_{12}}{E_1} & -\frac{\nu_{23}}{E_2} & \frac{1}{E_2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{2(1 + \nu_{23})}{E_2} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{G_{12}} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{G_{12}} \end{bmatrix}$$

(Note that if the axis of symmetry is in the “2” or “3” direction that the matrix is different.)

Later in this chapter we show which kinds of bone exhibit these forms of stiffness anisotropy. (Bone is also anisotropic in its strength properties, but there is no straightforward theory covering this behavior.) A much more complete and extremely lucid explanation of anisotropy and elasticity theory may be found in the treatise, *Theory of Elasticity of an Anisotropic Elastic Body*, by Lekhnitskii (1963).

7.3 Determinants of the Strength of a Whole Bone

7.3.1 Mechanics

Generally speaking, for a particular mode and rate of loading, the strength (or stiffness) of a structure is a product of (1) its shape and size, and (2) the strength (or stiffness) of the material within. For example, in compression, tension, or shear, the stress in a specimen of cross-sectional area A and bearing a load L is

$$\sigma = L/A \quad (7.2)$$

When this stress reaches the failure stress for the material, the structure will fail. Therefore, by simply rearranging this equation and using the subscript f to represent the failure condition, one obtains an equation for the compressive load that would fail a long bone diaphysis,

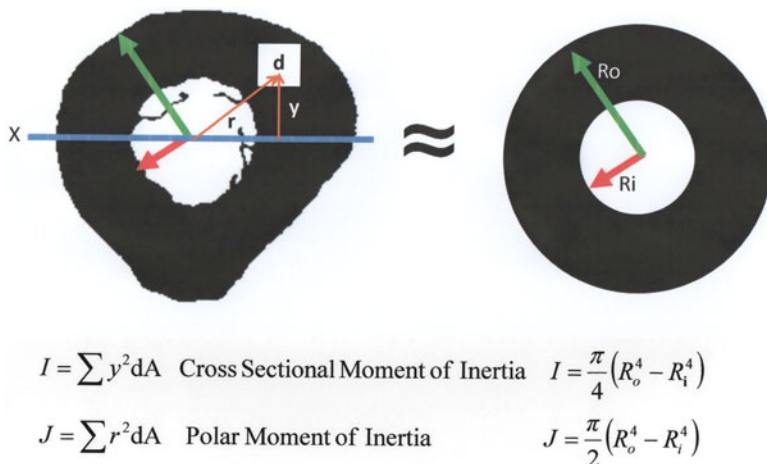


FIGURE 7.4. Cross-sectional and polar moments of inertia are defined for a bone cross section. The cross section is divided into many small elements of area, dA . The x -axis is assumed to be the neutral axis for bending. Equations are for I and J of a cylindrical tube.

$$L_f = A\sigma_f \quad (7.3)$$

where L_f is the failure load, and σ_f is the material strength in the appropriate mode of loading. In this case, the shape and size factor is simply the cross sectional area, A . Analogous equations describe the strength of the diaphysis in other modes of loading. When a circular cylindrical structure is twisted by a torque T , the shear stress varies linearly with distance, r , from the center, and is given by

$$\sigma_s = Tr/J \quad (7.4)$$

where J is the *polar moment of inertia* (Fig. 7.4). Rearranging this equation, one has, for the failure torque of a circular cylindrical long bone diaphysis loaded in torsion,

$$T_f = (J/r)\sigma_{sf} \quad (7.5)$$

where r is now the periosteal radius, and σ_{sf} is the torsional (or shear) strength of the bone material. The ratio J/r may be called the torsional “shape and size factor” for a circular cylinder.

For the hollow circular cylinder approximation of the femur the torsional stiffness is exactly “ GJ ” with “ G ” the shear modulus and “ J ” the polar moment of inertia. For a real bone with a cross-sectional shape other than a circle (or circular tube), however, the torsional rigidity is always *smaller* because the cross-sections “warp” in the direction of the axis when an axial

torque is applied. For *any* cross-section the equation relating torque (T_f), angle (θ), and length (l) is: $T_f = (GK)(\theta/l)$ where “ K ” is a factor dependent on shape and size of the cross-section. For a circular (or circular tubular) section like in Fig. 7.4, K is equal to the polar moment of inertia J . For any other cross-section the factor K is smaller than J and can actually be a very small fraction of J for some cross-sections. For an elliptical tube (*i.e.*, an elliptical tube with concentric elliptical inner and outer boundaries) with minor inner and outer radii b_0 , b and major inner and outer radii a_0 , a : $K = \left[\pi a^3 b^3 / (a^2 + b^2) \right] (1 - q^4)$ with $q = b_0/b = a_0/a$. The maximum stress is at the end of the minor axis on the outer surface and is: $\sigma_{sf} = 2T / \left[\pi a b^2 (1 - q^4) \right]$. For this elliptical tube $J = K \left[\frac{1}{2} \left(\frac{a}{b} + \frac{b}{a} \right) \right]^2$ so that J and K differ by less than 5 % if the ratio of the minor to major axis is between 0.8 and one. As a result, for a bone in torsion the mechanical response can often be approximated using a circular cross-section with errors accumulating if the bone isn’t straight or deviates greatly from a circle.

When a beam is bent by a moment M , there is tension on one side and compression on the other; the boundary between these regions is called the *neutral plane* or *axis*. The tensile and compressive stresses are given by

$$\sigma = Mc/I \quad (7.6)$$

where I is the *cross-sectional moment of inertia* and c is the distance from the point in question to the neutral plane. The failure moment is

$$M_f = (I/c)\sigma_{bf} \quad (7.7)$$

where σ_{bf} is the failure stress for bending and c is now the distance to the outermost fibers of the bone. Here, the shape and size factor is I/c .

Cross-sectional and polar moments of inertia are defined in Fig. 7.4. Consider that the cross-section of the bone cortex is divided into elements of area ΔA , of which one is shown. J is the sum of the products of each ΔA and the square of its distance from the centroid. I is similarly defined, except that the distance is measured from the neutral axis, which divides the tensile and compressive areas of a bent structure.

Equations of similar form would describe the stiffness of a bone: the primary difference is that a shape and size term is multiplied by the stiffness of the material (the elastic modulus in tension, compression, or bending; the shear modulus in shear or torsion).

7.3.2 Examples

Suppose a person’s femoral diaphysis is idealized as a *circular* cylinder with a periosteal diameter of 3 cm and an endosteal diameter of 1.2 cm. Calculate the diaphyseal failure load for uniaxial compression and the failure moments for bending and torsion.

7.3.2.1 Compression

Cross-sectional area = $A = \pi(R_p^2 - R_e^2) = \pi(0.015^2 - 0.006^2)$
 $= 5.91 \times 10^{-4} \text{ m}^2.$

From Table 7.1, use the ultimate stress in compression, $\sigma_f = 195 \text{ MPa}.$

TABLE 7.1. Typical mechanical properties for cortical bone

Property	Human	Bovine
Elastic modulus, GPa		
Longitudinal	17.4	20.4
Transverse	9.6	11.7
Bending	14.8 ^a	19.9 ^b
Shear modulus, GPa	3.51	4.14
Poisson's ratio	0.39	0.36
Tensile yield stress, MPa		
Longitudinal	115	141
Transverse	—	—
Compressive yield stress, MPa		
Longitudinal	182	196
Transverse	121	150
Shear yield stress, MPa	54	57
Tensile ultimate stress, MPa		
Longitudinal	133	156
Transverse	51	50
Compressive ultimate stress, MPa		
Longitudinal	195	237
Transverse	133	178
Shear ultimate stress, MPa	69	73
Bending ultimate stress, MPa	208.6 ^a	223.8 ^b
Tensile ultimate strain		
Longitudinal	0.0293	0.0072
Transverse	0.0324	0.0067
Compressive ultimate strain		
Longitudinal	0.0220	0.0253
Transverse	0.0462	0.0517
Shear ultimate strain	0.33	0.39
Bending ultimate strain	—	0.0178 ^b

Human data are for adult femur and tibia; bovine data are for femur
Data compiled from Cowin (1989, pp. 102, 103, 111–113), except as indicated
^aFrom Currey and Butler (1975); adult femur, three-point bending
^bFrom Martin and Boardman (1993); tibia, three-point bending

Then from Eq. (7.3) one has

$$L_f = A\sigma_f = (5.94 \times 10^{-4})(195 \times 10^6) = 1.16 \times 10^5 \text{ N}$$

For a 70-kg person, this is about $170 \times$ body weight.

7.3.2.2 Torsion

$$\begin{aligned} \text{Polar moment of inertia} &= J = \pi(R_p^4 - R_c^4)/2 = \pi(0.015^4 - 0.006^4)/2, \\ J &= 7.75 \times 10^{-8} \text{ m}^4. \end{aligned}$$

From Table 7.1, use the ultimate shear stress, $\sigma_{sf} = 69 \text{ MPa}$.

Then from Eq. (7.5) one has, letting $r = R_p = 0.015 \text{ m}$,

$$T_f = (J/r)\sigma_{sf} = (7.75 \times 10^{-8}/0.015)(69 \times 10^6) = 356 \text{ N} \cdot \text{m}$$

7.3.2.3 Bending

$$\begin{aligned} \text{Cross-sectional moment of inertia} &= I = \pi(R_p^4 - R_c^4)/4 = \pi(0.015^4 - 0.006^4)/4 \\ &= 3.87 \times 10^{-8} \text{ m}^4. \end{aligned}$$

$$\begin{aligned} \text{Cross-sectional moment of inertia} &= I = \pi(R_p^4 - R_c^4)/4 = \pi(0.015^4 - 0.006^4)/4, \\ I &= 3.87 \times 10^{-8} \text{ m}^4. \end{aligned}$$

From Table 7.1, use the ultimate stress in bending, $\sigma_{bf} = 208.6 \text{ MPa}$.

Then from Eq. (7.7) one has, letting $c = R_p = 0.015 \text{ m}$,

$$M_f = (I/c)\sigma_{bf} = (3.87 \times 10^{-8}/0.015)(208.6 \times 10^6) = 639 \text{ N} \cdot \text{m}.$$

If the femur were 42 cm long and supported at its ends, a 5130-N load in the middle would produce this moment. This would be about seven times the body weight of a 70-kg person.

It should be emphasized that the foregoing equations are restricted to fairly simple geometries of the cross section and load application. Bones may frequently depart from these conditions. For example, torsion of a bone having a cross section that departs significantly from axial symmetry, or bending in which the load vector does not pass through the centroid of the cross section, would not accurately follow these descriptions. Even so, it generally remains possible to separate the effects of material properties (which depend on histologic structure) from those of the shape and size of the structure (which depend on anatomy). That is the important point insofar as this discussion is concerned.

Another limitation of this approach is that it assumes that skeletal “shape and size” characteristics include only macroscopic anatomical features, and not microscopic features such as cancellous or osteonal architecture. It is assumed that the mechanical effects of microscopic features can be embodied in the material properties. This strategy, used in *continuum mechanics*, is

consistent with the approach taken when *finite element modeling* (FEA; see Box 7.3) is used to study skeletal stresses. The gross anatomy of the bone is modeled by the “shape” of the model, and variations in microscopic architecture are satisfied by assigning different material properties to individual regions (elements) within the model. The alternative is to attempt to formulate the shape and size of the microscopic structures, determine their behaviors and then assign properties to the macroscopic structure accordingly. Although not yet in widespread use, newer high-resolution imaging techniques combined with multi-scale FEA modeling now provide a platform for this computationally expensive approach. The more traditional approach favored here is to simply define material properties on a macroscopic scale, combining the mechanical effects of variations in microscopic architecture with those in microscopic material properties. This approach is also consistent with the common practice of measuring the material properties of bone using specimens small enough to define an anatomical region, but large compared to the size and distribution of microscopic features. For example, in the case of cancellous bone, test specimens are made large enough to contain perhaps 10–50 trabeculae (Brown and Ferguson 1980). For further discussion of this problem, see Fung (1977) and Harrigan et al. (1988). We return to this subject when we discuss the mechanical properties of cancellous bone later in this chapter.

Box 7.2 Mechanical Significance of Bone Fracture Patterns

When a person comes to the hospital with a fracture, the biomechanically knowledgeable physician can often deduce the kind of loading that caused the break from the pattern of the fracture surfaces. To do this, one must keep two things in mind. First, the ultimate stress for bone differs in various modes of loading: it is strongest in compression, weakest in shear, and intermediate in tensile failure (Table 7.1). Secondly, the stress inside a bone depends on orientation as well as position. If the cube in Fig. 7.2 is rotated to some new orientation, the stresses on its surfaces will change. Just as the cube can be rotated to minimize the shear stresses on its surfaces, there is another orientation that will maximize the shear stresses. Because the ultimate stress in compression is much higher than that in shear, when a bone is loaded in compression it tends to fail along the planes which carry the highest shear stresses, not those having the highest compressive stresses. These planes are typically at an angle of 45° to the load direction (Fig. 7.5).

When a bone is bent, failure often begins on the tensile side and propagates across the shaft because the tensile stresses on a cross section exceed the strength before the shear stresses on a 45° plane do. When the crack enters the compressed tissue, it again tends to travel along the 45° planes of maximum shear, causing ‘butterfly’ fragments or other types of *comminution*. When a bone is loaded in torsion the shear stress created by the twisting is across the fibers. In this case, compressive and tensile stresses lie at a 45° angle to the fibers. The tensile stress tends to separate the fibers, producing a fracture surface that spirals around the bone at a 45° angle. The crack then returns to its

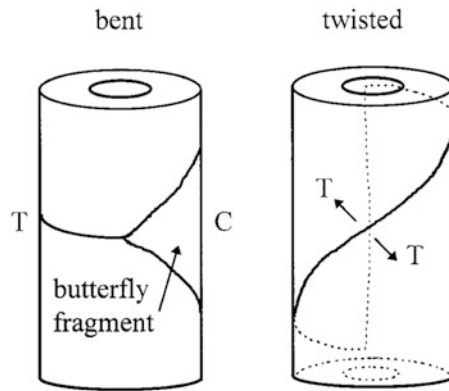


FIGURE 7.5. *Left:* In a bent bone, a transverse crack originates on the tensile side (T). On the compressed side (C) the fracture surface often follows a shear plane in one or two directions, producing a wedge-shaped fragment. *Right:* In a twisted bone, tensile stresses produce a spiral crack that winds around the cortex, and the bone breaks when the crack's ends are connected by a longitudinal fissure.

starting point along a longitudinal shear stress plane. Clinically, each of these fracture patterns is common, indicating that loading of long bones can occur by compression, bending, torsion, or a combination of these.

7.3.3 Mechanical Failure of Whole Bones

In considering the mechanical requirements of whole bones, several conflicting demands must be balanced. In normal daily life, the skeleton provides struts (bones) that the muscles can work against to create motion. The stiffer these struts are the more efficient is the action of the muscles. If bones are less stiff, then part of the mechanical output of the muscles is used in deforming the bones, rather than creating useful motion. This wastes metabolic energy. In this context, bones of greater stiffness are desirable. On the other hand, if the individual is in a situation where there is risk of bone fracture from a fall or a blow, less stiff (more compliant) bones would be preferred. This is because, for a given strength, the more compliant structure is able to absorb a greater amount of energy before reaching its failure load (Fig. 7.6). Another consideration is weight: mineralized tissues are twice as dense as other tissues and require considerably more metabolic energy to carry around. The mechanical characteristics of bones, therefore, represent a compromise between the need for stiffness to make muscle actions efficient, the need for compliance to absorb energy and avoid fracture, and the need for minimal skeletal weight. This balance between stiffness

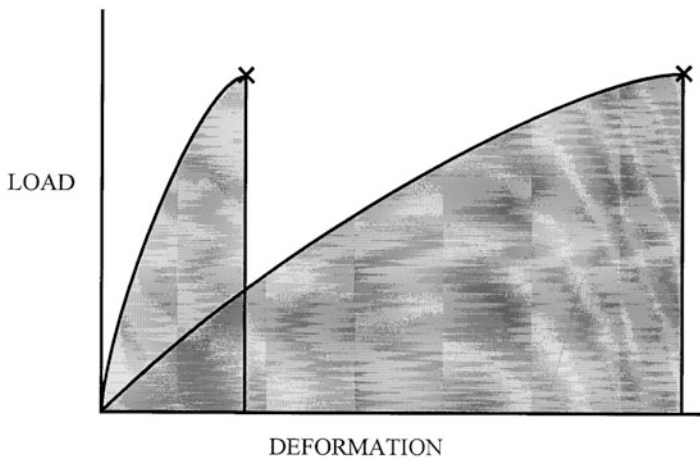


FIGURE 7.6. If two bones have the same strength (failure load), the more compliant one (at *right*) requires more energy to break it.

and the need for energy absorption to prevent fracture may be one explanation for the often inverse relationship found between stiffness and modulus of toughness in bone (2004).

Different bones have different functions, and the best solution to the conflicting mechanical demands need not be the same for every bone (Currey et al. 2007). The bones of the inner ear need to be extremely stiff to efficiently transmit sound waves; the fact that they could absorb very little energy before fracture is unimportant because they are protected from extrinsic loads. At the other extreme, the material in the bones of children is much more compliant than those of adults. Currey and Butler (1975) found the femurs of 2- to 8-year-old children to be 68 % as stiff, requiring 45 % more energy to break in bending, than those of 26- to 48-year-old adults. This increased compliance decreases the efficiency of children's muscle action, but it confers more protection against falling and crashing into things.

The energy required to break a bone is small compared with the energy encountered in daily activities. For example, the energy required breaking the shaft of an adult tibia or femur is only about 15 joules (J). The energy released when a 70-kg person falls to the ground from a standing position is about 500 J. The ability to absorb this energy using eccentric muscle contractions and deformations of soft tissues usually prevents fractures from occurring in such a fall. Soft tissue deformation absorbs energy during an accident in two ways. First, loads applied perpendicular to the tissues compress them and propagate stress waves up and down the body. The tissues effective in this way are skin, fat, and muscle (plus fur on animals and clothing on humans). Second, tissues such as fascia, tendons, ligaments, joint capsules, and contracted muscle brace bones against bending by supporting part of the

tensile forces. These tissues absorb energy as they are stretched, but they must be nearly as stiff as the bone to be effective. If these energy-absorbing mechanisms are impaired by surprise, restrictions, or incapacity, fracture occurs easily. The propensity of elderly persons to fracture their bones reflects not only the weakness of their bones but also their increased risk of falling and the inability of their muscles, ligaments, and soft tissues to adequately absorb the applied energy, which instead is transmitted to the bones.

Box 7.3 Finite Element Modeling

Finite element analysis, or modeling, is a process by which stresses and strains can be computed in a structure of irregular shape and varying internal properties. The name comes from the process in which the volume of the structure (e.g., the proximal end of a human femur, as shown in the figure) is divided into *elements*, each having a prescribed geometry and a set of mechanical properties (e.g., elastic modulus and Poisson's ratio). The model also includes *boundary conditions* that prescribe the forces applied to, or constraints on, portions of the structure. For example, in the case of a human femur, forces may be applied to some of the elements on the articular surface of the head and along the greater trochanter where the abductor muscles act. In addition, the elements at the distal end of the model may be constrained from moving in one or more directions. A computer is used to solve the equations that govern the ways in which the elements deform and the ways that they interact with one another as they deform. This method of analysis allows many problems to be solved that are far too complicated for a "closed form" solution (i.e., a single equation, such as Eq. (7.5), to give a simple example), either because of their irregular geometry or the variability of the material properties inside. Many commercial programs have been developed for finite element analysis, and the method has become very important in orthopaedics because of the irregular shapes of bones, their varying material properties, and the interest in studying the stresses produced when a metal implant is implanted in a bone.

Finite element modeling as a stress analysis tool is widely used in engineering and bioengineering. It is most powerful, however, when the stress analysis is part of an integrated approach that includes a theoretical model of the biological responses of the tissues to loading. Two examples of using finite elements along with empirical biological theories are the works of Hart (1989) and Beaupre and Carter (1992). Recently, some of the principles underlying an improved approach to integrated modeling of the skeleton were systematically presented by Professor Marco Viceconti (2012). The work of Viceconti goes beyond specific applications and puts the use of modeling in skeletal mechanics into the context of the scientific method, including making excellent distinctions between predictive (i.e., empirical) and mechanistic models that embody the theories of biology underlying system response. Mathematical analysis of bones and the skeleton is a small part of the large field of mathematical biology. A rich source and outstanding introduction to mathematical biology can be found in the book of JD Murray (2003).

7.4 Material Properties of Cortical Bone

It has been clearly demonstrated that osteonal remodeling of compact bone reduces its flexural fatigue strength and resistance to tension (Carter and Spengler 1978). Compression, shear, and bending strengths are also reduced. This section explores the reasons for this and otherwise examines the mechanical properties of osteonal bone. Osteons should affect the mechanical properties of the cortex in several ways: by the replacement of highly mineralized bone matrix with less calcified material, by increasing cortical porosity, possibly by altering collagen fiber orientation, and by the introduction of cement line interfaces that have different mechanical properties. We first consider the mechanical properties of single osteons, and then consider cortical bone as a composite of numerous such structural units.

7.4.1 *Properties of Individual Secondary Osteons*

In a remarkable series of papers beginning in 1964, Antonio Ascenzi, Ermanno Bonucci, and their colleagues presented a great deal of experimental information about the relationship between structure and mechanical function in individual secondary osteons. Taken together, their papers constitute the major part of our knowledge in this area, and we review them here in some detail.

7.4.1.1 **Tensile Properties**

In their initial work, Ascenzi and Bonucci (1964) examined the tensile properties of the secondary osteon. Undecalcified longitudinal sections, 20–50 μm thick, were prepared from bovine and human femurs. A sharp needle was used to dissect free the lamina between the Haversian canal and the cement line on one side of an individual osteon. This procedure produced an approximately rectangular segment of osteonal wall about 50 μm wide and 100–300 μm long. Figure 7.7 shows how the osteonal lamellae formed an approximately rectilinear system within the specimens. In their initial experiments, these osteonal segments were simply broken in quasistatic tension, with only the failure load recorded. Later, load-deformation curves were recorded when a microwave extensometer had been developed to measure the extremely small deformations encountered at this scale (Ascenzi and Bonucci 1967).

The specimens were categorized as either longitudinal (type L) or alternating (type A) in their lamellar structure (c or b , respectively, in Fig. 2.5). Microradiography was used to classify the degree of mineralization as either low or fully calcified. Both dry and wet bone were tested, but only the wet bone data are considered here. Figure 7.8 shows typical curves for secondary osteons with different lamellar orientations and degrees of mineralization. It was found that:

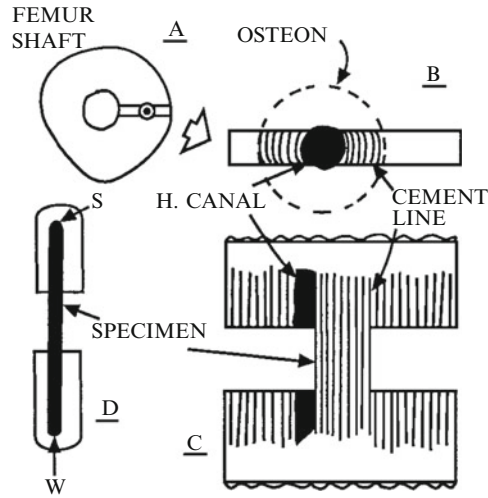


FIGURE 7.7. Diagram showing tension test of osteonal wall segment. (a), End view of test slab removed from femur cross section. (b), Relationship of a single osteon's geometry to the slab in an enlarged end view. (c), Side view of central portion of the test slab. Material has been removed to leave only the lamellae on the right side of the osteon in place. (d), The test slab supported at S and pulled in tension by weight W (Ascenzi and Bonucci 1964, 1967).

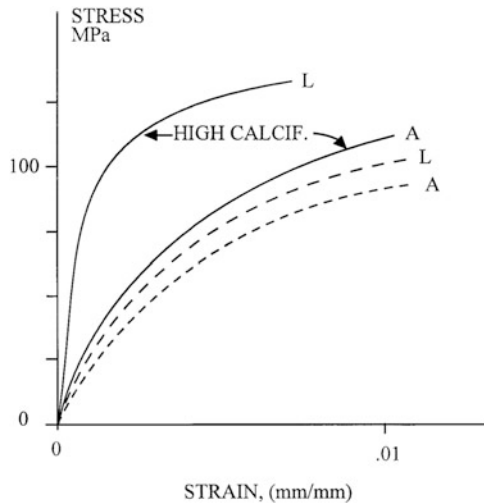


FIGURE 7.8. Typical stress-strain curves for osteon segments in tension. Solid curves, fully calcified osteons; dashed curves, osteons in the initial stages of calcification. L, type L osteon; A, type A osteon (see Fig. 3.5) (Ascenzi and Bonucci 1967).

- The tensile strength of osteon wall segments was about 114 MPa, a figure that lies in the midrange of data for larger test specimens composed of primary and secondary bone.
- Type L osteons had greater tensile strength and lower ultimate deformations than did type A osteons mineralized to the same degree.
- The degree of calcification of type A osteons had little effect on the tensile strength of their wall segments. The less mineralized type L osteons were more plastic and deformable than highly mineralized type L osteons, which were extremely stiff. The properties of poorly mineralized type L osteons and all type A osteons (regardless of how well mineralized) were nearly the same (Fig. 7.8).

No differences were found between human and ox osteons, but the authors remarked on the considerable variation in the results for apparently similar osteons. No specific effect of osteocyte lacunae could be mathematically formulated, but examination of the published photographs of failed specimens indicates that lacunae and other irregularities in the lamellar structure may have played an important role in the results. The specimens did not have the reduced central cross section that is preferred in tensile specimens. This factor surely led to stress concentrations at the points where the specimens were gripped, and the published photographs show breakage at these points.

7.4.1.2 Compressive Properties

Next, Ascenzi and Bonucci (1968) studied the compressive properties of single osteons using specimens from a human femur. They had by now developed a technique for cutting cylindrical segments from osteons contained in cross sections 500 μm thick. These minute cylinders were 180–200 μm in diameter and contained the Haversian canal and all but the outermost laminae of the osteon. The specimens were classified as to mineralization and lamellar structure as before, except that now osteons of circumferential structure (type T in Fig. 2.5) were also studied. In this case, they found that type T osteons were strongest and stiffest. Type A osteons were stronger than type L, but there was no difference in their elastic moduli. As before, the more fully mineralized osteons were more elastic, stiff, and brittle, regardless of their type. The specimens uniformly exhibited fracture planes at an angle of 30°–35° to the longitudinal direction, regardless of the lamellar structure or degree of calcification. Thus, the failure mode appeared to be shearing, regardless of the orientation of the collagen fibers.

7.4.1.3 Shear Properties

The next experiments examined the shear properties of Ascenzi and Bonucci (1972). A 160- μm -diameter cylindrical punch was used to push out the central lamellae from osteons in sections 300 μm thick. The section was supported from below by a plate containing an 800- μm -diameter circular

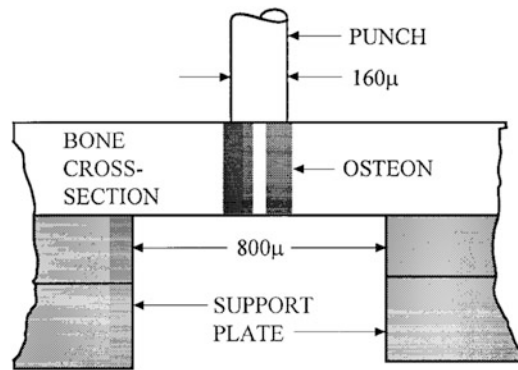


FIGURE 7.9. Diagram of relative dimensions of the specimen and test apparatus used in Ascenzi and Bonucci's (1972) shear experiments. Note that the $300\text{-}\mu\text{m}$ -thick bone section must have been subjected to significant bending loads.

hole, so that the test region involved the cement line. It should also be noted that the test involved significant bending as well as longitudinal shear (Fig. 7.9). The osteons did not fail at the cement line, but along the margins of the punch. Because there would have been a stress concentration at this boundary, one ought not necessarily accept Ascenzi and Bonucci's assertion that this proves that the cement line is stronger than the interstitial lamellae, but their results are not consistent with a weak cement line. Type L osteons were the weakest, and types A and T were of equal strength. Again, strength and stiffness increased with calcification in each type. The usefulness of the data from this experiment is limited by the fact that the shear test arrangement departed significantly from "simple shearing"; that is, the lateral distance between the edge of the punch and that of the support plate hole was relatively large. Unless one is willing to make a custom apparatus for each osteon, this condition is a necessary practical compromise.

There is another report of a similar experiment. Frost et al. (1961) arranged a simple shear test without regard to individual osteons or cement lines. Using $70\text{-}\mu\text{m}$ -thick sections, a $900\text{-}\mu\text{m}$ -diameter punch, and a "slightly larger" hole in the support plate, they pushed out areas containing several osteons and osteon fragments. The average stress for longitudinal shearing of bone from a human femur was about 74 MPa , considerably above the range reported by Ascenzi and Bonucci.

7.4.2 Combined Loading Effects

Ascenzi and Bonucci's analyses are theoretically instructive but over simplistic in considering only "pure" loading in compression, tension or shear. In reality, most bone is loaded in bending and incorporates combinations

of these types of stresses. Computational analyses under combined loading conditions (Pidaparti and Burr 1992) suggest that collagen fiber orientation has little effect on bending stiffness of bone if collagen fibers are oriented at $< 60^\circ$ with respect to the direction of load (*i.e.*, the longitudinal axis). Thus, bending stiffness is actually compromised to some extent by the presence of only transverse lamellae, or a predominance of Type T osteons. For bending, then, overall mechanical properties are maximized when collagen fibers are oriented at 30° or less to the axis of loading. Torsional stiffness, on the other hand, is maximized when collagen orientation is within 15° – 30° of either the longitudinal or transverse axis of the bone. In general and somewhat contrary to the results of Ascenzi and Bonucci, Type A osteons are not well adapted to resist torsional moments.

These analyses show that when loading conditions are unpredictable, as they often are, it is beneficial to the bone for collagen fibers to be oriented nearly longitudinally or at least within 15° of the loading axis. Axial stiffness is reduced by about 40 % as collagen fiber bundles alter orientation from 0° to 90° with respect to the long axis under conditions of bending. This is consistent with the analysis performed by Katz (1981) using the Hashin-Rosen model, and with the measurements by Rho et al. (2001). Although these analyses verified the measurements of Ascenzi under pure loading conditions, we can see that under more realistic conditions the adaptation of collagen fiber orientation is much more complex than the often quoted maxim that “longitudinal collagen fiber bundles are best in tension, and transversely oriented bundles are best in compression.”

7.5 Summary

The results of Ascenzi and Bonucci’s experiments are summarized in Table 7.2. Variations in the mechanical properties of osteons with different collagen fiber orientations suggest that they are individually adapted to locally enhance the ability of bone to support a particular type of stress. The lamellae of type L osteons are stronger in tension, but it is relatively easy for their inner lamellae to slip longitudinally relative to the outer lamellae. This observation fits nicely with the frequently reported tendency of osteonal bone tested in longitudinal tension to show “osteon pull-out” (Moyle et al. 1978; Moyle and Bowden 1984). Because type A and type T osteons are weaker in tension, but shear less readily in the longitudinal direction, it is not clear whether macroscopic specimens in which these osteonal types predominate should be weaker or stronger in tension. However, it is clear that in bending, Type L and Type A osteons with fibers oriented less than 60° to the loading axis will maximize the modulus.

As mentioned in Chap. 2, it is more likely that collagen fiber bundles are structured in a twisted plywood configuration, rather than completely in a

TABLE 7.2. Mechanical properties of different types of osteons

Mode of loading	Ultimate stress, MPa	Elastic modulus, GPa	Ultimate strain, %
Tension			
Type L	114 ± 17	11.7 ± 5.8	6.8 ± 2.9
Type A	94 ± 15	5.5 ± 2.6	10.3 ± 4.0
Type T	–	–	–
Compression			
Type L	110 ± 10	6.3 ± 1.8	2.5 ± 0.4
Type A	134 ± 9	7.4 ± 1.6	2.1 ± 0.5
Type T	164 ± 12	9.3 ± 1.6	1.9 ± 0.3
Shear			
Type L	46 ± 7	3.3 ± 0.5 ^a	4.9 ± 1.1 ^b
Type A	55 ± 3	4.1 ± 0.4 ^a	4.6 ± 0.6 ^b
Type T	57 ± 6	4.1 ± 0.4 ^a	4.6 ± 0.6 ^b

^aThe shear modulus is defined as the slope of the linear portion of the stress-strain curve

^bThe strain in the shear test is defined as the ratio of punch advancement to section thickness, expressed as a percent

parallel or alternating format between lamellae. From a mechanical standpoint, the twisted plywood configurations should enhance bone's ability to resist shear stresses developed in torsion and compression loading, because of their multidirectional organization. The twisted plywood configuration would be less than half as resistant as a parallel fiber arrangement to axial tensile forces (Ruggeri et al. 1997), but bone is rarely, if ever, subjected to purely axial tensile forces in-vivo. Moreover, even if it were, osteons with only longitudinally-oriented fibers would be highly susceptible to failure if the bone were loaded in any other orientation because the fibers would allow a preferential pathway for the growth of small fracture planes. Therefore, the twisted plywood organization may be the bone's adaptive response to withstanding unpredictable loads in many different directions. Bone's rather complex organization at the microscale provides an elegant example of evolutionary processes and the plasticity of bone at this scale of structure.

Box 7.4 Historical Note on Collagen Fiber Bundle Orientation in Osteons

Ever since van Leeuwenhoek identified lamellar patterns in bone in 1693, the structural patterning of osteonal lamellae, and its functional importance, has been a source of scientific observation, experimentation, and debate. By the mid-nineteenth century, investigators had identified "light" and "dark" lamellae both by light microscopy (Kölliker 1889) and using polarized light (Von Ebner 1875). Shortly afterward, Ranvier (1889) suggested that the lighter and darker lamellae were distinguished by fiber orientation. Recall that at about this time, Wolf was developing his ideas about the orientation of trabeculae as an adaptive response to mechanical inputs. As he put it,

“alterations of the internal architecture . . . occur as a consequence of primary changes in the . . . stressing of the bones.” Gebhardt (1901; Gebhardt 1905) applied this idea to a different level of structure by hypothesizing that collagen fibers within osteons were subject to similar kinds of organizing forces. By inference, then, collagen fiber orientation was determined by the predominant load directions on the bone. He was the first to suggest that longitudinally oriented fibers conferred better tensile properties, and transverse fibers were better adapted to compression. Moreover, he presaged the proposed twisted plywood model of Giraud-Guille (1988) by observing that the fibers within concentric osteonal lamellae are arranged in a spiral, and that it was these spirals that alternated between adjacent lamellae. This observation was apparently lost for a period of time, as scientists studied the functional implications of lamellae in which fibers were either unidirectional, or alternated between adjacent lamellae, but did not twist. It is fair to say that although we appear to have a better understanding in the twenty-first century about the three-dimensional characteristics of collagen fiber orientation in osteons, we have almost no idea about the physical and cellular mechanisms that control this organization, or the role if any of noncollagenous proteins in controlling that orientation. Students who are interested in the historical development of ideas surrounding lamellar orientation and its role in bone mechanics are referred to an excellent discussion of these topics in Ascenzi and Lomovtsev (2006). The Ascenzi, in this case, is not Antonio, but Maria-Grazia, his daughter.

7.5.1 *Effects of Osteons on Cortical Bone Mechanical Properties*

We first consider the effects of osteons per se, that is, the simple addition of osteons to primary bone, on mechanical properties. Subsequently, the effects of intraosteonal variations in collagen fiber orientation that is, of adding type L, A, or T osteons, are examined.

Walmsley and Smith (1957) suggested that Haversian systems weakened cortical bone. Soon after, Currey (1959) found that completely remodeled cow bone was 35 % weaker in tension than primary bone. This reduction in strength was attributed to the relative deficiency of mineral in more recently formed Haversian systems and the additional porosity caused by new Haversian canals. The former concept was reinforced when Hert and co-workers (1965) found that **fully mineralized** osteonal cow bone was as strong in compression as primary cow bone. A decade later, Reilly and Burstein (1974) tested Haversian and primary cow bone in both tension and compression, and found that osteonal remodeling decreased the strength about 20 % in both modes of loading. There was, however, no change in the elastic modulus. A year later Reilly and Burstein (1975) presented additional data showing that osteonal remodeling weakened cow bone about 11 % in tension and about 9 % in shear.

Margel-Robertson and Smith (1978) studied the effects of Haversian remodeling on the 3-point bending properties of bovine femur and the

compressive properties of porcine mandible. Although they did not present their data in enough detail to confirm statistically significant effects, they reported that remodeling reduced strength in both tests. Their bending data are unique and support the hypothesis that remodeling reduces the strength of cortical bone in all modes of loading.

Carter and colleagues (Carter et al. 1976; Carter and Hayes 1977a) found that osteonal remodeling weakened bovine bone with respect to both monotonic and fatigue loading. However, their analysis indicated that this was not caused simply by reduced bone density (that is, increased porosity and reduced mineralization). The osteonal bone matrix seemed to contain another source of weakness. This additional factor may be the cement line, which has been reported as either more compliant (Burr et al. 1988) or less compliant (Skedros et al. 2005) than surrounding bone tissue. Either of these cement line configurations could be a source of additional weakness, although either one is also likely to improve fatigue life under cyclic loading conditions (see Chap. 8).

Cortical bone exhibits substantial amounts of plastic deformation when loaded to failure in either tension or compression. The amount of plastic deformation is greatest when the bone is loaded parallel to the collagen fibers, in which case it typically exceeds the elastic deformation (Burstein et al. 1972). Osteonal bone exhibits more plastic deformation than primary bone (Reilly et al. 1974). It is reasonable to speculate that cement line slippage or yielding accounts for much of this plastic ductility, but no one has compared its effects to those of porosity and mineralization.

These studies of the effects of osteonal remodeling on bone strength all used nonhuman bone because it is very difficult to obtain human specimens of primary bone large enough to test. (There are standards for specimen geometry which require that the minimum dimensions of the specimen be several times larger than those of its internal structural features.) In the long bones of adults, and even adolescents, almost the entire cortex is occupied by several generations of secondary Haversian systems. Vincentelli and Grigorov (1985) attempted to overcome this problem by very selectively taking specimens from primary bone near the periosteal and endosteal surfaces of the human tibia. These specimens were tested in tension and compared with similar specimens taken from Haversian regions of the same bones. (Such equal-sized controls are the recommended way of handling cases in which the standard geometry cannot be achieved.) While there was some mixing of the two types of bone in various specimens, it was possible to analyze samples containing either more than 75 % or less than 25 % Haversian tissue. It was found that both the strength and the stiffness of the predominantly Haversian bone were less than that in primary bone (Table 7.3).

Ascenzi and Bonucci showed that the mechanical properties of individual osteons are affected by collagen fiber orientation. Does osteon type—that is, different arrangements of collagen fibers inside osteons—affect whole bone

TABLE 7.3. Tensile strength and stiffness of human primary and secondary bone

Tissue type	Ultimate stress, MPa	Elastic modulus, GPa
Primary bone	161 ± 11	19.4 ± 2.4
Haversian bone	130 ± 14	17.6 ± 2.0

From Vincentelli and Grigorov (1985)

strength as well? In 1969 Evans and Vincentelli reported that the **tensile** strength of tibial bone was positively correlated with the number of type L osteons that it contained (Evans and Vincentelli 1969). Because this work was done with embalmed material, they repeated the experiment with fresh human tibias (Vincentelli and Evans 1971). They again found significantly greater tensile strengths associated with the presence of type L osteons, in concert with the greater tensile strength of single type L osteons. They also found, however, that bone with greater numbers of Type T osteons deformed more as it broke. This result may be related to the capacity of these osteons to exhibit longitudinal shear between their lamellae, and may be consistent with the observation that axial stiffness is reduced as collagen fiber bundles alter orientation from 0° to 90°. Increased stiffness values were similarly correlated with Type L osteons, but not significantly so.

Later, this team reported the results of similar experiments on the *compressive* properties of osteonal bone (Evans and Vincentelli 1974). They did not find that compressive strength was inversely correlated to increased numbers of type L osteons, as they had expected from Ascenzi and Bonucci's work. Instead, they found that the greatest compressive strength was associated with osteons having a mixed dark–light appearance; these osteons were apparently intermediate between type L and type A. They also found that deformation at failure was greatest when more type A osteons were present, in contrast to the relative brittleness of single type A osteons when compared to type L (see Table 7.2). These results show that bone's macroscopic mechanical properties are not entirely predictable simply on the basis of osteon fiber orientation. Nevertheless, this feature appears to play an important role in determining mechanical properties.

In addition to studying the mechanical effects of osteons having different collagen fiber orientations, one may examine the effects of the **average** collagen fiber orientation throughout the bone matrix on the mechanical properties of primary or secondary cortical bone. Using birefringence in polarized light to estimate average collagen fiber orientation in a whole specimen cross section, it has been found that collagen fiber orientation affects strength and stiffness at least as much as do density, porosity, and mineralization (Martin and Ishida 1989; Martin and Boardman 1993; Martin et al. 1996).

If osteonal collagen fiber orientation affects bone strength, is this aspect of bone structure governed by mechanical loading? Gebhardt (1901), followed by Olivo and co-workers (Olivo 1937; Olivo et al. 1937) was perhaps the first to postulate that osteonal collagen fiber orientation is determined by mechanical stress. If this is the case, one would expect that osteons with different fiber orientations would be preferentially found in regions of bone carrying particular kinds of stress. Pursuing this hypothesis, Portigliatti-Barbos et al. (1983, 1984) demonstrated that the orientation of collagen fibers in the shaft of the human femur is not random, but is correlated with tensile and compressive stresses determined by biomechanical analysis. These investigators quantified the brightness of 1-mm-square elemental areas within a complete femur cross section when viewed in polarized light. Thus, their analysis averaged the osteon fiber orientation of intact osteons and interstitial lamellae, rather than characterizing osteons as belonging to one of three types. They found that regions carrying tensile stress tend to have more longitudinally oriented fibers and those carrying compressive stress contain more transversely oriented fibers. This is shown in Fig. 7.10, which contains computer-generated diagrams of human femur cross sections located 1 cm apart (upper and lower panels, respectively), showing collagen fiber orientation (left panels) and radiographic density (right panels). The diagonal lines approximate the position of the neutral axis for bending induced by loads at the hip (Pauwels 1980). The darker pixels of the left panels represent more longitudinally oriented collagen fibers and are concentrated in the lateroanterior regions carrying tensile stresses. The lighter pixels of the right panels represent greater volumetric mineralization and are concentrated in the medioposterior regions carrying more compressive stresses. Subsequently, others (Boyde and Riggs 1990; Batson et al. 2000) have correlated these variations in osteonal collagen fiber orientation to strain patterns determined in racehorses.

These observations are consistent with reports that porosity (Martin et al. 1980) and bone stiffness (Ashman et al. 1983) also vary in a regular fashion with angular position in the femoral cortex. They support the compound hypothesis that (a) osteon collagen fiber orientation and mineralization help determine cortical bone's mechanical properties and (b) osteonal remodeling adapts internal structure to functional stresses. However, because in healthy bone osteonal mineralization is only supposed to be a function of time since formation, it is not clear why radiographic density should be greater in the compressive cortex, as is shown in the images from Portigliatti-Barbos. Is mineralization of osteoid enhanced by compressive stress? Is the rate of remodeling greater on the tensile side, so that its osteons are, on average, younger? Perhaps an attractive hypothesis for the Portigliatti-Barbos data is that radiographic density reflected decreased porosity rather than mineralization of the tissue. On the other hand, work by Goodwin and Sharkey (2002) concluded that up to 20 % of the variance in the tissue modulus of interstitial lamellae within adult tibiae and

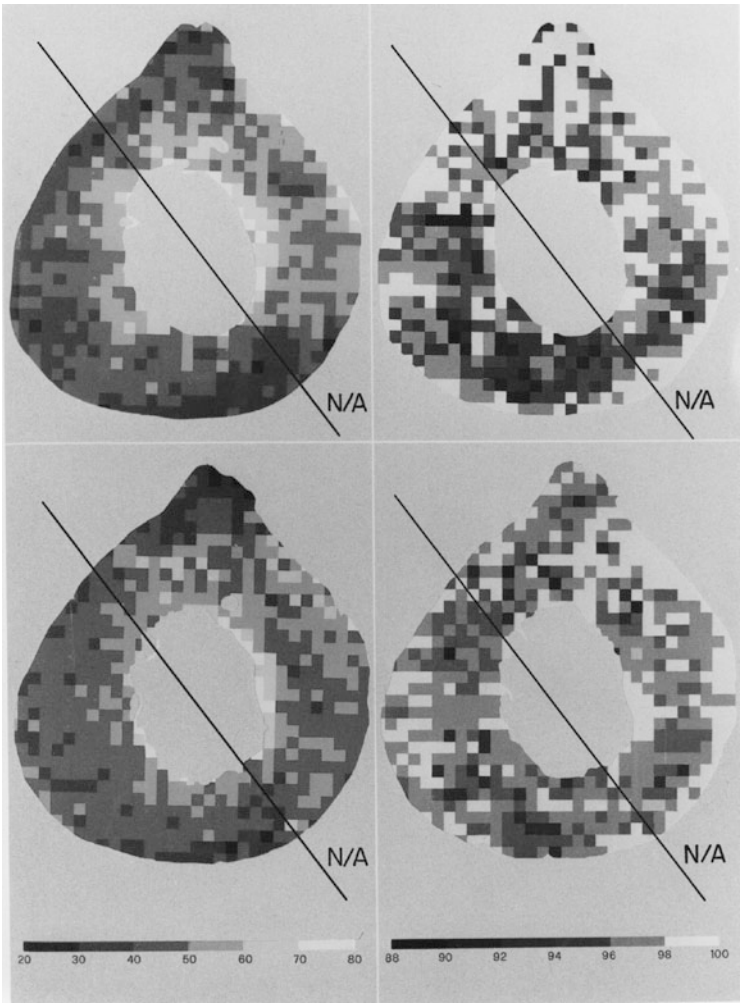


FIGURE 7.10. Plots of distribution of collagen fiber orientation (*left panels*) and radiographic density (*right panels*) in two cross sections from a human femur. These sections were located 1 cm apart in the central diaphysis. Superimposed on the plots are the approximate locations of the neutral axis (*N/A*) for bending produced by loading at the hip (Pauwels 1980). The *darker* pixels of the *left panels*, which have more longitudinally oriented collagen fibers and greater tensile strength, are concentrated in the lateroanterior regions, carrying more tensile stresses. The lighter pixels of the *right panels*, representing greater volumetric mineralization, are more common in the medioposterior regions on the compression side of the neutral axis (Portigliatti-Barbos et al. 1983). With kind permission of Springer Science+Business Media.

metatarsals could be explained by regional differences in the local strain environments routinely experienced by these bones. Moduli measurements on lamellae residing in regions experiencing the greatest compressive strains were 11 % greater than those made on lamellae in more tensile regions. Such differences could be due to differences in mineralization, collagen orientation or both.

7.5.2 Anisotropy of Cortical Bone Mechanical Properties

Cortical bone exhibits highly ordered internal arrangements of collagen fibers and mineral crystals; this is true for both the plexiform and circumferential lamellar forms of primary bone, as well as osteonal bone. X-Ray diffraction techniques show that the average orientation of the mineralized collagen fiber matrix is approximately parallel to the bone's long axis (Chen and Gundjian 1974). The fact that this causes marked differences in the mechanical properties associated with the longitudinal, radial, and circumferential directions was recognized long ago (Hulsen 1898). The radial and circumferential directions are visually distinguishable in both forms of primary bone (refer to Figs. 2.7 and 2.8), but heavily remodeled bone appears to the eye to be transversely isotropic (e.g., the osteonal region in Fig. 2.7). We now examine the effects of osteonal remodeling on the mechanical anisotropy of cortical bone.

7.5.3 Anisotropy Effects Determined by Load-Deformation Testing

Classic contributions in this area are the work of Reilly, Burstein, and Frankel (Reilly and Burstein 1975; Reilly et al. 1974). Using femurs, they measured the elastic moduli and strength of Haversian and primary bovine bone in different directions and compared them to those of human Haversian bone. Their data (summarized in Table 7.4) show that the strength and stiffness of bovine Haversian bone are similar to that of human bone when measured transverse to the long axis, but that cow bone seems to give consistently higher values in the longitudinal direction. In bovine bone, the ratios of elastic moduli in the longitudinal and transverse directions are similar in tension and compression and in primary and Haversian bone. These anisotropy ratios are less in human Haversian bone than in bovine Haversian bone (for both modes of loading) because the human bone is less stiff in the longitudinal direction. Both types of cow bone are stronger in compression than in tension, regardless of the loading direction. Human Haversian bone is stronger in compression when loaded transversely but not when loaded longitudinally. The strength anisotropy of Haversian bone

TABLE 7.4. Anisotropy of bovine and human bone

	Tension		Compression	
	Longitudinal	Transverse	Longitudinal	Transverse
Elastic modulus, GPa				
Human				
Haversian	17.9 ± 0.9	10.1 ± 2.4	18.2 ± 0.9	11.7 ± 1.0
Bovine				
Haversian	23.1 ± 3.2	10.4 ± 1.6	22.3 ± 4.6	10.1 ± 1.8
Primary	26.5 ± 5.4	11.0 ± 0.2	–	–
Ultimate stress, MPa				
Human				
Haversian	135 ± 16	53 ± 11	105 ± 17	131 ± 21
Bovine				
Haversian	150 ± 11 ^a	49 ± 7 ^a	272 ± 3	146 ± 32
Primary	167 ± 9	55 ± 9 ^a	–	–

^aStandard deviations approximate because groups were averaged (Reilly and Burstein 1975; Reilly et al. 1974)

from both species is more pronounced in tension than in compression, and it appears that bovine bone anisotropy is somewhat greater than that of human bone. Thus, Haversian cow bone is a reasonable model for human cortical bone, but it can be expected to be more anisotropic because of its enhanced strength and stiffness in the longitudinal direction.

Reilly and Burstein’s data suggest that osteonal remodeling has two important effects on mechanical anisotropy. First, secondary osteons reduce tensile strength and stiffness more in the longitudinal direction than transversely, suggesting that the differences between bovine and human Haversian bone result from the more extensive intracortical remodeling of human bone. Second, they serve to transform the bone from an orthotropic material into a more transversely isotropic one. The “transverse” direction referred to in Table 7.4 is the circumferential direction. Reilly and Burstein also tested bovine specimens in the radial direction. Recalling the structure of plexiform bone, one would expect to obtain different results in these two directions when this tissue is tested, and that is what was observed. Tensile strengths in the longitudinal, circumferential, and radial directions had proportionalities of about 3:1:0.4 in primary bone. In Haversian specimens the radial tensile strength was increased so that the ratios were about 3:1:0.7, and the compression strength ratios were equal in the two transverse directions: 3:1:1. Although distinct differences were found to persist in the deformability and energy to failure in the two transverse directions of bovine Haversian bone, these data show that osteons make whole bone transversely isotropic with respect to strength.

Knets (1978) determined the elastic constants of human tibial bone using conventional mechanical testing. He found that a statistically significant degree of orthotropy was present: $C_{11} = 11.6$ GPa, $C_{22} = 14.4$ GPa, $C_{33} = 22.5$ GPa. These values for human osteonal bone are in contrast to the ultrasonic data for bovine osteonal bone presented in the next section.

7.5.4 Anisotropy Effects Determined by Ultrasonic Measurements

Similar results have been obtained when ultrasonic techniques were used to measure stiffness values. In principle, the velocity of sound in a homogeneous continuum (i.e., a nonporous material) is proportional to the square root of its elastic modulus, so that measurements of sound velocity can be used to nondestructively determine material stiffness for very small deformations (Abendschein and Hyatt 1970; Ashman et al. 1984). (In composite materials such as bone, this relationship becomes complex when the wavelength of the sound approaches the dimensions of the voids.) Yoon and Katz (1976a, b) have described in detail how the anisotropic elastic properties of bone can be nondestructively determined by measuring the velocity of sound waves propagated through a specimen in various directions.

Van Buskirk et al. (1981), Bonfield and Tully (1982), and Katz and coworkers (Lipson and Katz 1984; Katz et al. 1984) have studied the effects of osteonal remodeling on the ultrasonic stiffness of bovine cortical bone. Their methods show that osteonal remodeling tends to convert plexiform bone from an orthotropic material to a transversely isotropic one, at least in the ultrasonic realm of very small strains (Katz et al. 1984). Table 7.5 shows

TABLE 7.5. Elastic coefficients in GPa

	Plexiform	Haversian
C_{11}	22.4 ± 0.6	21.2 ± 0.5
C_{22}	25.0 ± 1.0	21.0 ± 1.4
C_{33}	35.0 ± 2.0	29.0 ± 1.0
C_{44}	8.2 ± 0.4	6.3 ± 0.4
C_{55}	7.1 ± 0.3	6.3 ± 0.2
C_{66}	6.1 ± 0.2	5.4 ± 0.2
C_{12}	14.8 ± 0.8	11.7 ± 0.7
C_{23}	13.6 ± 0.7	11.1 ± 0.8
C_{13}	15.8 ± 0.8	12.7 ± 0.8

The C_{ij} subscripts refer to the following directions: 1 = radial direction; 2 = circumferential direction; 3 = longitudinal direction; 4 = circumferential-longitudinal shear; 5 = radial-longitudinal shear; 6 = radial-circumferential shear

From Katz et al. (1984)

that Haversian bone has very similar properties in the transverse (1 and 2) directions, but plexiform bone does not. Katz et al. (1984) have also demonstrated that osteoporosis has no effect on anisotropy, but osteopetrosis greatly reduces it. The latter finding was confirmed by Ashman and co-workers (1985a). Because osteopetrotic bone is largely calcified cartilage and contains few osteons, it is easy to see why it would be less anisotropic than normal bone.

In the inferior cortex of the canine mandible, which is a site of heavy osteonal remodeling, Ashman et al. (1985b) initially obtained values for E , G , and Poisson's ratio that were similar to those for osteopetrotic bone; that is, isotropic and less stiff than transversely isotropic bone. Later, Ashman and Van Buskirk (1987) reported additional measurements indicating that mandibular bone is transversely isotropic after all. There is a need for the mechanical properties of the bone inside the mandible to be determined with certainty. This bone has a complicated geometry and is loaded by a complex system of muscle and dental forces. It is of considerable interest to physical anthropologists, because its structure may provide information about the diet of its owner (Hylander 1981), and to dental clinicians for obvious reasons. Although long bones are usually used for testing theories of mechanical adaptation, the mandible may be a better choice in some respects.

Turner and Burr (1997) used an acoustic microscope (which produces images based on local differences in sound velocity) to investigate the source of cortical bone anisotropy. They measured sound velocity in the longitudinal and transverse directions in whole canine bone, then repeated the measurements after the specimens had been either decalcified or reduced to the mineral phase by dissolving their collagen. Table 7.6 shows "anisotropy ratios" (AR), defined as the ratio of the longitudinal elastic modulus to the transverse modulus, for the three types of specimens. It is seen that the mineral phase is as anisotropic as the whole bone but the organic phase is more nearly isotropic. However, the predominant direction of collagen orientation in canine bone, and likely in human bone as well, does not align with the longitudinal axis, so the AR measured in that study may not represent the maximum elastic anisotropy of the bone matrix, but some "off-axis" ratio. Additional data show that the AR in osteonal dog bone is similar to that in human and nonhuman primate bone, about 1.18. This is quite different from

TABLE 7.6. Anisotropy ratios in canine osteonal bone

Material	E_L/E_T
Whole bone	1.50
Mineral phase	1.50
Organic phase	1.18

From Turner and Burr (1997)

TABLE 7.7. Compressive strength and stiffness of cancellous bone

	Ultimate stress, MPa	Modulus, MPa
Human		
Yamada (1973) ^a	1.86–1.37	90–70
Neil et al. (1983) ^b	2.54 ± 0.62	272 ± 195
Kuhn et al. (1989b) ^c	5.6 ± 3.8	424 ± 208
Rohl et al. (1991) ^d	2.22 ± 1.42	489 ± 331
Canine		
Vahey et al. (1987) ^e	12.1 ± 5.7	434 ± 174
Kuhn et al. (1989b) ^c	7.12 ± 4.6	264 ± 132
Norrdin et al. (1990) ^f	9.60 ± 0.80	231 ± 22

^aMeans for people in their forties and sixties, respectively; lumbar vertebrae
^bLumbar vertebrae from men aged 54–90 years, 12-mm-diameter cylinders, 25–30 mm long
^c8-mm cubes from distal femur
^dProximal tibia, four men and three women, aged 42–76 years
^e5-mm cubes from head and neck of the femur in two dogs
^f21 vertebral specimens from seven dogs

the AR in non-osteonal rat bone, which is closer to 1.0 (Table 7.7). This suggests that collagen in osteonal bone has significant anisotropy (Turner and Burr 1997).

In summary, there are three important points concerning cortical bone anisotropy. First, the predominantly longitudinal orientation of collagen fibers and mineral crystals makes bone stronger and stiffer in this direction. Second, osteonal remodeling tends to convert bone from an orthotropic to a transversely isotropic material. Finally, certain metabolic diseases may render bone virtually isotropic, but it is clear that this is not functionally advantageous.

7.5.5 Determinants of Osteonal Bone Mechanical Properties

In addition to the presence of osteons, a number of other factors affect the mechanical properties of cortical and osteonal bone. This section summarizes the more significant of these.

7.5.5.1 Porosity

Obviously, the introduction of holes in a structure should weaken it, and this is certainly what happens in bone. It is one thing to demonstrate this effect, however, and quite another to explain in some detail the nature of the relationship. Empirical formulae have been developed relating porosity to

the mechanical properties of various porous materials, but a comprehensive theory useful over a broad range of porosity has not been developed. These problems are explored later in this chapter.

We can define porosity as the ratio of void volume to total volume, and note that volumetric porosity is equivalent to that measured two dimensionally on a cross section (Underwood 1970). Bone contains voids that range from a few to several hundred micrometers in diameter. Usually, however, the smallest of these (i.e., canaliculi and osteocyte lacunae) are not considered to affect its mechanical properties because, even though they probably do, they are rather uniformly distributed throughout all kinds of bone. They are therefore considered to be an intrinsic property of the bone material substance. In cortical bone, then, mechanical properties are considered to be affected by Haversian canals and related resorption cavities and vascular channels. (In trabecular bone they are affected by the marrow spaces between the trabeculae.) Several investigators have proposed empirical relationships between mechanical properties and porosity. For example, Schaffler and Burr (1988) found that the tensile elastic modulus (in GPa) of bovine cortical bone was related to porosity by the empirical equation

$$E = 33.9(1 - p)^{10.9} \quad (7.8)$$

(Note that $(1 - p)$ = bone volume fraction.) Currey (1988) obtained, for cortical bone from a wide variety of species loaded in tension,

$$E = 23.4(1 - p)^{5.7} \quad (7.9)$$

Comparing these results suggests that the modulus of bone becomes more sensitive to changes in porosity as its value decreases. Schaffler and Burr's maximum porosity value was 7.8 %; Currey's was 31 %. Pursuing this, McElhaney et al. (1970) found that the modulus of human skull bone (containing a very porous diploe layer) increased as the cube of $(1 - p)$. Similarly, modulus has been shown to increase as apparent density raised to the 7.4th power for cortical bone, as the cube of density for cortical and cancellous bone samples pooled together, and as the square of density when a sample containing only cancellous bone was analyzed (Schaffler and Burr 1988; Carter and Hayes 1977b; Rice et al. 1988).

Figure 7.11 shows the relationship between porosity and strength as compiled from several experiments. Curves of this shape pertain to other porous materials as well, including ceramics (Gannon et al. 1965). Clearly, the slope of the curve is quite different for cortical bone, on the left hand side of the curve where most of the material is solid matrix, and trabecular bone on the right hand side, where void space (filled with soft tissues, but not rigid) predominates. Presumably, different mechanical principles determine the slopes in these two regions, but they have not been elucidated.

In Fig. 7.11, porosity is implicitly regarded as a **material** feature of bone. Alternatively, one may think of porosity as a **structural** feature. If the

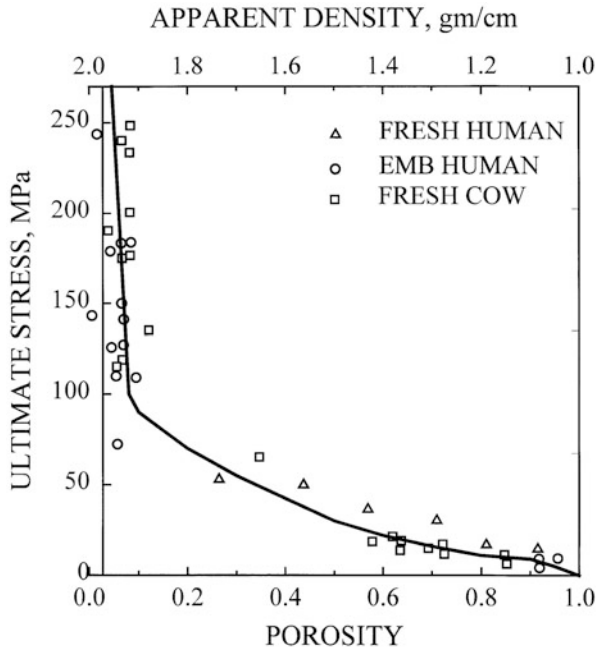


FIGURE 7.11. The relationships between ultimate compressive stress, porosity, and apparent density for fresh human bone (*triangles*) (Behrens et al. 1974), embalmed human bone (*circles*), and fresh bovine bone (*squares*) (Martin 1984). Very small porosity increases greatly affect cortical bone strength (Martin 1984).

porosity is uniformly distributed, the bone is loaded in simple tension or compression, and one is concerned only with macroscopic phenomena, the former point of view is convenient and acceptable. On the other hand, in a bent or twisted bone, voids located in the outer cortex, where stresses are high, weaken the bone more than those near the medullary canal. If there is a porosity gradient in the cortical wall, it makes a substantial difference whether the more porous region is in the outer, more highly stressed material, or near the medullary canal. Bones such as the femur are frequently characterized by porosity gradients across the cortical wall. Smith and Walmsley (1959) showed that gradients typical of those in elderly people can reduce the bending strength of a rectangular bone test specimen machined from a cortical wall by 20 % or so. However, the gradient is actually directed to enhance bone strength by placing most voids near the center of the bone. This is clear from Fig. 6.7, which compares the porosity distribution in the shaft of an elderly person's femur to that in a younger person. Not only are the void spaces concentrated near the medullary canal, but they are also preferentially located near the neutral axis.

7.5.5.2 Mineralization

As with porosity, it is intuitively clear that variations in the mineral content of bone should affect its mechanical properties, and experiments confirm this hypothesis. One may begin the analysis of this problem by breaking it into three parts according to the degree and cause of the mineralization difference. In the first place, because there is a delay in the mineralization of newly formed bone, bones that are being remodeled at a higher rate exhibit a reduced mineral content. This effect can be related to age, trauma, or metabolic disease. Reduced mineral content can also be caused by a lengthening of the osteonal refilling period because a delay in completion of new osteons means either that more osteoid or more void space is present for a given activation frequency. Much larger differences in mineralization may be produced by metabolic bone diseases, such as rickets or osteomalacia, in which the rate of turnover and refilling time are normal but the mineralization lag time is exaggerated.

It is very important to be clear about how “mineralization” is defined. One may define the amount of mineral per unit volume of whole bone as *volumetric mineralization*, and the amount per volume of bone matrix, exclusive of the volume of the void spaces in the structure, as *specific mineralization*. Clearly, volumetric mineralization is a function of both porosity and mineralization of the organic matrix; specific mineralization discounts the effects of porosity and is therefore more indicative of the mineralization process itself. Increased rate of bone turnover and increased refilling period, therefore, create both a volumetric and a specific mineralization deficit, while the inhibition of mineralization of osteoid creates only a specific mineralization deficit. In reading the literature one must be careful to understand which definition of “percent mineral” is being used. Similar comments could be applied to use of the term “density.”

Most noninvasive determinations of mineral content (e.g., photon absorptiometry and computed tomography) measure areal or volumetric mineralization in some fashion, and that is why these procedures are not generally capable of separating diseases with reduced areal/volumetric mineralization (e.g., osteoporoses) from those that involve reduced specific mineralization (e.g., osteopenias such as osteomalacia). Specific mineralization is usually measured by ashing the bone and calculating its ash fraction.

When Vose and Kubala (1959) correlated bending strength of human femurs with radiographic density, they were using a measure of specific mineralization because they pulverized and compressed the bone specimens before X-raying them. They found that small differences in matrix mineralization made very large differences in breaking stress: increasing the ash content from 63 to 71 % caused the strength to rise 3.7 fold. Currey (1969a) became interested in this relationship, and wondered why animals did not take better advantage of the apparently disproportionate enhancement of bone strength afforded by adding a few grams of mineral to each

bone. To explain this conundrum, he postulated that more highly mineralized bone is weaker at higher strain rates, and presented experimental data suggesting that the optimal specific mineralization for cortical bone is 66–67 %. The same year (Currey 1969b), he further hypothesized that the slope of the mineralization-stiffness (and strength) curve was especially steep because as mineralization became complete, the ends of the apatite crystals came into contact and fused. In a later publication, Currey and co-workers (Wainwright et al. 1976) concluded that 66–67 % mineralization was the point of maximum toughness. A few years later, Bonfield and Clark (1973) reached a similar conclusion (see following).

Currey (1986) related the elastic modulus of cortical bone from 17 vertebrate species, in both bending and tension, to porosity and specific mineralization (expressed in the form of calcium content). He found that these two variables together accounted for about 84 % of the stiffness variation, but he did not discuss their relative effects. He observed that such high correlations with porosity and mineralization left little room for dependency on such histologic features as collagen fiber orientation, but he also suggested that was probably not true for strength.

Finally, Schaffler and Burr (1988) found, for bovine cortical bone in tension,

$$E = 89.1A^{3.91} \quad (7.10)$$

where A is specific mineralization expressed as percent ash by mass. However, percent ash was not a statistically significant predictor of E , perhaps because A varied so little, mineralization being tightly regulated.

For a mixed sample of human cancellous and cortical bone samples Hernandez and coworkers (2001) found $E = 84.37(1 - p)^{2.58}\alpha^{2.74}$ where α was the ash fraction of the tissue. These results are similar to previous results, but differ in detail as discussed by the authors.

Box 7.5 Noninvasive Measurement of Bone Mass

There are several ways of noninvasively measuring bone mass, which is, the amount or density of bone in a particular part of the skeleton. Most of these methods depend on the absorption of photons by the bone, and because the mineral absorbs photons much more readily than the collagen, water, and soft tissues, these methods essentially measure bone mineral. However, because the variability of mineralization in adult bones is much less than that of porosity, these methods provide information about bone size and porosity as well.

The oldest and simplest of these methods is a simple radiographic film or “X-ray.” However, from variability in exposure and processing, a 30–50 % reduction in bone density may occur without being evident on a standard clinical radiograph. Calibrated *computed tomography* (CT) images (called quantitative CT, or QCT) are considerably more accurate, but both these methods involve significant radiation doses. In the 1970s, this problem was significantly reduced by the development of *photon*

absorptiometry by Mazess and Cameron (1972). This method involves scanning across a bone with a beam of photons emitted from a low-level radionuclide source such as ^{125}I . The source is fixed opposite a detector, the beam is scanned across the bone, and the intensity of the radiation received by the detector, I , is plotted as a function of scan position, x . The intensity of the radiation reaching the detector is

$$I = I_0 e^{-ky}$$

where I_0 is the intensity of the radiation from the source, y is the total thickness of the mineral in the beam's path through the bone, and k is an absorption coefficient. The area within the resulting absorption curve is proportional to the mineral content of the bone in the path of the beam. The width of the curve gives the width of the bone, and the dip in the middle is caused by the medullary canal. The most common site for such measurements has been the bones in the distal forearm. Results are reported as bone mineral content (BMC, g mineral/cm bone length) and bone mineral density (BMD, g mineral/cm³). Later, it was realized that by scanning with two radionuclides emitting photons with different absorption coefficients, correction could be made for the fact that soft tissues absorb some radiation; this is known as *dual energy photon absorptiometry* (Fig. 7.12).

In the 1990s, an alternative version of this known as *dual energy x-ray absorptiometry* (DEXA or DXA) was developed. This technique substitutes low-energy X-rays for the radionuclide source, and involves scanning back and forth along some length

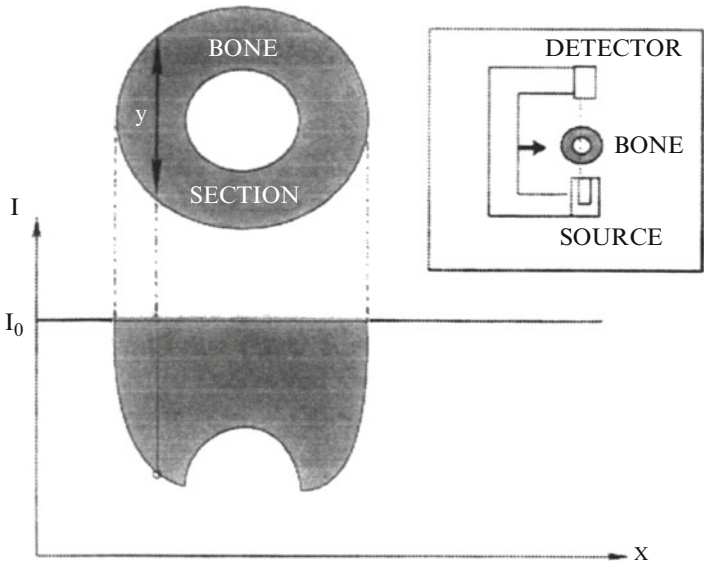


FIGURE 7.12. An absorption curve (I vs. x) in relationship to the cross section of the bone being scanned. The distance from the baseline (I_0) to the absorption curve is proportional to the distance y across the cortex and the volumetric density of the mineral along this path. Inset: schematic diagram of the scanner's source and detector passing across the bone.

of the bone to produce a two-dimensional image. Thus, the hip, the spine, and even the whole body can be scanned for bone mineral content. DEXA machines are now commonplace and used in a variety of clinical and research settings. Although primarily intended to aid in the diagnosis of osteoporosis and fracture risk, the technique also provides convenient and precise measurement of overall body composition, including lean and fat tissue mass. Absorptiometry methods are accurate to within 1–2 % and expose the patient to a fraction of the radiation dose of a standard chest X-ray.

7.5.5.3 Density

The material properties of bone are frequently discussed in terms of its “density.” Used alone, this word would be expected to refer to the specific gravity of the solid matrix. The term *apparent density*, on the other hand, refers to the mass per unit volume of a region of bulk bone, which includes Haversian canals, marrow spaces, and other soft tissue spaces. Therefore, apparent density is a function of both the porosity and mineralization of the bone material. Carter and Hayes (1977b) showed that the compressive strength (σ_{cf}) and elastic modulus of bone specimens having a wide range of apparent densities¹ ($d = 0.5\text{--}1.9$ g/ml) were correlated with the square and cube, respectively, of density in g/ml:

$$\sigma_{cf} = 68\dot{\epsilon}^{0.06}d^2 \text{ (MPa)} \quad (7.11)$$

$$E = 3790\dot{\epsilon}^{0.06}d^3 \text{ (GPa)} \quad (7.12)$$

In these equations, $\dot{\epsilon}$ is strain rate in s^{-1} . The latter relationship has been disputed by Rice et al. (1988). Through extensive multiple regression analysis of a large collection of data for cancellous bone, in which they attempted to control for differences in species, type of load, and load direction, Rice et al. found that both E and σ_{cf} correlate better with d^2 than d^3 . Other studies of cancellous bone have also obtained exponents closer to 2 than 3 (Bensusan et al. 1983; Ashman and Rho 1988). It appears that the reason for this discrepancy is that the data of Carter and Hayes included cortical as well as cancellous bone specimens, and these points increased the exponent from 2 to 3.

7.5.5.4 Histologic Architecture

Two regions of bone having similar porosity and mineralization may still have quite different material properties because the organization of the solid

¹ In this experiment, apparent density was measured after the void spaces had been emptied of soft tissues, so they contained only air. That is why the symbol d is used here instead of the ρ that was used for apparent density in Sect. 2.4. How would you change Eq. (2.6) to obtain an expression for d ?

matrix is highly variable. As noted in Chap. 2, compact bone may have a great variety of internal architectures. One may first distinguish between primary lamellar bone (mostly formed by modeling on periosteal or endosteal surfaces during growth) and lamellar bone composed of secondary osteons (formed by remodeling). We have seen that secondary bone is weaker than primary bone (Reilly and Burstein 1975). Primary cortical bone itself may also consist of two nonlamellar types of tissue. Plexiform bone is similar to lamellar bone but is laid down more rapidly in growing individuals (Enlow and Brown 1956, 1957, 1958). Much of the data on primary bone's mechanical properties is based on this material because it occurs in cows, a ready source of material for testing. How these properties compare to the "pure" lamellar bone usually found in humans is unclear because remodeling occurs so early in humans that it is difficult to obtain sufficient material to test.

7.5.5.5 Collagen Fiber Organization

Given a particular porosity and mineralization, and a particular trabecular or osteonal architecture, it remains possible for two regions of bone to have different material properties. There could be several reasons for this, but the next variable to be considered here is the orientation of constituent collagen fibers. This aspect of structure is apparently controlled both by cell functions and by extracellular physical processes. Frost (1986) has called these processes "micromodeling."

Lamellar bone, with lamellae of parallel collagen fibers, is stronger than woven bone in which the collagen fibers are more or less randomly oriented. Furthermore, the predominant orientation of the collagen fibers affects tissue strength in relation to the mode of loading: longitudinal fibers promote strength in tension, while transverse fibers are associated with strength in compression. This has been shown for individual osteons (Ascenzi and Bonucci 1967, 1968) as well as bone in bulk, and it was seen in Fig. 7.10 that fiber orientations vary with position relative to the neutral axis of the femur, so that regions in habitual tension have more longitudinal fibers and those in compression have fewer (Portigliatti-Barbos et al. 1983). In cortical bone, which generally exhibits little variability in porosity or mineralization, collagen fiber orientation may be the predominant factor in determining tensile strength (Martin and Ishida 1989). No studies have been found of collagen fiber orientation effects in cancellous bone, but it seems likely, for example, that trabeculae composed of woven bone must behave differently from those containing lamellar bone.

7.5.5.6 Fatigue Damage

Another factor that may affect the material properties of bone is microdamage. It is known that fatigue microdamage reduces the elastic modulus (Carter and Hayes 1977a; Schaffler et al. 1989; Launey et al. 2010), and as techniques for measuring such damage improve, this may be added to the list of those things studied as determinants of bone's mechanical properties.

When loaded repetitively in the laboratory, bone experiences fatigue damage and may fail from this like any other structural material (Carter et al. 1976; Gray and Korbacher 1974; Carter and Hayes 1976). The *in vivo* fatigue life of bone depends on the rate at which remodeling repairs fatigue damage. In Chap. 8 we consider the relationship between the creation of fatigue damage and its repair by remodeling and describe a model for stress fracture, which is what clinicians call the fatigue fracture that occurs when bone's repair process fails to keep up with damage formation.

7.5.5.7 Rate of Deformation

Cortical bone is a viscoelastic material; its mechanical properties depend considerably on the rate at which it is deformed. One may speculate that features of internal structure associated with remodeling and osteons could contribute to such behavior in several ways, including the presence of osteoid in varying amounts, the flow of fluid through Haversian canals, canaliculi, or the interstitial spaces in bone matrix, and the presence of cement lines. There is, however, very little information available regarding the effects of osteons on bone's viscoelastic behavior.

The classic paper on bone's viscoelasticity is that of McElhaney (1966). He studied bovine femoral bone over a wide range of deformation speeds, building a pneumatic cannon to achieve an upper strain rate of 1500 per second. In general, as shown in Fig. 7.13, cortical bone becomes stronger, stiffer, and more brittle as strain rate increases. Although he did not quantify the histologic structure of his specimens, McElhaney remarked that those loaded at high strain rates tended to fail along cement lines, while those

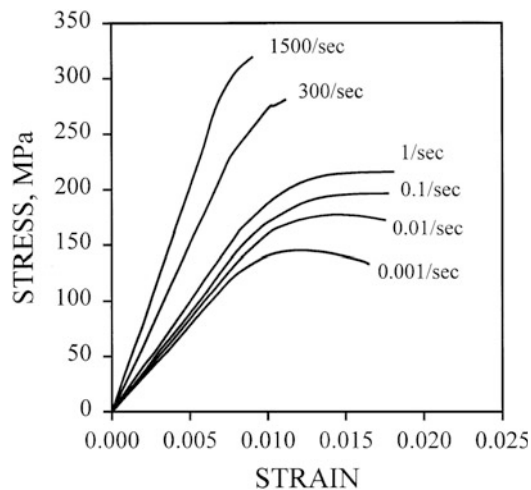


FIGURE 7.13. Stress-strain curves for cow bone loaded in compression at various strain rates (McElhaney 1966).

loaded at low strain rates failed along shear planes passing through osteons. It is interesting, and perhaps functionally significant, that the energy absorbed to failure is greatest not at the lowest or highest strain rates, but in the range 0.01–1.0 per second. Assuming a bone the size of an adult femur is loaded in bending, this corresponds to a mean impact velocity much less than 1 m/s or 2 miles/h, well below that expected in trauma, even before the advent of modern transportation and weapons. Margel-Robertson and Smith (1978) tested bovine femoral bone in compression at strain rates below this range and found a second local maximum in energy absorption at 0.00001–0.0001 per second. Others who have experimentally studied the viscoelastic properties of cortical bone include Smith and Keiper (1965), Sedlin (1965), Burstein and Frankel (1968), Panjabi et al. (1973), Crowninshield and Pope (1974), and Bargren et al. (1974). Sedlin proposed a rheological model for cortical bone consisting of a three parameter solid with a frictional element to account for plastic deformation at higher strains, but the data of Bargren and co-workers suggest that a two-element Kelvin model satisfactorily describes bone in the physiologic strain range. On the other hand, Laird and Kingsbury's (1973) tests on cow bone indicated that a three parameter model was inadequate to describe vibration phenomena in bone (frequencies of the order of kilohertz).

Ramaekers (1977, 1978, 1979a, b, c) published a series of papers on the rheological properties of bone. In one of these (Ramaekers 1978), he made the interesting suggestion that one of the viscoelastic functions of bone is to damp out the individual twitches of muscles. This hypothesis was prompted by the fact that one of the peaks in the shear damping frequency response corresponds to the twitch frequency of mammalian muscle.

Lakes, Katz, and Yoon have contributed a number of experiments and mathematical analyses of the effects of viscoelasticity on cortical bone mechanical properties (Lakes and Katz 1974, 1979a, b; Lakes and Saha 1980; Lakes et al. 1979, 1986). This work considers primarily the deformation of bone at sonic or ultrasonic frequencies and is related to Katz's ultrasonic studies of stiffness anisotropy. The work included torsional and biaxial loading modes, and the suggestion was made that peaks in the frequency response of the material properties could have histologic and physiologic significance. Johnson and Katz (1984) pointed out erroneous assumptions in these studies and attempted to correct them. They concluded that fluid flow through vascular channels is probably not important in cortical bone viscoelastic behavior; rather, it is the interactions of the fluids with the calcified matrix (which is not strictly solid, but contains minute water-filled spaces) that are important. Much remains to be done in extending these viscoelasticity studies to provide insights and predictive capabilities more relevant to bone fracture and remodeling.

In summary, three of the fundamental determinants of the mechanical properties of cortical bone's material include porosity and mineralization (which together determine apparent density), and collagen fiber orientation.

These variables are strongly affected by histologic structure, starting with distinctions between primary and osteonal bone structure. Clearly, these factors are largely controlled by remodeling, particularly in adults. Fatigue damage may also intrinsically affect mechanical properties, and finally, strain rate is an important extrinsic factor.

7.6 Material Properties of Cancellous Bone

We now turn our attention to cancellous bone. As one would expect from looking at its porous structure, recall Fig. 2.3, cancellous bone is substantially weaker and more compliant than cortical bone. Table 7.7 gives some typical values of cancellous bone's compressive strength and elastic modulus. Note the considerable variability in these values. Although bone porosity is bimodally distributed (i.e., either cortical or cancellous, with few regions between), it varies continuously, and there are not sharp geometric, densitometric, or mechanical distinctions among various examples of cancellous bone or between dense cancellous bone and porous cortical bone.

7.6.1 *Stress-Strain Curves for Cancellous Bone*

Figure 7.14 shows typical load-deformation curves for cubes or cylinders of cancellous bone loaded in compression. As the specimen compresses, it yields and reaches an ultimate stress, as does a cortical bone specimen. Thereafter, however, the load does not fall to zero because the specimen does not fracture into pieces. Instead, its void spaces collapse and its trabeculae compact against one another. As this process continues, the load is typically maintained, or may even rise far above the previous ultimate load.

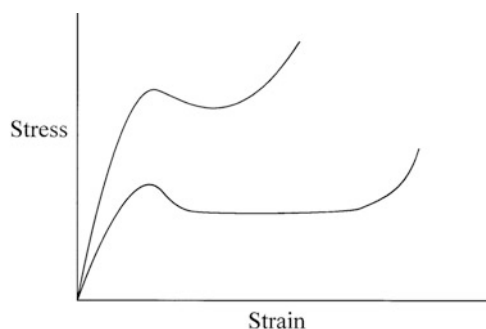


FIGURE 7.14. Load-deformation curves for two specimens of cancellous bone having different porosities. Yield and ultimate loads are apparent in the early portion of each curve. Subsequently, reduction and plateauing of the load occur. As void spaces collapse, the specimen densifies and the load rises steeply.

In addition to these behaviors, mechanical testing of cancellous bone is much more problematic than testing cortical bone. When a compression test is done between platens and strains are based on platen displacement, errors can occur because artifactual damage in the regions adjacent to the platens—damage caused by interruption of the trabecular network when the specimen is machined, friction against the platens, and stress concentrations. Other problems arise because supporting trabeculae have been removed along the sides of the specimen and the marrow in the void spaces is not confined as it is in the intact bone.

Keaveny and co-workers (1994) have developed testing methods designed to alleviate the problems associated with end effects at the platen interfaces. Cylindrical (40 mm long, 8.3 mm in diameter) specimens are cored from trabecular bone and their ends are cemented into brass end-caps. The specimen is then frozen and the central 8-mm region turned to a diameter of 6 mm on a lathe. Strain is measured using a device called an extensometer, which fastens to and measures the distance between two points on the central region of the specimen. These precautions eliminate or reduce the problems with end effects.

7.6.2 *Strength in Tension Versus Compression*

Carter et al. (1980), Bensusan et al. (1983), and Rice et al. (1988) all found that the failure stress of cancellous bone was the same in tension as compression. Kaplan et al. (1985), on the other hand, found that cancellous bone was stronger in compression for any given apparent density. Cowin (1989, p. 138), suggests that this may be because the specimens in the Kaplan experiment came from a region particularly well adapted to compressive loading in the tested direction.

7.6.2.1 **Caution: Continuum Limits Nearby**

As with cortical bone, we treat cancellous bone here as a porous material used to construct whole bones and try to identify and quantify the features (e.g., apparent density) that govern its mechanical behavior. However, in cancellous bone the situation is different and a bit tricky because of the relative sizes of the porosities in the bone.

The thickness of trabecular struts and plates is typically in the range of 100–300 μm , with the spacing between adjacent trabeculae of the order of 300–1500 μm . This is true regardless of the size of the animal being considered. Therefore, a line passing through the femoral head of an elephant would encounter hundreds of trabeculae, and a similar line through the femoral head of a mouse would encounter only a few. Although we can imagine taking 1 cm^3 of cancellous bone from the elephant and treating it as a material of which the strength and modulus can be measured, in the case of the mouse the cancellous bone of the femoral head must be modeled as a

structure rather than as a material, because slight geometric variations in shape and orientation of those few trabeculae could greatly alter the mechanical properties. One of the requirements of continuum mechanics is that the minimum dimension of a sample that is treated as a material must be significantly larger than the dimension of its structural sub-units (Fung 1969). In the case of trabecular bone, the minimum continuum length has been estimated to be about five trabecular spacings (i.e., 5–10 mm, depending on the sample) (Brown and Ferguson 1980; Harrigan et al. 1988). This is an order of magnitude greater than the continuum length for compact bone, where the porosity is much less and the void spaces are much smaller.

A practical implication of this is that in speaking of cancellous bone as a **porous material**, we need to be thinking in terms of samples of about 1 cm^3 (Linde et al. 1992). This leads to a conundrum. Cancellous bone can be quite heterogeneous, with its properties varying millimeter by millimeter, as shown, for example, by Fig. 2.3 and in penetration test² data (Hvid and Hansen 1985; Hvid 1988). A sample that is big enough to satisfy the requirements of the continuum length assumption, therefore, is quite likely to contain local heterogeneities in mechanical properties. When strength is measured on such a sample, the measurement often reflects the weakest portion of the sample. Consequently, local variations in mechanical properties tend to be overlooked in the results. Therefore, attempts to treat cancellous bone as a porous, continuum-level material must be evaluated very cautiously. With this reservation in mind, we can proceed to examine cancellous bone as a porous material.

7.6.2.2 The Three Determinants of Cancellous Bone Mechanical Properties

Insofar as we can treat cancellous bone porosity as a material property, we shall find that its mechanical properties are governed by three things: the magnitude of the porosity (or apparent density), the anisotropy of the trabecular architecture, and the material properties of the tissue in the individual trabeculae. We now discuss each of these, beginning with the latter.

7.6.2.3 Cancellous Bone Tissue Properties

The typical range of apparent density in cancellous bone is $1.0\text{--}1.4\text{ g/cm}^3$, compared to about $1.8\text{--}2.0\text{ g/cm}^3$ for compact bone. This decrease in apparent density primarily results from the presence of marrow-filled spaces; the density of trabecular bone's calcified matrix (or *tissue*) itself is similar to that

² A penetration test involves pushing a needle about 2.5 mm in diameter into the flat, machined surface of a region of cancellous bone. A graph of the resistive force vs. depth of penetration looks somewhat like a compression load-deformation plot. The correlation between the peak force seen in the penetration test and compressive strength is about $R = 0.9$ (Hvid and Hansen 1985).

of compact bone. Wolff (1892), in formulating his ideas about bone's mechanically optimized structure, assumed that cortical and trabecular bone were both made of the same kind of tissue, so that the differences in their mechanical properties only resulted from the varying sizes and shapes of the soft tissue spaces. This concept has received more critical attention in recent years, however. There are apparently some important differences in the microstructure of the calcified bone matrix in cancellous and in cortical bone, and these evidently are responsible for measured differences in material properties at the tissue level. Let us explore these differences.

The tissue in cancellous bone is usually made in the same way, and of the same lamellar material, as cortical bone. In both cases, osteoblasts sit on a bone surface and emit an organic matrix that calcifies to produce an anisotropic, lamellar composite. There is little evidence that, at the lamellar level, this basic structure is fundamentally different in cortical and in cancellous bone. Studies of osteocyte lacunar sizes and numbers in cancellous and cortical bone have not been able to establish any clear differences. However, the mineral content and tissue density of cancellous bone have been found to be less than in cortical bone (Gong et al. 1964; Dyson and Whitehouse 1968; Norrdin et al. 1977). Consistent with this, the water content of cancellous tissue has been reported to be greater than that of cortical bone. These results are consistent with the higher rate of turnover in trabecular bone; the average age of the bone matrix is simply less, so it is less mineralized.

Once one gets beyond the lamellar level and examines the geometry of the bone structural units created by the remodeling BMUs, another difference between cortical and cancellous bone tissues is apparent. Cortical bone osteons are parallel cylindrical structures that extend for several millimeters through the bone tissue. Their cement lines are closed surfaces. The trabecular equivalents, on the other hand, are dish-shaped "patches," or hemi-osteons, having cement lines that intersect the surface of the trabecular (Fig. 7.15). Therefore, cortical osteons act like reinforcing fibers running through the tissue, but trabeculae are composed of multiple dish-shaped patches. If a cement line is disrupted by stress in cortical bone, the disruption is contained within the bone, but in cancellous bone it may propagate to the surface, disconnecting the structural unit from the parent trabecula. Thus, remodeled cortical bone structure may be inherently more damage resistant than remodeled cancellous bone structure.

These mineralization and structural issues lead to the question of whether cancellous tissue has the same mechanical properties as cortical tissue. There have been two ways of approaching this question. The first is to measure the property (e.g., elastic modulus) in about 1-cm³ machined specimens of cortical and cancellous specimens and plot the results against the specimens' apparent density. If both kinds of specimen have the same tissue-level value of the property, they should fall on a single, smooth curve. Rho and co-workers (1993) did this with human bone, measuring the modulus ultrasonically (Fig. 7.16); they found that the modulus of the cortical bone

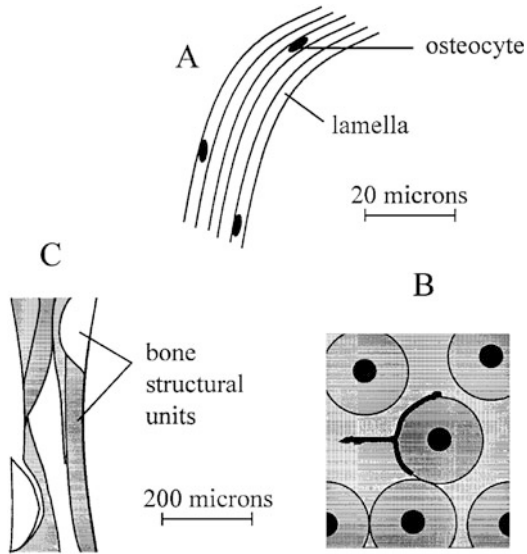


FIGURE 7.15. (a) At the level of lamellae, there is little to distinguish between osteonal and trabecular bone tissue. (b) In osteonal bone, the structural units produced by remodeling are largely intact, relatively long, fiberlike osteons. Their cement lines are internal to the bone; cement line disruptions are self-contained and do not compromise longitudinal loadbearing. (c) In trabeculae, the structural units produced by remodeling are much smaller and dish shaped with cement lines that, if disrupted, exfoliate the structural unit.

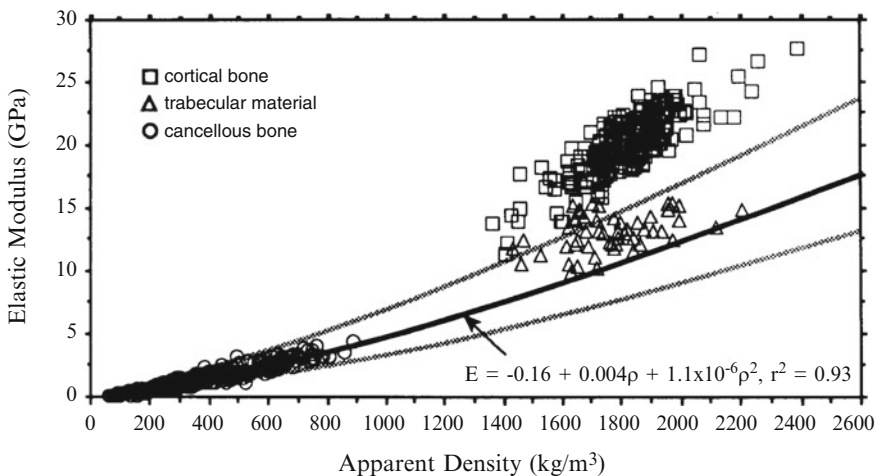


FIGURE 7.16. Ultrasonically determined elastic modulus vs. apparent density for continuum-level cancellous (*circles*) and cortical (*squares*) bone specimens. A linear regression line and its 95 % confidence limits are shown for the cancellous specimens. The fact that the cortical specimens have modulus values above the upper confidence limit for the cancellous data indicates that the cortical bone tissue has a greater modulus than that of cancellous bone. When the tissue modulus of cancellous bone was measured and plotted against the apparent density of the tissue, the data (*triangles*) fell inside the confidence limits (Rho et al. 1993).

specimens fell outside, and above, the confidence limits of the extrapolated regression line for the cancellous specimens. The triangular data points on the curve represent values ultrasonically measured in just the cancellous bone material. (This was done by increasing the frequency of the sound so it only propagated through the trabecular tissue.) These points fall within the extrapolated cancellous values. This graph provides strong evidence that cancellous bone has a reduced tissue-level modulus.

The other method for approaching the question is to test either individual trabeculae, or very small specimens machined from individual trabeculae, and compare the results with those from similarly sized specimens machined from cortical bone. Rho et al. (1993) also did this, measuring the modulus both ultrasonically and in tension tests in whole trabeculae and microspecimens of cortical bone. The tension tests showed the modulus of cancellous material to be 56 % of that of cortical bone (10.4 ± 3.5 vs. 18.6 ± 3.5 GPa). The ultrasonically measured values also showed the cancellous material was less stiff (14.8 ± 1.4 vs. 20.7 ± 1.9 GPa). Others have obtained different results when making the comparison between machined microspecimens of cancellous as well as cortical bone. In three-point bending, Kuhn et al. (1989a) found a larger modulus for cortical than cancellous bone (4.89 vs. 3.81 GPa), but it was substantially less than typical values for large cortical specimens (18–20 GPa). Choi et al. (1990) obtained similar bending results, and furthermore they found that the modulus of cortical specimens decreased steadily with the dimensions of the specimens, from 14.9 GPa for a beam with a square cross section 1000 μm on each edge to 5.5 GPa for a similar beam 100 μm on each side.

Taken together, these and other studies suggest that the mechanical properties of cancellous bone tissue are different from those of cortical bone tissue. The difference between the tension test results of Rho and the bending results of Kuhn and Choi, and the beam size results of the latter group, indicate that the elimination of intact osteons may reduce the modulus, and much more so in bending than in uniaxial tension. When cortical specimens as small as trabeculae are made, their osteons must be cut into fragments not unlike those of trabecular remodeling packets and this probably reduces the modulus, especially in bending. The reduced mineralization of cancellous bone tissue, at least in part caused by the increased rate of remodeling, may also explain some of the modulus reduction.

There may be a hierarchy of average, measured tissue moduli in descending order from primary cortical bone, to osteonal cortical bone, to cancellous bone tissue. *However, it is also important to bear in mind that there is no “correct” value of elastic modulus for any of these tissues because of normal variations in mineralization, collagen fiber orientation, and other factors.* It is important to understand the differences between cortical and trabecular bone tissue properties, but it is probably pointless to pursue increasingly “accurate” values for these properties.

Now let us move back up in scale once more, and consider how the architecture of a substantial group of trabeculae determines the mechanical properties of a region of cancellous bone a few millimeters across. There are two key aspects of this architecture: the **apparent density** and the **orientation** of the trabeculae.

7.6.2.4 Apparent Density

As we noted earlier, Carter and Hayes (1977b) found that, when elastic modulus is plotted against apparent density for cortical and cancellous bone specimens combined, E varies with d^3 (Eq. 7.12). However, Rice and co-workers (1988), in an extensive analysis of data from several experiments, found that the exponent was 2 when only cancellous bone was included. Typically, such regression analyses show that apparent density or porosity account for 70–80 % of the variability in elastic modulus if one is considering a particular species, type of stress (tension or compression), and direction of loading relative to anatomical axes. The percentage goes down if these factors vary.

Rice and co-workers also confirmed that the compressive strength of cancellous bone varies with d^2 , and this correlation accounts for about 73 % of the variability. Restrictions regarding species and directionality applied to this case as well. Thus, about three-fourths of the variability in the strength and stiffness of cancellous bone can be accounted for by apparent density. Presumably, the balance depends on other factors, one of which is thought to be the directionality of the trabeculae relative to the loading direction.

7.6.2.5 Anisotropy of Cancellous Bone Architecture: Mean Intercept Length

It is intuitively clear, and abundant experimental data support the notion that the orientation and spacing of trabeculae influence the elastic modulus and strength of cancellous bone as well as the anisotropy of these variables (Raux et al. 1975; Townsend et al. 1975; Goldstein 1987). Both trabecular and cortical bone are anisotropic. In cortical bone the solid matrix dominates the space, and anisotropy is primarily determined by lamellar and osteonal orientation. In cancellous bone, soft tissues dominate the space, and anisotropy is governed by trabecular orientation. The relative numbers and sizes of trabeculae oriented in different directions control the anisotropy of the macroscopically measured material properties. However, the nature of this relationship was slow to emerge. The development of this understanding began with the work of Whitehouse and Dyson (1974; Whitehouse 1974a, b), who developed a method for measuring the microscopic structural anisotropy of trabecular bone. They measured the average distance, along test lines laid out in various directions, from one trabecular surface to another through the marrow; this distance is called *mean intercept length*, or MIL (Fig. 7.17).

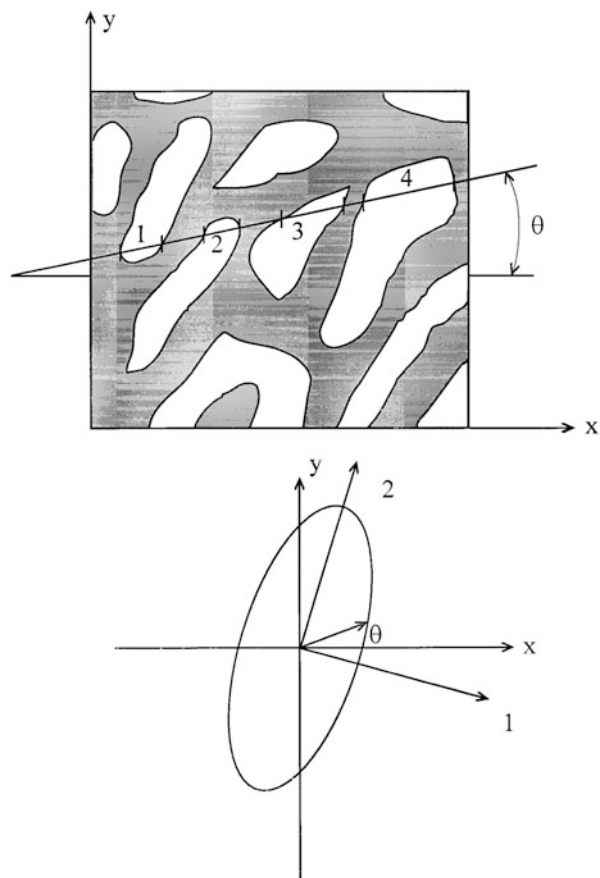


FIGURE 7.17. *Upper:* Schematic diagram of anisotropic cancellous bone structure as seen in a two-dimensional section also depicts an example of a *test line* used to measure mean intercept length. This *line* makes an angle θ with the *x*-axis and the intercept lengths measured along it are shown as *line segments* 1–4. *Lower:* Plot of mean intercept length (MIL) vs. θ is an *ellipse*. The vector for test lines having orientation θ is shown inside the *ellipse*. The *ellipse* has minor and major axes (labeled 1 and 2) aligned with the principal directions of the two-dimensional version of the MIL tensor. Note that the maximum principal axis of the MIL *ellipse* is aligned with the general orientation of the trabeculae in the *upper* diagram.

Whitehouse and Dyson found that, for a two-dimensional section, a graph of MIL vs. the measurement direction often approximately produced an ellipse whose eccentricity is a measure of anisotropy. Harrigan and Mann (1984) pointed out that if the three-dimensional MIL was an ellipse then it could be expressed by a second-rank tensor (the MIL tensor). For a material that has an ellipsoidal MIL Cowin (1985) derived an expansion for the elastic

constants of an anisotropic material expressed in terms of apparent density (or porosity) and a *fabric tensor* that is equal to the inverse square root of the MIL tensor.

The equations for the linear theory relating the elastic constants to porosity and the fabric tensor components were tested experimentally (Turner and Cowin 1987; Cowin and Mehrabadi 1989), and Turner et al. (1990) and accounted for 72–94 % of the variability in the elastic constants for cancellous bone.

Because the theory assumes that the solid matrix of the bone is homogeneous and isotropic, some of the unexplained variability probably resulted from organizational factors within individual trabeculae. For example, individual trabeculae more closely aligned with average principal **tensile** stresses may have collagen-mineral structures (and thus modulus values) that are different from those trabeculae closely aligned with average principal **compressive** stress directions. Models for mechanical properties that use fabric measurements continue to be used and have now been extended to model the nonlinear mechanical properties of cancellous bone (Charlebois et al. 2010).

7.6.3 Modeling Cancellous Bone as a Cellular System of Plates or Struts

Gibson (1985) explored the relationships between mechanical properties and apparent density of cancellous bone using open- and closed-cell porous structures like those shown in Fig. 7.18. She found that for a closed-cell structure (more representative of cortical bone or dense cancellous bone with unperforated plates), stiffness varied as d^3 , but for an open-cell structure (like most cancellous bone), stiffness varied as d^2 . This result is consistent with the different apparent density relationships for cortical and cancellous bone found by Carter and Hayes (1977b) and Rice et al. (1988). Gibson also

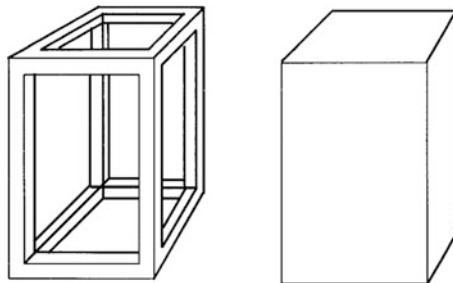


FIGURE 7.18. Types of open (*left*) and closed (*right*) cell structures used by Gibson (1985) to model cancellous bone behavior. (The “box” at *right* is hollow.) It is understood that in bone the *solid plates* forming the closed cell walls are actually perforated so that the marrow spaces communicate.

found that for bone arranged in columnar cells, both strength and stiffness should be linearly related to apparent density. This finding agreed with data from Williams and Lewis (1982) for such bone.

7.6.4 Invariance of Yield Strain Hypothesis

We have noted that both the strength and modulus of cancellous bone depend on apparent density squared. It follows that strength is linearly proportional to modulus. This concept has been confirmed in several studies (Brown and Ferguson 1980; Currey 1969a; Bensusan et al. 1983; Goldstein et al. 1983; Vahey et al. 1987) and has an important implication for defining a failure criterion for cancellous bone. If yield stress, σ_y , and modulus, E , each vary with apparent density but are related by a constant, k , then

$$\sigma_y = kE \quad (7.13)$$

One has from Hooke's law (which applies virtually up to the yield point),

$$\sigma_y = E\epsilon_y \quad (7.14)$$

where ϵ_y is the yield strain. Eliminating σ_y/E from these equations yields

$$\epsilon_y = k \quad (7.15)$$

That is, the yield strain is a constant, equal to the slope of the strength-modulus graph, and immune to the large variations in strength and stiffness caused by differences in apparent density.

Turner (1989) tested this concept using bovine bone. He found that compressive yield stress varied from 1.4 to 20 MPa, but yield strain only varied from 0.52 to 1.21 %. The value of k was 0.88 %, very close to the mean yield strain of 0.92 %. Furthermore, ϵ_y was independent of trabecular orientation as measured by MIL. Later, Change et al. (1994) confirmed the isotropy and relative constancy of ϵ_y . However, they found that yield strain was about 28 % higher in compression than tension. Shear yield strain is also isotropic and independent of apparent density, particularly when the specimen is twisted about its own longitudinal axis (Ford and Keaveny 1996).

There is a connection between Turner's (1989) work and Gibson's (1985) models. Gibson assumed that individual trabeculae could deform (and fail) either by bending of struts oriented perpendicular to the load direction or by axial compression (and buckling) of struts oriented parallel to the load direction. Cowin (1989) has pointed out that the structure would be much stiffer and stronger in the latter case, but would be capable of absorbing more energy if bending of trabeculae occurred. Therefore, this is an interesting issue for several reasons. Turner had found that there was a slight but statistically significant increase in yield strain with bone volume fraction (and, presumably, apparent density). He performed an analysis of Gibson's

model to show that this would be consistent with failure by buckling but not with failure by bending of trabeculae. Analysis of on axis loading of cancellous bone using finite element models (Bevill et al. 2009) resulted in a dependence of failure strain on bone volume fraction that was not reproduced for the same models loaded off axis. The tissue level damage in the models was quite different as a result of loading in different directions but could be explained (in general) by referring back to Gibson's (1985) models. Note that finite element modeling lets you break the same sample more than once and in more than one direction! This is a significant advantage over "real" bones that can only be broken once.

7.7 Predicting Material Properties: Bone as a Composite Material

A number of studies have attempted to mathematically model bone as a composite material. The most general approach has been to assume that bone is composed of two or more phases (for example, collagen and mineral, or soft and hard tissue), and that these phases contribute to a mechanical property P in direct proportion to the volume fraction F that they occupy. (Most commonly, P is the elastic modulus.) One may simplify the structure of such a model to consist of alternating slabs of the two phases; if two such slabs are arranged so that they have equal strains when they are loaded, one has the Voigt model (Fig. 7.19), in which the composite property is given by

$$P_c = P_1 F_1 + P_2 F_2 \quad (7.16)$$

Here the subscripts 1 and 2 refer to the two component phases and P_c designates the composite property. If the slabs have equal stresses, the result is known as the Reuss model at right in Fig. 7.19, described by

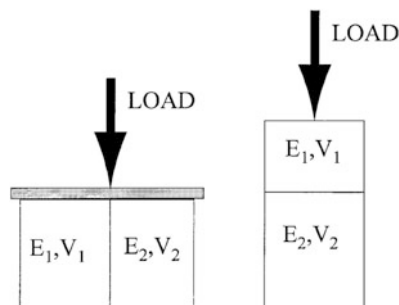


FIGURE 7.19. Voigt (*left*) and Reuss (*right*) models for composite material formed from a mixture of two materials. In the Voigt model, each material strains the same amount; in the Reuss model, each material carries the same stress.

$$P_c = \frac{P_1 P_2}{P_1(1 - F_1) + P_2(1 - F_2)} \quad (7.17)$$

It is known that the Voigt model provides an upper bound, and the Reuss model a lower bound, for the actual value of the composite property (Hill 1952, 1963). Unfortunately, these bounds are typically so far apart as to be virtually useless. The Hashin–Shtrickman model (Hashin and Shtrickman 1963) gives somewhat narrower bounds. Several authors (Currey 1964; Katz 1971; Bonfield and Li 1967; Piekarski 1973a, b) have assumed that compact bone consists of collagen (organic phase) and mineral (inorganic phase) and used these models to derive certain mechanical properties. These efforts have met with limited success, partly because the properties of the component phases are not well known, partly because their interactions are not well understood (Burstein et al. 1975), and perhaps because the void space was ignored. However, Martin (1984) had no better success when he assumed that bone consists of calcified matrix and voids and used the same methods to compute the elastic modulus. Others who have analyzed solid-void two-phase materials include Hirsch (1962) and Brown et al. (1964). Stech (1967) analyzed the anisotropy of compact bone using a model containing parallel cylindrical voids representing Haversian canals. Bonfield and Clark (1973) had some success in combining mixture theories to explain an age-related increase in the elastic modulus of rabbit bone that seemed to be caused by changes in both porosity and specific mineralization. They used a formulation by MacKenzie (1950)

$$E = E_0(1 - 1.9p + 0.9p^2) \quad (7.18)$$

where p is porosity and E_0 is the modulus of the calcified matrix, to account for porosity variation, and an equation similar to Eq. (7.10) to account for changes in mineralization. However, the analysis did not account for the entire increase in stiffness, supporting Currey's (1969b) hypothesis that bone apatite experiences age-related structural changes that increase its elastic modulus.

Many of these models contain constants that are more or less arbitrary and can be manipulated to fit various data. So far, however, no one has developed a theory of bone structure that can satisfactorily explain its mechanical behavior on the basis of its internal structure and composition.

7.8 Summary and Additional Reading

In this chapter, we have seen how the mechanical properties of cortical and cancellous bone are influenced by the histologic structures described and analyzed in Chaps. 2 and 3. Bone behaves as an elastic (and also viscoelastic), anisotropic, composite material. The feature of its histologic structure that dominates its mechanical properties is porosity, also expressed as bone

volume fraction (and closely related to apparent density). Mineralization and collagen fiber orientation are also important determinants of mechanical properties. In cortical bone, secondary osteons significantly alter anisotropy and strength, and in cancellous bone the orientations of trabeculae play a similar role. The elastic modulus and strength of cancellous bone are both heavily dependent on apparent density squared, and this results in yield strain being nearly isotropic and constant. Mixture theories may be used to formulate models for the mechanical behavior of bone, but the results have not been very satisfactory.

Recommendations for additional reading include the excellent book, *Bone Mechanics Handbook: Second Edition*, edited by Cowin (2001). This work contains chapters that provide a detailed mathematical presentation of elasticity and strength of materials theory, as well as compilations of data on the mechanical properties of cortical and cancellous bone. The chapter on bone's physical properties in Albright and Brand (1987) and the tutorial paper by Turner and Burr (1993) are excellent introductions to this subject for nonengineers. For an interesting read on the relationship of strength and toughness in a variety of biological and non-biological materials, we suggest Ritchie's (2011) review of this topic. Another resource intended for clinicians is the textbook *Fundamentals of Biomechanics: Second Edition* by Ozkaya and Nordin (2012). Alexander's monograph, *Optima for Animals* (Alexander 1996), contains a section on optimizing the dimensions of tubular bones with respect to strength and weight (pp. 13–23), and presents a succinct overview of optimization theory from the biological perspective.

7.9 Exercises

- 7.1. Show by analysis of the units of stress and strain that the area under a stress-strain curve has units of work/volume.
- 7.2. Suppose a human bone has a cylindrical shaft whose periosteal and endosteal diameters are 2 and 1 cm, respectively. Using the data in Table 7.1, calculate the yield and ultimate load for this bone when loaded in simple compression. (Ignore buckling as a possibility.) Then calculate the torques (twisting moments) required to yield and break the bone by twisting it. When the outer fibers of the bone have reached the ultimate shear stress, are the fibers on the endosteal surface above or below the shear yield stress?
- 7.3. Suppose the bone described in Exercise 7.2 has its shear strength diminished by 20 % by the accumulation of Haversian canals with age. Assume also that 1 mm of bone is lost from the endosteal surface during this period of time. How much must the periosteal diameter be increased to compensate for these changes? If all these changes occurred simultaneously over a 30-year period, what would the mean

periosteal apposition rate be, and how would it compare to typical values during the refilling of an osteonal BMU?

- 7.4. Suppose a long bone diaphysis has cleanly fractured straight across and is healing through the development of a fracture callus as described in Chap. 2. Sketch a cross section through the callus assuming circular geometry for it and the original bone. The outer diameter of the callus is 3 cm. Sketch in the original bone cross section, which had periosteal and endosteal diameters of 2 cm and 1 cm, respectively. Assume that no strength has developed between the cortices of the original bone, so that the strength of the healing fracture is entirely provided the periosteal and medullary calluses. Calculate the bending moment required to fracture the original, intact cortex, assuming that its ultimate bending stress was 200 MPa. Now calculate the cross-sectional moment of inertia of the periosteal and endosteal calluses and compare them, individually and together, to that of the original bone. Assume the material in these calluses is homogeneous. What would the ultimate stress of the callus material have to be for it to provide the same strength as the original bone? What percentage of this strength is provided by the periosteal callus?
- 7.5. Derive the Voigt and Reuss mixture theory equations (Eqs. 7.16 and 7.17) from the concepts shown in Fig. 7.19.
- 7.6. Assume that the two phases denoted by subscripts 1 and 2 in Eqs. (7.16) and (7.17) represent bone voids (soft tissue spaces) and calcified matrix, respectively. In that case, volume fraction F_1 becomes porosity (p) and $F_2 = 1 - p$. Let the material property P be elastic modulus, so that $P_2 = E_o$, the modulus of the calcified matrix. Try to plot Eqs. (7.8, 7.9, 7.16, and 7.17) on the same graph. You will have a problem with Eq. (7.17) if you assume the modulus of soft tissues to be 0, so choose a value small in proportion to E_o . What sort of insight into mixture theory and the effects of porosity on bone's mechanical properties can you obtain from your graph?
- 7.7. Assume that osteonal bone consists of parallel, longitudinally arranged secondary osteons in an interstitial matrix of primary bone and osteon remnants. Suppose that this composite is loaded in longitudinal tension so that the simple rule of mixtures (Eq. 7.16) applies to the elastic modulus. Schaffler (1985) has measured the elastic modulus of cow bone as a function of the volume fraction of secondary osteons, obtaining the following results:

Elastic modulus, GPa	Osteon volume fraction	Elastic modulus, GPa	Osteon volume fraction
19.90	0.00	20.70	0.21
26.20	0.00	18.10	0.41
24.00	0.00	21.50	0.21
24.90	0.00	14.90	0.37
24.50	0.00	15.70	0.40
26.00	0.08	21.00	0.27
21.80	0.00	19.10	0.32
22.10	0.00	18.10	0.47
23.30	0.00	25.20	0.05
22.80	0.03	25.20	0.03

Use Eq. (7.16) to derive a relationship between the elastic modulus of the composite bone (E_b), the osteon volume fraction (F_o), and the moduli of osteons (E_o) and the interstitial matrix (E_m). Based on this relationship, how are E_o and E_m related to the slope and intercept of a plot of E_b vs. F_o ? Obtain estimates of E_o and E_m from a plot of Schaffler's data. Which is greater? Why?

- 7.8. As described in Box 7.5, bone mass may be noninvasively measured by photon absorptiometry. The absorption curve shown in the box is a typical result for a long bone diaphysis. The area above the curve is proportional to the area occupied by mineral in the scanned section, and the distance across the curve is proportional to the periosteal diameter of the bone. It might be of mechanical interest to noninvasively determine the bending resistance of the bone's cross section. Derive a method for determining (a) the centroid of the cross-section's mineral in the scan direction and (b) the cross sectional moment of inertia of the mineral about an axis through the centroid and perpendicular to the scan direction.
- 7.9. Box 3.3 described the apparatus developed by Burstein and Frankel for testing the torsional strength of a long bone diaphysis under conditions of rapid loading. This machine uses the energy of a swinging pendulum to apply an impulsive torque to the specimen. The mass and length of the pendulum are 14 kg and 23 cm, respectively. To ensure that the impact will be sufficiently forceful, the energy in the swinging pendulum must be much greater than that required to break the bone. Assume each "arbitrary torque unit" in Fig. 3.7 is equal to 10 N-m. Estimate the energy to failure of the strongest of the specimens in this figure, and compare this energy with the potential energy of the pendulum at the start of the experiment. What percent of the pendulum's kinetic energy is required to break the healed rabbit bone, assuming no other energy losses?

- 7.10. Consider the stress in the neck of the femur of a 70-kg person during the stance phase of gait. Use what you learned in Chap. 1 about the magnitude of the hip joint force, and assume that it is directed at an angle of 20° to the neck axis. Calculate the compressive and bending stresses produced by the joint force components parallel and perpendicular to the neck axis, respectively. Choose reasonable values for the bony dimensions. Then estimate the total axial stress on the superior and inferior surfaces of the neck. How do these stresses compare to the compressive strengths of cancellous bone shown in Table 7.7? Discuss how you might account for the effect of a shell of cortical bone around the neck of the femur.

References

- Abendschein W, Hyatt GW. Ultrasonics and selected physical properties of bone. *Clin Orthop Relat Res.* 1970;69:294–301.
- Albright JA, Brand RA. *The scientific basis of orthopaedics.* New York, NY: Appleton; 1987.
- Alexander RM. *Optima for animals.* Princeton, NJ: Princeton University Press; 1996.
- Ascenzi A, Bonucci E. The ultimate tensile strength of single osteons. *Acta Anat.* 1964;58:160–83.
- Ascenzi A, Bonucci E. The tensile properties of single osteons. *Anat Rec.* 1967;158:375–86.
- Ascenzi A, Bonucci E. The compressive properties of single osteons. *Anat Rec.* 1968;161:377–91.
- Ascenzi A, Bonucci E. The shearing properties of single osteons. *Anat Rec.* 1972;172:499–510.
- Ascenzi MG, Lomovtsev A. Collagen orientation patterns in human secondary osteons, quantified in the radial direction by confocal microscopy. *J Struct Biol.* 2006;153:14–30.
- Ashman RB, Rho JY. Elastic modulus of trabecular bone material. *J Biomech.* 1988;21(3):177–81.
- Ashman RB, Van Buskirk WC. The elastic properties of a human mandible. *Adv Dent Res.* 1987;1(1):64–7.
- Ashman RB, Mayer DC, Cowin SC, Van Buskirk WC. Elastic properties of human and canine femora. *Trans Orthop Res Soc.* 1983;8:127.
- Ashman RB, Cowin SC, Van Buskirk WC, Rice JC. A continuous wave technique for the measurement of the elastic properties of cortical bone. *J Biomech.* 1984;17:349–61.
- Ashman RB, Van Buskirk WC, Cowin SC, Sandborn PM, Wells MK, Rice JC. The mechanical properties of immature osteopetrotic bone. *Calcif Tissue Int.* 1985a;37:73–6.
- Ashman RB, Rosinia G, Cowin SC, Fontenot MG. The bone tissue of the canine mandible is elastically isotropic. *J Biomech.* 1985b;18:717–21.
- Bargren JH, Bassett CAL, Gjelsvik A. Mechanical properties of hydrated cortical bone. *J Biomech.* 1974;7:239–45.

- Batson EL, Reilly GC, Currey JD, Balderson DS. Postexercise and positional variation in mechanical properties of the radius in young horses. *Equine Vet J.* 2000;32:95–100.
- Beaupre GS, Carter DR. Finite element analysis in biomechanics. In: Biewener AA, editor. *Biomechanics—structures and systems*. Oxford: Oxford University Press; 1992. p. 149–74.
- Behrens JC, Walker PS, Shoji H. Variations in strength and structure of cancellous bone at the knee. *J Biomech.* 1974;7:201–7.
- Bensusan JS, Davy DT, Heiple KG, Verdin PJ. Tensile, compressive, and torsional testing of cancellous bone. *Trans Orthop Res Soc.* 1983;8:132.
- Bevill G, Farhamand F, Keaveny TM. Heterogeneity of yield strain in low-density versus high-density human trabecular bone. *J Biomech.* 2009;42:2165–70.
- Bonfield W, Clark EA. Elastic deformation of compact bone. *J Mater Sci.* 1973;8:1590–4.
- Bonfield W, Li CH. Anisotropy of non-elastic flow in bone. *J Appl Physiol.* 1967;38:2450.
- Bonfield W, Tully AE. Ultrasonic analysis of the Young's modulus of cortical bone. *J Biomed Eng.* 1982;4:23–7.
- Boyde A, Riggs CM. The quantitative study of the orientation of collagen in compact bone slices. *Bone.* 1990;11:35–9.
- Brown TD, Ferguson AB. Mechanical property distributions in the cancellous bone of the human proximal femur. *Acta Orthop Scand.* 1980;51:429–37.
- Brown SD, Biddulph RB, Wilcox PD. A strength-porosity relation involving different pore geometry and orientation. *Am Ceram Soc J.* 1964;47:320.
- Burr DB, Schaffler MB, Fredrickson RG. Composition of the cement line and its possible mechanical role as a local interface in human compact bone. *J Biomech.* 1988;21:939–45.
- Burstein AH, Frankel VH. The viscoelastic properties of some biological materials. *Ann N Y Acad Sci.* 1968;146:158–65.
- Burstein AH, Currey JD, Frankel VH, Reilly DT. The ultimate properties of bone tissue: the effects of yielding. *J Biomech.* 1972;5:35–44.
- Burstein AH, Zika JM, Heiple KG, Klein L. Contribution of collagen and mineral to the elastic-plastic properties of bone. *J Bone Joint Surg.* 1975;57A:956–61.
- Carter DR, Hayes WC. Fatigue life of compact bone. I. Effects of stress amplitude, temperature and density. *J Biomech.* 1976;9:27–34.
- Carter DR, Hayes WC. Compact bone fatigue damage. I. Residual strength and stiffness. *J Biomech.* 1977a;10:325–37.
- Carter DR, Hayes WC. The compressive behavior of bone as a two-phase porous structure. *J Bone Joint Surg.* 1977b;59A:954–62.
- Carter DR, Spengler DM. Mechanical properties and composition of cortical bone. *Clin Orthop Relat Res.* 1978;135:192–217.
- Carter DR, Hayes WC, Schurman DJ. Fatigue life of compact bone. II. Effects of microstructure and density. *J Biomech.* 1976;9:211–8.
- Carter DR, Schwab GH, Spengler DM. Tensile fracture of cancellous bone. *Acta Orthop Scand.* 1980;51:733–41.
- Chang WCW, Christensen TM, Pinilla TP, Keaveny TM. Uniaxial yield strains for bovine trabecular bone are isotropic and asymmetric. *J Orthop Res.* 1994;17:582–5.

- Charlebois M, Pretterklieber M, Zysset PK. The role of fabric in the large strain compressive behavior of human trabecular bone. *Trans ASME J Biomech Eng.* 2010;132(12), 121006.
- Chen HL, Gundjian AA. Determination of the bone-crystallites distribution function by x-ray diffraction. *Med Biol Eng.* 1974;12:531–5.
- Choi K, Kuhn JL, Ciarrelli MJ, Goldstein SA. The elastic moduli of human subchondral, trabecular, and cortical bone tissue and the size-dependency of cortical bone modulus. *J Biomech.* 1990;23:1103–14.
- Cowin SC. The relationship between the elasticity tensor and the fabric tensor. *Mech Mater.* 1985;4:137–47.
- Cowin SC. Wolff's law of trabecular architecture at remodeling equilibrium. *J Biomech Eng.* 1986;108:83–8.
- Cowin SC. Bone mechanics. Boca Raton, FL: CRC Press; 1989.
- Cowin SC, editor. Bone mechanics handbook. 2nd ed. Boca Raton, FL: CRC Press; 2001.
- Cowin SC, Mehrabadi MM. Identification of the elastic symmetry of bone and other materials. *J Biomech.* 1989;22:503–15.
- Crowninshield RD, Pope MH. The response of compact bone in tension at various strain rates. *Ann Biomed Eng.* 1974;2:217–25.
- Currey JD. Differences in the tensile strength of bone of different histological types. *J Anat.* 1959;98:87–95.
- Currey JD. Three analogies to explain the mechanical properties of bone. *Biorheology.* 1964;2:1–10.
- Currey JD. The mechanical consequences of variation in the mineral content of bone. *J Biomech.* 1969a;2:1–11.
- Currey JD. The relationship between the stiffness and the mineral content of bone. *J Biomech.* 1969b;2:477–80.
- Currey JD. Effects of porosity and mineral content on the Young's modulus of bone. *Proc Eur Soc Biomech.* 1986;5:104.
- Currey JD. The effect of porosity and mineral content on the Young's modulus of elasticity of compact bone. *J Biomech.* 1988;21:131–9.
- Currey JD. Bones: structure and mechanics. Princeton, NJ: Princeton University Press; 2002.
- Currey JD. Incompatible mechanical properties in compact bone. *J Theoret Biol.* 2004;231:569–80.
- Currey JD, Butler G. The mechanical properties of bone tissue in children. *J Bone Joint Surg.* 1975;57A:810–4.
- Currey JD, Pitchford JW, Baxter PD. Variability of the mechanical properties of bone, and its evolutionary consequences. *J R Soc Interf.* 2007;4:127–35.
- Dyson ED, Whitehouse WJ. Composition of trabecular bone in children and its relation to radiation dosimetry. *Nature (London).* 1968;217:576–8.
- Enlow DH, Brown SO. A comparative histological study of fossil and recent bone tissues. Part I. Introduction, methods, fish and amphibian bone tissues. *Tex J Sci.* 1956;8:405–43.
- Enlow DH, Brown SO. A comparative histological study of fossil and recent bone tissues. Part II. Reptilian and bird bone tissues. *Tex J Sci.* 1957;9:186–214.
- Enlow DH, Brown SO. A comparative histological study of fossil and recent bone tissues. Part III. Mammalian bone tissues. General discussion. *Tex J Sci.* 1958;10:187–230.

- Evans FG, Vincentelli R. Relation of collagen fiber orientation to some mechanical properties of human cortical bone. *J Biomech.* 1969;2:63–71.
- Evans FG, Vincentelli R. Relations of the compressive properties of human cortical bone to histological structure and calcification. *J Biomech.* 1974;7:1–10.
- Ford CM, Keaveny TM. The dependence of shear failure properties of trabecular bone on apparent density and trabecular orientation. *J Biomech.* 1996;29:1309–17.
- Frost HM. *Orthopaedic biomechanics.* Springfield: Thomas; 1973.
- Frost HM. *Intermediary organization of the skeleton.* Boca Raton, FL: CRC Press; 1986.
- Frost HM, Roth H, Villanueva AR. Physical characteristics of bone. Part III. A semimicro measurement of unit shear stress. *Henry Ford Hosp Med J.* 1961;9:157–62.
- Fung YC. *A first course in continuum mechanics.* Englewood Cliffs, NJ: Prentice Hall; 1969.
- Fung YC. *A first course in continuum mechanics.* Englewood Cliffs, NJ: Prentice-Hall; 1977.
- Fyhrie DP, Carter DR. A unifying principle relating stress to trabecular bone morphology. *J Orthop Res.* 1986;4:304–17.
- Gannon RE, Harris GM, Vasilos T. Effect of porosity on mechanical thermal, and dielectric properties of fused silica. *Am Ceram Soc Bull.* 1965;44:460.
- Gebhardt FSMW. Über funktionell wichtige anordnungsweisen der feineren und groberen bauelemente des wirbeltierknochens. II. Spezieller Teil. 1. Der bau der Haversschen lamellensysteme und seine funktionelle bedeutung. *Roux Arch.* 1905;20:187–334.
- Gebhardt FSMW. Über funktionell wichtige anordnungsweisen der groberen und feineren bauelemente des wirbeltierknochens. I. Allgemeiner Teil. *Wilhelm Roux Arch fur Entwicklungsmeckanik der Organismen.* 1901;12:167–223.
- Gibson LJ. The mechanical behavior of cancellous bone. *J Biomech.* 1985;18:317–28.
- Giraud-Guille MM. Twisted plywood architecture of collagen fibrils on human compact bone osteons. *Calcif Tissue Int.* 1988;42:167–80.
- Goldstein SA. The mechanical properties of trabecular bone: dependence on anatomic location and function. *J Biomech.* 1987;20:1055–61.
- Goldstein SA, Wilson DL, Sonstegard DA, Matthews LS. The mechanical properties of human tibial trabecular bone as a function of metaphyseal location. *J Biomech.* 1983;16:865–969.
- Gong JK, Arnold JS, Cohn SH. Composition of trabecular and cortical bone. *Anat Rec.* 1964;149:325–32.
- Goodwin KJ, Sharkey NA. Material properties of interstitial lamellae reflect local strain environments. *J Orthop Res.* 2002;20:600–6.
- Gray RJ, Korbacher GK. Compressive fatigue behaviour of bovine compact bone. *J Biomech.* 1974;7:287–92.
- Harrigan TP, Mann RW. Characterization of microstructural anisotropy in orthotropic materials using a second rank tensor. *J Mater Sci.* 1984;19:761–7.
- Harrigan TP, Jasty M, Mann RW, Harris WH. Limitations of the continuum assumption in cancellous bone. *J Biomech.* 1988;21:269–75.
- Hart RT. The finite element method. In: Cowin SC, editor. *Bone mechanics.* Boca Raton, FL: CRC Press; 1989. p. 53–74.
- Hashin Z, Shtrikman S. A variation approach to the theory of the elastic behavior of multiphase materials. *J Mech Phys Solids.* 1963;11:127.

- Hernandez CJ, Beaupre GS, Keller TS, Carter DR. The influence of bone volume fraction and ash fraction on bone strength and modulus. *Bone*. 2001;29:74–8.
- Hèrt J, Kucera P, Vavra M, Volenik V. Comparison of the mechanical properties of both the primary and Haversian bone tissue. *Acta Anat*. 1965;61:412–23.
- Hill R. The elastic behavior of a crystalline aggregate. *Proc Phys Soc Lond*. 1952; A65:349.
- Hill R. Elastic properties of reinforced solids: some theoretical principles. *J Mech Phys Solids*. 1963;11:357.
- Hirsch TJ. Modulus of elasticity of concrete affected by elastic moduli of cement paste matrix and aggregate. *Proc Am Concr Soc*. 1962;59:427.
- Hulsen C. Spezifisches Gewicht, Elastizität und Festigkeit des Knochengewebes. *Bull Lab Biol St Petersburg*. 1898;1:7–37.
- Hvid I. Mechanical strength of trabecular bone at the knee. *Dan Med Bull*. 1988;35:345–65.
- Hvid I, Hansen SL. Trabecular bone strength patterns at the proximal tibial epiphysis. *J Orthop Res*. 1985;3:464–72.
- Hylander WL. Patterns of stress and strain in the macaque mandible. In: Carlson DS, editor. *Craniofacial biology*. Ann Arbor, MI: Center for Human Growth and Development; 1981.
- Johnson M, Katz JL. Some new developments in the rheology of bone. *Biorheol Suppl*. 1984;1:169–74.
- Kaplan SJ, Hayes WC, Stone JL, Beaupre GS. Tensile strength of bovine trabecular bone. *J Biomech*. 1985;18:723–7.
- Katz JL. Hard tissue as a composite material. I. Bounds on the elastic behavior. *J Biomech*. 1971;4:455–73.
- Katz JL. Composite material models for cortical bone. In: Cowin SC, editor. *Mechanical properties of bone*. New York, NY: ASME; 1981.
- Katz JL, Yoon HS, Lipson S, Maharidge R, Meunier A, Christel P. The effects of remodeling on the elastic properties of bone. *Calcif Tissue Int*. 1984;36:S31–6.
- Keaveny TM, Guo XE, Wachtel EF, McMahon TA, Hayes WC. Trabecular bone exhibits fully linear elastic behavior and yields at low strains. *J Biomech*. 1994;27:1127–37.
- Knets IV. Mechanics of biological tissues. A review. *Polym Mech* (translation of *Mekhanika Polimerov*). 1978;13:434–40.
- Köl liker A. *Handbuch der Bewebelehre de Menschen*. Leipzig, 1889.
- Kuhn JL, Goldstein SA, Choi KW, London M, Feldkamp LA, Mathews LS. Comparison of the trabecular and cortical tissue moduli from human iliac crests. *J Orthop Res*. 1989a;7:876–84.
- Kuhn JL, Goldstein SA, Ciarrelli MJ, Mathews LS. The limitations of canine trabecular bone as a model for human: a biomechanical study. *J Biomech*. 1989b;22:95–107.
- Laird GW, Kingsbury HB. Complex viscoelastic moduli of bovine bone. *J Biomech*. 1973;6(1):59–67.
- Lakes RS, Katz JL. Interrelationships among the viscoelastic functions for anisotropic solids: application to calcified tissues and related systems. *J Biomech*. 1974;7:259–70.
- Lakes RS, Katz JL. Viscoelastic properties of wet cortical bone. III. A nonlinear constitutive equation. *J Biomech*. 1979a;12:689–98.
- Lakes RS, Katz JL. Viscoelastic properties of wet cortical bone. II. Relaxation mechanisms. *J Biomech*. 1979b;12:679–87.

- Lakes R, Saha S. Long-term torsional creep in compact bone. *J Biomech Eng.* 1980;102:178–80.
- Lakes RS, Katz JL, Sternstein SS. Viscoelastic properties of wet cortical bone. I. Torsional and biaxial studies. *J Biomech.* 1979;12:657–78.
- Lakes RS, Yoon HS, Katz JL. Ultrasonic wave propagation and attenuation in wet bone. *J Biomed Eng.* 1986;8:143–8.
- Launey ME, Buehler MJ, Ritchie RO. On the mechanistic origins of toughness in bone. *Annu Rev Mater Res.* 2010;40:25–53.
- Lekhnitskii SG. *Theory of elasticity of an anisotropic body* (English translation). San Francisco: Holden-Day; 1963.
- Linde F, Hvid I, Madsen F. The effect of specimen geometry on the mechanical behavior of trabecular bone specimens. *J Biomech.* 1992;25:359–68.
- Lipson SF, Katz JL. The relationship between elastic properties and microstructure of bovine cortical bone. *J Biomech.* 1984;17:231–40.
- MacKenzie JK. The elastic constants of a solid containing spherical holes. *Proc Phys Soc.* 1950;B63:2–11.
- Margel-Robertson D, Smith DC. Compressive strength of mandibular bone as a function of microstructure and strain rate. *J Biomech.* 1978;11:455–71.
- Martin RB. Porosity and specific surface of bone. *Crit Rev Biomed Eng.* 1984;10:179–222.
- Martin RB, Boardman DL. The effects of collagen fiber orientation, porosity, density, and mineralization on bovine cortical bone bending properties. *J Biomech.* 1993;26:1047–54.
- Martin RB, Ishida J. The relative effects of collagen fiber orientation, porosity, density, and mineralization on bone strength. *J Biomech.* 1989;22:419–26.
- Martin RB, Pickett JC, Zinaich S. Studies of skeletal remodeling in aging men. *Clin Orthop Relat Res.* 1980;149:268–82.
- Martin RB, Lau ST, Mathews PV, Gibson VA, Stover SM. Collagen fiber organization is related to mechanical properties and remodeling in equine bone. A comparison of two methods. *J Biomech.* 1996;29:1515–22.
- Mazess RB, Cameron JR. Direct readout of bone mineral content using radionuclide absorptiometry. *Int J Appl Radiat Isot.* 1972;23:471–9.
- McElhaney JH. Dynamic response of bone and muscle tissue. *J Appl Physiol.* 1966;21:1231–6.
- McElhaney J, Alem N, Roberts V. A porous block model for cancellous bones. ASME Paper No. 70-WA/BHF-2, American Society of Mechanical Engineers, New York; 1970. p. 1–9.
- Moyle DD, Bowden RW. Fracture of human femoral bone. *J Biomech.* 1984;17:203–13.
- Moyle DD, Welborn JW, Cooke FW. Work to fracture of canine femoral bone. *J Biomech.* 1978;11:435–40.
- Murray JD. *Mathematical biology*. New York, NY: Springer; 2003.
- Murray PDF. *Bones: a study of the development and structure of the vertebrate skeleton*. (Copyright 1936 Cambridge University Press). Cambridge: Cambridge University Press. Reprinted 1985.
- Neil JL, Demos TC, Stone JL, Hayes WC. Tensile and compressive properties of vertebral trabecular bone. *Trans Orthop Res Soc.* 1983;8:344.

- Norrdin RW, Phemister RD, Jaenke RS, LoPresi CA. Density and composition of trabecular and cortical bone in perinatally irradiated beagles with chronic renal failure. *Calcif Tissue Res*. 1977;24:99–104.
- Norrdin RW, Histand MB, Sheahan HJ, Carpenter TR. Effects of corticosteroids on mechanical strength of intervertebral joints and vertebrae in dogs. *Clin Orthop Relat Res*. 1990;259:268–76.
- Olivio OM. Rispondenza della funzione meccanica varia degli osteoni con la loro diversa minuta architettura. *Bollentino della Societa Italiana Biologia Sperimentale*. 1937;12:400–1.
- Olivio OM, Maj G, Toajari E. Sul significato della minuta struttura del tessuto osseo compatto. *Bollentino della Societa Italiana Biologia Sperimentale*. 1937;109:369–94.
- Ozkaya N, Nordin M. *Fundamentals of biomechanics: equilibrium, motion and deformation*. 3rd ed. New York, NY: Springer; 2012.
- Panjabi MM, White AA, Southwick WO. Mechanical properties of bone as a function of rate of deformation. *J Bone Joint Surg*. 1973;55A:322–30.
- Pauwels F. *Biomechanics of the locomotor apparatus: contributions on the functional anatomy of the locomotor apparatus*. Berlin: Springer; 1980.
- Pidaparti R, Burr DB. Collagen fiber orientation and geometry effects on the mechanical properties of secondary osteons. *J Biomech*. 1992;25:869–80.
- Piekarski K. Analysis of bone as a composite material. *J Biomech*. 1973a;11:557–65.
- Piekarski K. Analysis of bone as a composite aggregate. *Int J Eng Sci*. 1973b;11:557.
- Portigliatti-Barbos M, Bianco P, Ascenzi A. Distribution of osteonic and interstitial components in the human femoral shaft with reference to structure, calcification and mechanical properties. *Acta Anat*. 1983;115:178–86.
- Portigliatti-Barbos M, Bianco P, Ascenzi A, Boyde A. Collagen orientation in compact bone: II. Distribution of lamellae in the whole of the human femoral shaft with reference to its mechanical properties. *Metab Bone Dis Relat Res*. 1984;5:309–15.
- Ramaekers JGM. The dynamic shear modulus of bone in dependence on the form. *Acta Morphol Neerl Scand*. 1977;15:185–201.
- Ramaekers JGM. The rheological behavior of cortical bone and cartilage of *Bos taurus*. The importance of maximal damping capacity in collagen. *Acta Morphol Neerl Scand*. 1978;16:55–67.
- Ramaekers JGM. The mechanical properties of mallard humeral bone in dependence on the form. *Acta Morphol Neerl Scand*. 1979a;17:173–80.
- Ramaekers JGM. The dynamic shear modulus and damping of compact bovine metacarpal bone in dependence on the topography along the bone shaft. *Neth J Zool*. 1979b;29:151–65.
- Ramaekers JGM. The rheological behavior of skeletal material originating from several classes of vertebrates. *Neth J Zool*. 1979c;29:166–76.
- Ranvier J. *Traite Technique d'Histologie*. Paris: Savy; 1889.
- Raux P, Townsend PR, Miegel R, Rose RM, Radin EL. Trabecular architecture of the human patella. *J Biomech*. 1975;8:1–7.
- Reilly DT, Burstein AH. The mechanical properties of cortical bone. *J Bone Joint Surg*. 1974;56A:1001–21.
- Reilly DT, Burstein AH. The elastic and ultimate properties of compact bone tissue. *J Biomech*. 1975;8:393–405.
- Reilly DT, Burstein AH, Frankel VH. The elastic modulus for bone. *J Biomech*. 1974;7:271–5.

- Rho J-Y, Ashman RB, Turner CH. Young's modulus of trabecular and cortical bone material: ultrasonic and microtensile specimens. *J Biomech.* 1993;26:111–9.
- Rho JY, Currey JD, Zioupos P, Pharr GM. The anisotropic Young's modulus of equine secondary osteons and interstitial bone determined by nanoindentation. *J Exp Biol.* 2001;204:1775–81.
- Rice JC, Cowin SC, Bowman JA. On the dependency of the elasticity and strength of cancellous bone on apparent density. *J Biomech.* 1988;21:155–68.
- Ritchie RO. The conflicts between strength and toughness. *Nat Mat.* 2011;10:817–22.
- Roesler H. The history of some fundamental concepts in bone biomechanics. *J Biomech.* 1987;20:1025–34.
- Rohl L, Larsen E, Linde F, Odgaard A, Jorgensen J. Tensile and compressive properties of cancellous bone. *J Biomech.* 1991;24:1143–9.
- Ruggeri A, Raspanti M, Martini D, Guizzardi S. Bone as a composite material: collagen fibers orientation and functional requirements. In: Motta PM, editor. *Recent advances in microscopy of cells, tissue, and organs.* Rome: Antonio Delfino Editore; 1997. p. 167–1970.
- Schaffler MB. Stiffness and fatigue of compact bone at physiological strains and strain rates. Doctoral dissertation, West Virginia University, Morgantown, WV; 1985.
- Schaffler MB, Burr DB. Stiffness of compact bone: effects of porosity and density. *J Biomech.* 1988;21:13–6.
- Schaffler MB, Radin EL, Burr DB. Mechanical and morphological effects of strain rate on fatigue of compact bone. *Bone (New York).* 1989;10:207–14.
- Sedlin ED. A rheological model for cortical bone. *Acta Orthop Scand.* 1965;34:1–78.
- Skedros JG, Holmes JL, Vada EG, et al. Cement lines of secondary osteons in human bone are not mineral-deficient: new data in a historical perspective. *Anat Rec A.* 2005;286:781–803.
- Smith RW, Keiper DA. Dynamic measurement of viscoelastic properties of bone. *Am J Med Electron.* 1965;4(4):156–60.
- Smith JW, Walmsley R. Factors affecting the elasticity of bone. *J Anat.* 1959;93:503–22.
- Stech EL. A descriptive model of lamellar bone anisotropy. In: Byars EF, editor. *Biomechanics monographs.* New York, NY: ASME; 1967. p. 236–45.
- Townsend PR, Raux P, Rose RM, Miegel RE, Radin EL. The distribution and anisotropy of the stiffness of cancellous bone in the human patella. *J Biomech.* 1975;8:363–7.
- Turner CH. Yield behavior of bovine cancellous bone. *J Biomech Eng.* 1989;111:256–60.
- Turner CH, Burr DB. Basic biomechanical measurements of bone: a tutorial. *Bone (New York).* 1993;14:595–608.
- Turner CH, Burr DB. Orientation of collagen in osteonal bone. *Calcif Tissue Int.* 1997;60:90.
- Turner CH, Cowin SC. Dependence of elastic constants of an anisotropic porous material upon porosity and fabric. *J Mater Sci.* 1987;22:3178–84.
- Turner CH, Cowin SC, Rho JY, Ashman RB, Rice JC. The fabric dependence of the orthotropic elastic constants of cancellous bone. *J Biomech.* 1990;23:549–61.
- Underwood EE. *Quantitative stereology.* Reading, MA: Addison-Wesley; 1970.
- Vahey JW, Lewis JL, Vanderby R. Elastic moduli, yield stress, and ultimate stress of cancellous bone in the canine proximal femur. *J Biomech.* 1987;20:29–33.

- Van Buskirk WC, Cowin SC, Ward RN. Ultrasonic measurement of orthotropic elastic constants of bovine femoral bone. *J Biomech Eng.* 1981;103:67–72.
- Viceconti M. Multiscale modeling of the skeletal system. Cambridge: Cambridge University Press; 2012.
- Vincentelli R, Evans FG. Relations among mechanical properties, collagen fibers, and calcification in adult human cortical bone. *J Biomech.* 1971;4:193–201.
- Vincentelli R, Grigorov M. The effect of Haversian remodeling on the tensile properties of human cortical bone. *J Biomech.* 1985;18:201–7.
- Von Ebner V. Über den feineren bau der knochensubstanz. *Sitzungsberichte der Deutschen Akademie der Wissenschaften Zu Berlin.* 1875;72:45–138.
- Vose GP, Kubala AL. Bone strength—its relationship to x-ray-determined ash content. *Human Biol.* 1959;31:261–70.
- Wainwright SA, Biggs WD, Currey JD, Gosline JM. Mechanical design in organisms. New York, NY: Halsted Press; 1976.
- Walmsley R, Smith JW. Variation in bone structure and the value of Young's modulus. *J Anat Lond.* 1957;91:603.
- Whitehouse WJ. A stereological method for calculating internal surface areas in structures which have become anisotropic as the result of linear expansions or contractions. *J Microsc (Oxford).* 1974a;101:169–80.
- Whitehouse WJ. The quantitative morphology of anisotropic trabecular bone. *J Microsc (Oxford).* 1974b;101(2):153–68.
- Whitehouse WJ, Dyson ED. Scanning electron microscope studies of trabecular bone in the proximal end of the human femur. *J Anat.* 1974;118:417–23.
- Williams JL, Lewis JL. Properties and an anisotropic model of cancellous bone from the proximal tibial epiphysis. *J Biomech Eng.* 1982;104:50–6.
- Wolff J. The law of bone remodeling (Translation of Wolff's *Das Gesetz der Transformation der Knochen* by P. Maquet and R. Furlong.). Berlin: Springer; 1892.
- Yamada H. Strength of biological materials. Huntington: Krieger; 1973.
- Yoon HS, Katz JL. Ultrasonic wave propagation in human cortical bone-I. Theoretical considerations for hexagonal symmetry. *J Biomech.* 1976a;9:407–12.
- Yoon HS, Katz JL. Ultrasonic wave propagation in human cortical bone. II. Measurements of elastic properties and microhardness. *J Biomech.* 1976b;9:459–64.

8

Fatigue and Fracture Resistance of Bone

Textbooks and Heaven only are ideal;
Solidity is an imperfect state.
Within the cracked and dislocated Real
Nonstoichiometric Crystals dominate.

John Updike in "The Dance of the Solids"

Abstract In Chap. 7, we examined the mechanical properties of bone when it is fractured by a load sufficient to exceed the failure stress of the material. Structures may also fail more gradually. There are two principal ways in which this can happen: creep and fatigue. In both, a stress less than the ultimate stress is applied, and damage from this stress grows and accumulates until failure occurs.

Keywords Fatigue • Creep • Fracture mechanics • Microcracks • Toughness • Targeted remodeling

8.1 Introduction

In Chap. 7, we examined the mechanical properties of bone when it is fractured by a load sufficient to exceed the failure stress of the material. Structures may also fail more gradually. There are two principal ways in which this can happen: creep and fatigue. In both, a stress less than the ultimate stress is applied, and damage from this stress grows and accumulates until failure occurs. In creep, the stress is applied continuously; in fatigue, the stress is applied periodically. Obviously, these phenomena may occur simultaneously in the components of an engineering structure like a bridge, where the weight of the structure provides creep loading and the periodic passage of vehicles produces fatigue. Similarly, in the skeleton, bones often support more or less constant loads for prolonged periods of time (such as the vertebral bodies in your spine as you sit reading this book), and equally often carry cyclic loads (such as when you walk to class). In this chapter we shall see that creep and fatigue are closely related and explore the ways in

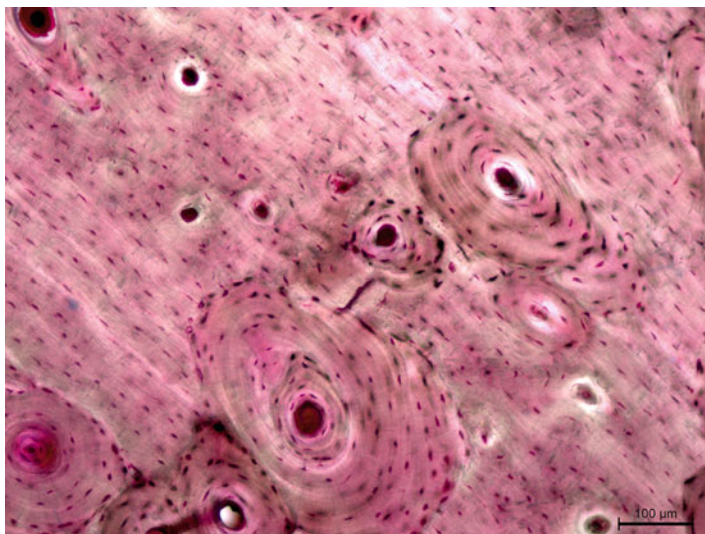


FIGURE 8.1. A typical microcrack (arrows) as seen in a 100- μ m-thick histologic cross section of human cortical bone. The specimen was stained in basic fuchsin en bloc to mark the surfaces of cracks existing before the section was cut. Artifactual cracks produced by the sectioning process are not stained (Burr and Stafford 1990). This crack is about 80 μ m long in the plane of the section, and its ends about the cement lines of two secondary osteons. Most microcracks in cortical bone are of this type.

which bone limits and repairs the damage that they produce. We also discuss what seems to happen when the limits of this capacity for damage control are exceeded and the bone fails by what is called a *stress fracture*.

At least some of the damage caused by creep and fatigue can be seen in histologic sections of normal, healthy bones (Fig. 8.1). These *microcracks* can, under continued loading, multiply and grow into “macrocracks,” eventually resulting in complete fracture. Ordinarily, microcracks (or other kinds of microdamage) are removed by remodeling. If fatigue damage in bone accumulates more rapidly than remodeling can remove it, fatigue failure or a “stress fracture” can occur. These fractures can be partial or complete, and tend to occur in soldiers and joggers, as well as ballet dancers and other elite athletes, including racehorses. In principle, inhibition of normal bone remodeling could cause an accumulation of unrepaired fatigue damage that, over time, could lead to fracture.

It is important to understand, however, that the prevention of fatigue failure is not entirely dependent on repair by remodeling. The growth of microdamage is also controlled by the lamellar structure of bone. In ways completely analogous to the crack deflection and blunting effects achieved in man-made composite materials, bone’s internal structure limits the development of fatigue damage. Much of the information contained in this chapter has been extrapolated from the extensive experimental data on the fatigue

behavior of nonbiologic composite laminates (e.g., graphite-reinforced epoxies, or polyester laminates). Knowledge of their behavior allows us to understand how various features of lamellar bone structure function to improve its fatigue properties. Before considering such crack-resistant materials, however, we should look at what is known about cracking in simpler materials.

8.2 Basic Fracture Mechanics

Fracture mechanics considers the conditions under which an existing flaw or crack in a load-bearing structure grows and causes catastrophic failure. In conventional engineering applications, the flaws of concern may be produced as the structure is manufactured or by overloading or repetitive loading of the structure thereafter. In the case of skeletal structures, “manufacturing flaws” are usually (i.e., in healthy people) of less concern than those arising from heavy usage. However, genetic errors can result in abnormal bone tissue (usually because of a collagen defect) that becomes damaged under normal loads. Sometimes damage is introduced by surgeons as they try to correct other difficulties. Whatever the cause of the cracking, bones have an exceptional ability to control it. The purpose of this section is to provide a brief summary of fracture mechanics to help us understand how cracks are propagated generally and how they are controlled in bone.

8.2.1 Linear Elastic Fracture Mechanics

Perhaps the simplest way to approach this topic is to consider it in the context of stress concentrations. Consider the stress concentration on the edge of an elliptical hole in an infinite plate (Fig. 8.2a). The plate bears a tensile stress, s , normal to the long axis of the ellipse. The stress at each end of this axis is

$$\sigma_A = s(1 + 2a/b) \quad (8.1)$$

where b and a are the minor and major radii of the ellipse. (For a circular hole, this reduces to 3 s .) However, this equation loses its validity if $a \gg b$ and the ellipse becomes crack-like (Fig. 8.2b). In that case, the stresses at a point given by polar coordinates (r , θ) relative to the end of the b -axis are

$$\sigma_x = s(a/2r)^{1/2} \cos(\theta) / [2(1 - \sin(\theta)/2 \sin(3\theta/2))] \quad (8.2a)$$

$$\sigma_y = s(a/2r)^{1/2} \cos(\theta) / [2(1 - \sin(\theta)/2 \sin(3\theta/2))] \quad (8.2b)$$

$$\sigma_{xy} = s(a/2r)^{1/2} \cos(\theta) / [2 \sin(\theta)/2 \sin(3\theta/2)] \quad (8.2c)$$

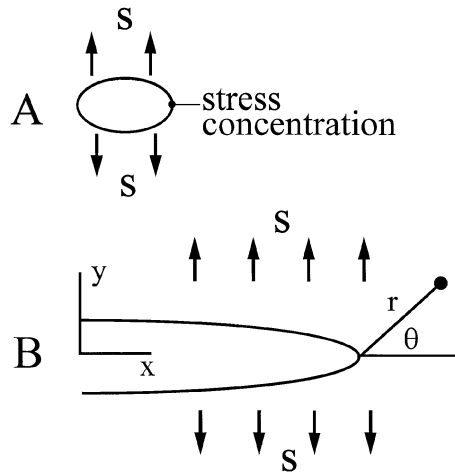


FIGURE 8.2. (a) An elliptical hole in a plate or solid structure bearing a tension stress s has a stress concentration several times s at the point shown. (b) A crack may be thought of as a very eccentric elliptical hole. The stress at a point near the crack tip (dot) is a function of its distance r and angular position θ .

The stress in the y direction along the $\theta = 0$ line is

$$\sigma_y = s(a/2r)^{1/2} \quad (8.3)$$

As one approaches the crack tip, r becomes very small relative to the crack length, a , and the stress concentration very large. At the crack tip itself, $r = 0$ and σ_y is singular (i.e., it is infinite). To avoid this difficulty, one may move the $2r$ term to the other side of the equation, obtaining $\sigma_y(2r)^{1/2}$, which is not singular at the crack tip but is still proportional to the “intensity” of the stress at the crack tip. One can then define:

$$\text{Stress Intensity} = \sigma_y(2r)^{1/2} = sa^{1/2} \quad (8.4)$$

That is, the stress intensity in the vicinity of the crack tip is proportional to the nominal stress, s , and the square root of the crack length, a . This stress intensity is what drives the crack forward.

Irwin (1957) used linear elasticity theory to obtain equations for the stresses near the tips of cracks with several different geometries and loading situations. He found that in general

$$\sigma = k(2\pi r)^{-1/2} + \text{other terms} \quad (8.5)$$

where k is a coefficient peculiar to the particular problem. As one approaches the crack tip, the “other terms” became less important and can be ignored.

Because of the π in these equations, a formal *stress intensity factor*, K , was defined so as to include π as well:

$$K = Cs(\pi a)^{1/2} \quad (8.6)$$

K has units of $\text{Pa m}^{1/2}$; compare Eq. (8.6) with Eq. (8.4). C is a dimensionless constant depending on the size and shape of the flaw and the size, shape, and mode of loading of the specimen, and s is a pertinent stress component in the specimen beyond the influence of the crack, for example, tensile stress perpendicular to the plane of the crack. K characterizes the state of stress near the crack tip; therefore it is known as a *characterizing parameter*. Its utility stems from the fact that **if K exceeds a critical value, the crack will propagate; otherwise, it will not.** There is no need to make calculations of the local stresses to gain this information.

To reiterate, the fundamental concept in linear elastic fracture mechanics is that K determines whether a crack will propagate or not, and if so, how fast. In practice, K bears a subscript denoting the type of cracking that is being considered.

8.2.2 Kinds of Cracks

Of the various ways in which a specimen may be loaded to extend cracks, three are commonly discussed in fracture mechanics (Fig. 8.3). Mode I cracks are propagated perpendicular to the line of action of a tensile load, as when a wedge splits a log lengthwise. Mode II cracks result when shearing forces are

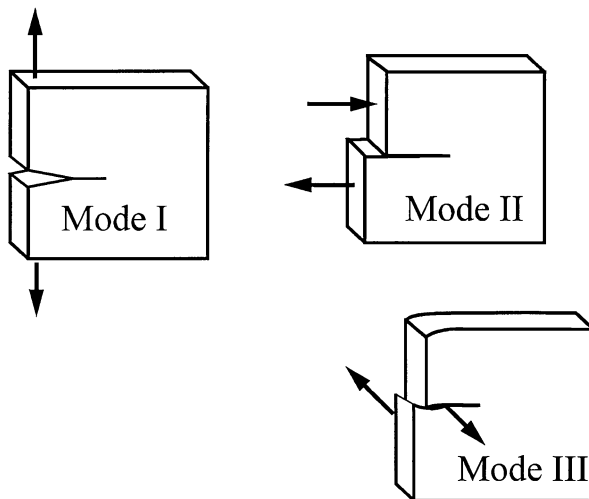


FIGURE 8.3. Modes of cracking.

applied parallel to the plane of the crack and parallel to the direction of its propagation. Mode III cracks are a “tearing” phenomenon; in this case, the shearing is perpendicular to the crack propagation direction. Mode I is the most commonly discussed type of cracking.

8.2.3 Modes of Cracking

As an example, consider the Mode I cracking that occurs when a plate of width b , containing a crack of length a on one edge, is loaded in tension. In this case, the C in Eq. (8.6) turns out to be:

$$C = [(\tan q)/q]^{1/2} \frac{0.752 + 2.02(a/b) + 0.37(1 - \sin q)}{\cos q} \tag{8.7}$$

where $q = \pi a/2b$. As shown in Eq. (8.6), when multiplied by $(\pi a)^{1/2}$ and the tensile stress in the plate, this C gives the stress intensity factor, K_I , for this situation. If this K_I is greater than a critical stress intensity factor, K_{Ic} , the crack propagates rapidly clear through the structure. K_{Ic} is also called the *fracture toughness*; it is a material property, like yield stress. Table 8.1 gives some typical values of K_{Ic} for metals, ceramics, polymers, and cortical bone (Pilkey 1994). By conducting many experiments and tabulating the results, the means for calculating K_I in various situations and comparing them to the critical values have been amassed for many conventional engineering applications. Figure 8.4 gives formulae for C in a few simple cases.

The measurement of K is beyond the scope of this book because it is complicated by the need to ensure that all the assumptions underlying its calculation are fulfilled. The basic process is simple, however, as illustrated in Fig. 8.5 for mode I cracking. A specimen with a standardized crack is manufactured, and the crack is propagated across the specimen by loading it in tension. The applied load and crack opening displacement are measured

TABLE 8.1. Fracture toughness for several materials

Material	K_{Ic} , MPa-m ^{1/2}
2024 Aluminum	20–40
4330V Steel	86–110
Ti-6Al-4V	106–123
Concrete	0.23–1.43
Al ₂ O ₃ ceramic	3–5.3
SiC ceramic	3.4
PMMA polymer	0.8–1.75
Polycarbonate polymer	2.75–3.3
Cortical bone	2.2–6.3

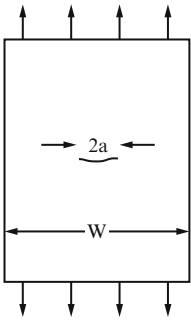
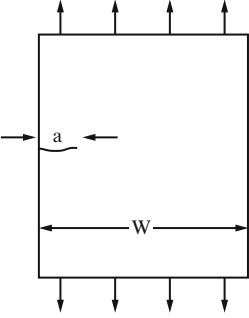
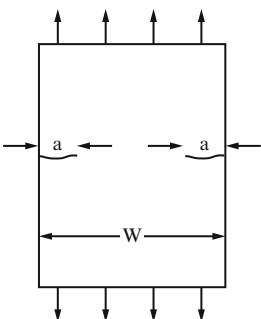
Crack Geometry	C
	Plate of width W in tension, Mode I crack of length $2a$ in center. $C = [1 - 0.1(a/W)^2 + 0.96(a/W)^4] [\sec(\pi a/W)]^{1/2}$
	Plate of width W in tension, Mode I crack of length a on one edge. $C = [2W/\pi a) \tan (\pi a/2W)]^{1/2} \cdot \frac{0.752 + 2.02(a/W) + 0.37[1 - \sin (\pi a/2W)]}{\cos (\pi a/2W)}$
	Plate of width W in tension, Mode I crack of length a on each edge. $C = [1 + 0.122 \cos^4 (\pi a/W)] [(W/\pi a) \tan (\pi a/W)]^{1/2}$

FIGURE 8.4. Equations for the constant C in three simple cases.

and plotted. A failure load is identified on the load-displacement curve using several rules. K_{Ic} is then calculated by multiplying the failure load by a formula derived from the length of the crack and the shape and dimensions of the specimen.

Several investigators have measured K_{Ic} for cortical bone (Table 8.2). Values for bone are an order of magnitude less than those for structural

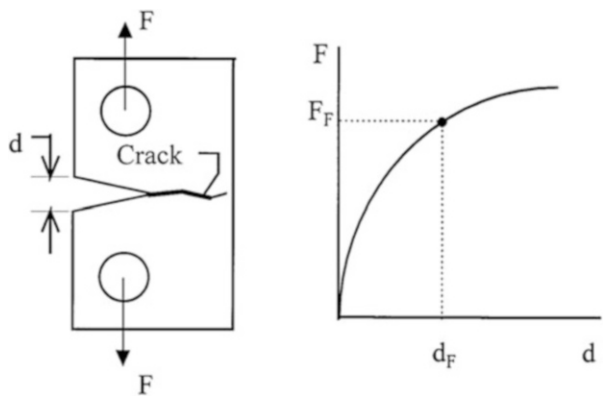


FIGURE 8.5. The basics of measuring K_{Ic} . A specimen with a notch is loaded in tension to drive a crack to the *right* as shown. The tensile force F is plotted against the crack opening displacement, d . A failure point, $P(F_F, d_F)$, on the curve is defined by certain rules, and K_{Ic} is calculated from F_F and d_F using a formula.

TABLE 8.2. Experimental values of K_{Ic} and G_c for bone (Behiri and Bonfield 1984; Norman et al. 1992, 1992; Melvin and Evans 1973; Bonfield and Datta 1976; Alto and Pope 1979; Wang et al. 2002)

Bone type	K_{Ic} , MPa m ^{1/2}	G_c , Jm ⁻²	Source
Mode I, transverse fracture			
Bovine femur	5.49	3100–5500	Melvin and Evans (1973)
Bovine tibia	2.2–4.6	780–1120	Bonfield and Datta (1976)
Equine metacarpus	7.5	2340–2680	Alto and Pope (1979)
Human tibia	2.2–5.7	350–900	Norman et al. (1992)
Human femur (young and old)	5.1 and 4.3	2262 and 1712 (derived from K_{Ic} and E)	Wang et al. (2002)
Mode I, longitudinal fracture			
Bovine femur	3.21	1388–2557	Melvin and Evans (1973)
Bovine tibia	2.8–6.3	630–2238	Behiri and Bonfield (1984)
Human femur	2.2–5.7	350–900	Norman et al. (1992)
Mode II			
Human tibia	2.2–2.7	365	Norman et al. (1992)

metals but similar to a tough ceramic like Al_2O_3 . In making these comparisons, it needs to be emphasized that the values for bone are for crack propagation parallel to the long axis of the bone, which is to say “with the grain.” If one tries to propagate a crack transversely, or perpendicular to the primary and secondary lamellar structure, the crack immediately turns and

goes “with the grain.” Thus, the fracture toughness of bone is decidedly anisotropic, and the values shown are for the direction offering the least resistance to cracking.

8.2.4 Maximum Tolerable Flaw Size

By rearranging Eq. (8.6) and setting the stress intensity factor equal to the critical value, one may solve for the flaw size that will precipitate unstable crack propagation. For human tibial compact bone, Norman et al. (1995) found that $K_{Ic} \approx 4 \text{ MPa m}^{1/2}$. For a semicircular crack of radius a at the surface of a semi-infinite body loaded in tension (e.g., on the tension side of a large bone loaded in bending), the stress intensity factor is Pilkey (1994).

$$K_I = 2(a/\pi)^{1/2}s \left[1.211 - 0.186(\cos \theta)^{1/2} \right] \quad (8.8)$$

Here, θ is as defined in Fig. 8.2 and s is the stress elsewhere in the body, far from the crack. Setting $\theta = 0$, the maximum tolerable crack radius would be

$$a_c = 0.238\pi K_{Ic}^2 / s^2 \quad (8.9)$$

Assuming that $E = 20 \text{ GPa}$ and the strain (ϵ) is a physiologic 2000 micro-strain, one has $s = E\epsilon = (20 \times 10^9) (2 \times 10^{-3}) = 40 \text{ MPa}$ and

$$a_c = 0.238\pi(4^2/40^2) = 0.0075\text{m} = 7.5\text{mm} \quad (8.10)$$

which is a rather large flaw. Remember also that the K_{Ic} used here is probably considerably smaller than the one for a crack propagating **across** a bone.

8.2.5 It Takes Energy to Propagate a Crack

Before Irwin and others developed the concept of a stress intensity factor, Griffith (1920) formulated the problem of crack propagation in brittle solids in terms of energy. He reasoned that energy was required to create the surfaces of the extended crack because all surfaces have “surface energy” associated with them. He conducted many experiments with glass rods to develop his idea into a theory.

Suppose it takes energy or work, dW_c , to propagate a crack for a distance, da . This energy usually comes from the work done in deforming the specimen via the strain energy, U , stored throughout its interior. Conservation of energy requires that

$$dW_c/da = dU/da \quad (8.11)$$

The right side of this equation is called the *strain energy release rate* and is denoted by G , after Griffith, who showed that for mode I propagation of a thin ellipsoidal crack the strain energy release rate is

$$G = dU/da = \pi as^2/E \quad (8.12)$$

where E is the elastic modulus. By setting up an experiment in which a crack can be slowly propagated in a material of modulus E , and its length (a) measured as a function of the applied stress (s), the strain energy release rate can be determined from Eq. (8.12).

The left side of Eq. (8.11) is called the *crack growth resistance*, R :

$$R = dW_c/da \quad (8.13)$$

For a crack to grow, strain energy must be produced and released at least as fast as it is consumed by breaking molecular bonds in the material to extend the crack. For the crack to grow precipitously, G must be greater than a critical value, which is R ; that is, $G \geq R$.

By comparing Eqs. (8.6) and (8.12), it is apparent that G is proportional to K^2 . For mode I cracking in a plane stress situation,

$$G_I = K_I^2/E \quad (8.14)$$

It is important to understand that a propagating crack is “nourished” by the elastic energy of the mass under stress. Finkel (1985) pointed out that although cracks are driven by the strain energy stored in the mass under stress, they are themselves without inertia, and thus capable of turning instantaneously in any direction that promises more elastic energy.

8.2.6 Beyond the Linear Theory: Real Cracks Have Ears

In most materials, and certainly in bone, linear elasticity theory does not hold when stresses become high, and that includes the situation at the tip of a crack. There is a zone of plastic deformation around the crack tip that is generally shaped as depicted in Fig. 8.6. Note the “Mickey Mouse ears” shape of the plastic zone. Finkel has called the plastic zone the “ears” of the crack, and refers to computations showing that these ears may “wiggle” as the crack propagates. The plastic zone at the leading edge of a propagating crack has an ear-like characteristic shape.

For mode I cracking, the distance directly ahead of the crack that is plastically deformed is given by

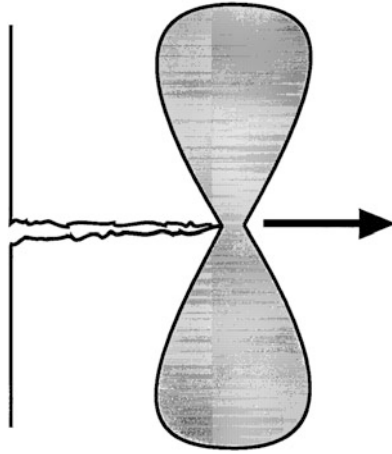


FIGURE 8.6. The plastic zone at the leading edge of a propagating crack has an ear-like characteristic shape.

$$r_p = \frac{1}{k\pi} \left(\frac{K_I}{\sigma_y} \right)^2 \quad (8.15)$$

where σ_y is the yield stress; $k = 2$ for plane stress or 6 for plane strain.¹ One way to accommodate the presence of a plastic zone in linear elastic fracture mechanics theory is to add this r_p (or some portion of it) to the crack length in the equations for K and G . Equivalently, when using the energy approach, another term must be added to Eq. (8.14): the rate at which energy is consumed by plastic deformation per unit length of crack extension.

Box 8.1 Dynamic Fracture Mechanics

Things break differently when they break extremely rapidly. Dynamic fracture mechanics is the study of cracking under these conditions (Freund 1990). When a bone is broken very rapidly in a situation involving a great deal of kinetic energy, the fracture is highly comminuted (i.e., the bone breaks into many pieces). The reasons for this are complex, but stem from two important phenomena. When a crack propagates slowly through a structure, it causes the adjacent material to be unloaded. This load change prevents nucleation of, or arrests, other cracks in the vicinity. When a structure is impulsively loaded, a stress wave passes through it at the speed of sound (about

¹ In a plane strain situation, the strains are all within a plane, i.e., deformations perpendicular to the plane may be assumed zero. Plane stress is similarly defined in terms of stress components perpendicular to the plane. Thanks to Poisson's ratio, the two situations are not the same.

3000 m/s for bone). This results in crack initiation and propagation at many sites simultaneously, before any one crack can arrest others. When these many cracks grow to intersect one another, many fracture fragments are produced. Emergency room physicians see this result in victims of high energy trauma.

The second phenomenon promoting fragmentation at high fracture rates is bone's viscoelastic nature. The more rapidly bone is loaded, the stiffer it is (Fig. 7.13). The velocity of the stress wave increases as the square root of the elastic modulus, so that the stress field more rapidly fills the bone, accentuating the effects just described. On the other hand, bone strength increases with loading rate. This effect is seen in other materials, too. The reasons for it are not clear.

In summary, when a bone is rapidly loaded, it is more fracture resistant, but if it breaks, the fracture is more fragmented. Also, logically enough, the severity of damage to the soft tissues tends to parallel the degree of damage to the bone; thus, the seriousness of severe fractures is compounded by impairment of the associated soft tissues, on which fracture healing depends.

8.2.7 *Beyond the Linear Theory: Real Cracks Are Complicated*

Propagation of a crack in bone is slowed (or accelerated!) by changes in the material ahead of or behind the tip of the propagating crack (Fig. 8.7). Crack toughening mechanisms *ahead* of the crack tip are called “Intrinsic mechanisms” and those that occur behind the crack tip are called “Extrinsic mechanisms” (Ritchie 2011). Intrinsic mechanisms involve changes in the material properties—for example, in a metal the material properties change from elastic to plastic in front of the crack tip. Extrinsic mechanisms act at the *location* of crack propagation (not in front of the crack) and, in bone, are strongly related to bridging of the crack by “ligaments” of unbroken tissue. For relatively large cracks (the kind in a fracture toughness specimen) a great deal of toughening comes from unbroken “ligaments” of bone tissue that bridge across the faces of the propagating crack. The ligaments hold the faces of the crack together creating a “closure stress”. Examination of the crack faces after failure of a piece of bone finds the remains of numerous ligaments that had to be broken in order to finally fracture the tissue.

There continues to be some debate about the role of microcrack formation in the toughening of bone tissue, but it is clear that younger/tougher bone forms much greater numbers of microcracks during failure than does older more brittle bone (Green et al. 2011). This is likely due to a more functional collagen matrix, a very important contributor to bone fracture toughness (Barth et al. 2011; Russell et al. 2012; Nyman et al. 2005; Nalla et al. 2003). The fracture toughness of bone tissue with the collagen matrix removed is extremely low and almost no ligaments are observed in the wake of the crack tip. Ligament bridging between the faces of microcracks by collagen fibers allows the microcrack to open, changing the shape of the surrounding tissue

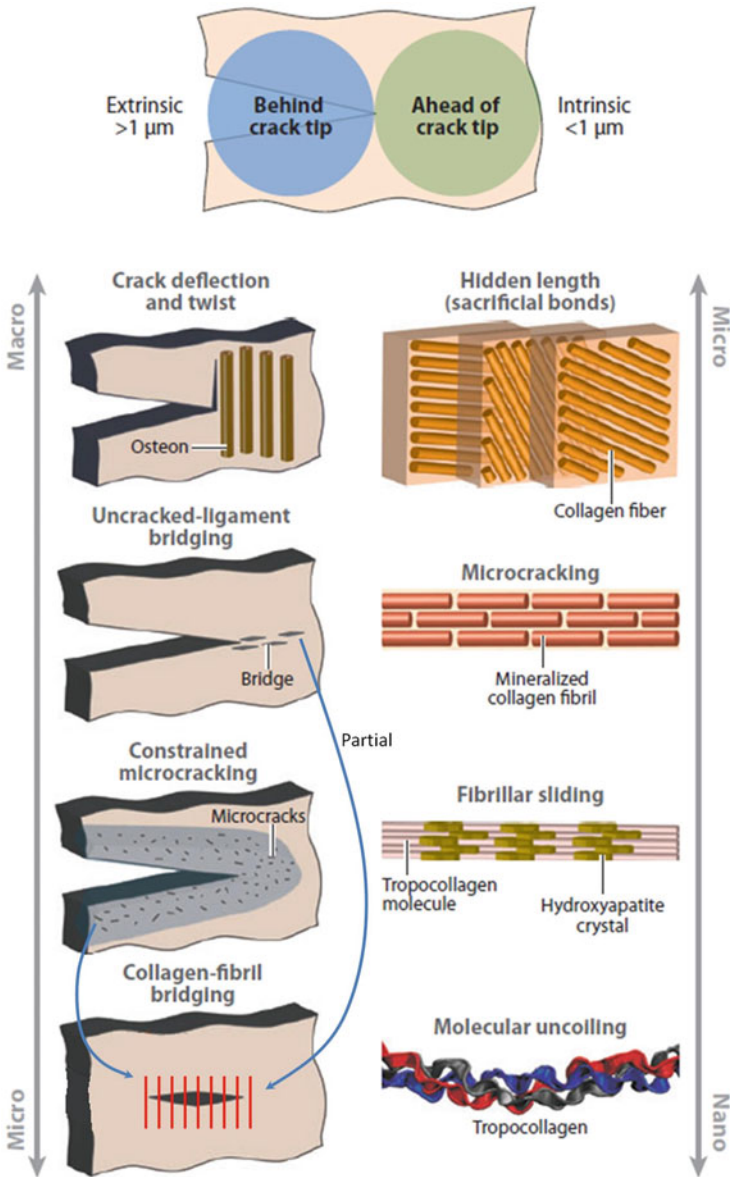


FIGURE 8.7. The extrinsic and intrinsic toughening mechanisms of bone. Crack bridging by uncracked ligaments (large cracks) and by the protein/collagenous bone matrix (small cracks) are very important to bone toughness. Removal of, or damage to, the bone matrix leaves an extremely brittle remnant mineral matrix. The energy absorbing mechanisms of the bone matrix include uncoiling of hidden length, fibrillar sliding and molecular uncoiling, all of which depend on the animal's genetic background, age and matrix cross-linking (Figure modified from Launey et al. 2010).

which consumes energy and thereby blunts propagation. One effect is that arrays of collagen bridged cracks at the micro level provide enough controlled deformation to enable tissue bridging across larger cracks at the macro level. This is an example of a hierarchical dependency in bone tissue where a macroscopic toughening mechanism (i.e. ligament bridging of large macroscopic cracks) depends on a more microscopic toughening mechanism (collagen matrix bridging of microcracks).

As a crack is propagated in a fracture specimen, the stress intensity to propagate the crack could decrease, stay the same or increase. Recall for linear elastic fracture mechanics that: $K = Cs(\pi a)^{1/2}$. Experimentally, the process is to determine “s” and “a” as the crack propagates and then calculate the stress intensity factor “K.” Clearly, the best situation is for the stress intensity factor to *increase* with crack length so that there is a tendency for the crack to stop propagating. Healthy bone has the property that K increases with crack length resulting in what is called a “rising R-curve” where “R” is for “Resistance” (Fig. 8.8). The name “R-curve” is probably the most obscure in all of bone mechanics. Fortunately the concept is simple: An R-curve plots the changing fracture propagation toughness of the tissue as a crack elongates.

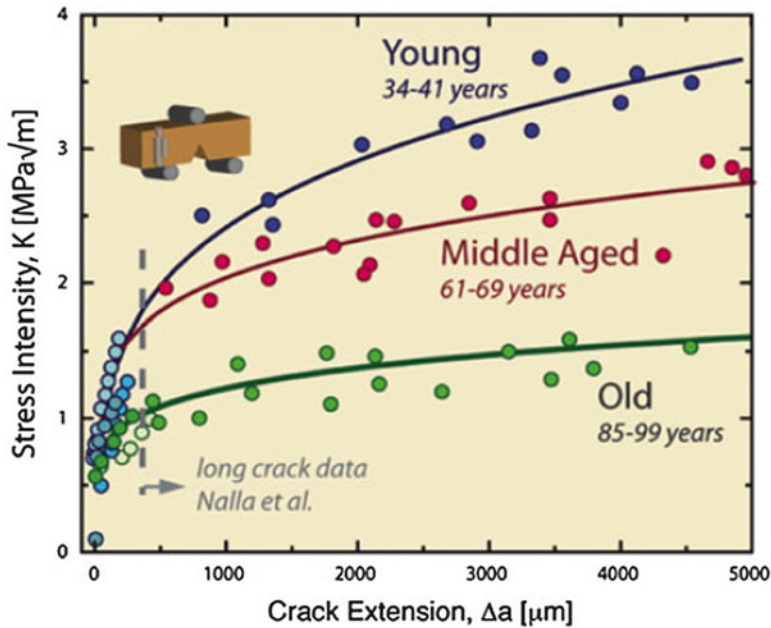


FIGURE 8.8. The increase in bending fracture toughness with crack length is strongly affected by age (Zimmerman et al. 2012).

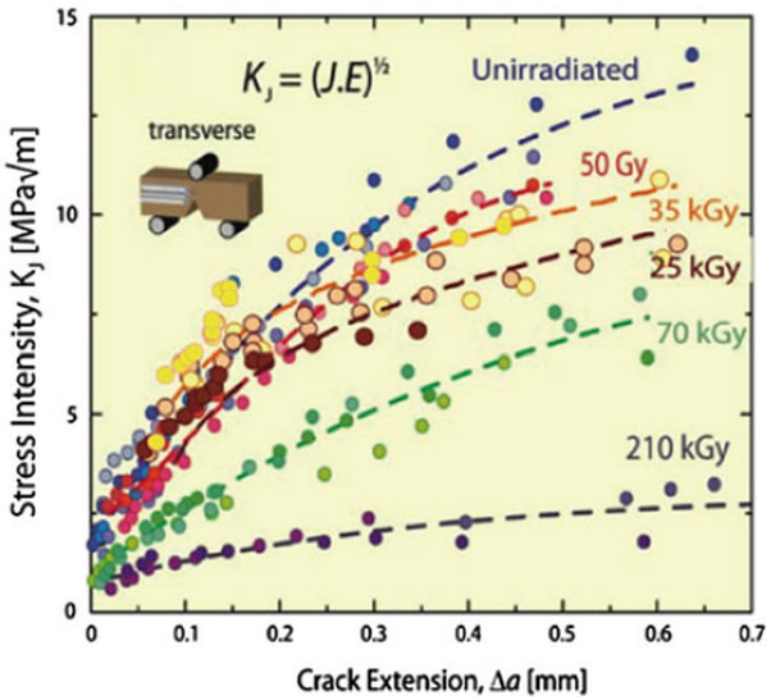


FIGURE 8.9. The increase in bending fracture toughness with crack length is strongly reduced by radiation damage to the collagen matrix (Zimmerman et al. 2012).

Both age and sequential degradation of the bone matrix using radiation correlate with a reduction of the propagation resistance (Fig. 8.9). The radiation experiment strongly shows the importance of the collagen matrix to fracture toughness. The implication of the *age related* decrease in the propagation resistance is that the collagen matrix is degrading with age in humans. There is a great deal of research currently being done to characterize how age and other conditions in humans affect both the bone matrix and fracture toughness. It is clear at the time of this writing that a significant portion of age related bone fragility in humans is caused by changes in the collagen matrix. What remains to be determined is *what* is changing in the matrix and whether it can be prevented or treated in vivo.

8.2.8 Crack Growth and Fatigue

Paris (1964) proposed that the number of cycles required for fatigue failure of a specimen with a crack could be related to the crack's stress intensity factor. His idea was that the crack length increase per load cycle, da/dN , would be

proportional to the range of K values experienced during the loading: $\Delta K = K_{\max} - K_{\min}$. A commonly suggested form for such a relationship is

$$da/dN = \xi(\Delta K)^\zeta \quad (8.16)$$

where ξ and ζ are constants. Unfortunately, the Paris equation has not proven to be a consistently valid relationship. It describes the results at intermediate crack growth rates better than at low or high values of da/dN . Within that range, if crack propagation data are presented as a log-log plot of da/dN vs. ΔK , the slope and intercept should yield the coefficients ζ and ξ respectively.

8.3 Fatigue Behavior of Bone

Depending on the fracture toughness of a material, and the magnitude of the applied loads, cracks start and grow more or less rapidly under repetitive loading. In this section we examine the basic fatigue and fatigue damage behavior of bone.

8.3.1 The S-N Curve

The basic means of presenting fatigue data is an S-N curve: a plot of the applied stress or strain vs. the number of cycles required for failure at that load level. A fatigue S-N curve was first measured for bone by King and Evans (1967). Subsequently, in the 1970s and 1980s, several other investigators (notably Carter and co-workers) reported similar experiments, establishing that bone behaves similar to engineering materials in fatigue. That is, its elastic modulus declines as fatigue progresses, the log of N is linearly proportional to the log of S , and cracks can be seen within fatigued bone which is apparently fatigue damage. The concept that the fatigue behavior of bone is similar to that of composite materials was also established in this period.

Fatigue life, which we now represent by N_F , is related to the applied stress in a very nonlinear way. When one regresses $\log N_F$ vs. $\log S$ for bone specimens fatigued in various modes of loading, one finds that

$$N_F = c/S^q \quad (8.17)$$

where c is a coefficient and the exponent q is a number between 5 and 15, depending on the type of bone and the mode of loading. Clearly, small changes in load magnitude (S), or the kind of bone and mode of loading (q), can result in very large changes in fatigue life. Caler and Carter (1989) and Pattin et al. (1996) have shown that the fatigue life of human femoral cortex is longer in compression than in tension. Testing in uniaxial tensile and

compressive fatigue at 2 Hz, Pattin et al. produced data that, when fitted to Eq. (8.17), had the following coefficients:

$$N_F = 1.445 \times 10^{53} / S^{14.1} \text{ (tension)} \quad (8.18)$$

$$N_F = 9.333 \times 10^{40} / S^{10.3} \text{ (compression)} \quad (8.19)$$

where the stress range, S , is from zero to a peak stress normalized by the initial elastic modulus. (Therefore, S is also the initial strain, and has units of microstrain.) These data were obtained at superphysiologic strains (2600–6600 $\mu\epsilon$). If one may extrapolate the results to a physiologic peak strain of 2000 $\mu\epsilon$, the fatigue life would be 4.1 million cycles in tension and 9.3 million cycles in compression.

8.3.2 Fatigue Damage in Bone

Although we typically think of microdamage in bone as linear cracks, damage caused by cyclic loading of bone actually comes in different flavors. Three distinct forms of damage can be identified in bone, distinguished by their morphology, their mechanical and physiological roles, and by the manner in which they are repaired. *Linear microcracks* (Fig. 8.10a) are typically about 40–100 μm “long” in cross-section, but may run for hundreds of microns or more than a millimeter longitudinally in bone (Burr and Martin 1993; Taylor and Lee 1998). They are less than 2 μm “wide” in cross-section and are typically repaired by normal bone remodeling processes in which the crack is resorbed by osteoclasts, and new bone is laid down in its place. If many linear microcracks accumulate, either because of high and repeated stresses or because remodeling is suppressed through disease or treatments for disease (e.g. bisphosphonates used to mitigate osteoporosis), they can degrade the elastic modulus of the bone to the extent that a stress fracture (as in athletes and soldiers), or a complete fracture (as in osteoporotic people who sustain atypical femoral fractures) can occur. A second form of microdamage, *diffuse damage* (Fig. 8.10b), is composed of a collection of very small cracks, most less than 10 μm long in cross-section. The cracks themselves can be difficult to detect with standard light microscopy, even on stained sections, but can be identified as individual cracks using confocal microscopy (Fazzalari et al. 1998; Vashishth et al. 2000). These cracks probably do not cause a significant degradation of the mechanical properties of bone, but, like linear microcracks, may be important in stress relief when bone is repeatedly loaded. They do not create the signals necessary to stimulate whole bone remodeling (Herman et al. 2010), so why do they not accumulate? Recent observational and experimental evidence suggests that they may heal through an intrinsic matrix repair process in which the defect is eliminated by filling with mineral (Boyde 2003), or in which non-collagenous proteins (perhaps

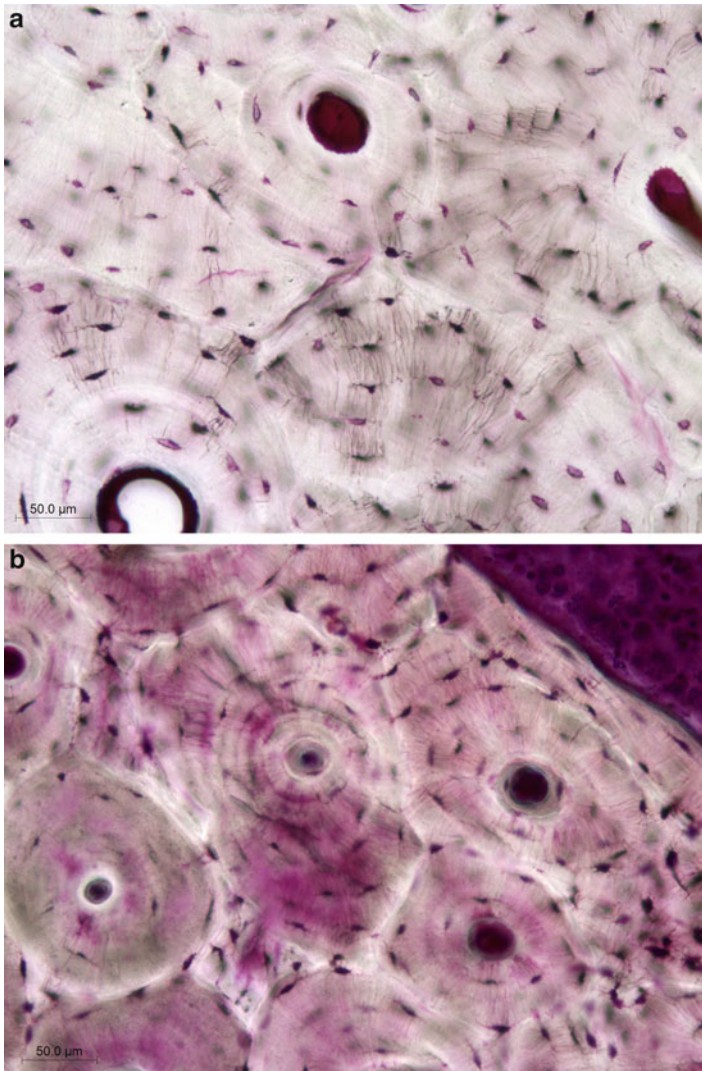


FIGURE 8.10. Microdamage (indicated by *arrows*) typical of (a) compressive and (b) tensile regions in a beam loaded in four-point bending. Human femoral cortex, 100- μ m-thick section stained with basic fuchsin. Field width, $\sim$ 500 μ m (Griffin et al. 1997).

osteopontin) act like a glue to bring the surfaces of bone back into contact (Seref-Ferlengez et al. 2012). *Microfractures* are a third type of microdamage in bone. These are not cracks at all, but full fractures of trabecular bone, which heal through callus formation and subsequent callus remodeling, just as a full fracture of a long bone would heal. Microfractures have a deleterious effect on the mechanical properties of whole bone as they disrupt the connectivity of the trabecular lattice.

Thus, fatigue damage is not a narrowly defined quantity in either bone or other materials. In the engineering literature, it has usually been defined theoretically as a parameter that varies from zero in virgin structures to unity in failed structures, based on changes in mechanical properties, but without any specific reference to cracks or other evidence of physical damage, which may be difficult to detect. For example, damage is often defined as the fractional reduction in elastic modulus. Alternatively, when cracks are apparent, as in certain kinds of composite materials, damage may be defined in terms of the numbers or lengths of microscopic cracks that appear after loading.

Cracks are relatively easy to find in normal bone. Frost (1960) first described such damage, and linked it to fatigue. Burr and Stafford (1990) counted 0.14 such cracks/mm² in the ribs of a 60-year-old man. The number of such cracks increases with age in the human femoral diaphysis and neck cortex (Schaffler et al. 1995), and in the trabecular bone of the femoral head (Mori et al. 1997). Microcracks accumulate exponentially after age 40 (Schaffler et al. 1995). Although the data suggest increased skeletal fragility caused by microcrack damage in aging women, they do not show why some older women fracture and others do not. Mori et al. (1997) found that older women have significantly more microcracks than younger women, but older women who fracture did not have more microcracks than older women who did not fracture.

Microdamage burdens may be higher in trabecular than cortical bone; Wenzel et al. (1996) found approximately 5 microcracks/mm² in trabecular bone from the spine, without a significant increase with age. In the spine, disk pathology may contribute to microdamage (Hasegawa et al. 1995), either because loads are improperly attenuated or because stress concentrations are created in the vertebral body.

If linear microcracks and diffuse damage both perform the function of stress relief, then why are two different kinds of damage created? This probably has to do with the nature of the stresses, and with the quality of the bone matrix. Linear microcracks are preferentially found in areas of higher compressive stress (Diab and Vashishth 2005), are associated with greater matrix deformation before the cracking occurs (Diab et al. 2006) and may be associated with shear strains/stresses in the matrix caused by the compression (Yeni et al. 2003). About 80–90 % of all linear microcracks in cortical bone are found in the interstitial matrix between osteons where mineralization is high (Schaffler et al. 1995; Norman and Wang 1997; Wasserman et al. 2005). Also, because bone from older individuals tends to be more highly cross-linked and may be more fully mineralized, linear microcracks are found more commonly in older people than younger people, and tend to be on average three times longer (136 μm in older people vs. 63 μm in younger folks) (Diab et al. 2006). This reduces the fatigue life

of bone by about 75 % between the ages of 40 and 80 years old. Diffuse damage, on the other hand, is more likely to be found in areas of high tensile stress, and more often seen than linear microcracks in bone from younger people (Diab and Vashishth 2005; Diab et al. 2006). Although stress is dissipated by linear microcracks that are attracted to and trapped by lamellar interfaces, the stress reduction associated with larger areas of smaller cracks initially increases the toughness of bone (Parsamian and Norman 2001), and is likely to be more effective in resisting a catastrophic fracture (Green et al. 2011). When linear cracks form, on the contrary, residual toughness declines. Although some have suggested that linear cracks are simply the natural sequelae to diffuse damage (Parsamian and Norman 2001; Vashishth et al. 1997), the very specific age- location- and stress-associated patterns of these two types of damage (Boyce et al. 1998) make it more likely that they are very distinct forms of damage that are not related by any temporal sequence.

8.4 Creep Behavior of Bone and Its Relationship to Fatigue

As we noted in Chap. 7, bone is viscoelastic. It therefore creeps, that is, it deforms for a prolonged period of time under a constant load. If a bone specimen is loaded to a constant stress well below the ultimate stress, it gradually deforms and eventually fractures. Figure 8.11a shows this behavior for cortical bone from a human femur. The initial strain is immediate and therefore elastic; then, as time goes by, the material deforms nonelastically as the result of damage processes. The strain-time curve is similar to the damage-cycle number curve for fatigue. After a period of rapid creep there is a plateau, and the rate of damage then increases again just before failure. If groups of specimens are loaded to various stress levels, σ , and the time to failure, t_F , is recorded, the log-log graph of σ vs. t_F is linear (Fig. 8.11b) and described by an equation similar to Eq. (8.17):

$$t_F = A/\sigma^B \quad (8.20)$$

where A and B are coefficients peculiar to the kind of bone and the mode of loading. Caler and Carter (1989) found that the exponent B was about 17 and similar for tensile and compressive creep. On the other hand, the coefficient A was greater for compression than for tension.

If the creep damage (D_c) occurs at a constant rate, its formation rate would be

$$dD_c/dt = \sigma^B/A \quad (8.21)$$

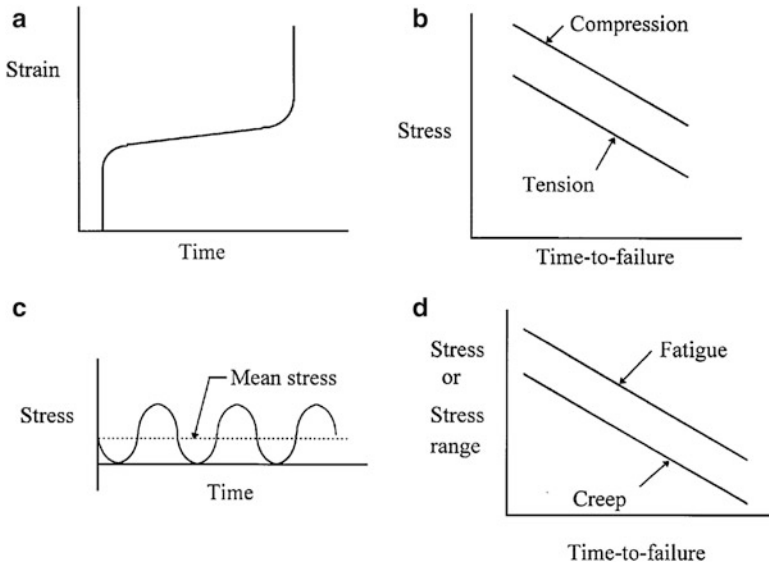


FIGURE 8.11. (a) Typical plot of strain vs. time for a cortical bone creep test to failure. The initial elastic response is followed by relatively constant strain during slow damage accumulation, eventually leading to rapid damage accumulation and failure. (b) Log-log plot of applied stress vs. creep time-to-failure is linear with different behaviors for compression and tension. (c) Cyclic loading from zero to a tensile or compressive stress produces a nonzero mean stress that superimposes creep effects on fatigue behavior. (d) Comparison of creep and fatigue responses for human femoral bone.

If the stress varies with time, the creep damage at any time, t , is

$$D_c = \int_0^t [\sigma^B/A] dt \quad (8.22)$$

In this formulation, failure occurs when $D_c = 1$, and the time required can be obtained by solving Eq. (8.22) for this condition. Carter and Caler pointed out that creep and fatigue often occur simultaneously. If a specimen is loaded in tensile fatigue, for example, the stress varies from zero to a peak tensile value. The loading is cyclic and therefore it is “fatigue loading,” but it also has a creep component because there is an average load value that is not zero (Fig. 8.11c). The creep component is only zero if the fatigue loading is “fully reversed,” that is, it cycles between peak tensile and compressive stresses that are equal in magnitude.

If fatigue loading occurs at a constant stress range, S , then the fatigue damage may be written as

$$D_f(t) = R_L t S^\beta / \alpha \quad (8.23)$$

where R_L is the loading rate (number of load cycles/day) and α and β are coefficients. If fracture occurs from fatigue damage alone, then the time to fracture is obtained by setting $D_F = 1$ and solving for $t = t_F$. When damage is accumulating in response to simultaneous creep and fatigue effects, it might be possible to formulate the result as a simple summation:

$$D_{total} = D_c + D_f \quad (8.24)$$

with $D_{total} = 1$ at failure. For cyclic loading with a non-zero mean stress, creep and fatigue damage combine to produce failure after N_F cycles and at time t_F , and

$$D_{total} = \int_0^{t_F} [\sigma(t)^B / A] dt + R_L t_F S^\beta / \alpha = 1 \quad (8.25)$$

For a zero-tension fatigue test in which stress is a haversine function of time and the peak stress is S , the stress history may be written as

$$\sigma(t) = (1 + \sin 2\pi R_L t) S / 2 \quad (8.26)$$

Substituting this function into Eq. (8.22) gives, at failure,

$$1 = (1/A)(S/2)^B \int_0^{t_F} (1 + \sin 2\pi R_L t)^B dt \quad (8.27)$$

which with two changes of variables can be solved to give

$$t_F = KA/S^B \quad (8.28)$$

for the time to failure in the zero-tension fatigue test. Here,

$$K = \frac{2^B}{\int_0^1 (1 + \sin 2\pi(t/t_F))^B d(t/t_F)} \quad (8.29)$$

Now compare Eq. (8.20), for time-to-failure in a creep test, to Eq. (8.28), for time-to-failure in a zero-tension fatigue test. The only difference is the coefficient K in Eq. (8.28). If we rewrite these equations in log-log form so as to place stress on the y -axis and time on the x -axis, as is customary, we have

$$\log S = B \log A - B \log t_F \quad (\text{creep}) \quad (8.30)$$

And

$$\log S = B \log KA - B \log t_F \quad (\text{zero-tension fatigue}) \quad (8.31)$$

The slopes of the two equations are the same but their intercepts are different. Figure 8.11d shows the results of creep and zero-tension fatigue experiments on human femoral bone performed by Carter and Caler (1983). Note that the fatigue data fall along a line with the same slope as the creep data, but displaced up and to the right. The specimens loaded in fatigue required longer to fail at a given stress level because the stress noted on the y -axis was only applied part of the time; in the case of the creep specimens, it was applied continuously.

In later papers, Carter and Caler (1989, 1985) showed that when fatigue stresses are high, so that N_F and t_F are low, creep damage dominates, and vice versa. This is significant because most fatigue experiments are done with superphysiologic stresses to keep the time necessary to bring the specimen to failure reasonable. See Exercise 8.3. Damage in these experimental specimens is mostly caused by creep. Conversely, bones are fatigued in vivo at relatively low stresses, so N_F and t_F are large; damage then is mostly caused by fatigue. Insofar as the mechanisms of damage caused by creep and by fatigue are different, the results of laboratory experiments cannot be extrapolated to the in vivo situation. On the other hand, differences between creep and fatigue damage have not been identified. These are important caveats to bear in mind as you read the remainder of this chapter and work through the exercises at the end.

8.5 Fatigue Behavior of Fiber-Reinforced Composite Laminates

Engineers have learned to manufacture materials with exceptional creep and fatigue resistance. These materials limit damage through the presence of lamellae or fibers in their structure. Because bone is a lamellar structure in which both ultrastructural (i.e., collagen fibrils) and micro-structural (i.e., osteons) fibers figure prominently, we can gain some insight about bone's damage resistance by examining the behaviors of these man-made materials.

8.5.1 The Birth and Growth of Cracks

The resistance of any material to fatigue failure is a function of its resistance to both the *initiation* and *propagation* of cracks. These two factors are in turn

highly dependent on the microstructural characteristics of the material. Materials that resist the initiation of microcracks often cannot effectively resist their propagation; on the other hand, those materials in which cracks are easily started are often difficult to propagate a crack through. Because the fatigue life of a material is a function of both crack nucleation and growth, a material in which cracks are easily started, but find it difficult to grow, can often provide greater resistance to fatigue failure than one in which cracks are difficult to initiate. In fact, many highly fatigue-resistant materials derive this property primarily from their resistance to crack growth rather than initiation. Bone is such a material.

8.5.2 Material Strength and Fiber Diameter

There are two primary microstructural factors that affect crack initiation and growth: *material strength* and *fiber diameter*. In general, decreased yield strength *increases* the potential for crack initiation, but it also *increases* the resistance to crack growth (Fig. 8.12). This is because plastic flow at the leading edge of the advancing crack blunts the tip, reducing the stress concentration and the strain energy to be found at this critical location. As discussed more fully later in this chapter, crack initiation and growth also depend on the size and number of fibers running through the material. (In cortical bone, osteons serve as large fibers and collagen fibrils as small fibers.) The interface between a fiber and the surrounding matrix is a source of weakness that can initiate cracks. On the other hand, the cracks tend to remain small and follow the fibers instead of propagating across the structure. Therefore, it is important that the fibers be oriented so as to lead the cracks in harmless directions; for example, along rather than across a beam.

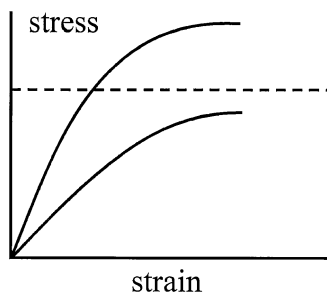


FIGURE 8.12. Stress-strain curves for materials having high and low yield stress. If the latter material's yield strength is less than the stress required to propagate a crack (*dashed line*), it may suffer from plastic deformation, but cracks cannot grow because the stresses never become high enough.

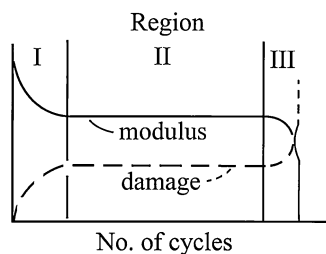


FIGURE 8.13. The road to fatigue failure for a typical composite material. In *Region I*, elastic modulus declines as damage begins under the initial loading. In *Region II*, the rates of change of damage and modulus are much less. In *Region III*, damage builds up rapidly to failure, and the modulus plummets.

8.5.3 The Road to Failure

The ability of fibrous, lamellar composites to keep cracks small and running in harmless directions gives them a characteristic, three-phase damage accumulation history as they fail. Consequently, the long-term fatigue behavior of fibrous composite laminates is often divided into three regions, as shown in Fig. 8.13 (Reifsnider 1990). We may think of the path to failure as a road running through these three regions.

8.5.3.1 Region I: Initiation

The first stage of fatigue is characterized by a relatively rapid but limited loss of stiffness coincident with the initiation of microdamage. The kind and amount of microdamage is determined by the properties of the composite structure. As illustrated in Fig. 8.14, in a cross-ply lamellar structure, cracks begin at structural discontinuities within a lamella and propagate to the interface with an adjacent lamella, where they are apt to stop, particularly if the fibers in the adjacent laminae are aligned in a different direction. These self-limiting cracks redistribute the stresses in the material, reducing the stress in the material alongside the cracks and increasing the stresses near the crack tips, at lamellar interfaces.

8.5.3.2 Region II: Benign Accumulation

The rate of damage accumulation and the stiffness loss quickly stabilize, generally within the first 25 % of the fatigue life, and this marks the transition to a period of benign damage accumulation. Stabilization occurs for two reasons. First, the crack tips are stopped at lamellar interfaces, so their progress across the specimen is halted. Second, as the cracks become more numerous, and the distance between their sides diminishes, the peak stresses between them become so low that additional cracks can no longer be initiated in the intact material. The level of damage at which this stabilization occurs is called the *characteristic damage state*. Because this intralamellar

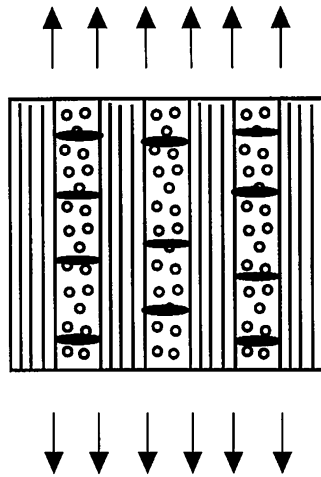


FIGURE 8.14. Typical cracking in a cross-ply laminate loaded in tension. Cracks open in the lamina having fibers perpendicular to the load direction, and propagate to the lamellar interface, where they stop. The material above and below each crack is unloaded because of the crack's free surfaces. When the cracks become so numerous that the peak stress between them is sufficiently low, no more cracks can form; a saturation damage state has been reached.

damage is now shut off, damage accumulation in region II shifts to interlamellar debonding or delamination. These cracks travel parallel to the lamellar interfaces which, if the design is right, are aligned with the principal stress directions. Consequently, these cracks cannot directly cause catastrophic failure and do not even change the elastic modulus very much. The surface area of the resulting secondary cracks is very large, serving as an energy sink and preventing the growth of a single, large, and more dangerous crack. Region I damage tends to be the same regardless of whether the material is loaded cyclically or monotonically; however, the kind of lamellar debonding damage typical of region II is characteristic of cyclic rather than monotonic loading.

8.5.3.3 Region III: Failure

Near the end of fibrous laminate fatigue life, stiffness declines rapidly. Eventually, debonding of lamellar interfaces, combined with matrix degeneration, increases the stress in the load-aligned fibers or lamellae, and they begin to fracture. Although the damage formed in region I is self-limiting because cracks stop at fiber or lamellar interfaces, the damage predominating in region III is self-feeding because as fibers or load-aligned lamellae break, they increase the stress in those elements that remain intact. This process usually occurs rapidly, constituting the final 10 % or so of the total fatigue life, and ends in ultimate failure of the structure.

8.6 Osteonal Bone as a Fibrous Lamellar Composite Material

We saw in Chap. 7 that Haversian bone is mechanically weaker than primary bone in monotonic bending and tension, in part because of its greater porosity, and probably because of the presence of cement lines as well. However, there is evidence to suggest that the unique microstructural arrangement of the various components of Haversian bone is adapted to control crack growth and thereby extend fatigue life. Bone tissue contains collagen fibers organized into lamellae; this is a fibrous laminate on a microscopic scale. In secondary bone, these fibrous lamellae are in turn organized into larger fibers (osteons) running through a matrix (primary bone or osteon fragments). An important hypothesis of bone mechanics is that this structure is responsible for the fact that osteonal bone subjected to cyclic loading behaves much like an engineered fibrous-lamellar composite.

8.6.1 Osteonal Bone's Road to Failure

In the upper part of Fig. 8.15, the tensile elastic modulus of human femoral bone declines over its fatigue life (Pattin et al. 1996). There is an early, rapid diminishment of stiffness, followed by a prolonged, much more gradual loss until, very suddenly, the modulus plummets, as failure occurs. This sort of modulus degradation curve is characteristic of osteonal bone loaded in tension or bending. The lower part of Fig. 8.15 shows that this behavior is substantially different for compression: the modulus falls slowly at first, but at an ever increasing rate until failure. Later in this chapter we attempt to explain this difference based on the kind of damage associated with each mode of loading.

There is evidence that if the imposed loads are physiologic in magnitude, region II of the road to failure may persist for many millions of cycles. Schaffler (1985) tested secondary and primary bovine cortical bone in tension at a strain range of 0–1200 microstrain and a frequency of 4 Hz. In both cases, there was an initial rapid stiffness loss over the first 1–3 million loading cycles (region I), after which no further stiffness loss occurred (region II), up to 45 million loading cycles. Thus, the results looked just like Fig. 8.13, but region III was not reached. It is believed that bone's region I is associated with multiple crack initiation events, just as in other composite materials. Much of this damage is apparently in the interstitial bone outside of the cement lines of *intact* osteons (i.e., those osteons not overlapped by more recent osteons). However, at physiologic strains and strain rates, region III and failure were beyond the capacity of Schaffler's experimental apparatus to keep cycling. Bone has excellent fatigue properties **at physiologic strains and strain rates**.

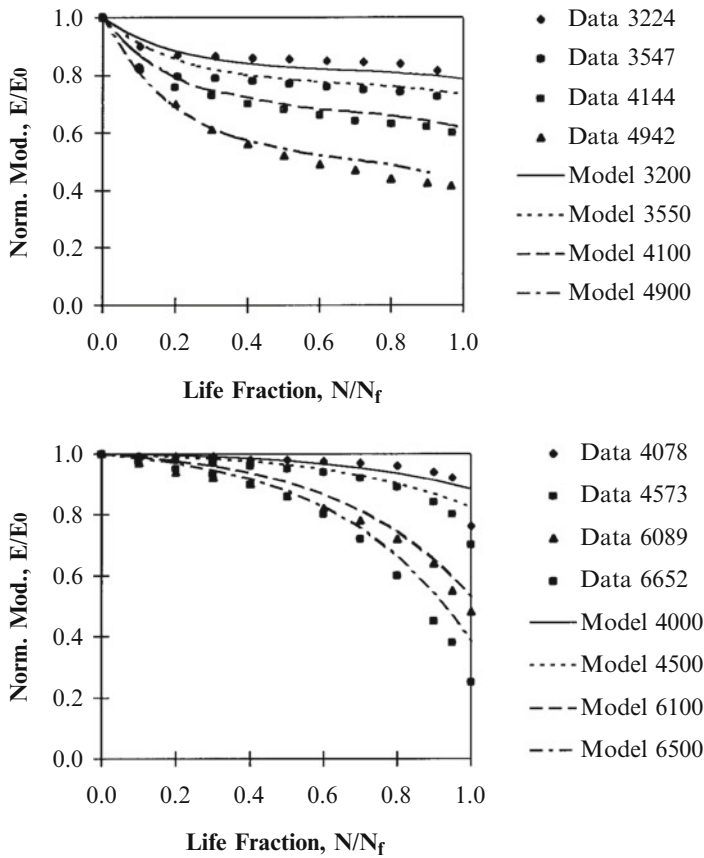


FIGURE 8.15. Degradation of elastic modulus during the fatigue life for human femoral bone specimens loaded in tension (*upper graph*) and compression (*lower graph*). Symbols are experimental data from Pattin et al. (1996); curves are from the theory described in Sect. 10.7. In both cases, numbers shown in the legend represent initial strain values in $\mu\epsilon$ (Pattin et al. 1996).

This fatigue behavior may be contrasted with that seen by Carter et al. (1976) when bovine bone was fatigued at superphysiologic strains and strain rates (Fig. 8.16). Under these conditions remodeled bone had less fatigue resistance than primary bone, and both had fatigue lives of fewer than a million cycles at 3300–5500 microstrain.

8.6.2 Bone Toughness

We have defined toughness in two ways: as the resistance to crack growth and as the amount of energy required to produce complete fracture. Clearly,

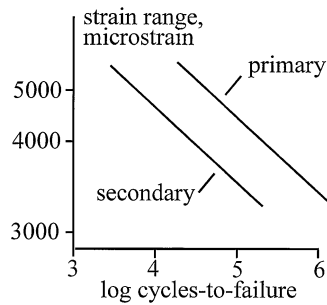


FIGURE 8.16. Effects of remodeling on fatigue life of bovine bone. For a given strain range, primary bone specimens had a longer fatigue life than those composed of secondary bone.

if more effort is required to propagate cracks through a specimen, more energy is required to produce complete fracture, too. Moyle and Bowden (1984) found that bone's work-to-failure was positively correlated ($r^2 = 0.31$) to the fractional area occupied by osteons. Haversian bone is very similar in its osteonal arrangement to very tough composite materials reinforced with discontinuous fibers. Fiber-reinforced composites are often highly anisotropic and have reduced transverse strength because of the directionality of the fibers. Aligned fibers provide high strength and stiffness in the predominant load direction at the expense of properties in the circumferential and radial directions. Secondary osteons, each only 5–10 mm long (Cohen and Harris 1958) and oriented nearly in the direction of principal strain (Lanyon and Bourn 1979), behave like oriented, discontinuous fibers within a matrix of more brittle material. The osteonal “fibers” each have a cylindrically lamellar structure inside and a relatively weak osteon–matrix interface. Failure at these weak interfaces by interlamellar debonding inside the osteon, or delamination at the cement lines, stops or deflects crack propagation. Crack deflection serves the dual purposes of energy absorption and reorientation so that the applied stresses no longer spread apart the crack tip. Both these mechanisms increase fracture toughness while at the same time minimizing damage to vascular and neural structures within the Haversian canal.

8.6.3 *The Conflict between Strength and Toughness in Bone*

Bone has evolved its structure under the demand to be strong and also to be tough in order to reduce the chances of catastrophic failure. Unfortunately, great strength and toughness are generally mutually exclusive (Ritchie 2011)

and, as a result, bone is a “lower-strength” material that also has sufficient toughness for its function. High strength is often associated with high material hardness but hard materials tend to be brittle. For high toughness, a material must be able to deform and reduce the stress concentrations that would cause fracture. The deformation mechanisms adding toughness to bone include the opening of bridged microcracks, sliding of mineralized collagen fibrils, motion along osteonal cement lines, uncoiling of collagen fibrils, breaking of sacrificial non-covalent bonds and others as discussed above in Fig. 8.7. Many of these toughening mechanisms reduce the *strength* of bone, but are essential to its proper function.

The existence of toughening mechanisms at many length scales that reduce strength results in a tissue that is both strong *enough* and also tough. The general motif in bone and some manmade materials that are both strong and tough is the combination of a nano/microstructure of a hard phase to provide stiffness (in bone this is the matrix of mineral crystals) that is incorporated in a strong softer matrix that has (1) mechanisms for alleviating local high stresses and (2) mechanisms such as ligament and collagen bridging to oppose crack opening.

8.6.4 Controlled Crack Propagation Studies

Controlled crack propagation may be accomplished by machining a wedge-shaped specimen (Fig. 8.17a). This specimen is loaded very slowly in three-point bending so that the apex of the central wedge is in tension, and the area under the load-deformation curve is taken as the work-to-failure, or

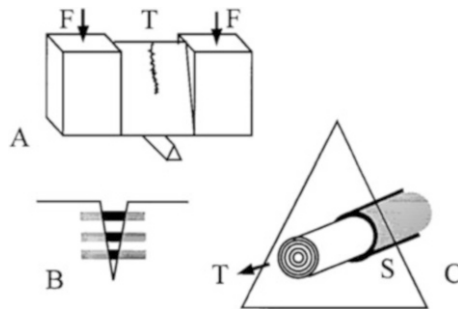


FIGURE 8.17. (a) Specimen geometry for controlled crack propagation studies. The *central region* is *wedge shaped* and loaded in bending. As the crack propagates from the tensile edge (*T*), it encounters more bone, and this helps control the crack velocity. (b) As the crack propagates through the bone, it tends to leave osteons in its path intact, bridging the opening. (c) An osteon bridging the crack experiences a tensile load (*T*) in the gap as well as a shear load (*S*) on the embedded portion.

toughness, of the specimen. The fact that the width of the crack front increases as it moves across the wedge helps keep the crack from accelerating. The first such experiment on bone was reported by Piekarski (1970), using bovine femurs. He showed that the microscopic appearance of the fracture surface depended a great deal on the rate at which the crack propagated across the wedge. Typically, cracks traveled slowly at first, and then suddenly accelerated to produce a catastrophic failure. The initial portion of the fracture surface was always rough and exhibited many pulled-out osteons; the surface produced by the subsequent rapid failure was consistently smooth, with osteons breaking across rather than pulling out. When the crack was advancing slowly it tended to enter and follow cement lines, thereby avoiding Haversian canals. Work-to-failure was very dependent upon the rate of crack propagation. The slow, controlled cracks required almost 60 kJ/m² of crack surface to propel them, while the catastrophic cracks required less than 1 kJ/m².

8.6.5 *Mathematical Analysis of Osteonal Pullout*

Applying techniques used to study reinforced plastics, Pope and Murphy (1974) mathematically analyzed the phenomenon of osteon pullout. Their analysis was directly applicable to the sort of experiment done by Piekarski and raised important questions about the mechanisms by which histologic structure may affect crack propagation. Consider a region of osteonal bone loaded in longitudinal tension as shown in Fig. 8.17b. A crack is propagating transversely through the cortex. As the leading edge of the crack envelops individual osteons, they experience two forces. First, they continue to be stretched by the tensile stress in the cortex. The tensile force from the left side of the crack attempts to pull the osteon out of the bone matrix on the right side of the crack, and vice versa. The tensile pull-out forces are resisted by the shear forces that the surrounding bone exerts on the walls of the embedded portions of the osteon. If the tensile strength of the portion of the osteon bridging the crack is great enough, these shear forces may exceed the shear strength of the cement line and the osteon will begin to pull out of the bone matrix. Subsequently, one would expect the osteon to fail not within the primary, transverse crack, but away from the primary crack, at the stress concentration at the end of the debonded region. If, on the other hand, the osteon is relatively weak in tension, it will break within the primary crack, without pulling out. Let us consider a simplified version of Pope and Murphy's mathematical formulation of this problem.

The tensile force acting on the osteon in the primary crack is

$$T = \pi R_0^2 \sigma_t \quad (8.32)$$

where σ_t is the tensile stress and $On.Rd$ is the radius of the osteon. The shear force on the walls of the embedded portion of the osteon on, say, the right side is given by

$$S = 2\pi(On.Rd)L\sigma_s \tag{8.33}$$

where σ_s is the shear stress in the cement line and L is the distance over which the surrounding bone grips the osteon. The osteon will pull out when the tensile force exceeds the resisting shear force; that is, when $T > S$. By substituting the foregoing equations into this inequality, one finds that pull-out occurs when

$$\sigma_{tf} > 2(L/On.Rd)\sigma_{sf} \tag{8.34}$$

where σ_{sf} is the shear failure stress for the cement line and σ_{tf} is the tensile failure stress of the osteon. Note that it is easier for pullout to occur if $On.Rd$ is relatively large. Conversely, smaller osteons are more likely to break off when forced to bridge a transverse crack.

8.6.6 Comparing Theory with Data

Moyle and co-workers (1978) performed an experiment similar to Piekarski’s on canine femoral bone. They confirmed the correlation between slow crack-ing and osteon pullout, and they also found that the work-to-failure for slow cracking was significantly higher in the lateral aspect of the dog’s femur than the medial side (Table 8.3). Because the bending caused by eccentric loading at the hip joint produces tensile stress on the lateral side of femurs and compression on the medial side, this difference in resistance to tensile cracking may be a manifestation of mechanical adaptation. Like Piekarski, Moyle’s group found that most specimens failed catastrophically, but some cracked slowly all the way across the wedge. These specimens turned out to have significantly *smaller* osteons than did the catastrophic failures (Table 8.4), in direct contradiction to the analytical result obtained above.

Osteon size and control of crack propagation are also important in fatigue resistance. In flexural fatigue experiments, Gibson et al. (1996) found that the dorsal region of the racehorse “cannon” (third metacarpal) bone had a longer fatigue life than the lateral region, which was monotonically stronger and stiffer. Fatigue resistance was strongly associated with signs of osteon pullout

TABLE 8.3. Work-to-failure in different regions of canine femur (Moyle et al. 1978)

Lateral side ($N = 10$)	$11,800 \pm 3000 \text{ J/m}^2$
Medial side ($N = 13$)	$6960 \pm 1350 \text{ J/m}^2$

TABLE 8.4. Effects of osteon diameter and density on crack propagation (Moyle et al. 1978)

Type failure	Mean osteon diameter, μm	Osteon density, mm^{-2}
Slow cracking ($N=23$)	148 ± 22	20 ± 7
Fast cracking ($N=45$)	202 ± 29	14 ± 5

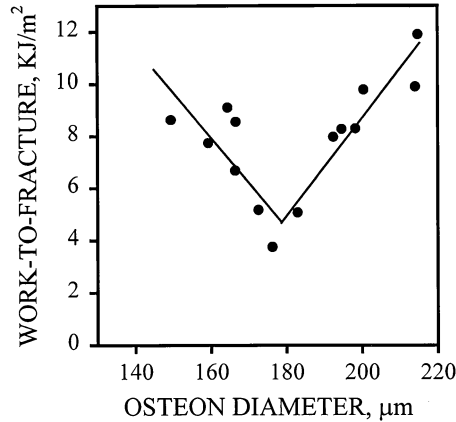


FIGURE 8.18. Relationship between work-to-failure and osteon diameter (Moyle and Bowden 1984).

on the fracture surfaces; there were no signs of pullout in the lateral specimens. The dorsal region had smaller osteons than the lateral region: 156 ± 19 vs. $179 \pm 13 \mu\text{m}$. The reason for this repeated disagreement with theory may be that osteon pullout depends as much on the number of osteons as it does on their diameter. In Chap. 3 it was shown that as secondary osteons accumulate in a section, their population asymptotically approaches a limit that is the reciprocal of their mean individual area. Thus, if the osteons are smaller, more of them can be packed into a section, and there are more to pull out.

Thus, the energy dissipated by osteon pullout should depend on both the likelihood of pullout (which increases with osteon diameter) and the number of osteons available to be pulled out (which decreases with diameter). Figure 8.18 shows data supporting such an argument (Moyle and Bowden 1984; Moyle et al. 1978). The left side of the graph shows that greater energy absorption can be achieved when there are many small osteons in the area of a developing crack. As osteon diameter increases, energy absorption declines. However, as shown on the right side of the graph, work of fracture again increases when osteon diameters are greater than $200 \mu\text{m}$. The reason might be that these larger osteons are more likely to pull out rather than break off, as predicted by Eq. (8.34). Corondan and Haworth (1986) presented

additional data which further demonstrate that fracture is inhibited by increased numbers of osteons as well as by larger osteons. Comparing the morphology of fractured surfaces of bone with that of sections adjacent to the fracture; they found that bone broke where there are fewer *and* smaller osteons. Bone *did not* break in adjacent regions where osteon packing *and* size were greater.

8.6.7 Crack Initiation

Persistent slip bands are “pre-failure” or plastic slip lines representing partial failure within a material, generally along maximum shear planes. These bands may represent the starting point for microcracks within a material. There are features in cyclically loaded bone that look like irreversible microscopic slip bands which occur in metals from plastic deformation (Fig. 8.19). Frost first demonstrated these in 1960. As in metals, slip bands in bone may form a focus for subsequent crack formation.

The relatively weak cement line interface between osteons and interstitial bone matrix may reduce the shear strength of osteonal bone (Saha 1977; Frasca 1981). At the same time, however, this interface may, by slipping, relax shear stresses, reduce strain energy, and slow crack propagation. Lakes and Saha (1979) suggested that cement line displacement may be a manifestation of shear failure in bone. Their observations of displacement at lamellar interfaces in bovine plexiform bone supported the notion that lamellar interfaces in bone reduce stress concentrations by reducing shear stress. In studying the viscoelastic creep behavior of human bone, Park and Lakes

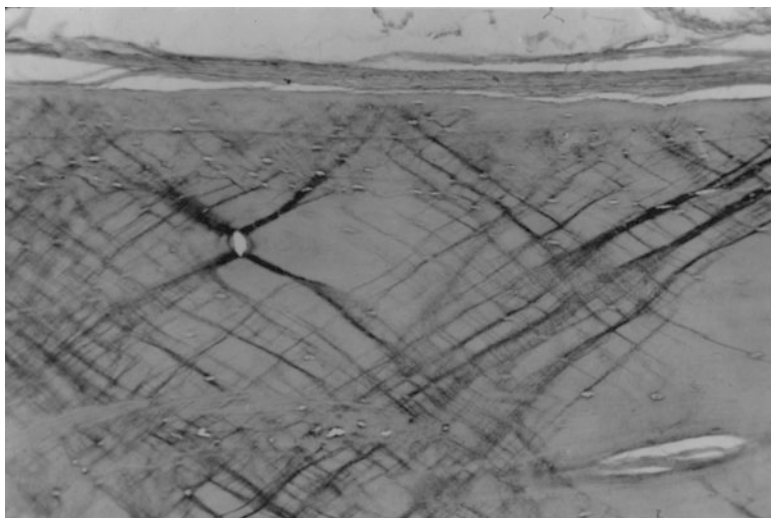


FIGURE 8.19. Slip bands in cortical bone.

(1986) observed displacement at the cement line, but it was one-third of what would be expected if it were the only source of the creep. This observation suggests that the osteonal cement line fails only after other damage (e.g., slip bands) has begun to grow in the matrix.

8.6.8 Crack Stopping

The fatigue life of Haversian bone is governed both by crack initiation and propagation. Because cracks can be removed through bone remodeling, it can be postulated that crack initiation is not so deleterious as the rapid propagation of cracks across the bone material. From this point of view, the most useful microstructural characteristics of bone tissue would be those able to slow, or alter the direction of, crack propagation, rather than those that inhibit crack nucleation. In fact, Haversian bone appears to be constructed in just this manner. The cement lines between osteons and the interstitial matrix are weaker than the bone matrix and when less mineralized are probably more compliant. On reaching a more compliant region, a crack will accelerate and then decelerate, and finally be arrested when it reaches another stiff region (Kendall 1975). Even if crack initiation occurs repeatedly during cyclic loading events, Haversian bone is postulated to stop the transverse growth of these cracks, diverting them along the bone axis and allowing them to expend the strain energy at the crack tip without seriously reducing structural integrity.

8.6.9 Strain Rate

The fracture toughness of bone may be reduced at high strain rates because crack velocity is generally increased under these conditions. When crack velocities exceed approximately 1 mm/s, the likelihood, of osteon pullout is reduced and brittle failure is more common (Behiri and Bonfield 1984). Increased crack velocity requires a higher stiffness ratio between bone matrix and the cement line to halt crack propagation. A change in mechanical properties of the cement line material from viscous to elastic at high strain rates would be expected to compound this problem (Balazs and Gibbs 1970; Gottesman and Hashin 1980). Even within the relatively narrow physiologic range of strain, Schaffler (1985) found that a strain rate of 0.03 s^{-1} caused significantly more microscopic damage and greater loss of stiffness than 0.01 s^{-1} .

Simkin and Robin (1974) found, in static tests at tensile strain rates of only 0.004 s^{-1} , that nearly two-thirds (59 %) of the microcracks were trapped either at cement lines or circumferential lamellae within an osteon; only 41 % cut across these boundaries. At higher strain rates (0.4 s^{-1}), osteons were capable of stopping fewer than 45 % of the microcracks. Maximum strain rates during moderate activity range between 0.0034 and 0.015 s^{-1}

(Lanyon and Baggott 1976; Rubin and Lanyon 1984). Thus, at strain rates experienced by sedentary individuals, osteons might be expected to stop close to two-thirds of the propagating microcracks coming their way. These observations are in concert with those of Norman and Wang (1997), who found that 67 % of all microcracks in cortical bone from human femurs were interstitial rather than intraosteonal, and most interstitial cracks abutted a cement line. The bones of more active individuals would be expected to experience higher strain rates and thus more intraosteonal cracking.

One must be careful to distinguish between strain rate and frequency effects. Ordinarily, strain rate increases as the frequency of loading increases, so that one might expect fatigue life to diminish as frequency increases. However, Caler and Carter (1989) found that the fatigue life of human femoral bone increased rather than decreased as the loading frequency increased from 0.02 to 2 Hz. This change is apparently related to creep effects. Even though the bone is more resistant to crack propagation at the low strain rate associated with a lower frequency, the amount of time that the bone is loaded for each cycle is longer at a low frequency. Thus, the number of cycles required for failure (i.e., the fatigue life) is less at a low frequency. Schaffler (1985) was careful to keep frequency constant (at 4 Hz) in his tests of the effects of strain rate; he did this by using a triangular waveform and varying the “off period” between “spikes” to vary frequency.

8.6.10 *Effect of Remodeling*

The effects of remodeling on fatigue life are unclear. Carter et al. (Carter et al. 1976, 1981; Carter and Hayes 1977a) tested bovine cortical bone in rotating bending fatigue at stress amplitudes of 65–108 MPa. Specimens with the most remodeling were significantly less fatigue resistant than primary bone specimens of equal density. Evans and Riolo (1970) found the opposite effect in human tibial specimens loaded in bending to 34 MPa. The source of this discrepancy could be the species tested, or it could be the different stress amplitudes. Assuming a modulus of 20 GPa, the Carter experiments produced superphysiologic strains (3250–5400 $\mu\epsilon$); the Evans–Riolo experiment did not (1700 $\mu\epsilon$). We have seen that much of the damage to bone at high strain magnitudes is caused by creep rather than fatigue, and osteons may affect fatigue and creep damage mechanisms differently. Strain rate may also have entered the picture: although Evans and Riolo did not specify their frequency, Carter et al. tested at 125 Hz. At the high strain rate associated with such a high frequency (and strain magnitude), the crack-stopping effects of osteons may disappear.

Schaffler's (1985) tensile fatigue tests at physiologic strains and strain rates revealed no differences between primary and secondary bone for either percent stiffness loss or microdamage accumulation. However, the modulus of the primary bone was greater than that of remodeled bone, both initially

and just before failure, and the primary bone's modulus held up better over the fatigue life. These experiments suggest that the remodeling does not substantially reduce the fatigue life of bone *at physiologic strain magnitudes*. There may be a strain rate at which crack resistance and fatigue life are optimal. At very low strain rates, creeplike conditions occur and damage is more severe. At very high strain rates, the effectiveness of cement lines and other tissue interfaces as crack stoppers is diminished. Perhaps the bone structure is adapted to crack stopping best at physiologic strain rates.

8.7 Modeling Fatigue Damage Effects in Osteonal Bone

Krajcinovic et al. (1987) developed a continuum damage theory for osteonal bone, which assumed that when a crack propagates transversely through the cortex, osteons become debonded, break, and pull out as described previously. He calculated the modulus degradation and residual strength of the specimen as this process continued. He was able to predict both bending and tensile elastic moduli and the relationships between elastic modulus and tensile failure properties.

Griffin and co-workers (1997) recently proposed a more detailed model for fatigue of osteonal bone, based on the observation that cortical bone fatigue behaviors are different in tension and compression. Fatigue lives are longer in compression than in tension, and the shapes of the modulus degradation curves are different (Fig. 8.15). In tension there are generally three distinct regimens of behavior with increasing numbers of cycles: the first is an initial reduction in modulus, the second is a plateau, and the third is a rapid modulus reduction before failure. In compression, on the other hand, the modulus diminishes very slowly at first, but at an ever increasing rate until failure.

Carter and Hayes (1977b) also observed differences in the kind of micro-damage associated with fatigue in tension and compression. In tension, most cracks were found in the interstitial matrix or in osteonal cement lines. In compression, the cracks tended to run from one vascular channel to another. Figure 8.20 shows similar behavior in the tensile and compressive regions of human femur specimens fatigued in four-point bending. Is it possible to mechanistically account for the differences in modulus degradation in tension and compression on the basis of the differences in the nature of the damage?

The theory of Griffin et al. (1997) models cortical bone as a fiber-matrix composite, the fibers being osteons and the matrix consisting of primary bone and remodeled osteon fragments. Damage is partitioned into components that independently affect the fibers and the matrix, and the “damaged

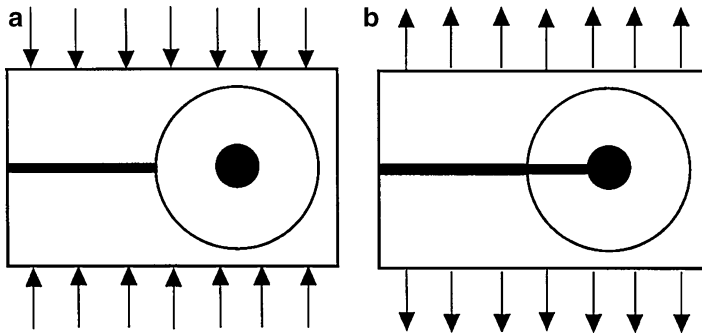


FIGURE 8.20. Hypothetical basis for cracking differences in tensile and compressive fatigue. (a) In tensile loading, Poisson strain is compressive, closing cracks and making them more likely to stop at cement lines. (b) In compressive loading, Poisson strain is tensile, opening cracks and making them less likely to stop at cement lines.

modulus” is obtained from the isostrain rule of mixtures given in Eq. (7.16). Here, we write this equation in the form

$$E = (1 - D_f)E_fV_f + (1 - D_m)E_mV_m \quad (8.35)$$

where E is the elastic modulus of the bone, V is volume fraction, and D represents damage. The subscripts m and f refer to matrix and fibers, respectively. D_f and D_m have values that range from zero, in an undamaged state, to one, in a failed state. When there is no damage, this equation is reduced to Eq. (7.16).

For ceramic matrix composites loaded in tension, the initial damage is caused by microcracks in the matrix. The number of these cracks does not increase indefinitely, but approaches a limit known as the saturation or characteristic damage state. It is postulated that for tensile fatigue of cortical bone, the stiffness loss is caused primarily by cracks that form in the interstitial matrix and are self-limiting in length because they consistently appear as seen in Fig. 8.1 or 8.10a. Therefore, damage to the matrix evolves as

$$dD_m/dN = K_1(D_s - D_m) \quad (8.36)$$

where D_m represents the matrix damage, N is the current cycle number, K_1 is a matrix damage rate parameter, and D_s represents the saturation level of matrix damage, a function of the applied stress.

Most matrix cracks intercept osteon cement lines, and can be expected to produce stress concentrations at the point of impingement. Matrix damage also results in load transfer from the matrix to the osteons. These factors should eventually lead to the tensile failure of osteons. As osteons break, the load is further redistributed to neighboring osteons, producing a positive

feedback and resulting in osteon failure at an ever increasing rate. Therefore, the fiber damage, D_f , is assumed to accumulate as

$$dD_f/dN = K_2D_s + K_3D_m \quad (8.37)$$

where K_2 is an osteonal damage rate coefficient and K_3 is an osteon–matrix damage interaction parameter. Equations (8.36) and (8.37) can be solved simultaneously to obtain

$$D_m = D_s \left(1 - e^{-AN'} \right) \quad (8.38)$$

$$D_f = D_s K_3 \left[K_2 \exp(-K_1 N') + K_1 \exp(K_2 N') - (K_1 + K_2) \right] / (K_1 + K_2) \quad (8.39)$$

where $N' = N/N_F$ is the normalized fatigue life in tension. Substitution of these two equations in Eq. (8.35) provides an expression for the reduction in modulus as a function of the normalized fatigue life.

In compression, microcracks appear less likely to stop at the osteon–matrix interfaces than they do in tensile loading. If so, then in bone that is fatigued in compression cracks may tend to grow in length and number without bound, degrading the material in a positive feedback manner, and damage is assumed to evolve as

$$dD_c/dN = K_4(D_1 + D_c) \quad (8.40)$$

where K_4 is a compression damage rate parameter and D_1 represents incipient damage in the form of stress concentrators, such as Haversian or Volkmann's canals, which initiate or extend D_c . Equation (8.40) is readily solved to obtain

$$D_c = D_1 \left[\exp(K_4 N'') - 1 \right] \quad (8.41)$$

where N'' is the normalized fatigue life in compression. For the compression case, D_c is used for both damage terms in Eq. (8.35) because there do not appear to be distinct damage modes in osteons and interstitial matrix in this mode of loading.

By fitting the resulting modulus degradation equations to Pattin's (Pattin et al. 1996) data, the constants K_1 , K_2 , K_3 , and K_4 were determined. Return to Fig. 8.15 to see examples of how well these results fit data for typical compression and tension specimens. These results support the hypothesis that the different shapes of these curves are caused by different damage modes in tension and compression. In particular, it appears that matrix

damage under tensile loading is self-limiting, with cracks being stopped by osteonal cement lines, but this may not be the case in compressive loading.

Griffin et al. (1997) went on to demonstrate that a mathematical model for fatigue of a beam loaded in four-point bending which incorporated the foregoing equations to describe the development of damage on the tensile and compressive sides was able to predict modulus degradation and fatigue life reasonably well.

8.7.1 *Why Would Cracks Be Stopped in Tension But Not in Compression?*

What could account for the observation that osteons stop cracks better in uniaxial tensile fatigue than in compression? Figure 8.21 offers a possible explanation. When a specimen is loaded in uniaxial tension, the Poisson stress acting across the fibers is compressive and tends to close a matrix crack of the kind shown in Fig. 8.1. Thus, it is more difficult to propagate the crack and easier for a cement line to stop it. On the other hand, when uniaxial compression is applied, the Poisson stress is tensile and tends to open the crack. This should make it more difficult for cement lines to stop the crack, consistent with the observation that such cracks travel from Haversian canal to Haversian canal, passing through cement lines.

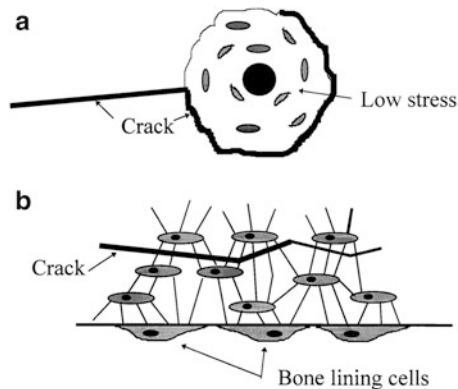


FIGURE 8.21. Alternative hypotheses regarding the mechanism for activation of remodeling by fatigue damage. (a) Debonding crack in cement line isolates osteon from stress, creating local disuse state that produces activation signal. (b) Damage to osteocytic network produces activation signal.

8.7.2 If Cracks Are Self-Limiting in Tensile Fatigue But Not in Compressive Fatigue, Why Is Fatigue Life Longer in Compression Than in Tension?

Although we have argued that the toughness and fatigue resistance of bone are based on resistance to crack propagation rather than crack initiation, perhaps this is an oversimplification. Consider the following possibility. When cortical bone is loaded in uniaxial tension, cracks are easily initiated but difficult to propagate substantial distances. When loaded in uniaxial compression, on the other hand, the opposite situation may obtain: difficult initiation but then easier propagation. This hypothesis is consistent with the possibility that the Poisson effect drives crack propagation in compressive loading. It is also in concert with the modulus degradation curves of Pattin and co-workers. In tensile fatigue, the modulus drops rapidly at first, then plateaus until just before failure. This finding is consistent with a form of damage that is easily initiated but self-limiting, followed eventually by accelerating osteon breakage. In compression, on the other hand, the early, self-limiting form of damage may not occur because crack initiation is more difficult in this mode of loading. However, if continued loading eventually breaks down the tissue's resistance to crack initiation, the cracks may be of a kind that grow relentlessly, and the modulus may then diminish rapidly (Fig. 8.13). To test this hypothesis, more work is needed to elucidate the relative ease of crack initiation and crack propagation in different modes of loading oriented in various ways to the lamellar structure.

8.8 The Role of Fatigue in Activating Bone Remodeling

We have asserted that one of the primary functions of remodeling is to repair fatigue damage. In this section we discuss the evidence for that. To begin, it is useful to compare the fatigue life of bone to the rate at which bone is renewed or turned over by remodeling. We have seen Schaffler's data indicating that bovine cortical bone can endure 30–45 million load cycles at 1200 microstrain without fatigue failure.² Pattin's data for the fatigue life of human femoral bone suggest that the fatigue life at a peak strain range of 2000 $\mu\epsilon$ is 4–10 million cycles. At 1200 $\mu\epsilon$, this might be substantially extended. How long would it take an average person to accumulate 30 million load cycles in, say, the femur? If a person's daily activities were the equivalent of walking 5 miles/day, each femur would receive about $5 \text{ miles} \times 5280 \text{ feet/mile} \div 6 \text{ feet/stride} = 4400 \text{ load cycles/day}$, assuming

² Unpublished data by Schaffler and co-workers extend this to 45 million cycles.

each step covers 3 feet. This works out to be roughly 1.5 million load cycles/year, so that 20 years would be required to accumulate 30 million cycles. The turnover rate for the cortex of a bone like the femur is about 3 % per year in adults (Parfitt 1983), but this rate should be higher in children and the elderly. Assuming an average value of 5 %, 20+ years is just about what would be required to replace all the bone in any given region. Thus, there is a rough correspondence between the fatigue life of cortical bone and the rate at which it is replaced.

8.8.1 *Random or Directed Repair*

Osteonal remodeling could replace microdamage in one of two ways: by directing the BMUs to the location of damage, or by simply randomly remodeling the cortex at a rate designed to keep up with the overall rate of damage accumulation. Random remodeling would not be ideal, however. Some regions can be expected to habitually have higher stresses than others (e.g., near points of muscle attachment). The process might be more efficient if osteonal activation occurred at sites where microdamage has actually developed.

It has been proposed that osteonal cement lines not only stop cracks, but that this event initiates the reparative remodeling process (Martin and Burr 1982). This could happen because debonding of the osteon during crack trapping produces a local disuse state, which activates a new BMU. This new remodeling unit could then resorb the damaged region and replace it with new bone (Fig. 8.21a). Alternatively, Frost (1973) postulated that damage triggers a response through disruption of the canalicular network. There is evidence that stress or strain provoke responses in osteocytes that are transmitted from cell to cell through their processes (Weinbaum et al. 1994; Cowin et al. 1995; Owan et al. 1997). The interruption of normal osteocyte signaling to bone lining cells could activate remodeling if these signals are inhibitory (Fig. 8.21b). These two theories provide very different mechanisms for activation of the repair of micro-damage in bone, and a contrast that should allow for experimental testing of their validity. They are similar, however, in that they both use the same pathway; in one case the signal is not generated because cement line debonding relieves stress, and in the other the signal is blocked by intraosteonal damage.

8.8.2 *The Physiological Basis for Targeted Remodeling*

The idea of targeted remodeling of bone damage has appeared in many guises over the years. Tschantz and Rutishauser (1967) suggested that microcracks in the compressed ulnar cortex of dogs stimulated remodeling. Chamay and

Tschantz (1972) proposed that loading stimulates bone remodeling. Rahn et al. (1972) showed that overload zones adjacent to fractures in the process of repair are also remodeled, resulting in the resorption of damage and the formation of more secondary osteons. Comparative primate morphology studies (Schaffler and Burr 1984) show that osteonal remodeling is more intense when bones are impulsively and repetitively loaded, that is, in very active animals, suggesting that remodeling may be initiated by such loads. Frost's (1960) earlier data on human ribs support the same concept: the rib is loaded many times each day by respiration. Typically, ribs have more microcracks and also a greater rate of remodeling than any other cortical bone site in the body. These observations are consistent with the hypothesis that bone remodeling is activated by microdamage, but they do not implicate microdamage itself as the direct cause of the remodeling process.

Burr and co-workers (Burr et al. 1985; Mori and Burr 1993) performed *in vivo* fatigue experiments to verify a cause-and-effect relationship between microcracks and BMU activation in cortical bone. The results showed that microcracks were associated with bone resorption spaces several times more often than expected by chance alone, even when the magnitude of the strain and the duration of loading were within reasonable physiological limits. They also demonstrated that cracks caused the resorption spaces, and not the other way around. However, the mechanisms behind this association have not been elucidated, and it remains to be shown how long remodeling takes to repair a "batch" of fatigue damage.

Over the past 10–15 years, experimental evidence has shown that microdamage has clear physiologic significance in signaling through osteocytes for targeted bone remodeling (Parfitt 2002; Burr 2002; Martin 2002). Bentolila and co-workers (1997) demonstrated that remodeling can be activated in association with fatigue damage in a bone that ordinarily lacks osteons and therefore has no cement lines. By cyclically loading the ulnas of rats, which do not ordinarily remodel their cortical bone, they were able to demonstrate microcracks in association with BMU resorption cavities. This experiment was important for two reasons. First, it suggested that cement line debonding, as proposed by Martin and Burr (1982), need not play a role in activation of BMUs by fatigue damage because the rat's ulnar cortex initially has no osteons. Instead, it supports Frost's proposal that damage acts directly on osteocytes, and perhaps through the disruption of osteocyte canaliculi. Second, it supports the hypothesis that the control of fatigue damage is an important function of remodeling, because in the normal rat ulna there is neither damage nor remodeling to be seen, but when damage is introduced, remodeling is activated within days. (Remodeling may also be activated in these animals by calcium deficiency, however.)

We now know that osteocyte signaling is indeed responsible for the initiation of the repair process (Schaffler and Kennedy 2012), and that this occurs through alterations in the RANKL-OPG signaling axis. Verborgt

et al. (2000) showed in the rat overload model that osteocyte apoptosis occurred following the induction of microdamage by loading, and that those osteocytes that became apoptotic were those that were closest to the damage. Moreover, osteocyte death occurred prior to the formation of resorption spaces, suggesting that the apoptosis was also temporally associated with the subsequent remodeling response. When osteocyte apoptosis is suppressed, no remodeling occurs even though damage is present (Cardoso et al. 2009). These experiments clearly showed that osteocyte apoptosis is *required* for the targeted repair of damage. At this time, the signals for the response had not been identified. Moreover, it was not clear whether osteocyte death was permissive (i.e., healthy osteocytes produce signals that prevent remodeling, but when the cells die, those signals are not present and allow the remodeling reaction), or whether the process of osteocyte death actively produced signals that enticed cells to begin the remodeling process. This question also now seems to have been resolved. Again using the same model, the RANKL/OPG ratio was increased significantly within three days following the induction of bone damage, and continued to increase over the following week (Kennedy et al. 2012). Apoptotic cells close to the damage produced less OPG, whereas healthy osteocytes away from the damage upregulated their release of RANKL. Thus, in damaged bone, there is more RANKL, which promotes the full differentiation of osteoclasts and is essential for the activation of resorption, and less OPG, which neutralizes the RANKL. Consequently, the initiation of the remodeling sequence and the activation of resorption is stimulated.

Several key experiments (Nakashima et al. 2011; Xiong et al. 2011) have been done to demonstrate that osteocytes are a key source of RANKL. Indeed, these studies showed that regulation of bone remodeling in adults is controlled by the osteocytes, and that the RANKL produced by osteoblasts probably does not contribute to the signaling for bone remodeling at all. There may be a dose-response element to the apoptotic response. Investigators attempting to simulate the effects of microcracks in vitro and studying osteocytic responses to larger cracks demonstrate that large cracks are associated with a greater increase in RANKL and a greater decrease in OPG than smaller cracks (Mulcahy et al. 2011). This would lead to a more robust remodeling and repair response.

All of these experiments provide compelling evidence that osteocytes help to regulate bone remodeling, and that the locations of the osteocytes undergoing apoptosis provide direction signals that permit specific areas of bone, those that have been damaged for instance, to be targeted for remodeling and/or repair (Fig. 8.22). This provides a significant new dimension in our understanding of the role that microdamage plays in normal bone remodeling. It may be very productive to begin to consider skeletal microdamage not only in mechanical terms, but very much in physiological terms.

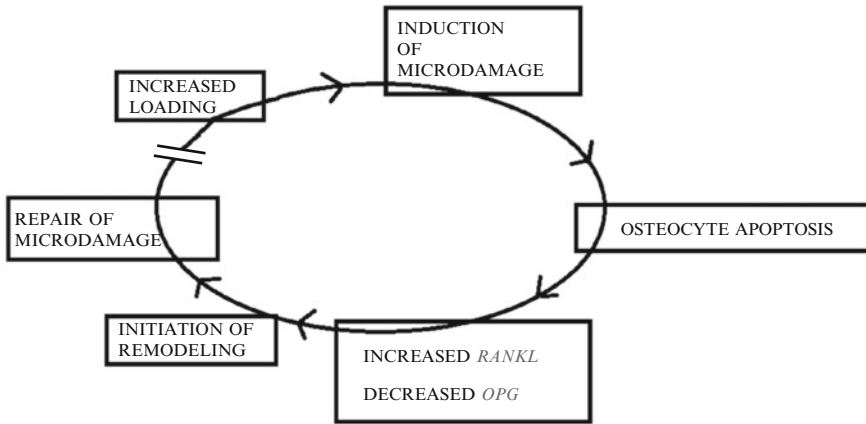


FIGURE 8.22. Steps in the targeted repair of microdamage.

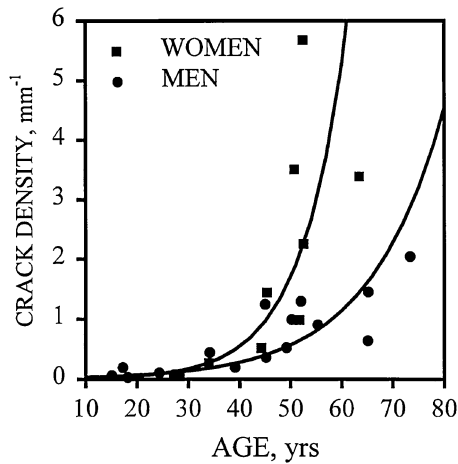


FIGURE 8.23. Microcrack density vs. age for men (*circles*) and women (*squares*) (Schaffler et al. 1995).

8.8.3 Microdamage and Bone Fragility in the Elderly

The idea that fatigue damage may be an important component of bone fragility in the elderly has gained credence among bone researchers (Heaney 1992). This idea is supported by data showing that microdamage increases exponentially with age in human femoral bone (Schaffler et al. 1995; Mori et al. 1997) (Fig. 8.23). This increase could occur because reduced bone mass increases bone strains, and fatigue damage is probably very sensitive to strain

magnitude. It follows from this view that many nontraumatic fractures in the elderly (i.e., osteoporotic fractures) could be essentially stress fractures. This hypothesis is supported by the fact that callus-like woven bone formations have been observed on periosteal surfaces adjacent to femoral neck fractures (Rutishauser and Majno 1950). It is clear that bone does indeed become more fragile with age, even apart from the loss of bone mass itself. The tissue of bone becomes less tough and more brittle with each ensuing decade (Burstein et al. 1976; Currey et al. 1996; Burr and Turner 1999). Bone from older individuals can sustain less plastic deformation, consequently the amount of energy it absorbs in the post-yield region of the stress-strain curve is significantly reduced and fracture occurs more quickly. There may be several reasons underlying this reduced energy absorption capacity beyond the well-documented accumulation of microdamage. These may include alterations in mineralization, changes in collagen cross-linking, increases in bone turnover that create resorption pits that serve as stress concentrations for damage, or some combination of these factors. Whether the reduced toughness that occurs with age really contributes to fractures in older people is still not known.

8.9 Modeling Stress Fractures

Stress fractures are the clinical manifestation of the accumulation of fatigue damage in bone. When cracks are not stopped and repaired, they can grow to become larger cracks visible radiographically (Figs. 8.24a, b). These “stress fractures” represent the imbalance between cyclic loading and bone adaptive responses to the loading that allow the formation and growth of cracks (Burr and Milgrom 2001). Stress fractures in humans occur most

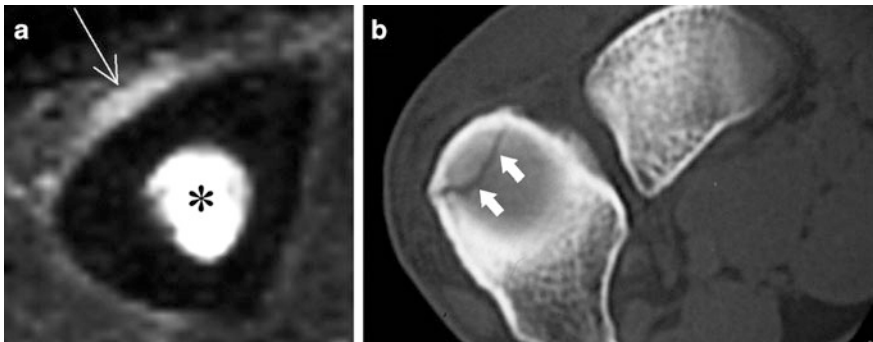


FIGURE 8.24. (a) Periosteal callus in the human tibia is indicative of the presence of a stress fracture, which cannot be seen. (b) Computed tomographic image of a navicular stress fracture. The arrows indicate the fracture (Gaeta et al. 2005; Warden et al. 2009).

commonly in the mid-distal tibia, metatarsals of the foot, the navicular, and occasionally the femur, although the specific incidence and location vary by age and activity (Hulkko and Orava 2001). They are an important medical problem among military recruits; ballet dancers, race horses, and other high performance athletes; and in recreational athletes. Their diagnosis and treatment is complicated by their capricious nature and the difficulty in predicting when bone pain is indicative of an incipient fracture. The earliest sign of a stress fracture is often not the fracture itself, but the bony reaction to it. The earliest radiographic evidence is generally the formation of a small periosteal callus, sometimes also referred to as a periosteal stress reaction, although the latter can also refer to localized soft tissue growth from the bone surface that is similar to the bridging callus formed during the repair of a complete bone fracture. The healing response to stress fracture increases the local metabolism of the tissue. Increased bone metabolism is measured using a radioactive bisphosphonate that collects at sites of bone formation. The localized radiation is detected using scintigraphy or positron emission tomography (PET). Stress fractures do not generally progress to full fracture of the bone if activity levels are reduced, although full fractures can develop in instances of metabolic disease (e.g., hypophosphatasia, which is represented by low levels of alkaline phosphatase that indicates insufficient bone formation) or in some cases in which repair is suppressed pharmacologically (as in atypical femoral fractures in those taking potent remodeling-suppressors for osteoporosis) (Shane et al. 2010).

The development of a stress fracture depends on the remodeling (repair) rate, but also on the material properties of the bone tissue, the magnitude (and perhaps direction) of strain, and the number of loading cycles. In racehorses, in which strains may be as high as 5000 $\mu\epsilon$, stress fractures can occur within 50,000 loading cycles. In military recruits and athletes in which lower strains are prevalent, stress fractures can occur with 100,000 or so cycles of loading. Therefore, it is important to minimize even brief periods of high strain or strain rate. There is some evidence that in athletic and military populations stress fractures are associated with muscular fatigue, as muscles are now known to control strain distributions in bone. When muscles fatigue, strains on bone can increase significantly (Fyhrie et al. 1998; Milgrom et al. 2007), and perhaps as importantly, strain distribution can be altered (Yoshikawa et al. 1994; Donahue et al. 2000). Thus the prevention of stress fractures in our bones is a delicate balance between the loading environment, the properties of the bone, the adaptive response of bone tissue to loading, and the ability of our muscles to maintain the loading environment within proper physiological limits.

We have argued for the importance of remodeling in repairing fatigue damage in bone to prevent stress fracture. We now attempt to learn by analysis more about how this repair system may work.

Nash (1966) made an early attempt to mathematically analyze the accumulation of fatigue damage in a self-repairing structure. Prendergast and Taylor (1994) used damage as a stimulus for adaptive remodeling on bone's external surfaces, but did not consider the problem of damage repair itself. Later, Martin (1995) studied the consequences of assuming that damage itself activates remodeling, and found that there could be negative as well as positive repercussions. When a bone experiences a "fatigue challenge," its mechanical properties may be changed not only by damage, but also by increased remodeling space (resorption cavities and refilling osteons) produced by the induced remodeling. It has been postulated that this may actually accelerate progression to a stress fracture (Scully and Besterman 1982). That is, increased remodeling introduces more porosity, which decreases the elastic modulus, elevates the strains produced by the applied loads, and further increases damage formation. Thus, internal remodeling is required to repair fatigue damage, but in the process it may make the bone more susceptible to damage. Modeling responses on the periosteal surfaces of a fatigued bone can offset the negative effects of internal remodeling. Sufficient fatigue damage can stimulate the proliferation of woven bone from the periosteal surface (Johnson et al. 1963; Uthoff and Jaworski 1985). Presumably, this external modeling response reduces strains in the weakened structure by increasing its cross-sectional area. It is of great interest to know what determines whether the remodeling and modeling responses to fatigue ultimately have a positive or negative effect.

Let us use some of the methods discussed in Chap. 3 and the approach of Martin (1995) to analyze this problem. Consider the fatigue damage and remodeling in a representative cross section of a long bone diaphysis uniformly loaded in simple axial compression. A sequence of calculations simulates daily mechanical and remodeling changes in the bone. For convenience, fatigue damage is defined as millimeters of observable crack length per square millimeter of bone cross section. The damage formation rate is assumed to be a function of the loading rate (R_L in cycles/day) and the resulting strain range (S). Based on Eqs. (8.16) and (8.23), the damage formation rate may be assumed to be proportional to the strain range raised to a power, q :

$$dD_F/dt = k_D S^q R_L \quad (8.42)$$

where k_D is a damage rate coefficient that depends on loading conditions and bone structure.

The BMU activation frequency within the section is $Ac.f$. As each BMU passes through the section, its resorption cavity expands to the diameter of an osteonal cement line, removing a portion of the cortical area (which is subsequently refilled) and any damage it contains. Assuming that fatigue damage can only be removed by this process, and that the rate of removal is proportional to the amount of existing damage (because the more there is,

the greater the probability that a BMU will remove it), the rate of damage removal (or repair) is

$$dD_R/dt = D(Ac.f)\pi(Cm.Ln.Rd)^2F_S \quad (8.43)$$

where D is the existing damage and $Cm.Ln.Rd$ is the osteonal cement line radius. (Note that $Ac.f\pi Cm.Ln.Rd^2$ is the amount of cross section removed by all the BMUs created in 1 day.) F_S is a *damage repair specificity factor*, which accounts for any mechanisms that serve to direct remodeling to fatigue-damaged sites, as opposed to strictly random remodeling. The net rate of accretion of fatigue damage is $dD/dt = dD_F/dt - dD_R/dt$. In the equilibrium state, $dD_R/dt = dD_F/dt$ and the corresponding burden of damage waiting to be repaired is

$$D_0 = (k_D S^q R_L) / \left(\pi (Cm.Ln.Rd)^2 Ac.f_0 F_S \right) \quad (8.44)$$

where the “0” subscripts indicate equilibrium values.

The “dose–response” relationship between damage and activation frequency is unknown, but many such curves in biology are sigmoidal, so that is a reasonable assumption. Mathematically, this may be expressed as

$$Ac.f_a = \frac{Ac.f_0 Ac.f_{\max}}{Ac.f_0 - (Ac.f_{\max} - Ac.f_0) \exp[k_R Ac.f_{\max} (D - D_0)/D_0]} \quad (8.45)$$

where $Ac.f_0$ and $Ac.f_{\max}$ are the minimum and maximum allowable activation frequencies and k_R is a coefficient that determines the steepness of the response curve. The next part of the analysis is a bit involved, but you have already explored similar problems in the Chap. 3 exercises. How does the bone’s porosity change as a result of the changes in $Ac.f$? This depends on the current numbers of resorbing ($N.Rs$) and refilling ($N.F$) BMUs in the section. In Exercises 3.5 through 3.6 you saw that the activation frequency could be integrated over past time intervals to obtain $N.Rs$ and $N.F$. Then, in Exercises 3.7 and 3.8 you saw that the net change in bone volume fraction could be obtained by calculating

$$Q = Q_B N.F - Q_C N.Rs \quad (8.46)$$

where Q_B and Q_C are the rates (mm^2/day) at which bone is added and removed by refilling and resorbing BMUs, respectively. Because the time rate of change of porosity is just $-Q$, porosity can be tracked by constantly recalculating $N.Rs$, $N.F$, and Q as $Ac.f$ changes.

The elastic modulus, E , is computed from porosity, Po , using a relationship similar to those seen in Chap. 7

$$E = E_0 (1 - Po)^3 \quad (8.47)$$

where E_0 is the modulus of the bone tissue. It is recognized that E may also be diminished by fatigue damage, but this effect is ignored for the moment.

Damage is also assumed to stimulate periosteal bone formation, increasing the area of the cross section. No dose–response data are available for this effect, but Turner et al. (1993) found that when rat tibias were cyclically loaded, surface bone formation was lamellar at lower loads and converted to woven bone at higher loads. Lamellar bone apposition rates were linearly proportional to load or strain magnitude, but woven bone production was an “all or nothing” response, that is, the rate of bone formation was not correlated with load. Mathematically, this may be represented by

$$MAR_P = k_p(D - D_0)/D_0 \quad \text{for } D < D_C \quad (8.48)$$

$$MAR_P = MAR_W \quad \text{for } D > D_C \quad (8.49)$$

where k_p is a coefficient and MAR_W is the apposition rate for woven bone formation. If MAR_P^* is the maximum apposition rate for lamellar bone formation, achieved when the damage reaches a critical value, D_C , then

$$k_p = MAR_P^*/[(D_C/D_0) - 1] \quad (8.50)$$

The daily change in periosteal radius is calculated as $\Delta P_s.Rd = MAR_P \Delta t$, and cortical area is calculated assuming the endosteal radius is constant. Strain is computed by dividing the applied load by the model’s cross-sectional area and elastic modulus.

All these calculations may be incorporated in a computer program and the behavior of the system under various conditions explored. The model’s parameters can be partially calibrated against the results of the Burr experiments (Burr et al. 1985; Mori and Burr 1993), in which remodeling was activated by fatiguing canine bones in vivo. The model can then be used to examine the general behavior of the system and its sensitivity to its various parameters. Suppose that R_{LE} cycles per day (CPD) of additional loading, initially producing 2500 $\mu\epsilon$, is superimposed on the equilibrium loading. When the periosteal response is disabled and R_{LE} is 90 CPD, the additional damage increases the remodeling rate and porosity, and the bone becomes less stiff, but a new equilibrium is reached. However, when R_{LE} surpasses a critical value, the remodeling never catches up with the increased damage formation rate. Porosity, strain, and damage increase at an ever increasing rate, without limit. It is reasonable to consider this instability to be fatigue failure in the model (i.e., a complete “stress fracture”). If the loading is only slightly supercritical, this instability can be quite insidious. Damage levels off at what seems to be a new equilibrium level. Then, suddenly, after what may be a long time, the damage “explodes,” rising rapidly at a very high rate.

Thus, without a periosteal response, the model can withstand only relatively small fatigue challenges because of the increased strain caused by the remodeling space. When a periosteal callus is included in the model, the results are quite different (Fig. 8.25). Now the porosity increases from

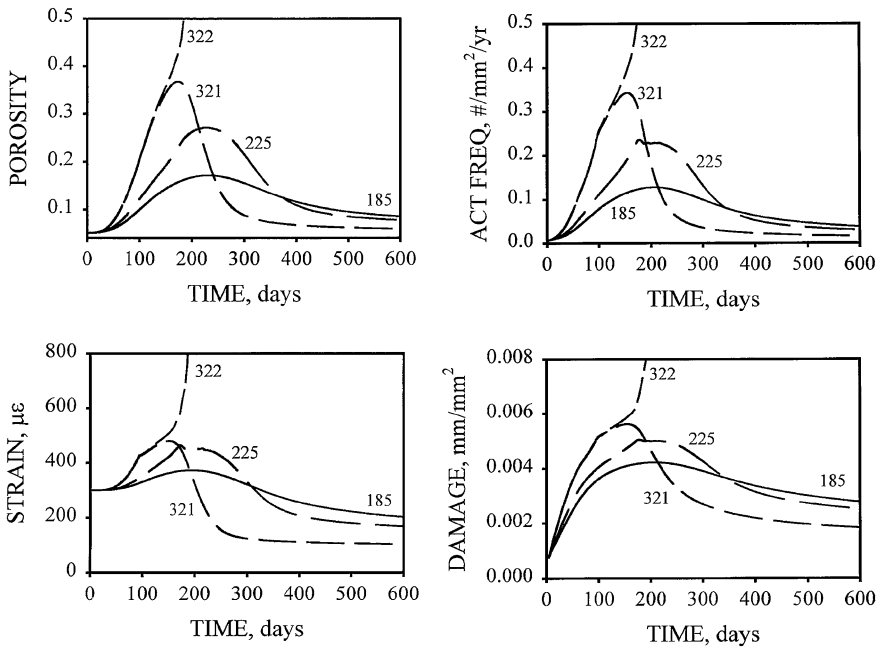


FIGURE 8.25. Stress fracture responses for model with periosteal modeling are shown for loading at four different loading levels: $R_{LE} = 185, 225, 321,$ and 322 cycles per day (CPD). The first of these rapidly produced a *stress fracture* in the model lacking a periosteal response. Now the model survives this loading easily, showing only transient responses in damage, activation frequency, porosity, and strain. As the loading increases, the responses increase in magnitude but recovery is quicker, until suddenly the system “blows up” (Martin 1995).

remodeling are offset by periosteal diameter changes, and the porosity comes back down as the bone’s cross-sectional area increases. The addition of external bone acts to reduce strain, and the system returns toward a new equilibrium level after a transient response lasting about a year increasing the loading causes larger transients, which end more quickly. Eventually, however, another critical load level is reached, and instability implying failure ensues once more. Again, the inherent instability of the system is remarkable. As loading increases, remodeling responds ever more vigorously, and the challenge is overcome more and more quickly. Then, when just one more cycle per day of loading is added, instead of damage simply not decreasing, the system precipitously fails.

Such instability is typical of feedback control systems that are nonlinear or contain delays in their responses. Bone remodeling in response to fatigue would have both these characteristics. It is clear that damage is a very nonlinear function of strain, and modulus is a nonlinear function of porosity.

The activation of remodeling may be a fairly linear function of damage, but there are delays in the development of porosity changes after Ac.f_a changes. These traits are fully consistent with the response of the model, particularly the fact that small changes in the loading conditions or the system parameters can produce large changes in the outcome if the system is near a critical state. This kind of behavior in the model, and the difficulty in predicting when someone will suffer a stress fracture and when they will not, are in concert with one another. However, much remains to be learned about the intriguing interactions between remodeling and mechanical behavior in the skeleton. We explore these interactions further in the next chapter.

8.10 Summary and Additional Reading

Bone is a natural composite material that shares with engineered composite materials many mechanisms for resisting crack propagation. Fracture mechanics teaches us that cracks are propagated through materials when the stress concentration at the crack tip becomes high enough to exceed a critical stress intensity factor, K_{Ic} . Because of the lamellar structure of bone, K_{Ic} is lower for longitudinal than transverse crack propagation, so cracks find it difficult to travel catastrophically across the diaphysis of a long bone. Cracks are driven by strain energy, and the cement lines and lamellar interfaces of bone serve to consume energy in the propagation of relatively harmless longitudinal cracks, preventing the propagation of dangerous transverse cracks. Both the numbers and the sizes of osteons appear to be important in maximizing fatigue resistance and the amount of energy required to fracture a bone.

Creep and fatigue damage of bone are closely related phenomena, and both kinds of loading are common in daily life. Experimentally, it is difficult to study fatigue damage mechanisms in bone because experiments at physiologic strain levels require very prolonged loading times. It appears that most of the fatigue failure data on bone involve primarily creep damage caused by superphysiologic stresses. Consequently, we still do not know very much about the differences between creep and fatigue damage in bone. At relatively high stress levels, crack damage appears to be different for cyclic tensile and compressive loading. During most of the tensile fatigue life cracks seem to be self-limiting, occurring primarily in interstitial regions between osteons. In compression, cracks seem to form later, running between Haversian canals and crossing cement lines. The differences in the shapes of the modulus degradation curves seen in tensile and compressive fatigue seem to be attributable to these differences in crack propagation characteristics.

Cyclic loading at physiologic stress levels produces microcracks in osteonal bone and activates increased intracortical remodeling; the mechanism behind this is unclear. If one assumes that bone remodeling is activated in proportion to fatigue damage, and that the increased porosity produced by this remodeling reduces the elastic modulus of the bone, increasing strains

and therefore damage, it becomes clear that limits exist on the ability of remodeling to control fatigue damage under increased loading. Mathematically modeling this system suggests that periosteal callus formation is essential to maintaining its stability. Even so, the nonlinearity of the system seems to make it critically stable, so that small differences in skeletal physiology may greatly alter the ability of individuals to avoid a stress fracture under conditions of increased physical activity.

Additional reading suggestions include the early clinical perspective of Devas (1975) on stress fractures and the more recent book edited by Burr and Milgrom (2001) that covers many aspects of stress fracture epidemiology and pathogenesis. A survey of the work of D.R. Carter between 1975 and 1990 will produce much of the literature on the mechanics of fatigue and creep in bone. A review by Donahue and Galley (2006) provides an excellent overview of microdamage and its implication for skeletal fragility and remodeling. Some information can also be found in the more recent paper “Why bones bend but don’t break” (Burr 2011).

8.11 Exercises

- 8.1. Equations (8.6) and (8.7) may be used together to calculate the stress intensity factor for a transverse crack in the edge of a plate loaded in tension. Plot a graph showing how both C and K/s change as the crack length increases from 1 to 90 % of the plate width $b = 10$ cm.
- 8.2. Cut a strip of paper 10 cm wide (x -dimension) and approximately 28 cm (11 in.) long (y -dimension). Mark the middle of its length. About 4 cm above the middle, cut a center slit (crack) 5 cm long parallel to the x -axis. About 4 cm below the middle, cut two edge slits, one from each side, each 1 cm long, again parallel to the x -axis. Do a calculation to determine where the strip of paper will fail if it is loaded in tension in the y -direction, see Fig 8.4. Then wrap the ends of the strip around pencils and do the experiment to check your results.
- 8.3. Equations (8.18) and (8.19) give the number of cycles to failure for fatigue of human cortical bone in longitudinal tension and compression, respectively. How many days would it take to achieve fatigue failure in each of these modes of loading if the normalized applied stress range, S , is selected to produce 2000 $\mu\epsilon$ in the specimen? Assume the loading rate is 2 Hz. What experimental problems would arise from these long fatigue times? How long would the experiment take if S produced a superphysiologic 5000 $\mu\epsilon$? Why would you not want to run the experiments at 20 Hz to reduce their duration?
- 8.4. Equation (8.44) states a relationship between remodeling variables and the damage burden (D_0) waiting to be repaired in an equilibrium state:

$$D_0 = (k_D S^q R_L) / \left(\pi (Cm.Ln.Rd)^2 (Ac.f_0) F_S \right).$$

If the loading increases, an increase in the damage burden can be avoided by increasing either the activation frequency or the diameter of the osteons, or both. Does one of these options have any advantages over the other? Think about the differences in the porosity increase that would occur during and after the remodeling, the metabolic energy required to support the cells doing the remodeling, and the effects of the altered osteonal structure on fatigue damage resistance. What are the limitations on increasing $Cm.Ln.Rd$? How about $Ac.f_0$?

- 8.5. Muscle fatigue may be an important cause of stress fractures (Yoshikawa et al. 1994; Devas 1975). We saw in Chap. 1 that there are more muscles to control our movements than we actually need. The muscle fatigue theory of stress fracture holds that the additional muscles not only make refinements of movement possible, but also reduce the stresses in the bones when the most common movements are performed. For example, there are six principal muscles that control the foot during the toe-off phase of walking. Sharkey et al. (1995) found that the compressive strain on the dorsal surface of the human second metatarsal bone at toe-off increases from about 1700 $\mu\epsilon$ to about 2190 $\mu\epsilon$ when just one of these (flexor hallucis longus) stops functioning. Using Eq. (8.17), find the percent decrease in fatigue life to be expected from this increase in strain.
- 8.6. The Pacific Crest Trail runs from Mexico to Canada through the Sierra Nevada and Cascade mountain ranges. Each year a few people succeed in walking the entire trail, starting in early spring and finishing in the fall. Many fail because of physical problems, including stress fractures. Using Eq. (8.42), consider the relative effects of the distance walked and the weight of a backpack on the rate of fatigue damage formation in one of your foot or leg bones if you undertook this walk. Suppose that you ordinarily walk 5 miles/day. Plot a graph of the percent increase in damage formation rate as a function of the weight of your pack, expressed as a percent of your body weight, assuming you continued walking just 5 miles/day. Try relatively high and low values of q . Then convert this graph to a three-dimensional plot of percent change in damage formation rate vs. pack weight and distance walked per day.
- 8.7. Show that Eq. (8.41) is the solution to Eq. (8.40).
- 8.8. Show that Eqs. (8.38) and (8.39) are the solutions to Eqs. (8.36) and (8.37).
- 8.9. Write a computer program that simulates the generation and repair of fatigue damage in a segment of cortical bone using Eqs. (8.42)–(8.47).

See if you can reproduce the responses described for the model **without** a periosteal response.

- 8.10. Using the following data, calculate the effect of grooved or ungrooved specimens on the stress intensity factor (K_c) and the strain energy release rate (G_c), where ug = ungrooved and g = grooved.

Crack length, a (mm)	Critical Load (P_g/P_{ug})	Compliance $(\delta C/\delta a)_{ug}/(\delta C/\delta a)_g$
7	1.309	1.789
9	1.103	1.499
10	1.317	1.752
11	1.278	1.734
12	1.314	1.639
13	1.412	1.635

See Norman et al. J Biomechanics 1992; 25:1489–1492 for solution (Norman et al. 1992)

References

- Alto A, Pope MH. On the fracture toughness of equine metacarpi. J Biomech. 1979;12:415–21.
- Balazs EA, Gibbs DA. The rheological properties and biological function of hyaluronic acid. In: Balazs EA, editor. Chemistry and molecular biology of the intercellular matrix. New York, NY: Academic; 1970.
- Barth HD, Zimmermann EA, Schaible E, Tang SY, Alliston T, Ritchie RO. Characterization of the effects of x-ray irradiation on the hierarchical structure and mechanical properties of human cortical bone. Biomaterials. 2011;32:8892–904.
- Behiri JC, Bonfield W. Fracture mechanics of bone—the effects of density, specimen thickness and crack velocity on longitudinal fracture. J Biomech. 1984;17:25–34.
- Bentolila V, Hillam RA, Skerry TM, Boyce TM, Fyhrie DP, Schaffler MB. Activation of intracortical remodeling in adult rat long bones by fatigue loading. Trans Orthop Res Soc. 1997;22:578.
- Bonfield W, Datta PK. Fracture toughness of compact bone. J Biomech. 1976;9:131–4.
- Boyce TM, Fyhrie DP, Glotkowski MC, et al. Damage type and strain mode associations in human compact bone bending fatigue. J Orthop Res. 1998;16:322–9.
- Boyde A. The real response of bone to exercise. J Anat. 2003;203:173–89.
- Burr DB. Targeted and nontargeted remodeling. Bone. 2002;30:2–4.
- Burr DB. Why bones bend but don't break. J Musculoskelet Neuronal Interact. 2011;11:270–85.
- Burr DB, Martin RB. Calculating the probability that microcracks initiate resorption spaces. J Biomech. 1993;26:613–6.
- Burr DB, Milgrom C, editors. Musculoskeletal fatigue and stress fractures, CRC Series in exercise physiology. Boca Raton, FL: CRC Press; 2001.
- Burr DB, Stafford T. Validity of the bulk-staining technique to separate artifactual from in vivo bone microdamage. Clin Orthop Relat Res. 1990;260:305–8.

- Burr DB, Turner CH. Biomechanical measurements in age-related bone loss. In: Rosen CJ, Glowacki J, Bilezikian JP, editors. *The aging skeleton*. London: Academic; 1999. p. 301–11. Chapter 20.
- Burr DB, Martin RB, Schaffler MB, Radin EL. Bone remodeling in response to in vivo fatigue microdamage. *J Biomech*. 1985;18:189–200.
- Burstein AH, Reilly DT, Martens M. Aging of bone tissue: mechanical properties. *J Bone Joint Surg*. 1976;58A:82–6.
- Caler WE, Carter DR. Bone creep-fatigue damage accumulation. *J Biomech*. 1989;22:625–35.
- Cardoso L, Herman BC, Verborgt O, Laudier D, Majeska RJ, Schaffler MB. Osteocyte apoptosis controls activation of intracortical resorption in response to bone fatigue. *J Bone Miner Res*. 2009;24:597–605.
- Carter DR, Caler WE. Cycle-dependent and time-dependent bone fracture with repeated loading. *J Biomech Eng*. 1983;105:166–70.
- Carter DR, Caler WE. A cumulative damage model for bone fracture. *J Orthop Res*. 1985;3:84–90.
- Carter DR, Hayes WC. Compact bone fatigue damage. I. Residual strength and stiffness. *J Biomech*. 1977a;10:325–37.
- Carter DR, Hayes WC. Compact bone fatigue damage: a microscopic examination. *Clin Orthop Relat Res*. 1977b;127:265–74.
- Carter DR, Hayes WC, Schurman DJ. Fatigue life of compact bone. II. Effects of microstructure and density. *J Biomech*. 1976;9:211–8.
- Carter DR, Caler WE, Spengler DM, Frankel VH. Uniaxial fatigue of human cortical bone. The influence of tissue physical characteristics. *J Biomech*. 1981;14:461–70.
- Chamay A, Tschantz P. Mechanical influences in bone remodeling. Experimental research on Wolff's law. *J Biomech*. 1972;5:173–80.
- Cohen J, Harris WH. The three-dimensional anatomy of haversian systems. *J Bone Joint Surg*. 1958;40A:419–34.
- Corondan G, Haworth WL. A fractographic study of human long bone. *J Biomech*. 1986;19:207–18.
- Cowin SC, Weinbaum S, Zeng Y. A case for bone canaliculi as the anatomic site of strain generated potentials. *J Biomech*. 1995;28:1281–97.
- Currey JD, Brear K, Zioupos P. The effects of ageing and changes in mineral content in degrading the toughness of human femora. *J Biomech*. 1996;29:257–60.
- Devas MB. *Stress fractures*. London: Churchill Livingstone; 1975.
- Diab T, Vashishth D. Effects of damage morphology on cortical bone fragility. *Bone*. 2005;37:96–102.
- Diab T, Condon KW, Burr DB, Vashishth D. Age-related change in the damage morphology of human cortical bone and its role in bone fragility. *Bone*. 2006;38:427–31.
- Donahue SW, Galley SA. Microdamage in bone: implications for fracture, repair, remodeling, and adaptation. *Crit Rev Biomed Eng*. 2006;34:215–71.
- Donahue SW, Sharkey NA, Modanlou KA, Sequeira LN, Martin RB. Bone strain and microcracks at stress fracture sites in human metatarsals. *Bone*. 2000;27:827–33.
- Evans FG, Riolo ML. Relations between the fatigue life and histology of adult human cortical bone. *J Bone Joint Surg*. 1970;52A:1579–86.
- Fazzalari NL, Forwood MR, Mathey BA, Smith K, Kolesik P. Three-dimensional confocal images of microdamage in cancellous bone. *Bone*. 1998;23:373–8.
- Finkel V. *Portrait of a crack* (translated from the Russian by Y. Nadler). Moscow: Mir Publications; 1985.

- Frasca P. Scanning electron microscopy studies of ground substance in the cement lines, resting lines, hypercalcified rings and reversal lines of human cortical bone. *Acta Anat.* 1981;109:115–21.
- Freund LB. Dynamic fracture mechanics. Cambridge: Cambridge University Press; 1990.
- Frost HM. Presence of microscopic cracks in vivo in bone. *Henry Ford Hosp Med Bull.* 1960;8:25–35.
- Frost HM. Bone remodeling and its relationship to metabolic bone disease. Springfield: Thomas; 1973.
- Fyhrie DP, Milgrom C, Hoshaw SJ, Simkin A, Dar S, Drumb D, Burr DB. Effect of fatiguing exercise on longitudinal bone strain as related to stress fracture in humans. *Ann Biomed Eng.* 1998;26:660–5.
- Gaeta M, Minutoli F, Scribano ER, Ascenti G, Vince S, Bruschetta D, Magaudo L, Glandino A. CT and MR imaging findings in athletes with early tibial stress injuries. Comparison with bone scintigraphy findings and emphasis on cortical abnormalities. *Radiology.* 2005;235:553–61.
- Gibson VA, Stover SM, Martin RB, Gibeling JC, Willits NH, Gustafson MB, Griffin LV. Fatigue behavior of the equine third metacarpus: mechanical property analysis. *J Orthop Res.* 1996;13:861–8.
- Gottesman T, Hashin Z. Analysis of viscoelastic behavior of bones on the basis of microstructure. *J Biomech.* 1980;13:89–96.
- Green JO, Wang J, Diab T, Vidakovic B, Guldborg RE. Age-related differences in the morphology of microdamage propagation in trabecular bone. *J Biomech.* 2011;44:2659–66.
- Griffin LV, Gibeling JC, Martin RB, Gibson VA, Stover SM. A model of flexural fatigue damage accumulation for cortical bone. *J Orthop Res.* 1997;15:607–14.
- Griffith AA. The phenomenon of rupture and flow in solids. *Philos Trans R Soc Lond.* 1920;A221:163–98.
- Hasegawa K, Turner CH, Chen J, Burr DB. Effect of disc lesion on microdamage accumulation in lumbar vertebrae under cyclic compression loading. *Clin Orthop.* 1995;311:190–8.
- Heaney RP. The natural history of vertebral osteoporosis. Is low bone mass an epiphenomenon? *Bone.* 1992;13 Suppl 2:S23–6.
- Herman BC, Cardoso L, Majeska RJ, Jepsen KJ, Schaffler MB. Activation of bone remodeling after fatigue: differential response to linear microcracks and diffuse damage. *Bone.* 2010;47:766–72.
- Hulkko A, Orava S. The role of age in the development of stress and fatigue fractures. In: Burr DB, Milgrom C, editors. *Musculoskeletal Fatigue and Stress Fractures*, CRC Series in exercise physiology. Boca Raton, FL: CRC Press; 2001. p. 35–54. Chapter 4.
- Irwin GR. Analysis of stresses and strains near the end of a crack traversing a plate. *J Appl Mech.* 1957;24:361–4.
- Johnson LC, Stradford HT, Geis RW, Dineen JR, Kerley E. Histogenesis of stress fractures. *J Bone Joint Surg.* 1963;45A:1542.
- Kendall K. Control of cracks by interfaces in composites. *Proc R Soc Lond.* 1975;341:409–28.
- Kennedy OD, Herman BC, Laudier DM, Majeska RJ, Sun HB, Schaffler MB. Activation of resorption in fatigue-loaded bone involves both apoptosis and active pro-osteoclastogenic signaling by distinct osteocyte populations. *Bone.* 2012;50:1115–22.

- King AI, Evans FG. Analysis of fatigue strength of human compact bone by the Weibull method. In: Jacobson B, editor. Digest of the 7th international conference on medical and biological engineering. The Royal Academy of Engineering Sciences, Stockholm; 1967.
- Krajcinovic D, Trafimow J, Sumarac D. Simple constitutive model for cortical bone. *J Biomech.* 1987;20:779–84.
- Lakes R, Saha S. Cement line motion in bone. *Science.* 1979;204:501–3.
- Lanyon LE, Baggott DG. Mechanical function as an influence on the structure and form of bone. *J Bone Joint Surg.* 1976;58B:436–43.
- Lanyon LE, Bourn S. The influence of mechanical function on the development and remodeling of the tibia. *J Bone Joint Surg.* 1979;62A:263–73.
- Launey ME, Buehler MJ, Ritchie RO. On the mechanistic origins of toughness in bone. *Annu Rev Mater Res.* 2010;40:25–53.
- Martin RB. A mathematical model for fatigue damage repair and stress fracture in osteonal bone. *J Orthop Res.* 1995;13:309–16.
- Martin RB. Is all cortical bone remodeling initiated by microdamage? *Bone.* 2002;30:8–13.
- Martin RB, Burr DB. A hypothetical mechanism for the stimulation of osteonal remodeling by fatigue damage. *J Biomech.* 1982;15:137–9.
- Melvin JW, Evans FG. Crack propagation in bone. In: Fung YC, Brighton JA, editors. Biomaterials symposium, AMD2. New York, NY: ASME; 1973.
- Milgrom C, Radeva-Petrova DR, Finestone A, Nyska M, Mendelson S, Benjuya N, Simkin A, Burr D. The effect of muscle fatigue on in vivo tibial bone strains. *J Biomech.* 2007;40:845–50.
- Mori S, Burr DB. Increased intracortical remodeling following fatigue damage. *Bone (New York).* 1993;14:103–9.
- Mori S, Harruff R, Ambrosius W, Burr DB. Trabecular bone volume and microdamage accumulation in the femoral heads of women with and without femoral neck fractures. *Bone (New York).* 1997;21:521–6.
- Moyle DD, Bowden RW. Fracture of human femoral bone. *J Biomech.* 1984;17:203–13.
- Moyle DD, Welborn JW, Cooke FW. Work to fracture of canine femoral bone. *J Biomech.* 1978;11:435–40.
- Mulcahy LE, Taylor D, Lee TC, Duffy GP. RankL and OPG activity is regulated by injury size in networks of osteocyte-like cells. *Bone.* 2011;48:182–8.
- Nakashima T, Hayashi M, Fukunaga T, et al. Evidence for osteocyte regulation of bone homeostasis through RANKL expression. *Nat Med.* 2011;17:1231–4.
- Nalla RK, Kinney JH, Ritchie RO. Mechanistic fracture criteria for the failure of human cortical bone. *Nat Mat.* 2003;2:164–8.
- Nash CD. Fatigue of a self-healing structure: a generalized theory of fatigue failure. ASME Publication #66-WA/BHF-3, American Society of Mechanical Engineers, New York; 1966.
- Norman TL, Wang Z. Microdamage of human cortical bone: incidence and morphology in long bones. *Bone (New York).* 1997;20:375–9.
- Norman TL, Vashishth D, Burr DB. Effect of groove on bone fracture toughness. *J Biomech.* 1992;25:1489–92.
- Norman TL, Vashishth D, Burr DB. Fracture toughness of human bone under tension. *J Biomech.* 1995;28:309–20.
- Nyman JS, Reyes M, Wang XD. Effect of ultrastructural changes on the toughness of bone. *Micron.* 2005;36:566–82.

- Owan I, Burr DB, Turner CH, Qui J, Tu Y, Onyia JE, Duncan RL. Mechanotransduction in bone: osteoblasts are more responsive to fluid forces than mechanical strain. *Am J Physiol*. 1997;273:C810–5.
- Parfitt AM. The physiologic and clinical significance of bone histomorphometric data. In: Recker RR, editor. *Bone histomorphometry techniques and interpretation*. Boca Raton, FL: CRC Press; 1983. p. 143–223.
- Parfitt AM. Targeted and nontargeted bone remodeling: relationship to basic multicellular unit origination and progression. *Bone*. 2002;30:5–7.
- Paris PC. The fracture mechanics approach to fatigue. In: *Fatigue—an interdisciplinary approach*. Proceedings, tenth Sagamore army materials research conference. Syracuse University Press, Syracuse, NY; 1964.
- Park HC, Lakes RS. Cosserat micromechanics of human bone: strain redistribution by a hydration sensitive constituent. *J Biomech*. 1986;19:385–97.
- Parsamian GP, Norman TL. Diffuse damage accumulation in the fracture process zone of human cortical bone specimens and its influence on fracture toughness. *J Mat Sci Mat Med*. 2001;12:779–83.
- Pattin CA, Caler WE, Carter DR. Cyclic mechanical property degradation during fatigue loading of cortical bone. *J Biomech*. 1996;29:69–80.
- Piekarski K. Fracture of bone. *J Appl Phys*. 1970;41:215–23.
- Pilkey WD. *Formulas for stress, strain, and structural matrices*. New York, NY: Wiley; 1994.
- Pope MH, Murphy MC. Fracture energy of bone in a shear mode. *Med Biol Eng*. 1974;12:763–7.
- Prendergast PJ, Taylor D. Prediction of bone adaptation using damage accumulation. *J Biomech*. 1994;27:1067–76.
- Rahn BA, Gallinaro P, Schenck RK. Compression interfragmentaire et surcharge locale de l'os. In: Boitzy A, editor. *Interarthrite de l'Epaule. Osteogenese et Compression*. Bern: Huber; 1972.
- Reifsnider KL. Damage and damage mechanics. In: Reifsnider KL, editor. *Fatigue of composite materials*. New York, NY: Elsevier; 1990. p. 11–77.
- Ritchie RO. The conflicts between strength and toughness. *Nat Mater*. 2011;10:817–22.
- Rubin CT, Lanyon LE. Regulation of bone formation by applied dynamic loads. *J Bone Joint Surg*. 1984;66A:397–402.
- Russell NA, Pelletier MH, Bruce WJ, Walsh WR. The effect of gamma irradiation on the anisotropy of bovine cortical bone. *Med Engng Phys*. 2012;34:1117–22.
- Rutishauser E, Majno G. Lesions osseuses par sur charge dans le squelette normal et pathologique. *Bulletin Schweizerischen Akademie der Medizinischen Wissenschaften*. 1950;6:333–41.
- Saha S. Longitudinal shear properties of human compact bone and its constituents, and the associated failure mechanisms. *J Mater Sci*. 1977;12:1798–806.
- Schaffler MB. Stiffness and fatigue of compact bone at physiological strains and strain rates. Doctoral dissertation, West Virginia University, Morgantown, WV; 1985.
- Schaffler MB, Burr DB. Primate cortical bone microstructure: relationship to locomotion. *Am J Phys Anthropol*. 1984;65:191–7.
- Schaffler MB, Kennedy OD. Osteocyte signaling in bone. *Curr Osteop Rep*. 2012;10:118–25.
- Schaffler MB, Choi K, Milgrom C. Aging and matrix microdamage accumulation in human compact bone. *Bone (New York)*. 1995;17:521–5.

- Scully TJ, Besterman G. Stress fracture—a preventable training injury. *Mil Med.* 1982;147:285–7.
- Seref-Ferlengez Z, Kennedy O, Schaffler MB. Diffuse microdamage induced in cortical bone in vivo decreases with time in the absence of bone remodeling. 34th Annual meeting of the American society for bone and mineral research, 2012.
- Shane E, Burr DB, Ebeling P, Abrahamsen B, Adler RA, Brown TD, et al. Atypical subtrochanteric and diaphyseal femoral fractures: report of a task force of the American society for bone and mineral research. *J Bone Miner Res.* 2010;25:1–28.
- Sharkey NA, Smith TS, Lundmark DC. Freeze clamping musculo-tendinous junctions for in-vitro simulation of joint mechanics. *J Biomech.* 1995;28(5):631–5.
- Simkin A, Robin G. Fracture formation in differing collagen fiber pattern of compact bone. *J Biomech.* 1974;7:183–8.
- Taylor D, Lee TC. Measuring the shape and size of microcracks in bone. *J Biomech.* 1998;31:1177–80.
- Tschantz P, Rutishauser E. La surcharge mécanique de l'os vivant. *Annales d'Anatomie Pathologique.* 1967;12:223–48.
- Turner CH, Forwood M, Rho J-Y, Yoshikawa T. Mechanical loading thresholds for lamellar and woven bone formation. *J Bone Miner Res.* 1993;9:87–97.
- Uthoff HK, Jaworski ZFG. Periosteal stress-induced reactions resembling stress fractures. *Clin Orthop Relat Res.* 1985;199:284–91.
- Vashishth D, Behiri JC, et al. Crack growth resistance in cortical bone: concept of microcrack toughening. *J Biomech.* 1997;30:763–9.
- Vashishth D, Koontz J, Qiu SJ, Lundin-Cannon D, Yeni YN, Schaffler MB, Fyhrie DP. In vivo diffuse damage in human vertebral trabecular bone. *Bone.* 2000;26:147–52.
- Verborgt O, Gibson GJ, Schaffler MB. Loss of osteocyte integrity in association with microdamage and bone remodeling after fatigue in vivo. *J Bone Miner Res.* 2000;15:60–7.
- Wang X, Shen X, Li X, et al. Age-related changes in the collagen network and toughness of bone. *Bone.* 2002;31:1–7.
- Warden SJ, Burr DB, Brukner PD. Repetitive stress pathology: bone. In: Magee DJ, Zachazewski JE, Quillen WS, editors. *Pathology and intervention in musculoskeletal rehabilitation.* St. Louis: Elsevier; 2009. p. 685–705.
- Wasserman N, Yerramshetty J, Akkus O. Microcracks colocalize within highly mineralized regions of cortical bone. *Eur J Morphol.* 2005;42:43–51.
- Weinbaum S, Cowin SC, Zeng Y. A model for the excitation of osteocytes by mechanical loading-induced bone fluid shear stresses. *J Biomech.* 1994;27:339–60.
- Wenzel TE, Schaffler MB, Fyhrie DP. In vivo trabecular microcracks in human vertebral bone. *Bone (New York).* 1996;19:89–95.
- Xiong J, Onal M, Jilka RL, et al. Matrix-embedded cells control osteoclast formation. *Nat Med.* 2011;17:1235–41.
- Yeni YN, Hou FJ, Ciarelli T, et al. Trabecular shear stresses predict in vivo linear microcrack density but not diffuse damage in human vertebral cancellous bone. *Ann Biomed Eng.* 2003;31:726–32.
- Yoshikawa T, Mori S, Santiesteban AJ, Sun TC, Hofstad E, Chen J, Burr DB. The effects of muscle fatigue on bone strain. *J Exp Biol.* 1994;188:217–33.
- Zimmerman EA, Barth HD, Ritchie RO. The multiscale origins of fracture resistance in human bone and its biological degradation. *JOM.* 2012;64:486–93.

Index

- Abductor muscle force, 9–11, 16, 27, 31
- Abductor muscles, 10, 11, 13, 373
- ACL (anterior cruciate ligament), 86, 192, 210, 213–215, 219
- Acoustic microscope, 388
- Activation, 12, 64, 86, 99–101, 107, 111, 118, 127, 134–139, 141, 154, 155, 165, 166, 229, 230, 287, 296–297, 310, 313, 320, 322, 329, 334, 335, 462–468, 470, 474
- Activation frequency
 - calculating, without in vivo labels, 162–164
 - changes in, 137
 - definitions of, 127, 134
 - estimating, 160, 163
 - measuring, 137
 - in rib, 139, 154, 162
- Activation, resorption, and formation (ARF), 99–100
- Activation signals, 462
- Acute inflammation, 86, 211
- Adaptation
 - biological, 283
 - term, 284
- Adaptive elasticity theory, 319–320, 332
- Adaptive finite element models, 323–324
- Age. *See also* Elderly
 - chronological, 80, 164
 - determining, of bone, 156
 - mechanical behavior of tendons and ligaments and, 227
 - skeletal, 80, 156, 162
- Age-at-death, 159, 161–162
- Aggrecan, 66, 69–71, 81, 232, 244
- Aggregate modulus, 233, 235, 237, 241–243, 247, 270
- Alizarin, 120
- Allografts, 214
- Alpha-actinin, 193
- Amenorrhea, 334
- Amontons' laws, 249
- Anisotropic materials, 363, 407
- Anisotropy
 - of bovine and human bone, 386
 - of cancellous bone architecture, 405–407
 - of cortical bone mechanical properties, 385
- Anisotropy effects
 - determined by load-deformation testing, 385–387
 - determined by ultrasonic measurements, 387–389
- Anisotropy ratios, 385, 388

- Anterior cruciate ligament (ACL),
86, 192, 210, 213–215, 219
- Aponeurosis, 5, 81, 181, 182,
192–194, 218
- Apoptosis, 60, 64, 78, 141, 315, 466
- Apparent density, 56, 323, 324, 326,
328, 330, 342, 390, 391,
395, 398, 400–403, 405,
407, 408, 411
- Apparent permeability, 235
- Apposition rates
bone, 61, 87, 472
declining, 151
mean, 145, 148
mineral, 74, 122, 126,
128, 132
models for, 152
during refilling process, 146–148
- ARF (activation, resorption, and
formation), 99–100
- A-R-F sequence, 99–100
- Arthrodesis, 22
- Articular cartilage
biphasic theory of, 240, 265
indentation testing of,
235–237, 269
organization of, 71–75
permeability of, 234–235
structure of, 233–234
tensile tests of, 237–239
- Ash fraction, 56, 392, 393
- Ash mass, 56, 393
- Australopithecus, 13–15, 29–32
- Autografts, 86, 214, 217
- Avian isolated ulna experiments, 324
- Basic multicellular units (BMUs), 60,
66, 99–108, 116, 118,
121–130, 134–146,
151–153, 155, 162,
164–167, 317, 320–322,
334, 342, 402, 412,
464–465, 470, 471
osteonal (*see* Osteonal BMUs)
- Bending, stress and, 298, 317, 338
- BFR (bone formation rate), 122,
126, 128, 129, 132,
139, 163, 292, 300,
302, 308, 324
- Biceps muscle, 5, 21
- Biglycan, 54, 81
- Biological adaptation, 283
- Bipedalism, cranial capacity and, 31
- Biphasic theory of articular cartilage,
240–243, 246
- Birefringence, 45, 382
- BMUs (basic multicellular units),
60, 66, 99–108, 116, 118,
121–130, 134–146,
151–153, 155, 162,
164–167, 317, 320–322,
334, 342, 402, 412,
464–465, 470, 471
- Bone(s)
breaking energy of, 372
as composite material, 387,
409–410, 438, 449–459
cortical (*see* Cortical bone)
creep behavior of, 442–445
determining age of, 156
fatigue and fracture resistance of,
423–477
fatigue behavior of, 438–442
fatigue damage in, 424, 438–441,
459–463, 465, 467,
474, 476
flat (*see* Flat bones)
fractured, self-straightening
of, 412
long (*see* Long bones)
longitudinal growth of, 75–80, 97
mechanical properties of, 54, 55,
96, 355–414, 423, 439
minerals in, 35, 36, 39, 53, 54, 60,
74, 102, 127, 167, 303, 307,
393–395
modeling of, 59, 95–168, 230,
386, 465

- porosity of, 43
- primary (*see* Primary bone)
- remodeling of, 95–168, 230, 465
- shapes of, 40, 95, 276, 373
- surgically overloaded, 212
- tensile strength of, 178
- water and, 393
- whole (*see* Whole bones)
- Bone apposition rate, 61, 88, 472
- Bone atrophy, posttraumatic, 118
- Bone biology, 35, 87
- Bone cells, 55, 59–65, 107, 280, 286,
 - 301, 306, 310–313, 315,
 - 323, 325, 330, 334
- Bone cement, 22, 403
- Bone chemistry, 35
- Bone composition, 56–57, 396
 - quantitative representation of,
 - 56–57
- Bone diseases, 65, 120, 124, 392
- Bone formation
 - intramembranous, 96, 109
- Bone formation rate (BFR), 122,
 - 126, 128, 129, 132, 139,
 - 163, 292, 300, 302,
 - 308, 324
- Bone fracture patterns, mechanical
 - significance of, 370–371
- Bone fragility in elderly, 467–468
- Bone histology, 59, 64–65
- Bone histomorphometry, standard
 - notation for, 119
- Bone lining cells, 64, 111, 134, 152,
 - 153, 329, 464
- Bone loss, senile, 107
- Bone mass, noninvasive
 - measurement of, 393–395
- Bone matrix, 37, 49, 54, 55, 61, 64,
 - 75, 101, 107, 112, 126, 134,
 - 152, 195, 197, 296, 313,
 - 319, 324, 329, 341, 374,
 - 381, 382, 388, 392, 397,
 - 402, 435, 437, 441, 453,
 - 456, 457
- Bone mechanical properties, effects
 - of osteons on, 385, 389–399
- Bone remodeling. *See* Remodeling
- Bone resorption, 64, 75, 96,
 - 100, 103, 104, 128–129,
 - 131, 139, 250, 309,
 - 310, 334, 465
- Bone structural units, 104–105, 402
- Bone tissue
 - types of, 43, 44
- Bone toughness, 53, 435, 450–451
- Bone tumors, 109
- Bone volume fraction, 57, 122, 324,
 - 390, 408, 409, 471
- Bony callus, 112, 115
- Bony union, 112, 113
- Boosted lubrication, 260, 263
- Borelli, G.A., 7–8, 17, 27
- Boundary conditions, 256, 373
- Boundary lubrication, 250–251,
 - 261, 262
- Bovine bone, anisotropy of, 385, 386
- Boyle's law, 361
- Brachialis muscle, 4, 5, 20, 21
- Brachioradialis muscle, 5, 20, 21
- Brush border, 60
- Cadaver specimens, 2, 15
- Calcification front, 62, 72
- Calcification, provisional, zone
 - of, 78
- Calcitonin, 294
- Calcium, 35, 63, 64, 78, 101, 186,
 - 287, 294, 313, 315, 334,
 - 393, 465
- Cambium layer, 96, 109
- Canaliculi, 57, 62, 63, 65, 313, 314,
 - 390, 397, 465
- Cancellous bone
 - material properties of, 399–409
 - modeling, as cellular system of
 - plates or struts, 407–408
 - as porous material, 400, 401
 - stress-strain curves for, 399–400

- Cancellous bone architecture,
 - anisotropy of, 405–407
- Cancellous bone mechanical
 - properties, determinants of, 401
- Cancellous bone tissue properties, 401–405
- Cancellous bone turnover rates, 106
- Canine disuse experiments, 295
- Cannon bone, equine, 24–26
- Carter–Beaupre approach, 323
- Carter–Wong chondral calcification theory, 336–337
- Cartilage
 - articular (*see* Articular cartilage)
 - composition of, 230
 - mechanical adaptability in, 336–339
 - mechanical properties of, 230–248
 - mechanical significance of, 70
 - remodeling and, 107
 - role of, in growth, 75
 - stress-strain curves for, 247
 - viscoelastic, 233
 - water and, 230–233
- Cell process, 63
- Cement line, 37, 48, 49, 50, 65, 101, 103, 121, 140, 141, 144, 148, 151, 153, 166, 374, 377, 381, 397–398, 402, 403, 424, 449, 451–454, 456–460, 462, 464, 465, 470, 471, 474
- Cement line radius, 140, 146, 153
- Center of rotation for joint, 20, 25, 268
- Characteristic damage state, 447, 460
- Characterizing parameter, 427
- Chelating agents, 64
- Chemotaxis, 141, 142
- Chondral calcification theory, 336–337
- Chondral modeling theory, 338–339
- Chondroblasts, 111, 112
- Chondrocalcin, 78
- Chondroclasts, 79
- Chondrocytes, 66, 67, 69, 70–74, 76–79, 82, 88, 195, 202, 230, 231, 234, 245, 250, 336, 339
- Chondrons, 66
- Chronological age, 80, 164
- Circumferential lamellar bone, 47–49
- Clear zone, 61
- Clinical problems, mechanical
 - adaptability and, 282, 303, 330
- Collagen
 - types of, 51, 68
- Collagenases, neutral, 84
- Collagen fiber organization, 396
- Collagen fiber orientations, 199, 234, 374, 378, 380–384, 393, 396, 398, 404, 411
- Collagen fibers, 37, 44–47, 52, 53, 55, 65, 67, 70, 72, 78, 81, 82, 84, 85, 101, 104, 175, 176, 190, 191, 195, 199, 206, 214, 216, 230, 233–234, 237, 246–247, 262, 376, 378–385, 393, 396, 434, 449
- Comminution, 370
- Compact bone, 37, 40–44, 47, 49, 50, 99, 105, 164, 374, 396, 401, 402, 410, 431
- Compliance, 189, 194, 200, 206, 356, 364, 365, 371, 372, 477
- Composite material, bone as, 409–410
- Compression, cracks in, 441, 461
- Compression creep test, confined, 240–242

- Compression relaxation test,
 - confined, 242–243
- Compression tests, effect of strain
 - rate on, 247–248
- Compressive load, 71, 267, 292, 327, 357, 460, 462, 474
- Compressive strain direction,
 - principal, steering in, 322
- Compressive strains, 239, 245, 290–295, 301, 306, 322, 385, 476
- Compressive strength, 53, 382, 389, 399, 401, 405, 414
- Computed tomography (CT) images, 57, 393
- Condylar process, 29
- Confined compression creep test, 240–242
- Confined compression relaxation test, 242–243
- Continuum damage theory for
 - osteonal bone, 459
- Continuum limits, 400–401
- Continuum mechanics, 319, 330, 369–370, 401
- Controlled crack propagation
 - studies, 452–453
- Coronoid process, 28
- Cortical bone
 - material properties of, 374–378
 - mechanical properties for
 - anisotropy of, 385
 - rheological model for, 398
 - slip bands in, 456
- Cortical bone osteons, 446
- Crack growth
 - fatigue and, 437–438
- Crack growth resistance, 432
- Crack initiation, 434, 446, 449, 456–457, 463
- Crack propagation
 - energy and, 431–432
 - osteon size and, 454
- Crack propagation studies,
 - controlled, 452–453
- Cracks
 - birth of, 445–446
 - kinds of, 427–428
 - matrix, 459, 460, 462
 - plastic deformation around, 432
 - self-limiting, 447, 448, 462, 463, 474
 - spiral, 371
 - in tension and compression, 459, 461
- Crack stopping, 457–459
- Cranial capacity, bipedalism and, 31
- Creep, 49, 180, 182, 183, 217, 218, 239–243, 247, 287, 335, 423, 424, 442–445, 456–459, 474
- Creep behavior of bone, 442–445, 456
- Creep damage formation rate, 442
- Crimp pattern, 82, 84
- Cross-sectional area occupied by
 - primary bone, 158
- Cross-sectional moment of inertia, 367, 369, 412, 413
- CT (computed tomography) images, 134, 393, 468
- Culmann, K., 276–279, 288, 335
- Culmann–Meyer stress-trajectory
 - hypothesis, 278
- Cytokines, 86, 107, 142
- Daily stress stimulus, 323
- Damage accumulation, 443, 447, 448, 458, 464
- Damage repair, mechanical
 - adaptability and, 335–336, 339
- Damage repair specificity factor, 471
- Darcy's law, 235
- Dead zones, 296, 326, 342, 343
- Decorin, 54, 55, 70, 81

- Deformation, rate of, 397–399
- Delayed union, 109, 112
- Density
 - apparent, 56, 323, 324, 326, 328, 330, 342, 390, 391, 395, 398, 400–403, 405, 407, 408, 411
- Diaphyseal modeling, 98, 319
- Diarthroidal joints, 227, 266
- Distal sesamoidian ligaments, equine, 24
- Distortional strain, 203, 205
- Disuse osteoporosis, 113, 335
- Dry mass, 56, 67, 83, 85, 231
- Dual energy photon
 - absorptiometry, 394
- Dual energy x-ray
 - absorptiometry, 384
- Dynamic fracture mechanics, 433–434
- Dynamics, 2, 17–21, 32, 44, 59, 87, 103, 119, 120, 123, 137, 140, 152–154, 165, 166, 181, 233, 244–247, 265, 267, 305–308, 314, 316, 338, 433–434
- Elastic cartilage, 66
- Elasticity, rubbery, 176, 198–199
- Elastic modulus
 - degradation of, 449
- Elastic response, 182–183, 443
- Elastohydrodynamic lubrication, 259–262
- Elbow joint
 - anatomical drawings of, 5
 - equilibrium problems with, 20
 - forces in, 4–9
- Elderly
 - bone loss in, 107
 - microdamage and bone fragility in, 467–468
- Electrical phenomena, fracture
 - healing, 119
- Electrostatics, 244, 265
- Endochondral ossification, 96, 109, 110, 112, 294
- Endosteal callus, 412
- Endosteal envelope, 129
- Endosteal remodeling, 106
- Endotenon, 85
- Energy
 - for breaking bone, 372
 - crack propagation and, 431–432
- Epifluorescent microscopes, 121
- Epiphysis, 76, 77, 79, 80
- Epitenon, 85
- Equal lamellar work areas principle, 148–150
- Equilibrium strain range, 295–297
- Equine fetlock forces, 23–26
- Erosion rate, 128
- Ethical concerns, 11
- Evolution, 17, 26, 31, 38, 40, 201, 264, 266, 278–280, 283–284, 300, 303, 308, 339, 355, 379
- Evolutionary adaptability
 - mechanical adaptability and, 280–284
 - somatic adaptation *versus*, 280–282
- Evolutionary effects, somatic effects *versus*, 282–284
- Exercise, passive motion and, 212
- Extensor tendons, 204–206, 209, 217
- Extracellular matrix, 53, 60, 66, 67, 69, 73, 74, 78, 81, 83, 185, 202, 230, 231
- Fabric tensor, 330, 407
- Failure
 - of osteonal bone, 449–450
 - regions before, 447–448
- Failure moment, 367
- Failure point, 356, 430

- Fairbairn crane, stress trajectories in, 277–279
- Fascicles, 85, 175, 184
- Fatigue
 - in activating remodeling, 463–468
 - crack growth and, 437–438, 449
 - creep behavior of bone and, 442–445
 - first stage of, 447
- Fatigue behavior
 - of bone, 438–442
 - of fiber-reinforced composite laminates, 445–448
- Fatigue challenge, 470, 472
- Fatigue damage
 - in bone, 424, 439–442, 468, 474
 - modeling effects of, in osteonal bone, 458–462
- Fatigue life
 - effects of remodeling on, 458–459
- Fatigue loading, 355, 381, 443
- Fatigue test, zero-tension, 444
- Feedback control systems, 473
- Femur midshaft cross sections, 298–299, 318
- Fetlock, 23–25, 29
- Fetlock forces, equine, 23–26
- Fiber diameter, material strength and, 446–447
- Fiber-reinforced composite laminates, fatigue behavior of, 445–448
- Fibrils, 39, 44, 52–55, 67–72, 77, 78, 81, 83, 175, 199, 231, 232, 234, 237, 445, 446, 452
- Fibroblasts, 60, 82–87, 104, 112, 195, 199, 202, 211, 262–263, 309
- Fibrocartilage, 66, 82, 194–196, 201, 203, 204
- Fibrocartilage disks, 66, 228
- Fibrochondrocyte, 202
- Fibromodulin, 70, 81
- Fibrous joint capsule, 227
- Fibrous lamellar composite material, osteonal bone as, 449–459
- Fine-fibered bone, 47
- Finite element modeling, 59, 203, 323–234, 325, 327, 328, 330, 336, 337, 342, 370, 373, 409
- Flat bones
 - modeling of, 99
- Flaw size, maximum tolerable, 431
- Flexor tendons
 - of hand, 216–217
- Flexural neutralization theory, 317, 322
- Fluid flows, 233, 242, 244, 246, 247, 260, 265, 313–316, 332, 335, 398
- Fluid flows, stress gradients and, 313, 332
- Fluorochrome labeling techniques, 120
- Fluorochromes, 65, 120, 121, 124, 126, 129, 142, 145–147, 162, 164
- Forces, 1, 35, 111, 175, 227, 276, 355, 427
- Forces in joints
 - high, clinical significance of, 13
 - reproducing, in laboratory, 15–17
 - static analysis of, 4–13
- Fracture healing
 - basic concepts of, 109
 - biological phases in, 111–113
 - biomechanical stages of, 114–116
 - clinical assessment of, 113–114
 - electrical phenomena and, 119
 - important tissues in, 109–111
 - inflammatory phase of, 111–112
 - primary, 118
 - remodeling phase of, 113
 - reparative phase of, 112

- Fracture healing (*cont.*)
 times for, 110
 torsional stiffness in, 115
- Fracture mechanics
 dynamic, 433–434
 linear elastic, 425–427, 436
- Fracture risk, minimization of, 284, 287, 302, 335, 336
- Fractures
 osteoporotic, 108, 468
 pathologic, 108, 109
- Fracture toughness
 fractured bone, self-straightening of, 434, 437
- Free-body diagram, 5, 6, 10, 25, 27, 29, 254, 358
- Friction, 23, 71, 85, 179, 206, 248–252, 258, 261, 264, 269, 398, 400
- Friction coefficients, 248–250, 261, 262
- Frost, H.M., 106, 120, 121, 129, 130, 135, 144, 153, 156, 164, 295, 297, 320, 334, 335, 338, 356, 377, 396, 441, 464
- Frost's chondral modeling theory, 338–339
- Frost's flexural neutralization theory, 322
- Frost's mechanostat theory, 296, 332
- Functional stimulus, 279, 280, 287
- GAGs (glycosaminoglycans), 67, 69, 71, 74, 232, 244–245
- Gap junctions, 62, 64, 66, 82
- Gender differences, 11
- Generalized Hooke's law, 362–365
- Gigapascals, 361
- Glycosaminoglycans (GAGs), 67, 69, 71, 74, 232, 244–245
- Golgi tendon organs, 61
- Ground reaction force, 25–26, 29, 199, 292
- Ground substance,
- Growth factors, 55, 67, 86, 107
- Growth plate, 66, 69, 76–80, 88, 96–98, 234, 317, 338
- Growth, role of cartilage in, 75
- Habituation, 294, 299–302
- Hand, flexor tendons of, 216–217
- Hashin–Shtrickman model, 410
- Haversian bone, 381, 382, 385–386, 388, 449, 451, 457
- Haversian canals (HC), 41, 44, 49, 50, 57, 88, 101, 102, 105, 108, 128, 135, 137, 146, 148, 153, 155–157, 166, 167, 296, 313, 374, 376, 380, 390, 395, 397, 410, 411, 451, 453, 462, 474
- Haversian envelope, 96, 108
- HC (haversian canals), 41, 44, 49, 50, 57, 88, 101, 102, 105, 108, 128, 135, 137, 146, 148, 153, 155–157, 166, 167, 296, 313, 374, 376, 380, 390, 395, 397, 410, 411, 451, 453, 462, 474
- Healing, fracture. *See* Fracture healing
- Helicoidal plywood structure, 45
- Hematocrit, 282
- Hematoma, 86, 110, 111
- Hematopoietic cells, 141, 142
- Hip joint
 forces in, 9–11
 in human ancestors, 13–17
 measurement of forces on, in vivo, 11–13
 surgical reconstruction of, 22–23
- Hip joint replacement prosthesis
 telemetry electronics inside, 12
- Histologic activation frequency, 134–138

- Histologic architecture, 395–396
- Histomorphometric measurement of
 - osteonal remodeling, 122–138, 164
- Histomorphometry, 119, 120, 123–125, 165
- Hooke, R., 360, 361
- Hooke's law
 - generalized, 362–365
- Hormonal activation, 154
- Horses, 23–26, 81, 177, 179, 199, 201, 208, 288, 290, 469
- Human bone, anisotropy of, 385, 386
- Hyaline cartilage, 66, 228, 337
- Hydraulic permeability
 - coefficient, 235
- Hydrodynamic lubrication, 252–263, 266
- Hydrostatic stress, 71, 202–205
- Hydroxyapatite crystals, 54
- Hypertrophic zone, 69, 78
- Hysteresis, 179, 184, 192, 217

- Impingement, 215, 460
- Indentation testing of articular cartilage, 235–237
- Index, 204, 281, 337
- Inflammatory phase of fracture healing, 111–112
- Insertions, 5–7, 12, 16, 18, 22, 24, 28, 29, 188, 190, 191, 194–197, 199–201, 204, 209–211, 213, 215, 218
- Integrin, 193, 286
- Intermediary organization, skeletal,
- Interstitial bone, 49, 449
- Intramembranous bone formation, 96, 109
- Intramembranous ossification, 96
- Involutional bone loss, 108
- Iron cross exercise, 29
- Isotropic materials, 363–365, 389

- Joint diseases, 229–230, 249, 258, 263
- Joint pain, 13, 22
- Joint problems, indeterminate, 19–23
- Joint reaction force, 3–4, 6, 7, 9–11, 16–17, 23, 25–27, 29
- Joints
 - center of rotation for, 25, 268
 - elbow (*see* Elbow joint)
 - forces in (*see* Forces in joints)
 - hip (*see* Hip joint)
 - knee (*see* Knee joint)
 - synovial (*see* Synovial joint)
- Junctions, 39, 62, 64, 66, 72, 73, 82, 87, 126, 152, 191, 193–197, 211, 234

- Kinematic, 2, 209
- Kinetics, 2, 205–206, 314, 413, 433
- Knee joint
 - walking up stairs,

- Label escape error, 129–130
- Lacunae, 55, 57, 62, 67, 69, 79, 146, 310, 313, 376, 390
- Lag time, mineralization, 101, 122, 126, 132, 392
- Lamellar bone, 44–50, 55, 109, 112, 113, 165, 195, 295, 333, 339, 396, 425, 472
- Lamellar pair width, 153
- Lamellar work areas principle, equal, 148–150
- Lamina splendens, 71
- Lazy zones, 296, 326
- Ligaments
 - age and mechanical behavior of, 206–208
 - determination of resting length of, 189

Ligaments (*cont.*)

- functional adaptation and specialization of, 198–206
 - functional considerations, 197
 - mathematical modeling of, 181–184
 - mechanical behavior of, 175, 178, 218
 - mechanical properties of, 175–220
 - mechanical testing of, 187
 - pathology and healing of, 210–211
 - quasistatic tensile properties of, 374
 - structure and composition of, 83–85
 - surgical repair of, 213
 - testing environment for, 184–186
 - viscoelastic properties of, 178–180, 182
- Ligament-to-bone attachments, 195
- Linear elastic fracture mechanics, 425–427, 433, 436
- Link proteins, 69, 232
- Load, compressive, 71, 267, 292, 327, 357, 460, 462, 474
- Load-deformation testing
- anisotropy effects determined by, 385–387
- Long bone diaphyses, 298, 315, 316, 320, 365, 366, 412, 413, 470
- Long bones, 36, 41–43, 80, 96–98, 107, 109, 116, 137, 154, 159, 290, 291, 300, 316, 317, 321, 322, 337, 339, 342, 363, 371, 388, 440, 474
- Lubrication
- boosted, 260, 263
 - boundary, 250–251, 258, 261–263, 266

- elastohydrodynamic, 259–262
 - hydrodynamic, 252–263, 266, 267
 - squeeze film, 251–252, 260, 264–265
 - of synovial joint, 248, 250, 253, 260–263, 266
 - types of, 250–260
 - weeping, 259–260, 263
- Lubricin, 263
- Macrophages, 60, 139, 141
- Manson–Waters theory, 148, 151
- MAR (mineral apposition rate), 74, 122, 126–129, 132, 139, 146–148, 151, 152, 472
- Marrow, 41–43, 50, 56–58, 61–64, 106–112, 132, 133, 142, 278, 313, 314, 330, 334, 390, 395, 400, 401, 405, 407
- types of cells in, 142
- Material strength, fiber diameter and, 446–447
- Mathematical models for remodeling, 119, 121, 123, 137
- Matrix cracks, 460, 462
- Matrix vesicles, 78
- Maximum tolerable flaw size, 431
- Mean intercept length (MIL), 289, 405–408
- Mechanical adaptability
- in cartilage, 336–339
 - clinical problems and, 283, 303, 331, 335
 - damage repair and, 335–336
 - evolutionary adaptability and, 280–284
 - of skeleton, 275–343
- Mechanical adaptability hypothesis, 328
- Mechanical demands, 35, 209, 283, 372

- Mechanical loading, 63, 73, 118, 119, 143, 218, 229, 233, 279, 280, 284, 287, 294, 297, 303, 310, 313, 316, 335, 337, 339, 364, 383
- Mechanical properties
 - of bone, 36, 38, 39, 54, 55, 96, 101, 109, 161, 297, 355–414, 423, 439, 440
 - of cancellous bone, 44, 288, 370, 401, 402, 407, 410, 411
 - of cartilage, 66, 230–248, 269, 337
 - of cortical bone, 368, 374, 380–385, 389, 390, 396–398, 410, 411
 - of ligaments and tendons, 175–220
- Mechanical significance of bone fracture patterns, 370–371
- Mechanics, solid, fundamentals of, 356–365
- Mechanoreceptors,
- Mechanostat theory, 295, 296, 332, 334
- Mechanotransduction, 63, 286, 287, 294, 340
- Medullary callus, 111, 112, 412
- Menisci, 66, 213, 228, 252, 264, 265
- Mesenchymal cells, 96, 112, 142
- Mesenchymal stem cells (MSC), 82, 142, 302
- Metaphyseal modeling, 97
- Metaphyseal trabeculae, development of, 79
- Metaphysis, 43, 76–80, 97, 98, 333
- Microcracks
 - stimulating remodeling, 439, 464
- Microdamage
 - in elderly, 467–468
 - fatigue, 396–397, 424, 447, 459
- Microdamage burdens, 441
- Microfibrils, 67, 83–84
- Micromodeling, 396
- Microstrain, 155, 288, 292, 306, 359, 439, 449, 450, 463
- Microtome, 36, 64, 65, 121
- MIL (mean intercept length), 289, 405–408
- Mineral apposition rate (MAR), 74, 122, 126–129, 132, 139, 146–148, 151, 152, 472
- Mineralization
 - specific, 392, 393, 410
 - volumetric, 384, 392
- Mineralization front, 62
- Mineralization lag time (Mlt), 101, 122, 126, 127, 132
- Mineral metabolism, 103
- Minerals in bone, 35, 36, 39, 53–55, 60, 61, 75, 102, 128, 167, 303, 307, 393–395
- Minimum effective strain, 295–296
- Mlt (mineralization lag time), 101, 122, 126, 127, 132
- MNPs (mononuclear phagocytes), 141
- Modeling
 - defined, 97–99
 - finite element, 59, 203, 323–328, 330, 336, 337, 342, 370, 373, 409
 - strength and, 107
 - stress fractures, 397, 468–474
 - term, 278
- Monocytes, 60, 79, 99
- Mononuclear phagocytes (MNPs), 141
- Monotonie, term, 355
- MSC (mesenchymal stem cells), 82, 142, 302
- Mucopolysaccharides,
- Muscles, 1, 2, 4–11, 13, 16, 17, 19–21, 24, 26–29, 31, 32, 42, 81, 83, 87, 107, 114, 175, 181, 182, 187–189, 192–195, 199–201, 206,

- 208, 213, 215–220, 229,
233, 264, 265, 283, 294,
309, 314, 337, 355,
371–373, 388, 398, 464,
469, 476
- abductor, 9–11, 13, 16, 27, 31,
32, 373
- Myotendinous junction,
193–194, 211
- Neutral collagenases, 84
- Neutral plane or axis, 367
- Noncollagenous proteins, 37,
107, 380
- Non-unions, 109, 112, 116, 119
- Normal loads, 108, 229, 358, 360
- OA (osteoarthritis), 9, 17, 69,
74–75, 87, 251–252, 259,
269, 297
- Olecranon, 4
- Optimal stiffness, 284
- Orthotropic materials, 288, 289,
363, 364, 386, 387
- Ossification
 - endochondral, 66, 96, 109, 110,
112, 294
 - intramembranous, 66, 109
- Osteoarthritis (OA), 9, 17, 69,
74–75, 87, 251–252, 259,
269, 297
- Osteoarthrosis, 69, 71
- Osteoblast precursors, 141–143,
150–151
- Osteoblasts
 - conversion of, to osteocytes, 152
 - generations of, 145, 150
 - origins of, 143
- Osteocalcin, 55, 62, 103
- Osteoclasts
 - lifetime of, 141, 144
- Osteocytes
 - conversion of osteoblasts to, 152
- Osteoid, 55, 61, 62, 65, 101, 102,
105, 107, 120–122,
126–128, 130–132, 141,
143, 144, 151, 152, 155,
383, 392, 397
- Osteomalacia, 102, 392
- Osteonal BMUs
 - rib, remodeling data for, 129
 - stresses around, 321–322
- Osteonal bone
 - continuum damage theory
for, 459
 - determinants of mechanical
properties of, 374
 - failure of, 460–461
 - as fibrous lamellar composite
material, 449–459
 - modeling fatigue damage effects
in, 459–464
- Osteonal damage rate
 - coefficient, 461
- Osteonal origins and trajectories,
- Osteonal overlapping, 155
- Osteonal pullout, mathematical
analysis of, 453–454
- Osteonal remodeling
 - histomorphometric measurement
of, 122–138
 - long-term effects of, 154–164
- Osteonal resorption, 88, 128, 146,
150–151, 322–323
- Osteonal tunneling directions, 320
- Osteon area ratio, 157
- Osteonectin, 55
- Osteon-matrix damage interaction
parameter, 461
- Osteon population density
measurements, 159
- Osteons
 - compressive properties of, 376
 - cortical bone, 380–385, 404, 446
 - effects of, on bone mechanical
properties, 38

- individual secondary, properties
 - of, 374–377
- lifetime of, 37, 144
- mechanical properties of different
 - types of, 379
- primary, 47–49
- refilling, 130, 470
- secondary, 37, 49, 51, 101, 104,
 - 105, 126, 154, 159, 167,
 - 197, 374–377, 386, 411,
 - 412, 424, 451, 455, 465
- shear properties of, 376–377
- tensile properties of, 374–376
- Osteon size, crack propagation
 - and, 454
- Osteon types, 46, 378, 381
- Osteopetrosis, 142, 285, 388
- Osteophytes, 229
- Osteopontin, 55, 314, 439–440
- Osteoporosis
 - disuse, 113
 - posttraumatic, 118
- Osteoporotic fractures, 108, 468
- Osteotomy experiments, 341

- Pain, joint, 13, 22
- Paratenon,
- Parathyroid hormone (PTH), 64,
 - 107, 297, 335
- Passive joint stabilizers, 217
- Passive motion, exercise and,
 - 212–213
- Pastern joint, equine, 25
- Pathologic fractures, 108–109
- Pauwels' stress magnitude theory,
 - 317–318
- PCA (physiologic cross-sectional
 - area), 20, 21, 218
- Peak isometric muscle force,
 - 181, 218
- Peak strains, 289–290, 292,
 - 305–308, 310, 312,
 - 324, 463
- Pendulum experiment,
 - 260–261, 269
- Penetration test, 401
- Periodontal ligament, 143, 144
- Periosteal callus, 110–113, 118, 412,
 - 468, 469, 472, 475
- Periosteal envelope, 108
- Periosteal membrane, 96
- Periosteal remodeling, 106
- Periosteum, 65, 109–112, 142, 195,
 - 300, 314
- Permeability
 - apparent, 235
 - of articular cartilage, 234,
 - 239, 242
- PGs (proteoglycans), 54, 55, 65–72,
 - 74, 78, 81, 83, 85, 199,
 - 201–202, 231, 232, 237,
 - 247, 262, 265
- Phagocytes, mononuclear
 - (MNP), 141
- Photon absorptiometry, 392,
 - 394, 413
- Physiologic cross-sectional area
 - (PCA), 20, 21, 218
- Physiologic testing fluids, 185
- Physis
 - closure of, 80, 88
 - growth of, 80, 97
- Piezoelectric molecules, 119
- Plastic deformation around cracks,
 - 432, 446
- Plexiform bone, 49, 50, 332–333,
 - 388, 396, 456
- PMMA (polymethyl methacrylate),
 - 22, 23, 428
- Poisson's ratio, 203, 235–237, 243,
 - 270, 359, 363, 364, 368,
 - 373, 388, 433
- Polar moment of inertia, 366,
 - 367, 369
- Polymer mechanics, 247,
 - 265–266

- Polymethyl methacrylate (PMMA),
22, 23, 428
- Porcine exercise experiment, 381
- Porosity
of bone, 43, 56, 57, 128, 157,
166, 295, 305, 374, 383,
390–393, 395, 396,
398–401, 410–412, 472
strength and, 390, 393
- Porous material, cancellous bone
as, 401
- Posttraumatic osteoporosis, 118
- Primary bone
cross-sectional area occupied
by, 158
- Primary fracture healing, 118
- Primary Haversian canal, 49
- Primary mineralization, 101, 126
- Primary osteons, 48
- Primary spongiosa, 79
- Principal strain directions, 155, 288,
291, 315, 320, 360
- Principal stress directions, 276,
278, 288, 289, 329,
330, 360
- Proliferation, 55, 77–78, 86, 87,
107, 112, 211, 306,
314, 470
- Proliferative zone, 77
- Proportional limit, 356, 360
- Proprioception, 83, 217
- Prosthesis, hip joint replacement
telemetry electronics inside, 12
- Proteoglycans (PGs), 54, 55, 65–72,
74, 78, 81, 83, 85, 199,
201–202, 231, 232, 237,
247, 262, 265
- Provisional calcification, zone of, 78
- Provisional callus, 112, 115
- Proximal sesamoid bones, equine, 24
- Pseudoarthrosis, 109, 112
- PTH (parathyroid hormone), 64,
107, 297, 335
- Racehorses, 23, 49, 155, 186, 219,
290, 383, 424, 454, 469
- Radius, 4, 5, 7, 23, 81, 99, 106, 110,
140, 144, 146, 149, 153,
159, 197, 235, 243, 268,
288, 290, 291, 312,
314, 343, 366, 431,
454, 471, 472
- Random remodeling, 464, 471
- Random remodeling assumption,
155–156
- RAP (regional acceleratory
phenomenon), 108,
118, 141
- Rapid loading torsion tester,
116–117
- Rat tibia bending experiments, 308
- Reduced relaxation function, 183
- References, 76, 106, 125, 165, 276,
290, 317, 319, 335, 441
- Refilling osteons, 130, 470
- Refilling process
apposition rates during, 147–148
- Regional acceleratory phenomenon
(RAP), 108, 118, 141
- Remodeling
analysis of, 95–168
cartilage and, 107
data for rib osteonal BMUs, 129
defined, 96
details of, 138–154
effects of, on fatigue life, 451, 458
fatigue in activating, 463–468
mathematical models for,
121, 123
microcracks stimulating,
464–465
osteonal (*see* Osteonal
remodeling)
random, 464
strength and, 374
theoretical approach to, 162–164
use of word,

- Remodeling cycle duration, 102–104
- Remodeling phase of fracture healing, 113
- Reparative phase of fracture healing, 112, 115
- Reserve/resting zone, 77
- Residual strain, 180
- Resorption, 41, 44, 49, 57, 61, 64, 74, 75, 79, 88, 95–105, 107, 108, 118, 120, 122, 123, 125, 128–129, 131, 132, 135, 138, 139–144, 146, 150–151, 152, 154, 156, 158, 165, 166, 250, 286, 295, 306, 309–311, 315, 316, 318, 319, 321–323, 329, 332–334, 341, 342, 390, 465, 468, 470
- Resorption cavities, 44, 57, 118, 135, 323, 390, 465, 470
- Resting length, determination of, 189
- Reuss model, 409
- Reynolds' equation, derivation of, 254
- Rheological model for cortical bone, 398
- Rib, activation frequency in, 154, 162–163
- Rib osteonal BMUs, remodeling data for, 129
- Rigidity, 112, 303, 356, 366
- Rodan theory,
- Rosette strain gauges, 288
- Rotation, center of, for joints, 25, 200, 268
- Rotator cuff, 87, 215–216, 229
- Rubbery elasticity, 176, 198, 199
- Ruffled border, 60, 61
- Scalars, 6, 319, 323, 328, 332, 336–337
- Secondary bone, 47–51, 160, 376, 382, 396, 449, 451, 458
- Secondary Haversian systems, 381
- Secondary healing, 118
- Secondary mineralization, 101
- Secondary osteons, 37, 49, 51, 101, 104, 105, 126, 157, 159, 167, 197, 374–377, 386, 411, 412, 424, 451, 455, 465
- Secondary spongiosa, 79
- Secretory vesicles, 61
- SED (strain energy density), 205, 219, 323, 325–329, 330, 332, 339, 342
- Self-limiting cracks, 447, 460–463
- Self-regulation, 276, 279, 280
- Self-straightening of fractured bone, 434, 437
- Self-trabeculating models, 326–331
- Senile bone loss, 108. *See also* Elderly
- Septa, 79
- Septoclasts,
- Sharpey's fibers, 195, 197
- Shear loads, 336, 358, 360–361, 452
- Shear modulus, 246, 360–361, 363, 366–368, 379
- Shin splints, 290
- Skeletal age, 80, 156, 162
- Skeletal biology, 35–89, 123
- Skeletal envelopes, 96, 108
- Skeletal intermediary organization,
- Skeletal tissue mechanics, course called, 355
- Skeletal tissue mechanics, introduction to, 355
- Skeletal tissues
 - biologic role of, 1
- Skeleton
 - mechanical adaptability of, 275–343
 - somatic change in, 282

- Slip bands in cortical bone, 456
- Slipped capital femoral epiphysis, 79–80
- Slippery tissues, gripping, 187–189
- S-N curve, 438–439
- Solid mechanics, fundamentals of, 356–365
- Somatic adaptation, evolutionary adaptation *versus*, 280–282
- Somatic change in skeleton, 282
- Somatic effects, evolutionary effects *versus*, 282–284
- Sound velocity, 114, 387, 388
- Specific mineralization, 392, 393, 410
- Specific surface,
- Specimen cross-sectional area, measurement of, 186
- Specimen elongation, measurement of, 190–192
- Spiral cracks, 371
- Spongy bone, 43, 79
- Sprains, 86, 195, 210
- Squeeze film lubrication, 251–252, 260, 264–265
- Standard notation for bone histomorphometry, 165
- Static equilibrium, equations for, 5
- Statics, 2, 4–13, 15–20, 26, 27, 29, 40, 175–178, 245, 249, 250, 265, 305–306, 308, 314, 338, 358, 457
- Stem cells, 41, 42, 82, 88, 108, 111, 112, 142, 143
- Stereology, 38, 57, 124, 132, 138, 168
- Stiffness
 - optimal, 230
- Strain(s)
 - compressive, 239, 245, 290–295, 301, 306, 322–323, 385, 476
 - distortional, 203, 205
 - measurement of, in living animals, 287–288
 - peak, 155, 289–290, 292, 305–308, 310, 312, 324, 463
 - tensile, 55, 238, 288, 290–295, 306, 312–313, 457
- Strain direction(s)
 - principal, 155, 288, 291, 315, 320, 360
 - principal compressive, steering in, 322–323
- Strain energy density (SED), 205, 219, 323, 325–329, 330, 332, 339, 342
- Strain energy release rate, 432
- Strain gauges
 - rosette, 288, 291
- Strain rate
 - effect of
 - on compression tests, 247–248
 - on tensile tests, 246–247
- Strain-time curve, 307, 442
- Streaming potentials, 119, 286, 311
- Strength
 - compressive, 53, 71, 75, 203, 204, 267, 321–323, 338, 341, 370, 383, 391, 414, 441, 443, 459
 - modeling and remodeling and, 468–474
 - porosity and, 390
 - of whole bone, determinants of, 365–373
- Stress
 - around osteonal BMU, 321–322
 - bending and, 317, 338, 367, 458
 - hydrostatic, 71, 202–204, 336, 337
- Stress components, 358, 362, 427, 433
- Stress concentration, 152, 234, 246, 320, 321, 376, 377, 400, 425, 426, 441, 446, 452, 453, 456, 460, 468, 474

- Stress directions, principal, 276, 278, 280, 288, 329, 330, 339, 360, 448
- Stress fractures
 - modeling, 468–474
- Stress-generated potentials, 119
- Stress gradients, fluid flows and, 313
- Stress intensity, 426, 436
- Stress intensity factor, 427, 428, 431, 436, 437, 474, 475, 477
- Stress magnitude theory, 317–318
- Stress relaxation, 180, 183, 184, 239, 242, 244
- Stress stimulus, daily, 323
- Stress-strain curves
 - for cancellous bone, 399–400
 - for cartilage, 238, 247
- Stress trajectories in Fairbairn crane, 276, 277
- Structure, 4, 36, 96, 177, 229, 275, 355, 423
- The Structure, function and adaptation of cortical bone, 283
- Subfibrils, 84, 175
- Superficial tangential zone, 234
- Supraspinatus tendon, 215–216
- Surgical reconstruction of hip joint, 22–23
- Suspensory apparatus, equine, 24
- Suspensory ligament, equine, 24
- Synovial joint
 - diversity of architecture of, 263–265
 - functions of, 228–229
 - lubrication of, 260–263
- Synovial joint mechanics, 227–270
- Synovial membrane, 206, 227, 250
- Synovial sheath, 85, 206, 217
- Talin, 193
- Tangential zone
 - superficial, 234
- Targeted remodeling, 464–467
- Technical constants, 363, 364
- Telemetry electronics inside hip prosthesis, 12
- Temporalis muscle, 28, 29
- Temporomandibular joint (TMJ), 28, 29, 32, 264
- Tendons
 - age and mechanical behavior of, 206–208
 - determinants of architecture of, 200–201
 - determination of resting length of, 189
 - flexor (*see* Flexor tendons)
 - as flexors and extensors, 204–206
 - functional adaptation and specialization of, 198–206
 - functional considerations, 193–197
 - mathematical modeling of, 181–184
 - mechanical behavior of, 206–208
 - mechanical properties of, 175–220
 - mechanical testing of, 187, 217
 - pathology and healing of, 210–217
 - quasi static tensile properties of, 175–178
 - sites of compression of, 201–204
 - structure and composition of, 83
 - surgical repair of, 211–213
 - testing environment for, 184–186
 - viscoelastic properties of, 178–180, 182
- Tendon slack length, 181
- Tendon-to-bone attachments, 218
- Tenocytes, 85, 87
- Tensile fatigue tests, 458
- Tensile strains, 55, 238, 288, 290–295, 306, 312–313, 457
- Tensile strength of bone, 178

- Tensile tests
 - of articular cartilage, 237, 238
 - effect of strain rate on, 246–247
- Tension, cracks in, 428, 459, 462–463
- Tetracycline, 120, 140, 144, 162, 163, 167
- Thixotropic fluid, 253, 258
- Three-force rule,
- Tibiofemoral joint at knee, 265
- Tidemark, 68, 72, 195, 196, 234
- Tissue mechanics, introduction to, 38, 266
- TMJ (temporomandibular joint), 28, 29, 32, 264
- Tolerable flaw size, maximum, 431
- Torsional shape and size
 - factor, 366
- Torsional stiffness in fracture
 - healing, 115
- Torsion tester, rapid loading, 116–117
- Total hip arthroplasty, 22
- Toughness
 - bone, 53, 55, 285, 372, 435, 442, 450–451
 - fracture, 285, 428, 431, 434, 436–438, 451, 457
- Trabeculae
 - metaphyseal, development of, 79
- Trabecular bone, 51, 56, 104, 111, 121, 140, 144, 151, 154, 276, 283, 313, 316, 323, 333, 339, 340, 390, 400–405, 440, 441
- Trabecular envelope, 105
- Trabecular remodeling, 105–106, 333, 404
- Transitional zone, 68, 71, 72
- Transversely isotropic materials, 365, 389
- Trapeziometacarpal joint of
 - thumb, 264
- Tribology, 266
- True activation frequency, 137
- Ulna, 4, 5, 7, 114, 189, 288, 290, 291, 293, 302, 305–307, 312, 324, 333, 343, 464, 465
- Ultimate load, 356, 399, 411
- Ultrasonic measurements,
 - anisotropy effects
 - determined by, 387–389
- Ultrasonic wave, 113
- Vector quantities, 6
- Vinculin, 193
- Viscoelastic cartilage, 233, 247
- Viscoelasticity, 65–67, 116, 178–189, 217, 233, 236, 239, 240, 244, 245, 247, 265, 397, 398, 410, 434, 442, 456
- Vitronectin, 193
- Void volume fraction, 390
- Voigt model, 342, 409, 410
- Volkmann's canals, 44, 57, 63, 137, 461
- Volume fraction measurements, 128
- Volumetric mineralization, 384, 392
- von Meyer, G.H., 276, 277, 289, 340
- Walking, 1–4, 9, 10, 12–14, 23, 32, 192, 199, 288, 291, 307–309, 423, 463, 476
- Water
 - bone and, 51, 55, 402
 - cartilage and, 65, 67, 71, 230–234, 242, 243, 251, 259, 260, 265, 294
- Wear, 23, 214, 228, 246, 250, 258
- Wedge-shaped fragments, 371
- Weeping lubrication, 259, 260
- Whole bones
 - determinants of strength of, 365–373
 - mechanical failure of, 371–373

- Wolff's law, 275, 276, 278,
302–303, 321–322, 330,
339–340, 343
historical context of, 276–280
- Work, 8, 13, 21, 26, 27, 29, 36, 38,
60, 95, 99, 105, 111, 119,
121, 134, 137, 138, 144,
145, 148–150, 152, 153,
164, 165, 176, 179, 186,
187, 192, 204, 206, 229,
236, 247, 252, 253, 262,
281, 283, 285, 287, 289,
292, 294, 299, 304, 305,
309–310, 313, 316–317,
332, 337, 343, 356,
361–362, 371, 373, 374,
382, 383, 385, 398, 405,
408, 411, 431, 445,
451–455, 463, 464, 475
- Woven bone, 44–47, 109–113,
293, 295–297, 306, 312,
333, 334, 343, 396, 468,
470, 472
- Yield point, 176, 177, 285, 356,
357, 408
- Yield strain, invariance of,
408–409
- Young's modulus, 192, 208, 236,
237, 243, 360, 361,
364–365
- Young, T., 361
- Zero-tension fatigue test,
444, 445
- Zone of Ranvier, 79
- Zygomatic arch, 28–29