BIOLOGICAL ANTHROPOLOGY
OF THE HUMAN SKELETON
To our families

to Steve and Marty
M. Anne Katzenberg

to Victor, Rob, and Barb
Shelley R. Saunders
# CONTENTS

PREFACE TO THE SECOND EDITION—M. ANNE KATZENBERG AND SHELLEY R. SAUNDERS  xi
PREFACE TO THE FIRST EDITION—M. ANNE KATZENBERG AND SHELLEY R. SAUNDERS xv
ACKNOWLEDGMENTS  xxi
CONTRIBUTORS  xxiii
FOREWORD—JANE E. BUIKSTRA  xxxiii

## PART I  THEORY AND APPLICATION IN STUDIES OF PAST PEOPLES  1

1  Bioarchaeological Ethics: A Historical Perspective on the Value of Human Remains  3  
   Phillip L. Walker

2  Forensic Anthropology: Methodology and Diversity of Applications  41  
   Douglas H. Ubelaker

3  Taphonomy and the Nature of Archaeological Assemblages  71  
   Ann L. W. Stodder
PART II MORPHOLOGICAL AND DEVELOPMENTAL ANALYSES 115

4 Juvenile Skeletons and Growth-Related Studies 117
Shelley R. Saunders

5 Histomorphometry of Human Cortical Bone: Applications to Age Estimation 149
Alexander G. Robling and Sam D. Stout

6 Biomechanical Analyses of Archaeological Human Skeletons 183
Christopher B. Ruff

7 Morphometrics and Biological Anthropology in the Postgenomic Age 207
Benedikt Hallgrimsson, Miriam Leah Zelditch, Trish E. Parsons, Erika Kristensen, Nathan M. Young, and Steven K. Boyd

8 Reading Between the Lines: Dental Development and Subadult Age Assessment Using the Microstructural Growth Markers of Teeth 237
Charles M. FitzGerald and Jerome C. Rose

9 Dental Morphology 265
G. Richard Scott

PART III PREHISTORIC HEALTH AND DISEASE 299

10 Dental Pathology 301
Simon Hillson

11 Analysis and Interpretation of Skeletal Trauma 341
Nancy C. Lovell

12 Light and Broken Bones: Examining and Interpreting Bone Loss and Osteoporosis in Past Populations 387
Sabrina C. Agarwal

PART IV CHEMICAL AND GENETIC ANALYSES OF HARD TISSUES 411

13 Stable Isotope Analysis: A Tool for Studying Past Diet, Demography, and Life History 413
M. Anne Katzenberg
14 Bone Chemistry and Trace Element Analysis 443
   James Burton

15 DNA Analysis of Archaeological Remains 461
   Anne C. Stone

PART V QUANTITATIVE METHODS AND POPULATION STUDIES 485

16 Metric Analysis of Skeletal Remains: Methods and Applications 487
   Michael Pietrusewsky

17 Nonmetric Trait Variation in the Skeleton: Abnormalities, Anomalies, and Atavisms 533
   Shelley R. Saunders and Dori L. Rainey

18 Advances in Paleodemography 561
   George R. Milner, James W. Wood, and Jesper L. Boldsen

19 Method and Theory in Paleodemography, with an Application to a Hunting, Fishing and Gathering Village from the Late Eastern Woodlands of North America 601
   Richard S. Meindl, Robert P. Mensforth, and C. Owen Lovejoy

Index 619
The preface to the first edition sets out the goals we hoped to accomplish by preparing the volume titled, *Biological Anthropology of the Human Skeleton*. In this second edition we endeavor to maintain those goals, to update recent developments in skeletal studies and particularly, to emphasize information that provides the reader with a basic understanding of the various techniques and methods of investigating bones and teeth. Many chapters include examples, set off from the main body of the text, that illustrate or offer more detail about the particular analysis under consideration. We also provide six new chapters on topics not covered in the previous edition. These topics include taphonomic factors affecting burial assemblages, nonmetric traits of the skeleton and dentition, trauma, osteoporosis, and new developments in morphometric analysis. It is our hope that the book will be used in upper level undergraduate and graduate courses in human skeletal studies (e.g., advanced human osteology) as well as practical guidelines, applications, and critical reviews of research approaches, including a wealth of selected references for additional reading.

The book is divided into five parts, although considerable overlap exists and some chapters could have easily appeared in one or another section. Part I is titled “Theory and Application in Studies of Past People” and includes three chapters. As before, the first chapter, on ethical considerations of working with human skeletal remains, is presented by Phillip Walker. Walker has updated and expanded the scope of his chapter to include worldwide examples of problems and solutions to working with human remains. The second chapter, by Douglas Ubelaker, provides current perspectives on the interrelationship between forensic anthropology and more traditional studies in human osteology, arguing that these are complementary fields of inquiry. The third chapter, by Douglas Ubelaker, provides current perspectives on the interrelationship between forensic anthropology and more traditional studies in human osteology, arguing that these are complementary fields of inquiry. The field of forensic anthropology has gained prominence since the publication of the first edition of this book, with an increase in the number of academic positions in the field, and increased participation by forensic anthropologists in medico-legal investigations. The third chapter
in the first section is a new contribution by Ann L. Stodder, on taphonomy and human skeletal remains. Stodder draws from her experience in working with burials and specific burial contexts in several different regions of the world to offer a comprehensive review of the various postmortem factors that affect the integrity of the skeleton after death.

Part II is newly titled “Morphological and Developmental Analyses” and includes five chapters on development and modeling of bones and teeth. “Juvenile Skeletons and Growth Related Studies,” by Shelley Saunders, examines the problems of studying juvenile skeletal remains from archaeological sites. This chapter has been updated with new examples of applications from Saunders’ extensive work with historic cemeteries. Alexander Robling and Sam Stout have revised their previous contribution on histomorphometry and, once again, provide helpful appendices, including a worked example of age determination from cortical bone histology and a compilation of various histomorphometric techniques for age determination from various skeletal elements. Christopher Ruff has updated his chapter on biomechanical analyses, providing new examples and illustrations. Benedikt Hallgrimsson and colleagues present a new chapter on the “new morphometrics” and the importance of understanding the interface between morphometric studies in the biological sciences and those studies in biological anthropology in the context of a more solid understanding of genetic mechanisms and their role in determining phenotypic variation. The chapter on dental microstructure, by Charles FitzGerald and Jerome Rose, retains the clear descriptions of the microscopic structure of teeth and the events recorded in dental microstructures, including evidence for stress and for age determination. Examples of more recent applications and new technological developments have been added. The final chapter in this section is a new contribution by Richard Scott on dental morphology, specifically dental nonmetric traits. The chapter provides descriptions of various morphological variants of the teeth and very practical advice on how to recognize and record dental crown traits so that they can be used in population studies.

Part III, “Prehistoric Health and Disease,” includes three chapters. Simon Hillson has updated his previous chapter on dental pathology, retaining a protocol for data collection and updating the state of our understanding of the causes and implications of pathological conditions of the teeth and supporting structures. Nancy Lovell provides a new chapter on skeletal trauma. This chapter focuses on fractures but also includes more general information on responses of bone to trauma and diagnostic procedures for evaluating trauma in the past. The third chapter is also new to the second edition. Sabrina Agarwal presents information on osteoporosis in past populations with both cross-cultural and historical perspectives. She offers a very useful comparison of the advantages and disadvantages of the different methodological techniques for obtaining information on bone mass, density, and quality and their relevance to the study of past populations.

Part IV, “Chemical and Genetic Analyses of Hard Tissues,” as in the previous edition, includes three chapters. Anne Katzenberg describes methods and applications of stable isotope analysis that are used to reconstruct diet, estimate the duration of nursing, and determine residence and migration patterns of the past. This field has expanded considerably since the previous edition. James Burton discusses bone chemistry and the trace elements of bone that have been used to reconstruct past diet as well as studies focusing on postmortem alterations of bone chemistry. Anne Stone provides background and examples of ancient DNA studies from human remains. She illustrates the significant challenges of working with ancient DNA, the fact that it is highly subject to destruction and contamination and expensive to analyze. She also points out that hypotheses about the genetics of populations in the past must be consistent with what is known about modern populations and that...
research projects require the coordination of expertise from many different fields from paleopathology to archaeology.

Part V includes four chapters on quantitative methods and population studies. Michael Pietrusewsky has updated his contribution on metric analysis, focusing on craniometric studies for population reconstruction. He provides an example from his extensive work in Polynesia. Shelley Saunders and Dori Rainey provide a new chapter on skeletal non-metric traits. They critically review the background of such studies and include illustrations of many of the more commonly used traits. They offer suggestions for future areas of research in this field, including study of the ontogenic development of specific traits and the relationship between the prevalence of traits in past populations and information on other skeletal criteria such as DNA studies. As in the previous edition, there are two chapters on paleodemography. George Milner, James Wood, and Jesper Boldsen have updated their previous contribution focusing on both the promises and the limitations of paleodemographic studies. A new chapter by Richard Meindl, Robert Mensforth, and Owen Lovejoy provides a detailed example of a particular paleodemographic study, from the Libben site, in northern Ohio.

As with the previous edition, we hope that this volume will provide useful information for both current and future biological anthropologists interested in the latest research on human skeletal and dental remains.

M. ANNE KATZENBERG
SHELLEY R. SAUNDERS
What’s Bred in the Bone, a novel by Robertson Davies, begins with the proverb, “What’s bred in the bone will not out in the flesh.” The story is about a man who supposedly reflects his “breeding” since his behavior and characteristics are direct reflections of what he has inherited from his family. Although biological determinism may work in fiction, it is anathema to the biological anthropologist. The cornerstone of biological anthropology is the interaction of culture and human biology. What is manifested in the physical and behavioral characteristics of any living being is a result of the intertwining of an inherited genome with environmental factors. Human osteologists have struggled with this concept from the earliest beginnings of skeletal studies and continue to struggle with it today. Ancient DNA studies suggest that we ultimately want to know the “inherent” properties coming out of the bones. If we could read the genome, we would “know” the person. But of course, we understand that, as living tissues, bones and teeth are influenced by environmental forces. Bones respond to mechanical forces, and thus, they alter in response to activities and stresses. Cranio- metric studies attempt to study population relationships, assuming that cranial shape and size reflect inherited features, but we know that cranial shape and size can be altered purposefully (head binding) or unintentionally (chewing stresses). It is the job of the human osteologist to study the interactions between inherited characteristics and their modification by the environment in order to understand, not just what is “bred” in the bone but also what bones can tell us about the flesh, that is, the lives of earlier peoples.

Each of the following chapters deals with a specific type of advanced analysis of bones and teeth. The original plan for the book was to be a second edition of our earlier edited book, Skeletal Biology of Past People: Research Methods. However, as work progressed, it seemed that with five additional chapters and many new contributors, it is really something different. The differences are directly related to changes that have occurred in the analysis of human skeletal and dental remains over the past few years. Most notably these changes include heightened ethical concerns about studying the skeletal remains of aboriginal peoples in many countries where those people are no longer the dominant culture. These concerns and the resulting legislation in some jurisdictions have radically changed the way physical anthropologists and archaeologists carry out their work. A second change is the
The rise of forensic anthropology and the fact that research in forensic anthropology, while still overlapping with more traditional approaches, now includes topics not central to studies of archaeological skeletons. We begin this book with chapters on the ethics of studying human remains and forensic anthropology.

An important theme that is found throughout the book is the progress of new methods. We were training to become anthropologists in the 1970s when many new research areas were emerging in physical anthropology. The earlier practice of providing descriptive osteological reports either as stand-alone works or, more commonly, as appendices to archaeological site reports was fading out and more problem-oriented research was emerging. Biological distance studies using both metric and non-metric traits on human bones and teeth were carried out in order to investigate prehistoric migration and relatedness through time and space. Paleopathology was emerging as a means of addressing questions about prehistoric adaptations in contrast to the earlier emphasis on unusual cases of specific diseases. Paleodemography, similarly, addressed questions of adaptation of earlier populations. Since the initial enthusiastic studies all of these topics have undergone criticism and have emerged as, perhaps, humbled, but also strengthened, by the critiques. The same is true of the more recently introduced methods involving biochemical analyses of bones and teeth. These methods include analyses of trace elements, stable isotopes, and ancient DNA.

Each of these methods has undergone a series of stages that may be characterized as follows:

- Discovery—either entirely new or new to physical anthropology, a new method is discovered and the potential applications are explored.
- Applications to questions of interest regarding reconstructing past peoples.
- Critique, introspection, experimentation.
- Emergence in a stronger, more reasoned form.

NAGPRA (Native American Graves Protection and Repatriation Act) and similar legislation in other countries have led to a reconfiguring of how skeletal studies of past peoples are carried out. Some of these changes can be viewed in a positive light. For example, standards have been developed in the expectation that collections will not be curated indefinitely. These standards were needed even before the prospect of reburial emerged. In addition, an interesting configuration of events happened in the 1990s. As some Native Americans voiced their disapproval of skeletal studies, expanding urban development led to archaeological excavations of several large, historic cemeteries dating from the eighteenth and nineteenth centuries. These cemeteries contained the remains of Euro-americans and African Americans as well as other groups. At the same time, the growing number of students trained in human osteology provided a pool of individuals to excavate and study these remains. Debates about excavation and study continued but in many cases some period of time was allowed for proper scientific study. One special example of the cooperation between scientists and concerned descendants is the work being conducted at Howard University on a large African-American slave cemetery discovered in New York City. In Europe, there is a long history of excavating historic cemeteries and the increasing number of trained human osteologists has led to larger scale studies (the St. Brides’ skeletal collection in London, England is a good example). The increased scientific study of skeletons from historic cemeteries has also provided opportunities for testing methods. In many cases, the identities of at least some individuals are known from legible coffin inscriptions or detailed cemetery maps. It has been possible to investigate the accuracy of methods of determining sex and age at death and to detect biases in mortality samples that are directly related to causes of death.

This book is organized into five parts. Part I, theory and application, features two chapters that describe recent shifts in skeletal studies. Walker’s chapter provides information on how
humans have regarded the dead over time and across cultures. He grapples with the issues surrounding the ethics of skeletal research, the clash with cultural beliefs about treatment of the dead, and the politics of communities. Taking a clearly anthropological approach to these questions, he shows us that there is a tremendous diversity of attitudes about the physical remains of the dead. He makes a strong case for the value of and the justification of scientific research. Ubelaker focuses on the development of forensic anthropology with its roots in descriptive osteology and its current form as an applied specialization of human osteology. He discusses the major comparative collections used for establishing standards, including the recently developed forensic data bank. He then takes the reader through the various steps in forensic anthropology, including recovery, identification, sex and age determination, stature estimation, and positive identification. He concludes with information on training opportunities and professional organizations dedicated to forensic anthropology.

Part II includes chapters on morphological analyses of bones and teeth and age changes. Four of these contributors prepared chapters for our earlier book, and although the topics are similar, each chapter includes contributions and advances that have occurred throughout the 1990s. Ruff describes biomechanical analyses of bones and the applications of such studies to understanding past human behavior ranging from fossil hominids through to early historic human groups. He draws from his own extensive research to provide examples of how biomechanical studies have improved our understanding of past activity patterns. Examples include changes in robusticity throughout human evolution, the relationship between subsistence and bone strength, and the relationship between gender roles and their biological manifestation in bone structure. Mayhall covers dental morphology highlighting newer methods of characterizing tooth size and shape, and the applications of such studies to biological and behavioral characteristics of past peoples. He emphasizes the importance of achieving precision of observations of both dental measures and dental morphological traits. He also argues for maintaining simplicity in our methodological approaches. Both of these aspects of the research process are absolutely necessary for us to make meaningful comparisons of the results obtained by different observers. Mayhall shows that knowledge in the field of dental morphology remains limited because the precision necessary for properly evaluating population variability has still not been achieved. Saunders covers the various types of studies that are specific to subadults, focusing on age determination but also considering sex determination and variations in growth and development. One problem with proceeding to studies of growth and development is that of sampling. Differential burial practices, differential preservation, and biases related to cause of death can all cause problems in assessing past growth patterns from subadult burials. Some of these problems have been addressed in studies of a large historic cemetery where parish records are available for comparison. This cemetery has also provided opportunities for assessing historic variation in growth and development as well as for testing methods of age determination. Saunders and her students have demonstrated how careful study of historic samples can not only tell us more about those particular people but also can help us to evaluate methods used on prehistoric samples. FitzGerald and Rose present information on age determination for subadult remains through dental microstructure analysis. The use of newer image analysis techniques (which are now easy to install in most anthropology laboratories) improves precision and relieves the tedium of collecting these data. This research shows great promise. If we can get a clearer picture of the amount of inter- and intrapopulation variation in dental development, we will know more about how tissue growth is buffered from stress and whether meaningful population differences really do exist. As these authors explain, it is only very recently that the investigation of microstructural growth markers in dental tissues has become
accepted as appropriate for estimating tooth crown formation times. Robling and Stout provide details as well as examples of adult age determination based on bone histomorphometry. They review the principles of bone modeling and remodeling as a prelude to explaining how cortical bone microstructure is used in age determination. Variations caused by activity, sex, disease, and population affinity are discussed. Appendices to their chapter allow one to practice the methods of histological age determination on photomicrographs from a femur and a rib.

Part III is titled “Prehistoric Health and Disease” and includes three chapters. As in Part II, the sequence of chapters is as follows: studies based on gross observations of bones, gross observations of teeth, and microscopic studies. Lovell focuses on paleopathology and diagnosis of bony lesions. She provides detailed information on various diagnostic methods, including radiology and microscopy. Steps toward diagnosis are discussed with emphasis on accurate description and consideration of the distribution of lesions within an individual skeleton as well as the distribution within skeletal samples. Hillson presents methods for analyzing and describing dental pathology, with detailed information on the underlying causes of various conditions. He stresses the importance of careful observation, demonstrating how different ways of scoring pathological changes can dramatically alter determinations of disease prevalence. If care is taken with observations, so that the surviving jaws and teeth in skeletal collections really do represent what was buried, then the distribution of dental disease can tell us a lot about the diets and activities of past populations. Then we can seek correspondence between dental data and data from stable isotopes, faunal and botanical assemblages, and artifacts used in daily life. Pfeiffer covers the subject of bone histology with respect to healthy bone turnover and various disease states. This chapter ties in nicely with those of Ruff and Robling and Stout in that it covers information on bone structure at the histological level and the factors that account for variation. Her work includes variation in bone histology over recent human evolution with examples drawn from Neandertals to recent European immigrants to Canada. Procedures for preparing bones in thin sections are reviewed with cautions regarding diagenetic alteration.

Part IV, “Chemical and Genetic Analyses of Hard Tissues,” includes chapters on stable isotope analysis, trace element analysis, and ancient DNA. Katzenberg provides background information on stable isotope studies and examples of applications to questions regarding paleodiet, migration, and life history. She demonstrates how isotopic analysis of archaeological tissues has advanced dramatically over a relatively short time span. Rather than simply confirming information that was already available from other sources, she shows how this field has called into question various archaeologically hypotheses about subsistence adaptations as well as adding to our understanding of human ecology. She discusses three areas of research that are particularly promising because of their implications for a more detailed reading of the past. These areas include reconstructing infant feeding practices, detecting pathological changes in bones, and the management of animal and plant species by earlier human populations. Sandford and Weaver provide information on the current status of trace element studies. These studies include attempts to control for postmortem changes. They focus their discussion on the dietary indicators, strontium and barium, and the toxic element, lead. This chapter nicely illustrates the stages of new methods, discussed early in the preface. Sandford and Weaver have labeled these “Inaugural” (discovery and early applications), “Intermediate” (reevaluation and testing), and “Modern” (emphasis on experimental and simulation studies). The chapters on stable isotope analysis and trace element analysis both emphasize the importance of training in the physical sciences. Stone discusses advances in the isolation and analysis of ancient DNA. A great wave of excitement was ushered in with the first developments in the
extraction and amplification of ancient DNA. If we can retrieve fragments of genes from long deceased humans, surely we can reconstruct the evolutionary and population history of past human groups. But the early claims for the retrieval of ancient DNA from dinosaurs and other fossils were cast aside when it was shown that the amplified DNA came from modern contaminants. The promise of ancient DNA research lost some of its luster. Yet, more recently, Stone was part of the research team able to offer clear evidence for the sequencing of Neanderthal DNA. Nevertheless, she cautions us about the difficulties of proving positive results and warns us that the promise is there, but the road ahead is still difficult.

Part V, “Quantitative Methods and Population Studies,” includes three chapters. Pietrusewsky discusses metric techniques and their applications to biological distance studies. He takes the reader through the various statistical procedures used to visualize biological relationships. These procedures include a range of multivariate statistics such as clustering techniques, multidimensional scaling, and Mahalanobis’s generalized distance. Cranio metric analysis has been one of the transitional realms of osteological research. Pietrusewsky shows how this approach is still appropriate for the investigation of widespread museum collections, where destructive analyses are prohibited. Furthermore he demonstrates by using examples from his own extensive research in the Pacific, that multiple lines of evidence, including cranio metric, dental, linguistic, and molecular data are all necessary to contribute to our understanding of human population history. Jackes tackles the problem of adult age determination and evaluates recent attempts to circumvent some problems. She surveys and evaluates all of the different approaches to age-at-death estimation, including single methods, such as metamorphosis of the pubic symphysis and cranial suture closure, as well as complex methods. She emphasizes the difficulties of dealing with the biases of reference samples and the effects of skeletal preservation on efforts to produce age distributions for archaeological samples. She takes the position that statistical investigation and manipulation cannot substitute for the necessity of having accurate biological age estimates. Finally, Milner, Wood, and Boldsen evaluate the current status of paleodemography by focusing on some questions that have fueled past debates within the field. They address problems of sampling, age and sex estimation, nonstationarity, heterogeneous risk, and selective mortality. Paleodemography draws from many of the types of studies covered in previous chapters and attempts to tie together the success of populations based on factors such as diet, disease experience, activity patterns, growth and development, and population interactions. Milner and colleagues provide a frank view of the potential and the limitations of achieving the goal of being able to determine the level of adaptation of past populations.

All of these chapters have the common theme of determining information about past peoples from their skeletal and dental remains. Adult age determination is an important theme that appears in many chapters. Similarly, postmortem change, sampling, and the relationship between cemetery samples and living populations recurs throughout the book. Ethical considerations have had a major impact on all topics discussed. It is our hope that this information will provide both breadth and depth for advanced studies in human osteology and will serve as a guide to more intensive study.

M. ANNE KATZENBERG
SHELLEY R. SAUNDERS
ACKNOWLEDGMENTS

The editors wish to thank the contributors to this volume for their efforts to present their areas of research expertise to students and professionals. We are grateful to contributors who have revised their work from the previous edition as well as those to who have prepared new contributions for this edition. It has been a pleasure to work with everyone. We also thank our editors at Wiley, Thom Moore and Karen Chambers and our editorial assistant Ian Collins. Kristine Hepple and Charles FitzGerald provided excellent editorial assistance during the final phases of manuscript preparation.
CONTRIBUTORS

Sabrina C. Agarwal, received her M.Sc. and Ph.D. (2001) from the University of Toronto, working in the Department of Anthropology and the Samuel Lunenfeld Research Institute of Mount Sinai Hospital, Toronto. She completed a two-year Social Sciences and Humanities Research Council of Canada Postdoctoral Fellowship in the Department of Anthropology, McMaster University. She is currently assistant professor of anthropology and faculty associate at the Archaeological Research Facility, University of California, Berkeley. Her research interests focus broadly on the study of skeletal health and disease, with particular focus on age, sex, and gender-related changes in bone quantity and quality. She is particularly interested in the application of biocultural and evolutionary approaches to the study of bone maintenance and fragility. Her work has examined patterns of cortical bone microstructure, trabecular architecture, mineral density, and biomechanical properties in several archaeological populations in the Old World, including Medieval and Post-Medieval Britain, and Imperial Rome, as well as a pioneer population from nineteenth-century Upper Canada. She is currently examining the effects of growth and development and lifestyle on bone maintenance and fragility over the lifecycle in an Anatolian Neolithic archaeological population from Çatalhöyük, Turkey. She is also examining the long-term effect of parity and lactation on the maternal skeleton with a large-scale project on bone maintenance in the nonhuman primate (monkey) model. Her publications include a co-edited volume, Bone Loss and Osteoporosis: An Anthropological Perspective (Kluwer Plenum Academic Press) with S.D. Stout, published in 2003, and “Medieval Trabecular Bone Architecture: The Influence of Age, Sex and Lifestyle” American Journal of Physical Anthropology, 2004, with M. Dimitriu, and M.D. Grynpas.

Jesper L. Boldsen, received his Ph.D. in biology from the Department of Theoretical Statistics, Aarhus University (Denmark) in 1983. He is an associate professor with the Institute of Forensic Medicine, University of Southern Denmark, Odense. He is also adjunct professor of anthropology at the University of Utah and associate editor of the Journal of Biosocial Science and Journal of Applied Oral Science. Research interests revolve around human population
biology, epidemiology, demography, and evolution. In recent years, he has concentrated much of his research effort on analyzing the structure of the medieval population of Denmark. This work has been based on studies of extensively excavated medieval cemeteries. He has published widely in biological and anthropological journals. Several recent publications include “Early childhood stress and adult age mortality—a study of dental enamel hypoplasia in the medieval Danish village of Tirup,” American Journal of Physical Anthropology, 2007; “Outside St. Jørgen: Leprosy in the medieval Danish city of Odense,” American Journal of Physical Anthropology, 2006; “Leprosy and mortality in the Medieval Danish village of Tirup,” American Journal of Physical Anthropology, 2005; and “Testing conditional independence in diagnostic palaeodemography,” American Journal of Physical Anthropology, 2005.

STEVEN K. BOYD, received his Ph.D. from the University of Calgary in 2001. He then completed a postdoctoral fellowship at the Swiss Federal Institute of Technology, Switzerland, before joining the faculty at the University of Calgary with a joint appointment in the Department of Mechanical and Manufacturing Engineering (Schulich School of Engineering) and the Faculty of Kinesiology. Utilizing high-resolution computed tomography (micro-CT), his research has focused on the area of orthopaedic biomechanics, focusing on adaptive changes to tissues that occur after a joint injury or disease, with particular interest in bone. Recent publications include “Monitoring individual morphological changes over time in ovariectomized rats by in vivo micro-computed tomography,” Bone, 2006, with P. Davison, R. Müller, and J.A Gasser, and “Establishment of an architecture-specific experimental validation approach for finite element modeling of bone by rapid prototyping and high resolution computed tomography,” Medical Engineering and Physics, 2006, with R. Su and G.M. Campbell.

JAMES BURTON, received his Ph.D. from the University of Arizona in 1986. He is the senior scientist and associate director of the Laboratory for Archaeological Chemistry at the University of Wisconsin-Madison (Department of Anthropology). His research interests include the development of new archaeometric methods, particularly the use of chemical and isotopic methods for provenience studies. Current projects include exploration of alkaline-earth elements and various isotopic systems in the study of human mobility and the development of nondestructive methods to characterize historical materials; studies of the effect of marine resources on bone levels of barium and strontium; the origins of the Gila Polychrome tradition; and the eastern, selvatic origin of unusual pottery found at Late Formative sites in the Ecuadorian Andes. Recently he co-authored with T.D. Price and V. Tieseler “Early African diaspora in colonial Campeche, Mexico: Strontium isotopic evidence” American Journal of Physical Anthropology, 2006, and “Interpreting the trace-element components of bone—a current perspective from the Laboratory for Archaeological Chemistry” with T.D. Price in Deciphering Ancient Bones: The Research Potential of Bioarchaeological Collections, edited by G. Grupe and J. Peters (2003).

CHARLES M. FITZGERALD, received his Ph.D. in biological anthropology from the University of Cambridge in 1996. In 2000, after completing a postdoctoral fellowship at McMaster University, he worked with Simon Hillson at the Institute of Archaeology in University College London on a three-year project funded by the Natural Environment Research Council of the UK. This study’s objective was to clarify the nature of the dental reduction that is observed in Middle and Upper Palaeolithic, Mesolithic, and Early Neolithic hominids and to test the
major hypotheses that account for it. The study included the use of an approach to assess tooth size that overcame the problem of tooth wear, which was severe in these early modern humans. He returned to Canada where he was appointed the research coordinator for the Anthropology Hard Tissue and Light Microscopy Laboratory at McMaster University. The focus of much of his research has been on the validation and application of odontochronological techniques, but in addition to growth and development, his interests embrace several other areas of skeleto-dental biology and palaeoanthropology. Recent publications include two papers in the American Journal of Physical Anthropology: “Health of infants in an Imperial Roman skeletal sample: perspective from dental microstructure,” 2006, with S.R. Saunders, L. Bondioli, and R. Macchiarelli, and “A Test of histological methods of determining the chronology of accentuated striae in deciduous teeth,” 2005, with S.R. Saunders.

Benedikt Hallgrimsson, received his Ph.D. from the University of Chicago. He is an associate professor in the Department of Cell Biology and Anatomy, Faculty of Medicine, University of Calgary, and an adjunct professor in the Department of Archaeology. He serves as the associate dean, Undergraduate Science Education and is a member of the Alberta Bone and Joint Institute and an associate member of the Institute of Maternal and Child Health. Research interests include the developmental-genetic basis for phenotypic variation and the developmental-genetic basis for variation in canalization, morphological integration, and developmental stability. Dr. Hallgrimsson has published extensively in medical, biological, and anthropological journals, including a co-edited volume, Variation: A Central Concept in Biology with B.K. Hall (Academic Press, 2005). Recent articles include: “The brachymorph mouse and the developmental-genetic basis for canalization and morphological integration,” Evolution & Development, 2006, with J.J.Y. Brown, A.F. Ford-Hutchinson, H.D. Sheets, M.L. Zelditch, and F.R. Jirik, and “Canalization and developmental stability in the Brachyrhine mouse,” Journal of Anatomy, 2006, with K.E. Willmore, M.L. Zelditch, N. Young, A. Ah-Seng, and S. Lozanoff.

Simon Hillson, received his Ph.D. in London in 1979. He is a professor of bioarchaeology at the Institute of Archaeology, University College London. Dr. Hillson’s research interests focus on the biology of past human populations, ranging from the most recent post-Medieval populations of London to Predynastic Egyptians, and remains from Upper Palaeolithic contexts. In particular, he works with dental remains, studying their morphology, microstructure, development, pathology, and evolution in Neanderthals and modern humans. As well as numerous journal articles, he has published three books: Dental Anthropology (Cambridge University Press, 1996), Mammal Bones and Teeth (Institute of Archaeology, UCL, 1992) and Teeth (Cambridge University Press, 2005, second edition).

M. Anne Katzenberg, received her Ph.D. in anthropology from the University of Toronto in 1983. She holds a university professorship at the University of Calgary (Department of Archaeology). Her research interests include diet and health in past peoples, and in particular, she explores the various applications of stable isotope analysis to reconstructing paleodiet, palaeodemography, and ecology. She serves on the editorial board of the Yearbook of Physical Anthropology and is the vice-president of the Paleopathology Association. In 2003, she was elected to the Royal Society of Canada. She serves as a consultant in forensic anthropology for the Medical Examiner of Alberta (southern division). Recent publications include “Identification of historical human skeletal remains: a case

ERIKA KRISTENSEN, received her B.Sc. in biological engineering from the University of Guelph. She is completing her M.Sc. in the Department of Mechanical and Manufacturing Engineering at the University of Calgary and plans to start on her Ph.D. in late 2007. Her research interests include applications of mechanical engineering principles to questions in developmental biology, such as the relationship between craniofacial shape and mandibular strength in growth-hormone-deficient mice. She recently submitted a paper titled “A novel high-throughput morphological method for phenotypic analysis,” with T.E. Parsons, B. Hallgrímsson, and S.K. Boyd.

C. OWEN LOVEJOY, received his Ph.D. from the University of Massachusetts. He is a university professor in the Department of Anthropology at Kent State University, as well as serves in many other professional capacities, such as technical advisor in biological anthropology, Cuyahoga County Coroner’s Office, Cleveland; adjunct professor of anatomy, Department of Human Anatomy, NEOUCOM, Department of Orthopaedic Surgery, School of Medicine, Case Western Reserve University). His research covers a broad spectrum of areas in human biology, from human paleoanthropology to the genetics of development. Much of his research focuses on bipedal locomotion and its evolution. Many of his publications involve the human postcranium and that of our ancestors. Recent articles include a series in Gait and Posture (2005–2006) on “The natural history of human gait and posture”; and the article “Of muscle-bound crania and human brain evolution: The story behind the MYH16 headlines,” Journal of Human Evolution, 2006, with M.A. McCollum, C.C. Sherwood, C.J. Vinyard, and F. Schachat.

NANCY C. LOVELL, received her Ph.D. from Cornell University in 1987 and is a professor of anthropology at the University of Alberta. She specializes in bioarchaeology, paleopathology, and mortuary archaeology. Current research projects include the history of cemeteries in western Canada in order to assess the fit between archaeological interpretive models for prehistoric cemeteries with the documented evidence for burial practices, and the analysis of skeletal remains from ancient Egypt and Mesopotamia to examine health and biological relationships in complex societies. Recent publications include “Nonmetric traits of the deciduous dentitions from Bronze Age Tell Leilan, Syria,” 2007, and “Nonmetric analysis of the permanent dentition of Bronze Age Tell Leilan, Syria,” 2006, both with S. Haddow and appearing in the International Journal of Dental Anthropology, and “A comparative experiment in the consolidation of cremated bone,” International Journal of Osteoarchaeology, 2004, with D. Rossi and S. De Gruchy.

RICHARD S. MEINDL, received his Ph.D. from the University of Massachusetts. He is professor and chair of the Department of Anthropology at Kent State University. He conducts research in archaeology and osteology, forensic sciences, primate and human evolution, and historic demography and population structure. Most recently, he has studied the paleodemography of various archeological areas, including the Shell Mound Region of western Kentucky, a Late Woodland population from northern Ohio, and a Late Classic site in lowland
Yucatan. These areas provide special opportunities to examine the rise of complex societies, pre-agricultural sedentary living, and human demographic evolution. Publications include the review “Recent Advances in Method and Theory in Paleodemography,” *Annual Review of Anthropology*, 1998; and the co-edited book *Archaic Transitions in Ohio and Kentucky Prehistory*, with O.H. Prufer and S.E. Pedde (Kent State University Press, 2001).

**Robert P. Mensforth**, received his Ph.D. from Kent State University. He joined the faculty in the Department of Anthropology at Cleveland State University in 1991. His research interests include skeletal biology, hominin paleontology, forensic anthropology, paleodemography, paleopidemiology, medical anthropology, and human gross anatomy. Having completed the first phase of a forensic study on patterns and types of adult craniofacial trauma that characterize a very large sample of early twentieth century individuals from the Hamann–Todd collection at the Cleveland Museum of Natural History, his current project is the analyses of the relationships among various forms of trauma and the age, sex, and ancestry of affected individuals. Among his publications are “Paleodemography of the Carlston-Annis (Bt-5) late Archaic skeletal population,” *American Journal of Physical Anthropology*, 1990, and “Warfare and trophy taking in the Archaic Period,” *Archaic Transitions in Ohio and Kentucky Prehistory*, 2001, edited by O.H. Prufer, S.E. Pedde, and R.S. Meindl.


**Trish E. Parsons**, received her BSc. in biological anthropology from The Pennsylvania State University. She is currently a student in the Biological Anthropology Graduate Program at the University of Calgary. Her research interests include the developmental basis of facial length and the study of basic mammalian developmental processes in the skull through mouse models and primate evolution. She has been a student of morphometrics since her days as an undergraduate at Penn State. Her most recent publication is “Microstructure of trabecular bone in a mouse model for Down syndrome,” *Anatomical Record: Advances in Integrative Anatomy and Evolutionary Biology*, 2007, with T.M. Ryan, R.H. Reeves, and J.T. Richtsmeier.

**Michael Pietrusiewsky**, received his Ph.D. in anthropology from the University of Toronto in 1969. He is currently professor of anthropology, University of Hawaii, Honolulu. His
research focuses on studies of human skeletal remains from Australia, the Pacific, East Asia, and Southeast Asia, including the application of multivariate statistical procedures to cranial data for investigating historical–biological relationships. Recent publications include “Ban Chiang, A prehistoric village site in Northeast Thailand II: The human skeletal remains,” *The University Museum, University of Pennsylvania, Memoir Series*, 2002, with M.T. Douglas; “Multivariate comparisons of female cranial series from the Ryukyu Islands and Japan,” *Anthropological Science*, 2004; “Matrilocality during the prehistoric transition to agriculture in Thailand,” *Antiquity*, 2005, with R.A. Bentley, M.T. Douglas, and T.C. Atkinson; and “A multivariate craniometric study of the prehistoric and modern inhabitants of Southeast Asia, East Asia, and surrounding regions: a human kaleidoscope?” *Bioarchaeology of Southeast Asia*, edited by M.R. Oxenham and N. Tayles (Cambridge University Press, 2006). He is a member of the American Association of Physical Anthropologists, fellow of the American Academy of Forensic Sciences, diplomate of the American Board of Forensic Anthropologists, and member of the Editorial Boards of *Anthropological Science, Anatomical Science International*, and the *Journal of Island and Coastal Archaeology*. He has been a visiting professor/scholar to Australia, Canada, France, Germany, Japan, New Zealand, People’s Republic of China, and Taiwan.

**Dori L. Rainey**, received her MA from the University of Western Ontario (working with the skeletal remains of the Huron of Southern Ontario under Dr. Michael Spence) and is currently working on her Ph.D. in the Department of Anthropology at McMaster University. With a broad focus on skeletal biology, including palaeopathology and forensic anthropology, her research attempts to determine social organization in archaeological populations through nonmetric traits of the infracranial skeleton. More specifically, the research endeavors to address the variability observed in nonmetric traits and to integrate these data with pathological conditions, occupational markers, and the archaeological context.

**Alexander G. Robling**, received his Ph.D. from the University of Missouri in 1998 and completed a postdoctoral fellowship in the Department of Anatomy and Orthopedic Surgery at Indiana University—Purdue University, Indianapolis (IUPUI). He is currently an Assistant Professor in the Department of Anatomy & Cell Biology and Biomedical Engineering at the IUPUI. Specializing in the effects of mechanical loading on bone at the organ, tissue, cell, and molecular levels, his current interests are focused on the signal transduction cascades involved in bone cell mechanosensation, including the genetic regulation of mechanosensitivity. Recent publications include the co-authored article, “Biomechanical and Molecular Regulation of Bone Remodeling” *Annual Review of Biomedical Engineering*, 2006; and the chapter, “Histomorphology, geometry, and mechanical loading in past populations” with S.D. Stout in *Bone Loss and Osteoporosis: an Anthropological Perspective*, (edited by S.C. Agarwal and S.D. Stout, (Kluwer Academic/Plenum Publishers, Inc.).

**Jerome C. Rose**, received his Ph.D. in anthropology/biological anthropology from the University of Massachusetts, Amherst, in 1973. He is currently chair and professor of anthropology at the University of Arkansas, Fayetteville, and a member of the Department of Anthropology, Institute of Archaeology and Anthropology, Yarmouk University, Irbid, Jordan. His early research interest was dental anthropology with a particular focus on dental histology, including the study of enamel microdefects and enamel hypoplasias. Methodological
specialties include light and electron microscopy and thin sectioning technology. He has conducted bioarchaeological excavations in Illinois, Arkansas, Texas, Egypt, and Jordan. Research on the origins of agriculture and the economic determinants of health provided the motivation to begin a specialty in the bioarchaeology of the Middle East. He is currently working in Jordan on the excavation and analysis of Byzantine tombs and in Egypt at the site of Amarna. Recent publications include “Saad: A Late Roman/Byzantine Site in North Jordan,” co-edited with D.L. Burke (2004); and “The Backbone of History: Health and Nutrition in the Western Hemisphere,” co-edited with R.H. Steckel (Cambridge University Press, Cambridge).

CHRISTOPHER B. RUFF, received his Ph.D. in biological anthropology from the University of Pennsylvania in 1981 and is currently the director of the Center for Functional Anatomy and Evolution, the Johns Hopkins University School of Medicine. His research interests are in biomechanics and primate locomotion, evolution of the hominoid postcranium, skeletal growth and development, osteoporosis, and skeletal remodeling and behavioral reconstruction in human populations. Recent publications include the co-authored “Body size, body shape, and long bone strength of the Tyrolean ‘Iceman,’” Journal of Human Evolution, 2006; “Gracilization of the modern human skeleton—The latent strength in our slender bones teaches lessons about human lives, current and past,” American Scientist, 2006; and “Who’s Afraid of the Big Bad Wolff?: Wolff’s Law and Bone Functional Adaptation,” American Journal of Physical Anthropology, 2006, with B. Holt and E. Trinkaus.

SHELLEY R. SAUNDERS, received her Ph.D. in anthropology from the University of Toronto in 1977. She is currently a professor in the Department of Anthropology, McMaster University, and the director of the McMaster Anthropology Hard Tissue and Light Microscopy Laboratory. Her research covers microscopic and macroscopic human dental and skeletal growth and development, dental pathology, and methods of sex and age estimation from teeth and bones. She holds a Canada Research Chair in Human Disease and Population Origins and is the founder of the McMaster Ancient DNA Centre. She is also North American editor of the International Journal of Osteoarchaeology and was elected to the Royal Society of Canada in 2001. Recent publications include the co-authored article, “Health of infants in an Imperial Roman skeletal sample: perspective from dental microstructure,” American Journal of Physical Anthropology, 2006, and “Sexual dimorphism of the dental tissues in human permanent mandibular canines and first premolars,” American Journal of Physical Anthropology, 2007, with A.H.W. Chan, B. Kahlon, H.F. Kluge, and C.M. FitzGerald.

G. RICHARD SCOTT, holds a Ph.D. from Arizona State University (1973). He is currently an assistant professor in the Department of Anthropology at the University of Nevada, Reno, and is professor emeritus of anthropology, University of Alaska Fairbanks. Research interests include dental anthropology and skeletal biology, with special emphasis on tooth morphology and bioarchaeology. He has worked with groups in the American Southwest, Arctic, North Atlantic (Greenland, Iceland, Norway, and Denmark), and Spain. He co-authored “Physical anthropology of the Arctic,” in The Arctic: Environment, People, Policy, edited by M. Nuttall and T.V. Callaghan (Harwood Academic Publishers, Amsterdam, 2000); “The Anthropology of Modern Human Teeth,” with C.G. Turner II (Cambridge University Press, Cambridge, UK, 1997); and “Dentition,” also co-authored with Christy

**Ann L. W. Stodder**, received her Ph.D. from the University of Colorado (1990). She is a research associate in the Department of Anthropology at the Field Museum as well as an adjunct faculty member in the Department of Anthropology at the University of Wisconsin, Milwaukee. She also serves as treasurer of the Paleopathology Association. Her study of human remains focuses on paleopathology and the use of life histories in understanding human adaptation in the past, and on the taphonomic signatures of mortuary practices that reflect the human relationship with death and with the deceased. She is the author of “Skeletal biology: Southwest,” in the *Handbook of North American Indians*, edited by D.H. Ubelaker (Volume 3, 2006), and several taphonomic studies of human remains from New Guinea, including “The bioarchaeology and taphonomy of mortuary ritual on the Sepik Coast, Papua New Guinea,” in *Interacting With the Dead: Perspectives on Mortuary Archaeology for the New Millennium*, edited by G. Rakita, J. Buikstra, L. Beck, and S. Williams (2005). Dr. Stodder is also the editor of the forthcoming volume “Reanalysis and Reinterpretation in Southwestern Bioarchaeology” (Anthropological Research Paper No. 59, Arizona State University).

**Anne C. Stone**, received her Ph.D. from Pennsylvania State University in 1996. She is currently an associate professor in the School of Human Evolution and Social Change at Arizona State University. She specializes in anthropological genetics and specifically applications of population genetics to questions concerning the origins, population history, and evolution of humans and the great apes. Her current research focuses on the biological history of Native Americans and on the evolution and genetic diversity of the genus *Pan*, which includes chimpanzees and bonobos. She is also investigating the evolutionary history of tuberculosis in ancient samples dating prior to and after the Age of Exploration. Current publications include the co-authored article, “Independent evolution of bitter-taste sensitivity in humans and chimpanzees,” in *Nature*, 2006, and “Comparative analyses reveal a complex history of molecular evolution for human MYH16,” *Molecular Biology and Evolution*, 2005.


**Douglas H. Ubelaker**, received his Ph.D. degree from the University of Kansas in 1973. Currently, he holds the position of

**PHILLIP L. WALKER**, received his Ph.D. from the University of Chicago in 1973. He is currently a professor of anthropology at the University of California, Santa Barbara. His interests include bioarchaeology, paleopathology, forensic anthropology, and human evolution. Professor Walker is currently working on several bioarchaeological projects involving human skeletal remains from various parts of the world, including North America, Asia, and Europe. He is the co-director of an archaeological project in Iceland that includes the excavation of a settlement period cemetery and church. Walker is a principle investigator on a large NSF-funded collaborative project entitled “A History of Health in Europe from the Late Paleolithic Era to the Present.” This project involves researchers from many different European countries. Its goal is to measure and analyze the evolution of skeletal health by combining data from human remains with information gathered from sources in archaeology, climate history, geography, and history. Some of his recent publications include the co-authored “The evolution of treponemal disease in the Santa Barbara Channel Area of Southern California,” in *New World Treponemal Disease*, edited by M.L. Powell and D.C. Cook (University of Florida Press, 2005); “Greater sciatic notch morphology: Sex, age, and population differences,” *American Journal of Physical Anthropology*, 2005; and “Bioarchaeological evidence for trophy taking in prehistoric Central California,” *American Journal of Physical Anthropology*, 2005, with V.A. Andrushko, K. Latham, D. Grady, and A. Pastron.

**JAMES W. WOOD**, received his Ph.D. from the University of Michigan in 1980. He is currently professor of biological anthropology and demography at Pennsylvania State University. His research spans several areas of human population biology, including biodemography, historical demography, population ecology, paleodemography, reproductive biology, and infectious disease dynamics. He has carried out extensive research on fertility and reproductive physiology and has conducted fieldwork in Papua New Guinea on birth-spacing patterns, the contraceptive effects of breastfeeding, fecundability, and pregnancy loss. He was also involved in a long-term prospective study of the endocrinology of menopause in a large cohort of U.S. women. Dr. Wood’s more recent interests include paleodemography, where he has made important statistical and analytical contributions and has collaborated with several colleagues working on a large collection of medieval Danish skeletons. He also has a long-standing interest in the demographic effects of infectious diseases and is currently directing a large multidisciplinary study of the fourteenth-century Black Death. His other current research interests focus on the historical demography and


In 1953, Lucile E. Hoyme published an article entitled “Physical anthropology and its instruments: An historical study.” She asserted, “Measurement is the oldest and most distinctive hallmark of the physical anthropologist” (Hoyme, 1953: 409). By this she meant absolute and relative size measurements of the human body, which she argued was core methodology within physical anthropological investigations.

The roots of this core methodology, when applied systematically to human remains, can be traced to comparative studies that began at the turn of the eighteenth century at the University of Edinburgh. Following such early precursors, scholarship associated with the influential Société d’Anthropologie de Paris (1859) included Paul Broca’s prolific descriptions of new instruments and techniques for studying living and skeletalized humans. After Broca’s death, methodological innovation shifted to Germany where von Török, for example, developed a composite instrument that combined calipers, a goniometer, and a craniometer to facilitate recording the 5300 measurements he advocated (Hoyme, 1953). Thus, the technical ancestry of contemporary physical anthropological investigations of humans and their remains includes measurement via a remarkable range of instruments. Most nineteenth century applications of these techniques were simply descriptive and comparative craniology, although some research questions extended, for example, to the estimation of stature (Rollet, 1888). Some workers used measurements for the estimation of sex, including Matthews et al. (1893), whose methods followed those of European scholars.

A second persistent theme in physical anthropology’s methodological heritage is visual observation, without direct measurement, of human and skeletal morphology to describe and interpret variation. Systematic cranial shape comparisons and constructing typologies generated through such methods have a history as long as that of cranial measurement, extending into the eighteenth century (Cook, 2006). During the nineteenth century, careful observations without measurement also characterize reports by medical doctors and anatomists of osseous cultural modifications and evidence of ancient disease. As Cook and Powell (2006) emphasize, early nineteenth-century
scholars such as Warren and Morton, focused on cranial morphology rather than on disease per se, but as the century matured, so did an interest in pathological conditions. Other types of nineteenth-century observations used to infer ancient life ways through the direct observation of human osseous and dental tissues include the use of nonmetric traits, such as the Inca bone, to estimate population relationships. Tibial flattening served as a basis for inferring load bearing and movement across landscapes (Matthews et al. 1893). In dental studies, associations among wear, dental disease, and diet were well established during the nineteenth century (Rose and Burke, 2006), with early American applications, including comparative studies by Matthews et al. (1893). Morphological dental variation was overshadowed by the study of cranial morphology during the nineteenth century, but it assumed prominence during the early twentieth century (Rose and Burke, 2006).

During the late nineteenth and early twentieth centuries, new methods, some invasive, supplemented those traditionally employed by researchers who studied skeletonized remains (Buikstra, 2006). Joseph Jones (1876), for example, used histological techniques to examine the internal structure of diseased bones recovered from Tennessee’s stone box graves. He also experimented with hydrochloric acid as a means of assessing antiquity and relative age of these remains. Ancient teeth had been studied as early as 1892 (Tomes, cited in Rose and Burke, 2006). Within a year of Wilhelm Röntgen’s discovery (1895), X-rays were used to investigate mummies of a cat and human child (Aufderheide, 2003).

Thus, by the early years of the twentieth century, basic techniques included direct measurement and morphological observations, which were frequently applied to the reconstruction of population histories. Less commonly were they directed toward issues such as estimating sex, diet, and behavior. Technically complex methods, such as radiography and histology, had been applied to ancient materials, but studies were limited in scope and number.

Another important observation made by Hoyme during the mid-twentieth century involved the degree to which physical anthropologists were themselves creating the instruments used within the discipline or borrowing them from other fields. Prior to 1900, instrumentation was commonly generated internally, a pattern that changed within the twentieth century.

Among the instruments and methods introduced by mid-century, Hoyme mentioned the anthropological use of stereographic technology adapted from orthodontics, as well as X-rays and chemical tests used in studies of physiology. In addition, she asserted, “Like an individual, a science comes of age when it is in a position to contribute to society” (Hoyme, 1953: 423).

Fast forward a century to the twentieth/twenty-first-century transition and we see that many of the same topics that were of at least passing interest by the close of the nineteenth century remained highly visible: genetic relationships, diet, disease, and behavior. A distinctive twentieth-century development was paleodemography. With roots in Hooton’s studies, interest in a demography of the past largely post-dates 1960 (Frankenberg and Konigsberg, 2006). Although distinguishing juvenile from adult remains had resolved “pigmy race” allegations by mid-nineteenth century (Morton, 1841), systematic study of age-related changes developed within the twentieth century, beginning with the landmark contributions of the anatomist T. Wingate Todd (1920). Methods for estimating both age-at-death and sex have been refined over the twentieth century, influenced by the increasing sophistication of forensic anthropologists.
Although many of the questions addressed by those studying ancient skeletal material have nineteenth-century roots, the pace of methodological advancement has increased markedly in recent years. Within the ~40 years that I have been studying remains from archaeological and forensic contexts, increasingly sophisticated biochemical, bimolecular, and quantitative methods have virtually revolutionized the manner in which we can measure, for example, genetic relatedness and diet. As expected, some false starts have occurred, as in the enthusiastic over-interpretations of trace elements as dietary indicators, but there are doubtless many significant advances on the horizon for many approaches, especially those involving biomolecules. The rapid tempo of such changes is illustrated by a comparison of the three volumes edited by Katzenberg and Saunders: Saunders and Katzenberg (1992), Katzenberg and Saunders (2000), and the current book. The preface to Katzenberg and Saunders (2000) recognizes this phenomenon, as the editors comment: “The original plan for the book was to be a second edition of our earlier edited book, Skeletal Biology of Past People: Research Methods. However, as work progressed it seemed that with five additional chapters and many new contributors, it is really something different. The differences are directly related to changes that have occurred in the analysis of human skeletal and dental remains over the past few years.”

The Saunders and Katzenberg (1992) edited volume clearly illustrated the fact that skeletal biological and dental studies frequently used techniques developed in other fields. Bone and tooth histology, biomechanical analyses, light or optical spectrometry and X-ray fluorescence for trace element analysis, mass spectrometry for isotopes studies, electron microscopy, computerized tomography scans and other imaging techniques, immunological and molecular methods in the study of disease, and complex quantitative methods amply illustrated the use of innovative methods and state-of-the-art instrumentation developed elsewhere. By 2000, Katzenberg and Saunders had added a full chapter on DNA analysis of archaeological remains, a discussion of paleohistology in the study of paleopathology, and contributions defining increasingly sophisticated methodological and quantitative approaches to paleodemography. In this contribution, the chapter entitled “Morphometrics and biological anthropology in the post-genomic age” explicitly addresses the startling opportunities that technological advances in computer and molecular methods offer for phenotypic analyses. These new developments are revolutionary and they are applied to the core problem of physical anthropology, the measurement of human morphological variation, as Hoyme emphasized in 1953.

Although such technological advancement anchors many current research trajectories, the direct observation of remains continues to be an essential aspect of our craft, complementing complex methods, as it did in the nineteenth century. Although the subjective cranial typologies of that period have fallen away, rigorous and replicable assessments of dental morphology and disease, evaluated through visual inspection, remain significant venues within physical anthropology. Analyses of bone pathology, for example, are also grounded in careful observations that differentiate healthy from diseased bones.

Other changes noted in comparison of the three volumes included a shift away from the exclusive focus on the methods and topics appropriate to the study of archaeological samples that dominated Saunders and Katzenberg (1992). As emphasized in the preface to Katzenberg and Saunders (2000), these include chapters added to reflect influences both internal and external to physical anthropology. Increasingly,
ethical concerns about the excavation and study of human remains now influence the manner in which excavation and analysis of mortuary contexts is performed. These approaches are explicitly addressed in the first chapter of the two more recent Katzenberg and Saunders volumes. The second chapter in each describes the growth of forensic anthropology and its important contributions to, and synergism with, skeletal biology. Forensic anthropology clearly influences the range of topics presented in the current volume, including the addition of chapters on taphonomy and skeletal trauma. Returning to Hoyme’s remarks about societal contributions as a measure of the maturation of a discipline, the growth of forensic anthropology would suggest that our field is mature indeed.

Katzenberg and Saunders, in their three volumes, have masterfully represented both the maturity and the methodologies of biological anthropology. Thus, they have both enriched our science and advanced our field.

REFERENCES


Figure 7.10  (a) shows images of mean shapes for 20 C57BL/6J and 20 A/WySnJ mouse skulls. (b) shows shape variation represented as a color map of the magnitude of the image gradient for both strains. Areas of high shape variation within each group are confined to the incisors of both strains as well as to the lateral mandibular body of the A/WySnJ strain.

Figure 7.11  A comparison of shape between the C57BL/6J and the A/WySnJ mouse skulls is determined by calculating the surface-to-surface distances between the group average images (GSIs). The magnitude of the shape differences is displayed on the C57BL/6J mouse skull GSI by a scalar color map representing the relative differences to the A/WySnJ group.
PART I

THEORY AND APPLICATION IN
STUDIES OF PAST PEOPLES
CHAPTER 1

BIOARCHAEOLOGICAL ETHICS: A HISTORICAL PERSPECTIVE ON THE VALUE OF HUMAN REMAINS

PHILLIP L. WALKER

INTRODUCTION

The rapidity of technological and cultural change in current times is forcing us to confront a myriad of moral dilemmas over issues as wide ranging as the ethics of cloning humans, the ownership of our genetic material, and the rights of animals relative to those of humans. These ethical issues concern the very nature of what it means to be human and our relationships, not only to other people, but also to the plants and animals that sustain us.

The enormous strides we have taken toward human equality during this century mean that formerly disenfranchised and enslaved members of minority groups are beginning to gain power and control over their lives. In many countries there has been a decline in the political dominance and moral authority of organized religions. Notions of multiculturalism and a growing acceptance of the moral principle of not discriminating against people based on gender, ethnicity, or religious beliefs mean that there is no longer a shared set of cultural values we can use for guidance in dealing with moral issues (Cottingham, 1994).

This increased tolerance of cultural diversity poses ethical dilemmas because, as the range of value systems and religious beliefs that are considered socially acceptable increases, so does the probability of social conflict. To deal with these issues, many scientific associations are beginning to reconsider ethical principles that underlie their research activities. The field of bioarchaeology is especially problematic in this respect, positioned as it is between medicine with its ethical focus on generating scientific knowledge for use in helping individual patients and anthropology with its ethical principles that stem from deep belief in the power of cultural relativism to overcome ethnocentrism and encourage tolerance.

It is in this context that skeletal biologists are increasingly being forced to adapt their activities to the value systems of the descendants of the people they study. Human skeletal remains are more than utilitarian objects of value for scientific research. For many people, they also are objects of religious veneration of great symbolic and cultural significance (Sadongei and Cash Cash, 2007). Over the past 30 years, formerly disenfranchised groups such as Native Americans...
and Australian Aborigines have increasingly been able to assert their claims of moral authority to control the disposition of both the remains of their ancestors and the land their ancestors occupied (Howitt, 1998; Scott, 1996; Walker, 2004). This trend toward repatriating museum collections and granting land rights to indigenous people can only be understood within a broader social and historical context.

To provide this historical perspective, I will describe the evolution of religious beliefs about the proper treatment of the dead and conflicts that have arisen over the centuries between these beliefs and the value scientists place on the empirical information that can be gained through research on human remains. It is followed by a discussion of the generally accepted ethical principles that are beginning to emerge in the field of bioarchaeology. Finally, some practical suggestions are offered for dealing with conflicts that arise when these ethical principles conflict with those of descendant groups.

**THE HISTORY OF BELIEFS ABOUT THE DEAD**

Early in our evolutionary history people began to develop a keen interest in the remains of their dead comrades. At first this was undoubtedly simply a response to the practical considerations of removing the decaying remains of a dead relative from one’s domicile or preventing scavengers from consuming their body. More elaborate patterns of mortuary behavior soon began to develop. Cut marks on the crania of some of the earliest members of our species show that as early as 600,000 years ago people living at the Bodo site in Ethiopia were defleshing the heads of the dead (White, 1986). It has been suggested that such practices reflect a widespread belief among our ancestors concerning the role of the brain in reproduction (La Barre, 1984).

By 50,000 to 100,000 years ago mortuary practices had evolved into elaborate rituals that involved painting bodies with red ochre and including food or animal remains with the body as offerings. Through time these cultural practices became associated with increasingly complex religious beliefs that helped people cope with the uncertainties of death. Depositing utilitarian items and valuables such as ornaments in graves became commonplace in the Upper Paleolithic period. Such practices suggest continued use of these items was anticipated in the afterlife. Expressions of such beliefs can be found in some of the earliest surviving religious texts. The Egyptian Book of the Dead, for instance, provides spells and elaborate directions for use by the souls of the deceased during their journeys in the land of the dead (Allen, 1960; Ellis, 1996b).

The belief that the soul persists in an afterworld has deep roots in Western religious traditions. The ancient Greeks held elaborate funeral rituals to help a dead person’s soul find its way across the River Styx to a community of souls in the underworld. Once in the underworld there was continued communion between the living and the dead. For example, the soul of a dead person could be reborn in a new body if their living family members continued to attend to their needs by bringing them honey cakes and other special foods on ceremonial occasions (Barber, 1988). By medieval times most people continued to view death as a semi-permanent state in which the living and the spirit of the dead person could maintain contact with each other. Folktales about ghosts and corpses coming to life were widespread and contributed to the idea of the dead functioning in society with the living (Barber, 1988; Caciola, 1996). The issue of integrity of the corpse and the relationship of this to the afterlife dominated medieval discussions of the body: Salvation became equated with wholeness, and hell with decay and partition of the body (Bynum, 1995:114).

After the Reformation, conservative Protestant groups continued to emphasize the profound significance of a person’s physical remains after death. In fact, one of the more troublesome issues facing Protestant reformers after the
abolition of purgatory in the early sixteenth century was the need to provide a rational explanation for the status of body and soul in the period intervening between death and resurrection (Spellman, 1994). One strategy for dealing with this vexing problem is provided by the constitution for the Old School Presbyterian Church, published in 1822, which asserts that the bodies of deceased members of the church “even in death continue united in Christ, and rest in the graves as in their beds, till at the last day they be again united with their souls . . . the self same bodies of the dead which were laid in the grave, being then again raised up by the power of Christ (Laderman, 1996:54).”

Such beliefs in the continuance of life after death remain prevalent in modern Western societies (Cohen, 1992). Recent surveys show that 25% of European adults report having contact with the dead (Haraldsson and Houtkooper, 1991), and a significant number of Americans believe in reincarnation (Donahue, 1993; Walter, 1993). About half of the people in the United States believe that hell is a real place in which people suffer eternal damnation (Marty, 1997). In another survey, 80% of the North American population believes in some kind of an afterlife (Goldhaber, 1996; Tonne, 1996). Among Canadians, 40% believe in the Devil and 43% in Hell (Belief in the Devil, 1995).

Surveys also show that, despite speculation about the secularizing effects of education and academia, most highly educated people including professors and scientists are about as religious as other Americans. Anthropologists are one of the few groups that deviate significantly from the majority view that individual human beings continue to exist in some kind of an afterlife. Compared with faculty in the physical sciences, anthropologists are almost twice as likely to be irreligious and to never attend church, and one in five actually declare themselves “opposed” to religion (Iannaccone et al., 1998). This is significant in the context of the ethical issues considered in this chapter because it means that the values of the anthropologists who do skeletal research will often differ dramatically from those of the descendants of the people they study.

Although the prevalence of conviction in an afterlife seems to have changed relatively little during the twentieth century, the cultural context in which it occurs has been dramatically transformed. The familiarity with death that characterized earlier societies in which people were forced to confront the dead directly on a daily basis has been replaced by avoidance of the dead. With the commercialization of the burial process by the “death-care” industry in wealthy countries, traditions such as wakes and ritual preparation of the dead for burial by family members have been replaced by the processing of the dead in remote settings (Badone, 1987; Horn, 1998; Rundblad, 1995).

This cultural trend toward lack of contact with the dead has greatly increased the cultural gulf between a public that has little familiarity with death and skeletal researchers, such as bioarchaeologists, who confront the dead on a daily basis.

**THE HISTORY OF RESEARCH ON HUMAN REMAINS**

Ambivalence toward scientific research on human remains has deep roots in Western societies. From its onset, scientific research on the dead has been the domain of physicians who were often forced to work under clandestine conditions on the bodies of social outcasts. The earliest recorded systematic dissections of a human body were conducted in the first half of the third century B.C., by two Greeks, Herophilus of Chalcedon and Erasistratus of Ceosc. These studies were performed in Alexandria, a city where traditional Greek values were weakened by Ptolemic influences, and probably involved vivisection and the use of condemned criminals (Von Staden, 1989:52–53; Von Staden, 1992). In the ancient world, scientific research of this kind was extremely problematic because it violated Greco-Roman, Arabic, and early
Judeo-Christian beliefs about the afterlife, impurity, and pollution (Bynum, 1994; Eknoyan, 1994; Von Staden, 1992). In the Christian world, anatomical studies of the dead were especially troublesome because many people feared resurrection would be impossible if their body had been dissected. This belief derived from the conviction that at resurrection the actual body is reconnected with the soul. People thus feared that dissection would somehow interfere with this process and leave the soul eternally wandering around in search of lost parts (Bynum, 1994; also see Edgerton, 2003 for similar beliefs held by enslaved Africans in the American South).

During the Renaissance the strength of religious sanctions against dissection began to weaken and, by the sixteenth century, surgeons in Protestant countries such as England were officially given the authority to take the bodies of hanged criminals for use in their anatomical studies. This practice had the dual purpose of furthering the healing arts and serving as a deterrent to criminals who feared the desecration of their bodies (Humphrey, 1973; Wilf, 1989). The repugnance of being dissected was so great that riots sometimes erupted after executions over the disposition of the bodies. Samuel Richardson observed one of these spectacles: “As soon as the poor creatures were half-dead, I was much surprised, before such a number of peace-officers, to see the populace fall to hauling and pulling the carcasses with so much earnestness, as to occasion several warm encounters, and broken heads. These, I was told, were the friends of the person executed, or such as, for the sake of tumult, chose to appear so, and some persons sent by private surgeons to obtain bodies for dissection. The contests between these were fierce and bloody, and frightful to look at (Richardson, 1987).”

As appreciation for the medical value of the information that could be gained through dissection increased, so did the need for anatomical specimens. Soon the demand for bodies for use in teaching and research outstripped the legal supply of executed criminals, and physicians increasingly began to obtain cadavers through robbing graves and hiring body-snatchers who were referred to as “resurrectionists” (Hutchens, 1997; Millican, 1992; Schultz, 1992). This practice was widespread and still persists at medical schools in some economically disadvantaged countries (Ochani et al. 2004). The desire for bodies even led to the series of infamous murders committed by William Burke and William Hare in Edinburgh in the 1820s, with the aim of supplying dissection subjects to Dr. Robert Knox, the anatomist. Hare turned king’s evidence, and Burke was hanged for his crimes, and the incident led to controlling legislation in Britain.

Grave-robbing activities sometimes met with violent public resistance. In 1788, for example, New Yorkers rioted for three days after some children peered through windows of the Society of the Hospital of the City of New York and discovered medical students dissecting human cadavers, one of whom turned out to be their recently deceased mother. A mob of 5000 eventually stormed the hospital and the jail where several doctors had taken refuge. The militia had to be called in and finally dispersed the crowd by firing muskets into it.

To avoid these problems the professional body-snatchers hired by medical schools concentrated on robbing the graves of the poor and powerless. The cemeteries of almshouses were favorite targets and, in the United States, African-American graveyards were favored as places to plunder. Upon visiting Baltimore in 1835, Harriet Martineau commented that the bodies used for dissection were exclusively those of African Americans “because the whites do not like it, and the coloured people cannot resist” (Martineau, 1838:140).

Although much of the early anatomical research focused on resolving issues concerning physiology and surgical anatomy, from the beginning, skeletal studies with a decidedly anthropological flavor were done to answer questions related to human variation and adaptation. As early as 440 B.C., Herodotus (484–425 B.C.) reported on an investigation
into the effect of the environment on the strength of the skull:

On the field where this battle was fought I saw a very wonderful thing which the natives pointed out to me. The bones of the slain lie scattered on the field in two lots, those of the Persians in one place by themselves, as the bodies lay at the first—those of the Egyptians in another place apart from them. If, then, you strike the Persian skulls, even with a pebble, they are so weak, that you break a hole in them; but the Egyptian skulls are so strong, that you may smite them with a stone and you will scarcely break them in. They gave me the following reason for this difference, which seemed to me likely enough: The Egyptians (they said) from early childhood have the head shaved, and so by the action of the sun the skull becomes thick and hard (Herodotus, 1990).

Much of the early anatomical work on human variation had its roots in the belief of Aristotle and his contemporaries that Nature was organized hierarchically as a continuous chain. He was certain that all other animals existed for the sake of Man. This view of the world provided a useful framework for comprehending the enormous complexity of the natural world and also had the appeal of rationalizing the stratified nature of Greek society with powerful rulers and a social elite at the top and the slaves at the bottom (Clutton-Brock, 1995).

By the Middle Ages this hierarchical view of the world had been transformed into the Christian doctrine in which the world was seen as a perfect expression of God’s will that descended in continuous succession through a “Great Chain of Being” from the perfection of the creator to the dregs of things at the very bottom of creation. This perspective permeated much of the work of early natural historians such as John Ray who developed the doctrine of “natural theology,” in which he argued that the power of God could be understood through the study of his creation, the natural world (Ray, 1692). In this context, the description of biological variation, including that found among humans, was a frankly religious activity in which the exploration of the fabric of the natural world at both its macroscopic and its microscopic levels was seen as a way of revealing the “divine architect’s” plan for the universe.

The expanded view of biological diversity provided by the specimens brought back by Columbus and other early European explorers stimulated a frenzy of species description and the first detailed anatomical studies of the differences between apes and humans. Through his careful dissections of a chimpanzee, Edward Tyson (1650–1708) was able to debunk myths based on the reports of classical authors such as Homer, Herodotus, and Aristotle that humankind contained several species, including “satyrs,” “sphinges,” and “pygmies,” and in 1779, Charles Bonnet (1720–1793) wrote a detailed account of the orangutan in which he noted a close relationship to us, albeit with the “lowest races” of our species (Bonnet, 1779; Clutton-Brock, 1995; Tyson, 1966).

After resolving the issue of whether humans and apes are members of the same species, Enlightenment scholars were still faced with the problem of interpreting the previously unsuspected extent of human biological and cultural diversity revealed by European colonial expansion into remote areas of the world. Linnaeus, for example, recognized five divisions of our genus, which included “Homo monstrosus,” a catchall category for a variety of mythical creatures reported by early explorers. The debate soon took on a strong religious flavor and began to focus on how the empirical facts of human variation could be made congruent with biblical accounts of Adam and Eve and the Tower of Babel. Interpretations of human diversity became sharply divided between the adherents of the theory of monogenesis, which traced all humans to a single origin in the Garden of Eden, and the adherents of polygenesis, who rejected the criteria of interfertility as the basis for the identification of biological species and took the unorthodox position that Europeans,
Africans, Asians, and Native Americans were derived from different ancestral forms.

By the end of the eighteenth century, evidence obtained from human skeletal remains began to assume an increasingly important role in these debates over the origins and significance of human biological and cultural differences. Cranial evidence (a total of 82 skulls), for instance, figured prominently in the famous M.D. thesis of Johann Friedrich Blumenbach (1752–1840) in which he argued that modern human diversity had arisen as a consequence of the degeneration of a primordial type (varietas primigenia) whose closest living approximation could be found in the people of the Caucasus Mountains (Blumenbach et al., 1865). Such studies generated considerable interest in human cranial variation, and soon systematic efforts were begun to assemble research collections of human skeletal material from throughout the world.

In the United States, research on population differences in cranial morphology was dominated by Samuel George Morton (1799–1851), a physician from Philadelphia. Morton studied medicine at the University of Edinburgh where he was influenced by theories of polygenism and the hereditarian views of phrenologists that were in vogue at the time (Spencer, 1983). Underlying Morton’s careful craniometric research was the basic theoretical assumption of phrenology: Differences in skull shape corresponded to differences in the shape of the brain and consequent differences in brain function. To test these theories, Morton amassed a large collection of human crania from all over the world that he compared using cranial measurements. From this collection he derived a hierarchy of racial types with Blacks at the bottom, American Indians at the middle, and Whites at the top (Morton, 1839).

Morton’s craniometric approach to understanding human variation set the stage for much of the osteological research done by physical anthropologists during the rest of the nineteenth century. Most of this work was typological in orientation and focused on the classification of people into broad categories such as brachycephalic (round-headed) or dolichocephalic (long-headed) based on ratios of measurements. Although acceptance of the monogeneticists’ theory that all humans trace their ancestry to a single origin gradually increased, especially after the publication of Darwin’s theory of natural selection, a typological, craniometrically oriented approach emphasizing taxonomic description and definition over functional interpretation persisted well into the middle of the twentieth century in the work of influential skeletal biologists such as Aleš Hrdlička (1869–1943) and Ernest Hooton (1887–1954).

There are several reasons for the remarkable tenacity of the typological emphasis in research on human skeletal remains. First, there is the idea that human variation can be adequately accommodated by a few, fundamentally different racial types, which conveniently coincides with beliefs in racial inferiority and superiority that continue to persist in modern societies. The idea of a straightforward relationship between the shape of a person’s skull and their genetic makeup also was seductive to physical anthropologists because it meant that cranial differences could be used as a powerful tool to further one of anthropology’s principle goals: producing detailed reconstructions of population movements and historical relationships. Finally, there is a practical consideration behind the persistence of the typological orientation of skeletal research. Until recently, the computational problems of someone attempting to statistically compare quantitative observations made on skeletal collections of any meaningful size were practically insurmountable. The typological approach, with all of its simplifying assumptions and loss of information on within-group heterogeneity, offered a cost-effective alternative to this practical dilemma.

The last point is nicely illustrated by the anthropometric work of Franz Boas (1858–1942), the founder of American anthropology, and a strong opponent of simplistic hereditarian interpretations of human variation. Through his anthropometric studies of Europeans who
immigrated to the United States, Boas showed that the shape of the cranial vault, a trait nineteenth-century racial typologists had fixated on, is highly responsive to environmental influences and thus of limited value in taxonomic analysis (Boas, 1912). Boas realized the potential of anthropometric research for elucidating the cultural and biological history of our species and from 1888 to 1903 worked to assemble anthropometric data on 15,000 Native Americans and 2000 Siberians (Jantz et al., 1992). In contrast to Hrdlička and many of his other contemporaries, Boas realized the necessity of statistical analysis for understanding the variability within these samples. Unfortunately, the computational capabilities of the data processing tools that were available at the beginning of the nineteenth century (i.e., pencil and paper) made impossible meaningful analysis of the information on human variation contained within this monumental collection of anthropometric observations (Jantz, 1995). Consequently, almost nothing was done with these data until a few years ago when availability of computers with adequate data storage and processing capability made their analysis possible.

During the past 30 years, physical anthropology has finally escaped from the methodological and conceptual shackles of nineteenth-century racial typology. Research on the skeletal remains of earlier human populations has entered a vibrant new phase in which the great potential Boas saw in studies of human variation as a source of insights into the biological and cultural evolution of humankind is beginning to be realized. This paradigm shift has involved replacing the futile nineteenth-century preoccupation with drawing stable boundaries around populations, whose biological and cultural makeup is constantly in flux, with new evolutionary ecological approaches that recognize the complexity and adaptive significance of interactions between genetic variability and developmental plasticity. This theoretical reorientation has resulted in a new bioarchaeological approach to the analysis of skeletal remains from earlier human populations that uses cultural, biological, and paleoenvironmental evidence to illuminate the processes of human adaptation (Larsen, 1997). With this new approach has come an increasing appreciation for the many ways the remains of our ancestors can help us to both better understand and devise solutions to the many seemingly intractable problems of violence, disease, and social inequity that we currently face.

THE SOURCES OF SKELETAL COLLECTIONS

To fully appreciate the concerns that modern indigenous people have about the collections of human skeletons, it is necessary to understand the historical and social context in which skeletal collections have been made throughout history (Walker, 2004). The practice of collecting human skeletal remains as war trophies and for religious purposes has deep historical roots. It has been argued that taking the heads of the dead to obtain their power is among the earliest of ritual practices (La Barre, 1984). In the past, the taking of heads, scalps, and other body parts during warfare was a widespread practice, especially among Native Americans and Melanesians, and can nearly be considered a cultural universal (Driver, 1969; Harner, 1972; Olsen and Shipman, 1994; Owsley et al., 1994; White and Toth, 1991; Willey and Emerson, 1993). Although suppressed in modern societies, such practices continue in the form of the collection of “trophy skulls” from battlefields by modern soldiers (McCarthy, 1994; Sledzik and Ousley, 1991).

Among Christians, the belief that proximity to the bones and other body parts of saints could bring miracles was common as early as the fourth century A.D. This use of human remains as objects of religious veneration gradually resulted in the accumulation of substantial skeletal collections. By the ninth century the remains of martyrs had become so valuable that competition between religious centers created a regular commerce that sometimes degenerated to the point of melees between
monks attempting to seize the bodies of martyrs by force of arms (Gauthier, 1986; Geary, 1978; Thurston, 1913). The belief that the miraculous powers of important religious figures could be accessed through their bones stimulated a lively market in human remains. At one point 19 churches claimed to possess the mandible of John the Baptist (Collin de Plancy, 1821). Philip II (1556–1598) of Spain, a zealous Catholic, commissioned an envoy to collect the remains of as many saints and martyrs as he could, and he assembled a collection of 11 complete skeletons along with thousands of skulls, long bones, and other miscellaneous skeletal elements at his residence, the Escorial near Madrid (Wittlin, 1949). Belief in the magical powers of human remains was not limited to those of Catholic saints. When an Egyptian mummy was obtained by Leipzig, Germany, in 1693, it soon became a tourist attraction owing to the common belief “that it pierceth all parts, restores wasted limbs, consumption, hecksicks, and cures all ulcers and corruption” (Wittlin, 1949).

Until the middle of the eighteenth century, Europe had no museum collections in the modern sense. Instead, there were vast collections held by monarchs and the Catholic Church that functioned as reliquaries, storehouses, and treasuries. During the Enlightenment, a strong belief in the power of empirical investigations of the natural world as a method for the discovery of God’s laws brought with it a need for museums whose purpose was the preservation of historical artifacts and natural objects for scientific scrutiny. At first these collections took the form of “curio cabinets” maintained by wealthy aristocrats for their personal research and the edification of their friends. Many of these early collectors were physicians, and owing to their professional interest in human anatomy, they naturally included human skeletons and preserved anatomical specimens in their cabinets. For example, the large collection amassed by Sir Hans Sloane (1660–1753), the personal physician to Queen Anne and King George II, included several human skeletons. Upon Sloane’s death, these skeletons and the rest of his collection were bequeathed to the British Parliament at a nominal sum and served as the nucleus of the British Museum’s natural history collection. In America, scholarly associations such as The Library Company of Philadelphia, which was formed in 1731 by Benjamin Franklin and his colleagues, began to maintain collections that included anatomical specimens, and around the same time, the Pennsylvania Hospital in Philadelphia established its teaching cabinet with the acquisition of a human skeleton and a series of anatomical models (Orosz, 1990:16–17).

These collections of skeletons and anatomical specimens were of great value because they made it possible to provide instruction in surgical anatomy without offending Christians who had religious objections to the dissection of cadavers. During the last half of the eighteenth century, the inadequacies of the old system of learning anatomy by studying models and occasionally observing a demonstrator dissect a criminal’s body became increasingly apparent. With the growth of medical knowledge, aspiring surgeons began clamoring for more hands-on experience so they could avoid the horrifying prospect of learning their trade through the butchery of their first living patients. This desire was reinforced by a growing public recognition of the value of being operated on by someone with practical experience in dissection.

These social pressures resulted in an exponential increase in the demand for cadavers. To meet this need, “anatomical acts” were eventually passed that expanded the legal sources of cadavers to include the victims of duels, suicides, and most importantly unclaimed bodies. The demand was so great that this new legal supply of bodies was often inadequate, and throughout the nineteenth-century, medical schools were still enlisting the services of body-snatchers to obtain their instructional materials (Blake, 1955; Blakely et al., 1997; Newman, 1957).

Although the increase in dissections opened the possibility of increasing the scope of skeletal collections, this potential was not fully realized.
Collections were made of specimens with interesting anomalies and pathological conditions but, as a rule, the rest of the dissected person’s skeleton was disposed of in what often seems to have been a cavalier fashion (Blakely and Harrington 1997:167). From what can be discerned from the remnants of nineteenth-century medical school collections that survive today, little effort was made to create carefully documented skeletal collections of known age and sex for use in assessing the normal range of human variation. The failure to create such systematic collections probably stems in part from the prevalence of racist views that minimized the importance of variation within groups and exaggerated the significance of population differences.

The immensity of the carnage brought by the Civil War profoundly affected attitudes toward the dead, in the United States (Laderman, 1996). The war desensitized people to death, and this made it possible to view corpses with increasing detachment. At the same time, the logistic problems the military faced in preserving the bodies of so many dead soldiers for transportation back to their families turned corpses into commodities that needed to be processed by professionals such as doctors and undertakers. In this context of mass slaughter, rising professionalism, and growing rejection of religious beliefs in the resurrection of the body, surgeons struggling to devise standardized treatments for the sometimes horrifying injuries they faced began to view autopsies and other medical research on dead soldiers as an ethical imperative. To accommodate this research the Army Medical Museum was founded in 1862 as a repository for thousands of skeletal specimens, preserved organs, photographs, and other medical records obtained during the treatment and autopsy of military casualties (Barnes et al., 1870; Otis and Woodward, 1865).

At the close of the Civil War, Army doctors shifted the focus of their collecting activities toward medical concerns arising from the Indian Wars in the Western United States, such as the treatment of arrow wounds (Bill, 1862; Parker, 1883; Wilson, 1901). One aspect of this work involved the collection of Native American crania and artifacts from battlefields and cemeteries. This was implemented through a letter from the Surgeon General’s Office, dated January 13, 1868, that stated: “Will you allow me to ask your kind interposition in urging upon the medical officers in your departments the importance of collecting for the Army Medical Museum specimens of Indian Crania and of Indian Weapons and Utensils, so far as they may be able to procure them.” Other documents make it clear that these collections were made under the protest of the Indians whose graves were being raided and that such activities could even result in further hostilities with the Indians (Bieder, 1992). Although government-sanctioned grave robbing of this kind eventually stopped, it understandably continues to provoke outrage among the descendants of the people whose bodies were stolen (Riding In, 1992).

Beginning in the middle of the nineteenth century, large public natural history museums began to be established whose goals were both popular education and scholarly research (Orosz, 1990). These museums provided an institutional framework within which the large skeletal collections could be consolidated from the smaller private collections of physicians and wealthy amateur archaeologists. These new museums had the resources necessary to maintain staffs of professional research scientists and to augment their osteological collections through purchases from private collectors and the sponsorship of archaeological expeditions throughout the world.

In the United States, the most important natural history museums from the perspective of collections of human skeletal remains are the Smithsonian Institution founded in 1846, the Peabody Museum of Archaeology and Ethnology founded in 1866, the American Museum of Natural History founded in 1869, the Harvard Peabody Museum of Archaeology and Ethnology founded in 1866, the Columbian Museum of Chicago (now the Chicago Field Museum) founded in 1893, the Lowie
Museum of Anthropology (now the Phoebe Hearst Museum) founded in 1901, and the San Diego Museum of Man founded in 1915. During the twentieth century the number of museums with significant holdings of human skeletal remains rapidly increased, and by 1998, about 700 federal and private institutions possessed skeletal remains from an estimated 110,000 individuals.

The research value of these collections varies enormously depending on the conditions under which they were collected. Because of the cranial typological orientation of nineteenth-century physicians, most of the material collected before the beginning of the twentieth century consists of isolated crania, lacking associated mandibles or infra-cranial remains. Because of the predisposition of these researchers to interpret human variation within a framework of stable types that were comparatively immune to environmental influences, most of them lack adequate provenience information and are simply labeled in terms of preconceived racial categories or broad geographical regions. All of these factors greatly reduce the value of such collections for research purposes. Fortunately, most of the skeletal material in museums derives from the work of professional archaeologists and is associated with at least some contextual information that allows the individual to be placed in a meaningful historical, environmental, and cultural context. This type of information is essential for modern bioarchaeological research, which relies heavily on contextual information to reconstruct the cultural ecology of earlier human populations.

During the first half of the twentieth century, several visionary anatomists realized the value of having skeletons from individuals of known age, sex, and ethnic background for use in anthropological and forensic research on the effects that environmental and genetic factors have on health, disease, and morphological variation. Working in conjunction with the teaching programs of medical schools, these researchers carefully recorded anthropometric data, vital statistics, health histories, and other relevant information on the people scheduled for dissection. Afterward they prepared their skeletons for curation in research collections. Three of the largest of these dissection room collections were established in the United States, at the Washington University School of Medicine in St. Louis, the Western Reserve University in Cleveland, and Howard University in Washington, D.C.

A central figure in the creation of these collection efforts is William Montague Cobb (1904–1990). Cobb, an African American, who was an acknowledged activist leader in the African-American community, realized the value that empirical data on human variation has as an antidote to racism. After receiving his M.D. at Howard University, he did postgraduate studies at the Western Reserve University where he helped T. Wingate Todd (1885–1938) assemble that university’s skeletal collection. After writing a Ph.D. dissertation on anthropological materials, which included information on the geographic and ethnic origins of the people who contributed their skeletons to the Western Reserve collection, Cobb returned to Washington where he created a similar collection at Howard University (Cobb, 1936). A prolific author and dedicated teacher of anatomy, Cobb used his understanding of human biology, which in part was derived from dissections and skeletal research, to improve the health and reinforce the civil rights of African Americans (Cobb, 1939, 1948; Rankin-Hill and Blakey, 1994).

In Great Britain and Europe, a different approach has been taken to the creation of known age and sex skeletal collections for use in anthropological research. The crypts outside Saint Bride’s Church, London, were disturbed through bombing during World War II. Restoration of the church has resulted in a documented collection of skeletal remains dating from the mid-eighteenth century (Huda and Bowman, 1995; Scheuer and Bowman, 1995). Similar collections of people of known age and sex from historic cemeteries have
been established in Coimbra, Portugal (Cunha, 1995), Lisbon, Portugal, Geneva, Switzerland (Gemmerich, 1997), and Hallstatt, Austria (Sjøvold, 1990, 1993). However, a great many anatomical collections of skeletons of nineteenth- and twentieth-century individuals exist in anatomy departments and medical schools throughout Europe, Britain, and other countries.

THE VALUE OF HUMAN SKELETAL REMAINS

In the ongoing debate over the disposition and scientific analysis of ancient human remains in museum collections, there is a tendency for the ethical issues surrounding skeletal research and the maintenance of skeletal collections to be reduced to simplistic oppositions: science vs. religion, right vs. wrong, and so on. Although framing the complex social issues underlying the debate in this way may be politically expedient, it is counterproductive for anyone seeking a solution that balances the concerns of descendants against those of the scientific community.

From my brief discussion of the evolution of beliefs about human remains, it is obvious that the details of the rituals people have devised for the treatment of the dead have varied enormously among the cultures of the world through time. The practice of funeral rites by friends and relatives and the use of a method of disposing of the body seen to be human universals, but beyond that, there is little uniformity (Brown, 1991; Murdock, 1945). This diversity of beliefs about how the dead should be treated poses ethical dilemmas for bioarchaeologists when their scientific work conflicts with the beliefs of the descendants of the people whose remains they study.

One approach to resolving disputes over research on ancient skeletal remains is to view such disagreements as cultural issues arising from competing value systems (Goldstein and Kintigh, 1990). Conceiving of disputes over the treatment of the dead as products of conflicting value systems avoids polemics and self-righteous posturing in which each side battles for moral superiority and instead promotes communication and mutual understanding. This can eventually result in the discovery of solutions that are consistent with the value systems of both parties in the dispute.

The only justification for the study of skeletal remains from earlier human populations is that such research yields information that is useful to modern people. Although the value of skeletal research seems self-evident to the people who conduct it, many indigenous people feel such work is not only useless, but also extremely harmful because of the damage it does to them and the spirits of their ancestors (Sadongei and Cash Cash, 2007). This conflict between the values scientists and descendant groups attach to human remains is central to the most important ethical dilemmas bioarchaeologists face. Since mutual understanding is a prerequisite for finding a common ground between these apparently incommensurable world views, it is useful to briefly describe the values scientists and descendant groups attach to ancient human remains.

Bioarchaeologists focus their research on ancient human skeletal remains, not out of idle scientific curiosity, but instead because they believe that the information contained within the remains of our ancestors is of great value to modern people (Larsen and Walker, 2004). Human skeletal remains are a unique source of information on the genetic and physiological responses our ancestors made to the challenges posed by past natural and sociocultural environments. Consequently, they provide an extremely valuable adaptive perspective on the history of our species.

Most of what we know about our recent history is based on inferences derived through analysis of artifacts, documents, oral histories, and other products of human cultural activity. Because of their symbolic content, such cultural artifacts are difficult to interpret and often consistent with multiple, sometimes contradictory views of the past. The subjective aspects of attempting to interpret cultural artifacts, from the perspective of our current cultural milieu
are well recognized: Historical works often reveal more about the cultural values and political biases of the historian than they do about the reality of the historical event being described. All historians are products of the culture in which they live, and they are always selective in what they report.

Because of its biological basis in the physiological processes of growth, development, and acclimatization to environmental change, the information about interactions with past environments encoded in human remains provides an extremely valuable comparative basis for evaluating interpretations of the past based on artifacts, documents, and other culture-based sources. The historical data provided by skeletal studies are of such great value because the methodological problems inherent in extracting evidence from a skeleton are completely different from those historians face when they attempt to interpret the historical significance of the cultural products with which they work. The only way we can reduce the cultural biases that distort our understanding of past events is through collecting a diversity of evidence from sources that are susceptible to different types of interpretative error. The greater the diversity of the evidence we have about the past, the easier it is to rule out alternative interpretations that are unlikely to reflect actual events. By using a series of data sources that, standing alone, would be open to many different interpretations, it is in this way possible to triangulate on what really happened in the past.

The unique perspective that skeletal evidence provides on the history of our species makes it a potent weapon against cultural relativists and historical revisionists who view the past as a source of raw materials they can exploit to refashion history into whatever narrative is currently considered au courant or politically expedient. In some schools of postmodernist thought, history is viewed as a symbolic construct devoid of any objective truth: All we are left with is an endless process of constructing conflicting narratives about the past that are all of equal merit or are only of merit because they are different. In some rarified corners of the humanities, the possibility of knowing with certainty that voluminously documented historical events such as the Holocaust actually occurred is actively debated (Braun, 1994; Friedman, 1998; Jordan, 1995; Kellner, 1994). In the world of these theorists, people interested in discovering what happened in the Holocaust are doomed to an academic life of continuously revisualizing and recontextualizing subjective impressions of subjective descriptions of the slaughter of millions of people into new, contradictory, and from their perspective, more meaningful imaginations of the past.

In contrast to the symbolic problems inherent in historical reconstructions based on written records and oral histories, human skeletal remains provide a direct source of evidence about the lives and deaths of ancient and modern people, that is, at a fundamental level, free from cultural bias (Walker, 1996, 1997). The skeletons of the people buried row upon row at concentration camps such as Terezin, the racks of skulls from the Cambodian killing fields at Tuol Sleng Prison, and the cut marks on the skeletons of the hundreds of massacred prehistoric Native Americans unceremoniously buried at the Crow Creek site in South Dakota speak volumes about real historical events that ended the lives of real people.

In certain respects, bones do not lie. To give a specific example from my own research, the presence of lesions indicative of severe, repeated physical abuse in the skeletons of children murdered by their parents says something very specific about a history of traumatic experiences that a child suffered during its short life (Walker, 2001; Walker et al., 1997). Although multiple “narratives” can be constructed based on the presence of such lesions (the child was extraordinarily clumsy or accident prone, the child’s parents repeatedly beat him over a prolonged period until he died, and so on) at a fundamental level, such skeletal evidence says something indisputable about a physical interaction that took place between the dead child and his or her physical environment. Unlike written records or oral histories, human
remains are not culture-dependent symbolic constructs. Instead they provide an extraordinarily detailed material record of actual physical interactions that occurred between our ancestors and their natural and sociocultural environments. As such, human remains are extremely valuable sources of evidence for reconstructing what actually happened in the past.

This esoteric view that bioarchaeologists hold concerning the central role that collections of human skeletal remains play in helping us to obtain an objective view of history is not widespread. Most of the world’s population views human remains with a mixture of morbid fascination and dread because they serve as such vivid reminders of one’s own mortality and impending death. The symbolic saliency of directly confronting a dead person has been deftly exploited for a diversity of religious, political, and economic purposes. Throughout the world, in many different settings, human remains are placed on public display and used in ways designed to foster group cohesion and legitimize religious or political authority. During times of social instability, it is common for these same remains to be destroyed or humiliated to weaken and disrupt the group solidarity they once fostered (Cantwell, 1990). The controversy of the continued display of Lenin’s remains in Red Square and the disposition of the recently discovered remains of Czar Nicholas II and his family provide good examples of how human remains can be used as tools to advance or suppress political ideas and facilitate or disrupt social cohesion (Caryl, 1998; Fenyvesi, 1997).

The strong symbolic power of human remains has encouraged people to devise an amazing number of uses for them. Throughout the world, displays of human remains are among the most effective tools for luring people into museums (Brooks and Rumsey, 2007). At the British Museum, for example, postcards of mummies rival the Rosetta Stone in public popularity (Beard, 1992). In many places displays of human remains are such popular tourist attractions that they have become the mainstays of local economies. The Museo de los Momias in Mexico, where the naturally mummified bodies of poor people who could not afford to purchase permanent graves are on display, is touted as Mexico’s second most popular museum, bested only by the anthropological museum in Mexico City (Osmond, 1998). Two similar examples are the awe-inspiring creativity of displays of thousands of human bones disinterred from a cemetery near Kutna Hora in the Czech Republic and in the Church of the Capuchins in Rome (Fig. 1.1).

In some cases the symbolic value of retaining human remains for display is sufficient to override religious sanctions against it. Medieval Chinese Ch’an Buddhists practiced mummification of eminent priests as demonstrations of the relationship between spiritual attainment and the incorruptibility of the body even though they espoused a religious doctrine that accorded little value to the corpse. A similar example is the recent decision that the value of the display of bones from Khmer Rouge victims at the Tuol Sleng Prison Museum as evidence of the Cambodian genocide outweighed Buddhist religious beliefs that mandate cremation (Erlanger, 1988; Peters, 1995). The denial of burial in Christian countries as a form of posthumous punishment and object lesson for the living has already been mentioned. In England, the heads of people such as Oliver Cromwell were displayed on poles erected on the roof of the Great Stone Gate of London Bridge, and gibbets containing the rotting bodies of famous pirates such as Captain Kidd were strategically placed along the banks of the Thames to greet sailors as they returned from the sea. During the nineteenth century, the heads of Miguel Hidalgo and three other leaders of the Mexican war of independence met a similar fate when they hung on public display in cages for 10 years as grim reminders of the folly of revolution. Ironically, these same skulls of Mexico’s founding fathers have recently been resurrected and again put on public display for the opposite purpose: They rest next to each other under glass on red velvet in a dimly lit crypt where they remind
school children of the heroism of the country’s founders (Osmond, 1998).

As is illustrated by the case for Hidalgo’s skull, the strong symbolic value of human remains endows those who control the remains with a powerful tool that can be used to vividly express multiple, sometimes contradictory, meanings. Because of this great symbolic power, it is not surprising that issues surrounding the control, treatment, and disposition of human remains pose some of the most vexing ethical dilemmas skeletal biologists face. Bioarchaeologists do not view human remains primarily as symbols. Instead they value them as sources of historical evidence that are key to understanding what really happened during the biological and cultural evolution of our species. This lack of concern with symbolic issues is in stark contrast to the richness of the symbolic connotations human skeletons have for most people.

This conflict in worldviews is especially acute in areas of the world that were subjected to European colonization. In North America, Hawaii, and Australia, where the indigenous people suffered the greatest devastation at the hands of European colonists, ancient human remains have assumed great significance as symbols of cultural integrity and colonial oppression (Sadongei and Cash Cash, 2007:98). In this postcolonial world, gaining control over ancestral remains is increasingly considered essential to the survival and revitalization of indigenous cultures.

That the views of indigenous people concerning this issue have changed dramatically during the past 40 years is amply illustrated by archaeological reports that describe the enthusiastic participation of Native Americans in the excavation of burials, some of whose study by bioarchaeologists is currently under dispute (Benson and Bowers, 1997; Brew, 1941; Fewkes, 1898; Hewett, 1953; Hrdlička, 1930a, 1930b, 1931; Hurt et al., 1962; Judd, 1968; Neuman, 1975; Roberts, 1931; Smith, 1971; Smith et al., 1966). As late as the 1960s, Inuit people in the Northwest Territory of Canada, with whom I worked, seemed little
concerned about the excavation of ancient skeletal remains. In fact, they were extremely cordial to the members of the expedition I was on and assisted us in any way they could. Although they expressed mild concerns about carrying human skeletons in their boats, they otherwise were supportive of and expressed considerable interest in our bioarchaeological work.

To comprehend the urgency of the current concerns Native Americans have about the treatment of their ancestral remains, it is necessary to understand the magnitude of the recent disruptions of their cultures. Beginning at the end of the nineteenth century, systematic attempts began to be made to separate Native American children from their families, suppress their Native identities, and inculcate them with Christian values (Ellis, 1996a; Lomawaima, 1993). Simultaneously, the isolation that formerly characterized life on the remote reservations in marginal areas that the government relegated to them began to break down because of the development of interstate highways, radio, television, and the intrusions of tourists. These developments have had such a devastating effect on the transmission of traditional beliefs and practices that the remnants of earlier times preserved in museums have increasingly become a cultural focus. Control over these collections is an important political issue for Native Americans because, by gaining control over the biological and cultural remains of their ancestors, they can begin to reassert their cultural identity within the dominant Euro-American culture.

When viewed within this context of cultural marginalization and repression, it is easy to see why many indigenous people see little value in what to them are the very nebulous goals of bioarchaeologists. Zimmerman (Ubelaker and Grant, 1989) presents evidence supporting the depth of Indian concern about the retention of museum collections. He cites an unpublished survey that John S. Sigstad conducted in 1972 of Indian tribes in the BIA Aberdeen region. All respondents agreed that human remains in museums should be reburied, 95% indicated bones should not be displayed in museums, and only 35% of the respondents believed that human remains should be excavated for scientific research (Ubelaker and Grant, 1989).

Some indigenous people have the erroneous belief that only the remains of their ancestors are studied and cite this as a reflection of the racist attitudes of the European colonists who robbed them of their land (Tobias, 1991; Vizenor, 1986). They feel that such research degrades them by singling them out to be “made fun of and looked at as novelties” (Mihesuah, 1996; Walters, 1989). Bioarchaeologists respond to this charge by pointing out the vast collections of non-Native-American skeletal remains in European museums and by arguing that it would be racist not to have collections of Native-American remains in New World museums, since this would imply that knowledge of the history of the indigenous people of the New World had nothing to contribute to the understanding of our common past (Ubelaker and Grant, 1989).

Some indigenous people reject the epistemology of science, at least as it applies to their history and cultural affairs, and instead they prefer to view the past as it is revealed through traditional ways of knowing such as oral history, legend, myth, and appeal to the authority of revered leaders. For people with this perspective, scientific research directed toward documenting the past is not only superfluous, but also potentially culturally subversive, because of the capacity of scientific evidence to conflict with traditional beliefs about the past and, in this way, undermine the authority of traditional religious leaders. From this perspective, scientific investigations into the history of indigenous cultures are simply another manifestation of the attempts of an oppressive imperialist colonial power to control and weaken the belief systems of indigenous people so that they will be easier to exploit (Bray, 1995; Dirlik, 1996; Riding In, 1996).

In academia, this position clearly resonates with radical postmodernist theorists of the humanities who believe that reconstructing
history as an objective reality is a hopeless endeavor and instead argue that history is a symbolic weapon that ethical people should use to help the marginal political and cultural constituencies of the world in their struggles against the holders of power (Hodder et al., 1995).

This tension between traditional and scientific views of the past has recently been brought into sharp focus through the controversy over the disposition of the 9300-year-old human remains found at the Kennewick site on the banks of the Columbia River in Washington (Hastings and Sampson, 1997; Lemonick, 1996; Morell, 1998; Petit, 1998; Preston, 1997; Slayman, 1997; Watkins, 2004). Scientists who have examined these remains say they possess characteristics unlike those of modern Native Americans. They believe that research into reasons for this difference has the potential to make an important contribution to our understanding of the history of humankind. Members of the five Native-American tribes that have claimed the skeleton, on the other hand, believe that the question of the cultural affiliation of this individual has already been resolved by their elders who tell them that they have lived in the area in which the skeleton was found since the beginning of creation. The complexity of this dispute increased further when members of the Asutru Folk Assembly, a traditional European pagan religion, sued for the right to use scientific research to decide whether this individual is one of their ancestors. They claim that, “It’s not an accident that he came to us at this time and place . . . Our job is to listen to (the bones) and hear what they have to say” (Lee, 1997).

Modern indigenous people often frame such disputes over the power to control the interpretation of tribal history in spiritual terms. It is a common pan-Indian religious belief that all modern Native Americans are spiritually linked to all other Indian people living and dead (Walters, 1989). Another widely held belief is that space is spherical and time is cyclical (Clark, 1997). All living Indians thus have a responsibility for the spiritual well-being of their ancestors that requires them to assure that their ancestors are buried in the ground where they can be reintegrated into the earth and complete the circle of life and death (Bray, 1995; Halfe, 1989). Contemporary Native Americans who hold these beliefs argue that, so long as ancestral spirits are suffering because their bones are not buried in the earth, living people will continue to suffer a myriad of adverse consequences. Thus, any activity inconsistent with reburial, such as excavation, study, museum curation, and storage, is considered an act of desecration and disrespect. For indigenous people with such views, there is no middle ground on which scientific research can be conducted on human skeletal remains and associated artifacts. These remains are of great spiritual and psychological importance, and their reburial is required to heal the wounds of colonial oppression (Emspak, 1995; Murray and Allen, 1995; Sadongei and Cash Cash, 2007).

**ETHICAL RESPONSIBILITIES OF SKELETAL BIOLOGISTS**

Given these sharply polarized views concerning the value of scientific research on human remains, what are the ethical responsibilities of skeletal biologists? On the one hand, we have bioarchaeologists who believe that the historical evidence obtained from human remains is critical for defending humankind against the historical revisionist tendencies of repressive, genocidal political systems, and on the other hand, we have indigenous people who believe that the spirits of their ancestors are being tortured on the shelves of museums by racist genocidal, colonial oppressors. If we can accept the relativist perspective that both of these views have some validity, then it is possible to envisage a compromise that gives due recognition to both value systems.

Although there is still a broad spectrum of perceptions of what is right and what is wrong among modern people, with the precipitous decline in cultural diversity that has occurred because of the expansion of modern
communication systems, we are seeing a worldwide convergence of values, at least concerning certain areas of human affairs (Donaldson, 1992). These shared values are developing as part of the evolution of the transnational political and economic systems that are beginning to unite the world’s disparate cultures. The Declaration of Human Rights of the United Nations, for example, provides a generally accepted set of rules for ethical human behavior that most people can accept in principle, if not in practice. They include recognition of the right to equality, freedom from discrimination, freedom from torture and degrading treatment, freedom from interference with privacy, and freedom of belief and religion (UN, 1948). Other attempts to devise a set of ethical rules that encompass what some people believe is emerging as a culturally universal system of moral principles include widespread humanistic values such as the recognition that it is wrong to be indifferent to suffering, that tolerance of the beliefs of others is good, and that people ought to be free to live as they choose without having their affairs deliberately interfered with by others (Hatch, 1983).

The cultural values expressed by the assertion of basic human rights and universal moral principles such as these can be criticized as hegemonic attempts to use Western cultural ideas as tools for gaining power and political control for transnational business interests. For example, the Chinese government has recently criticized allegations concerning its suppression of the rights of political dissidents, as insensitive to unique Chinese cultural values such as obedience to authority, collectivism, family, and other dispositions (Li, 1998).

This issue of developing universal, government-sponsored standards of ethical behavior is of more than theoretical interest to bioarchaeologists since it is commonly asserted that the maintenance of skeletal collections for use in scientific research is a violation of a fundamental human right. For example, Article X of the draft of the “Inter-American Declaration on the Rights of Indigenous Peoples” approved by the Inter-American Commission on Human Rights of the Organization of American States in a section entitled “Spiritual and religious freedom” specifically states that when “sacred graves and relics have been appropriated by state institutions, they shall be returned” to indigenous people (IACHR, 1995).

At the opposite end of the spectrum of political inclusiveness and governmental authority from the UN and OAS statements on human rights are the ethics statements that professional associations develop for their members to use as guides for the decisions they make during their everyday activities. The decline in the capacity of organized religions and other traditional social institutions to impose a unifying set of ethical principles acceptable to modern multicultural societies and the constant stream of ethical challenges posed by new technological developments has stimulated enormous interest in the formulation of standards for ethical conduct in many areas of professional activity (Behi and Nolan, 1995; Bulger, 1994; Fluehr-Lobban, 1991; Kruckeberg, 1996; Kuhse et al., 1997; Kunstadter, 1980; Lynott, 1997; Muller and Desmond, 1992; Navran, 1997; Parker, 1994; Pellegrino, 1995; Pyne, 1994; Salmon, 1997; Scanlon and Glover, 1995; Schick, 1998).

Many professional associations and governmental agencies have developed ethical guidelines for use by researchers in the biomedical and social sciences that contain information directly relevant to resolving the ethical dilemmas bioarchaeologists face when they work with ancient human remains (AAA, 1986, 1997; AIA, 1991, 1994; CAPA, 1979; MRCC, 1998; NAPA, 1988; NAS, 1995; SAA, 1996; SOPA, 1976, 1983; UNESCO, 1995).

Although only a few of these statements deal specifically with issues surrounding the study of human remains, a comparison of the principles for ethical behavior they espouse suggests considerable agreement on a few fundamental rules that can be used to guide researchers who work with ancient human remains: (1) human remains should be treated with dignity and respect, (2) descendants should have the authority to control the disposition of the remains of their
relatives, and (3) because of their importance for understanding the history of our species, the preservation of collections of archaeological collections of human remains is an ethical imperative.

Each of these principles is based on a complicated set of value judgements whose implications for the real-world practices of skeletal biologists depend in many ways on the cultural lens through which they are viewed. For example, what is considered the dignified treatment of human remains varies widely depending on a person’s cultural background. These ethical principles also contain an inherent contradiction since recognizing the rights of descendants may at times conflict with the preservation ethic.

Respect for Human Dignity

The ethical principle that human remains should be treated with respect and dignity is consistent with, and can be seen as an extension of, respect for human dignity, which is the cardinal ethical principle for modern research on human subjects in the biomedical and social sciences (Margareta, 1996; MRCC, 1998; Teague, 2007; UNESCO, 1995). This ethical principle is based on the belief that it is unacceptable to treat human remains solely as a means (mere objects or things), because doing so fails to respect the intrinsic human dignity of the person they represent and thus impoverishes all of humanity. Although an argument can be made that since the remains of dead people are just that, “decaying organic matter” that “feels nothing, conceptualizes nothing, has no interests, and cannot suffer,” in other words, that there is no person here to respect or disrespect, the respect is not for the body, but the ante-mortem person from whom the remains are derived (Lynch, 1990). Although it is true that, for most skeletal biologists, human remains are viewed as depersonalized and desanctified, there is still general agreement that they are nevertheless highly meaningful and should be treated with dignity and respect (Buikstra, 1981; Ubelaker and Grant, 1989).

A skeptic might question the wisdom of extending the concept of human dignity to the dead: What does the treatment of human remains have to do with human rights or human dignity? In view of the atrocities currently being perpetrated on helpless people by repressive governments throughout the world, would it not be more productive to focus the fight for human rights on living people who could actually benefit from the results? In my view, a convincing argument can be made that, although the human being that skeletal remains are derived from no longer exists, their former intimate association with a living person is more than sufficient to earn them respectful treatment. The logic of this argument is similar to that used by animal rights activists who admit that, although animals by definition do not have human rights, their ill treatment does demean humans and thus has implications for human behavior (Man’s Mirror, 1991; McShea, 1994). In the same way it can be argued that disrespectful treatment of human remains is morally repugnant because of its potential to desensitize people in a way that is likely to encourage a lack of respect for and consequent ill-treatment of the living (Grey, 1983:105–153).

If we accept the premise that it is unethical to treat human remains with disrespect, we are still faced with the problem that respectful treatment is a highly subjective concept. The cultures of the world have devised an enormous variety of ways of respecting the dead that include hanging the skulls of close relatives from the rafters of huts, using skulls of parents as pillows, and letting vultures feed on dead relatives. Some modern people believe that pumping dead relatives full of chemicals, dressing them up, and burying them in the ground is respectful. Others believe that incinerating them, grinding up what’s left in a mill, and putting the resulting bone meal in a cardboard box is respectful. In the cultural context of scientific research, respect for human remains derives not only from their association with a person
who was once alive, but also from an appreciation of the information about the past they can yield. To a scientist, respectful treatment of human remains includes taking measures to ensure the physical integrity of the remains and the documentation associated with them, avoiding treatments that will contaminate or degrade their organic and inorganic constituents, and so on (Alfonso and Powell, 2007).

These convoluted academic arguments about the definition of and justification for treating human remains with respect, of course, seem bizarre to indigenous people who view ancestral remains, not as inanimate objects devoid of life but instead as living entities that are imbued with ancestral spirits. From the perspective of some Native Americans, for example, ancient human skeletons are “not just remains, they’re not bone to be studied, you’re dealing with spirits as you touch those remains” (Augustine, 1994). As Rachel Craig, a Native Alaskan put it, “I feel an obligation to give back to them, to speak for them. Our grandmothers have told us the importance of the spirit world. The spirits of those people cannot rest and make their progress in the spirit world unless they know that those bones are put back in the earth where they belong. That is our teaching” (Craig, 1994). This same view of the retention of skeletons in museums as interfering with the afterlife and separating the spirits of the dead from the community of the living is forcefully expressed by William Tallbull, a member of the Northern Cheyenne tribe: “We talk about people coming home. When the people came home from the museum and are buried at home, they all go and visit every house. This is where the joy comes in. They are home. They are here. They walk around through the village and become part of us again. That’s all we are asking” (Tallbull, 1994).

**Descendant Rights**

Since disputes over who should have the right to control the disposition of ancient human remains are central to many of the ethical dilemmas bioarchaeologists face, it is useful to consider this issue as broad a perspective as possible. Giving close relatives authority to make decisions about the disposition of the remains of the recent dead seems to be a cultural universal. Only in exceptional circumstances, such as the special dispositions mandated for the bodies of executed criminals as part of their punishment, and the control that coroners are given over bodies that might yield evidence relevant to legal proceedings, is the right of close relatives to decide the disposition of a body denied. Many cultures have special rules governing the disposition of the bodies of people who die under unusual circumstances, and some of these make exceptions to the rule of kin control over the dead. Herodotus, for example, observed that the Egyptians gave special treatment to the bodies of people who drown in the Nile or were eaten by crocodiles: “No one may touch the corpse, not even any of the friends or relatives, but only the priests of the Nile, who prepare it for burial with their own hands—regarding it as something more than the mere body of a man—and themselves lay it in the tomb” (Herodotus, 1990).

Considering the universal recognition of the rights of relatives, it not surprising that this is one issue on which, as far as I know, all bioarchaeologists agree: If skeletal remains can be identified as those of a known individual for whom specific biological descendants can be traced, the disposition of those remains, including possible reburial, should be decided by the closest living relatives.

Many of the ethical dilemmas that skeletal biologists face arise, not out of a disagreement over this fundamental principle of ethical behavior but, instead, over how the rights of descendants should be recognized in real-world situations. The first problematic area concerns how the rights of relatives with different relationships to the dead person should be balanced against each other. In modern legal systems, authority over the dead is judged using a rigid hierarchy of rights. For example, the Uniform Anatomical Gifts Act establishes the following order of priority for people authorized to make decisions about the authorization.
of removal of body parts: (1) the spouse, (2) an adult son or daughter, (3) either parent, (4) adult brother or sister, (5) the person’s legal guardian at the time of death, and (6) any other person authorized to dispose of the body. Even here, there is considerable room for cultural variation in rules governing control over the dead. In China, for example, because of its pervasive patriarchal family structure, authority of the wife regarding funeral arrangements is likely to be less than that of the male members of his patriline (Cooper, 1998).

In contrast to the agreement about giving lineal descendants control over the disposition of the remains or close relatives, there is no consensus concerning the question of the appropriate way to decide the disposition of human remains that are distantly related to living people. What is the ethical way to decide the disposition of the remains of people who are many generations removed from any living person? How shall we weigh the many attenuated genetic and cultural ties that link large numbers of living people to ancestors who lived thousands, hundreds of thousands, or even millions of years ago? Which living individuals should be granted the moral authority to decide the disposition of our ancient ancestors?

The basic elements of the dilemma can be better understood from a scientific perspective by considering how the genetic and cultural connections that link modern people and earlier generations vary as a function of time. The first problem is that the more distant an ancestor is from a descendant, the more descendants there are sharing the same genetic relationship to that ancestor. The variables that influence the number of shared ancestors that living people have are complex. However, one fact is indisputable: As we probe more deeply into our family tree, the probability of discovering an ancestor we share with a large number of other living people increases dramatically. In a lineage of people who each had two children and did not marry relatives, it would take seven generations, or about 250 years, to produce over five billion modern descendants. People, of course, tend to marry relatives and not everyone has the same number of children. Even if we account for these complicating variables, the fact remains that many living people are likely to be related to an individual who lived many generations ago.

If we really believe that relatives should decide the disposition of ancestral remains, how can we identify those descendants and allow them to make a collective decision about the proper treatment of their relative’s bones? The problem of linking modern people to our hunter-gatherer ancestors is complicated by the highly mobile lifestyle of such populations. This decreases the likelihood that the ancestral remains of a modern group will be found in the territory in which that modern group currently resides. In situations of population replacement, it is in fact more likely that the modern people who now live in an area were directly responsible for the extermination of the ancient people who formerly occupied that same territory.

Even in cases where it is clear that descendants continue to occupy the land of their ancestors, there is still the problem posed by the expansion of living descendants with increasing genealogical remoteness. In an area such as Europe, with a relatively stable gene pool, someone who died more than a few hundred years ago is likely to be related to hundreds of thousands, if not millions, of living people. For instance, DNA studies conducted on the 5000-year-old mummified body recently found in the Tyrolean Alps suggests a genetic relationship between this person and the 300 million or so contemporary people living in central and northern Europe (Handt et al., 1994). This of course does not include many millions of additional people living in North America and elsewhere with ancestral ties to northern Europe.

In the Western Hemisphere, the problem of assigning rights for the control of ancestral remains to living descendants is complicated by gene flow between indigenous Americans and the people of Europe, Africa, and Asia. For example, geneticists estimate that 31% of
the contemporary gene pool of people identified as Hispanic or Mexican Americans is derived from their Native-American ancestors (Gardner et al., 1984; Hanis et al., 1991). These Native-American descendants are thus numerically a very significant component of the New World population and, if demographic trends continue, are likely to replace Non-Hispanic Euro-Americans as the ethnic majority in the United States in less than one life span (Edmondson, 1996; Nicklin, 1997).

If we believe that descendants should have a right to decide the disposition of the remains of their ancestors, then we need to find a way to incorporate the views of Hispanic Americans into the process through which the disposition of ancient American remains is decided.

Some people see focusing on genetic relationships in this way as a myopic and misguided biological reductionism. After all, isn’t a person’s cultural background more important than the genetic links that tie them to earlier generations? From this perspective, there are two types of ancestors: genetic and cultural, and it is the cultural link that a person feels they have with the people who lived in the past that counts. Although the idea of limiting authority to make decisions about the disposition of ancient human remains to people who share the deceased person’s cultural identity makes some sense, applying this ethical principle is extremely problematic in real-world situations. If the strength of a modern person’s belief in their cultural link to an earlier person’s remains is to be the measure of moral authority, how are we to evaluate the relative validity of such beliefs?

To give a specific example, many Native Americans see the intrusions of the “New Age” movement into their cultural identity as the appropriation of Native-American spiritual traditions by outsiders who are destroying Indian spirituality and contributing to white racism and genocide (Geertz, 1996; Hernandez-Avila, 1996; Jocks, 1996; Johnson, 1996b; Kehoe, 1996; Smith, 1991; Specktor, 1989). Is it ethically acceptable to give the same authority to the beliefs of people who received their cultural identity during a psychotherapy session, in which it was revealed to them that they are the reincarnation of an Inca princess, that we give to descendants with demonstrable genetic links to earlier populations? This is where the rejection of scientific evidence and the unconditional acceptance of cultural relativism can become problematic (Goldstein and Kintigh, 1990:587–588).

It is also fair to ask at what point does a living person’s cultural connection to a dead person become so attenuated that it merges into the common cultural heritage of all people, and thus no longer provides a moral basis for special rights and control. Several cultural variables could be considered relevant here: a shared language, common religious practices, and so on. The difficulty is weighing the significance of such disparate cultural traits, especially in the context of ancient remains and cultural evolution.

This issue of cultural continuity is a contentious one, in part, because when indigenous cultures are marginalized, disrupted, and driven to the brink of extinction, remnants of the past, including ancestral human remains, become increasingly important as symbols of cultural oppression and survival. This inverse relationship between concern over ancestral remains and cultural continuity is illustrated by the differences between Latin America and North America in concern over ancestral remains and repatriation issues. In Latin-American countries where a strong sense of “Indianness” has been integrated into the national identity, human remains are excavated and displayed without opposition in museums. In this context, they serve as symbols of a national past that is shared by and important to all citizens (Ubelaker and Grant, 1989). The government of the United States, in contrast, has historically considered Native Americans as outsiders to be dealt with by isolating them on reservations and suppressing their indigenous languages and beliefs to facilitate converting them into functional members of the dominant Euro-American culture. These government policies
have devastated Native-American cultures and contributed enormously to the hostility Indian people feel over issues related to the control of ancestral remains.

In the United States, a legislative attempt has been made to use a combination of biological and cultural continuity as the basis for giving modern indigenous groups the rights over ancient skeletal remains. The Native American Graves Protection and Repatriation Act (NAGPRA) gives federally recognized tribes that can demonstrate a “cultural affiliation” to ancestral remains the authority to control their disposition (Lovis et al., 2004; Mc Laughlin, 2004; Ousley et al., 2005; Richman, 2004). In this legal context, cultural affiliation means “a relationship of shared group identity which can be reasonably traced historically or prehistorically between a present day group and an identifiable earlier group.” In this statute, cultural affiliation is established when “the preponderance of the evidence—based on geographical, kinship, biological, archeological, linguistic, folklore, oral tradition, historical evidence, or other information or expert opinion—reasonably leads” to the conclusion that a federally recognized tribe is culturally affiliated with an “earlier group.”

Although NAGPRA has benefited many federally recognized tribes and has had the positive effect of increasing communication between Native Americans and bioarchaeologists, its exclusion of Native Americans who lack federal recognition raises serious ethical issues. It is derided by some Native Americans who see it as another step in the long history of attempts to define “Native-American groups” in ways that facilitate their control and manipulation by oppressive governmental agencies. In California, for instance, many groups that by any even-handed definition are authentic “tribes” have failed to receive official recognition by the federal government, or have had their federal recognition removed, and thus are denied full access to the provision of NAGPRA (Goldberg, 1997; Walker, 1995).

Again, these legalistic considerations and academic concerns over how to establish a connection between the living and the dead seem strange to those indigenous people whose religious beliefs resolve such issues for them. Many indigenous people are creationists who reject the idea that all modern people share a common ancestor. Instead, some believe that their tribe is the result of a special creation and that they have lived in the area currently occupied by their tribe since the beginning of time. Such beliefs remove any uncertainties regarding ancestral relationships and result in acrimonious disputes between scientists and tribal members such as those that have occurred over the Kennewick skeleton (Hastings and Sampson, 1997; Johnson, 2002; Lemonick, 1996; Lovis et al., 2004; Morell, 1998; Petit, 1998; Preston, 1997; Slayman, 1997).

The Preservation Ethic

The final universally accepted principle of bioarchaeologists is the preservation ethic. Human remains are a source of unique insights into the history of our species. They constitute the “material memory” of the people who preceded us and thus provide a direct means through which we may come to know our ancestors. Because we believe that the lessons that the remains of our ancestors can teach us about our common heritage have great value to modern people, it is an ethical imperative to work to preserve as much as possible of this information for future generations. This position is championed by governments throughout the world who support archaeological research, encourage the conservation and preservation of archaeological resources, and discourage unnecessary destruction of archaeological sites (Knudson, 1986:397; Richman and Forsyth, 2004).

As caretakers of this fundamental source of information on the biological history of our species, we need to promote the long-term preservation of skeletal collections and in this way ensure that future generations will have the opportunity to learn from them and in this way know about and understand that history (Turner, 1986). Prehistoric research, including
osteological study, is one way that our common heritage can be fully revealed (White and Folkens, 1991:418–423). This position is forcefully expressed in the Society for American Archaeology Statement Concerning the Treatment of Human Remains:

WHEREAS human remains constitute part of the archaeological record and provide unique information about demography, genetic relationship, diet, and disease which is of special significance in interpreting descent, health and nutritional status in living and ancient human groups; and

WHEREAS education and research in the anthropological, biological, social and forensic sciences require that collections of human skeletal remains be available to responsible scholars; and

WHEREAS the study of humankind’s past should not discriminate against any biological or cultural group:

THEREFORE BE IT RESOLVED that the Society for American Archaeology deplores the indiscriminant reburial of human skeletal remains and opposes reburial of any human skeletal remains except in situations where specific lineal descendants can be traced and it is the explicit wish of these living descendants that remains be reburied rather than being retained for research purposes; and that no remains should be reburied without appropriate study by physical anthropologists with special training in skeletal biology unless lineal descendants explicitly oppose such study.

AND BE IT FURTHER RESOLVED that the Society for American Archaeology encourage close and effective communication with appropriate groups and with individual scholars who study human remains that may have biological or cultural affinity to those groups (SAA, 1984).

The preservation ethic is based on the scientific premise that there are aspects of our shared reality that have the potential to be brought into sharper focus through the examination of ancient human skeletal remains. The fact that each person sees the world as it is refracted through a slightly different cultural lens does not mean that it is impossible to translate between these different experiences to find a common basis of understanding (Colwell-Chanthaphonh and Ferguson, 2006). The physical facts that we have for deciding what happened in the past are not infinitely plastic and this places material constraints on our culturally biased interpretations.

The progressive aspect of creating the more accurate view of reality that we strive for is an important justification for the preservation of skeletal collections. Most scientists recognize the cultural influences that focus their observations on certain aspects of reality and color the inferences they make based on those observations (Glock, 1995; Tomaskova, 1995; Wylie, 1989). Although we know that our conclusions are to some extent distorted by our cultural biases, we take comfort in the fact that these distortions will be detected and corrected through future research by others, with different cultural perspectives.

For this self-correcting aspect of the scientific method to be operative, the evidence on which our conclusions are based must be available for scrutiny by future researchers. In experimental fields such as physics, this is accomplished through repeating experiments. In historical sciences such as bioarchaeology, our reconstructions of what happened in the past are refined and corrected through the reexamination of collections using new analytical techniques and theoretical perspectives.

During the past 20 years, the rate at which this self-correcting process operates has increased markedly as a result of the restudy of skeletal collections in museums using newly developed analytical techniques that have greatly expanded the types of information we can retrieve from ancient human remains. Especially exciting are new chemical techniques that provide precise information on the types of food people ate (Hult and Fessler, 1998; Stott and Evershed, 1996; Tuross and Stathoplos, 1993) and procedures for reconstructing ancestral relationships through DNA
analysis (Hagelberg et al., 1994; Stone and Stoneking, 1993; Von Haeseler et al., 1996). New techniques are also being developed for reconstructing the disease histories of human populations through the analysis of pathogen-specific bone proteins (Drancourt et al., 1998; Hoffman, 1998; Ortner et al., 1992).

The development of these new, enormously informative, analytical techniques underscores how valuable human remains are as a source of insight into the history of our species. The information content of a cultural product such as stone tool is very meager in comparison with the wealth of biological and cultural information that can be extracted from a human skeleton. The historical information an artifact yields is limited to data on the activity patterns and mental processes that can be inferred from its physical properties, form, and archaeological context. As Carver (1996) has pointed out, there is a subjective aspect to the identification of the artifacts of human cultural activity that are measurable, historically meaningful entities within the corpus of mud and stones that humans have left as the traces of their past activities. Through archaeological research, what was once muck is transformed into monuments and the thunderstones of one generation become the flint axes of the next.

The information contained within the structure of the human skeleton, in contrast, is of a different sort. It is not a culture-dependent symbolic construct. Skeletal remains instead have their basis in adaptive physiological and demographic processes operating at the individual and species levels. Encoded within the molecular and histological structure of skeletal tissues is a detailed record of the person’s childhood development and adult history of metabolic responses to the challenges encountered in his or her natural and sociocultural environment. This information can be supplemented by an equally rich record of ancestral relationships and the evolutionary history of our species recorded in the structure of the DNA molecules preserved within a skeleton. The information about historical events encoded in the skeletons of our ancestors can be thought of as a complex message from the past that we can decode through bioarchaeological research. Each skeleton has a unique story to tell about that individual’s life as well as the evolutionary events that constitute the history of our species. By working to preserve ancient skeletal remains, we ensure that future generations will be able to gain access to the important historical information they contain.

**SOURCES OF CONFLICT**

The ethical principles described above have an inherent potential for conflict. The preservation ethic, with its basis in the belief that the information that skeletal studies can yield is of great value to all people, can easily conflict with the ethical principle that the descendants should have the right to decide the disposition of their ancestor’s remains. If we recognize the validity of the interests of both descendants and scientists in human skeletal remains, how do we deal with the ethical problems that arise when the preservation ethic conflicts with the desires of descendants?

When the remains of close relatives are involved, there is unanimity among bioarchaeologists that the concerns of descendants should override any scientific interests in those remains. Ethical dilemmas, however, frequently do arise when the ancestor-descendant relationship is less clear-cut. How do we balance the scientific value of very ancient skeletal remains against the concerns of modern people who are remotely related to those same individuals?

In balancing the scientific value of archaeological collections against descendant rights, most scientists see the strength of the ancestor-descendant relationship as a continuum that becomes attenuated with succeeding generations. At one end of this continuum we have remains of people with living children and grandchildren who have an undisputed right to determine the disposition of their close relative’s remains. At the other we have the remains of very distant relatives, such as the
earliest members of our species, to which all modern people are equally related. From this evolutionary perspective, descendant rights are viewed as decreasing as the number of generations separating the living and the dead increases. At some point, claims by one modern group of descendants to decide the disposition of ancient human remains is counterbalanced by the right of all people to have access to the unique source of evidence on the history of our species that human skeletal remains provide. How do we decide when the scientific value of skeletal evidence is sufficient to override the concerns of remotely related descendants?

There is no easy answer to the question of how to balance descendant rights against the right of all people to know about the past because the values skeletal biologists and descendants attach to human remains are essentially incommensurable. Part of the problem arises from the fact that many modern indigenous people do not accept the idea that the ancestor–descendant relationship becomes attenuated with time. Instead they see the spirits of their ancestors, no matter how distant, as an integral part of the modern community of the living (Sadongei and Cash Cash, 2007). Nor do they see themselves as closely related to the rest of humanity. Instead they believe that they are the products of a special creation that occurred in the area their tribe currently occupies, and this is an issue of faith about which scientific evidence is irrelevant (Johnson, 1996a). For instance, Armand Minthorn, a member of the Umatilla tribe, which claims the 9500-year-old Kennewick skeleton, made this point when he stated: “We know how time began and how Indian people were created. They can say whatever they want, the scientists” (The Invisible Man, 1996). The implication of such beliefs is that all human remains, no matter how ancient, if they are from the area in which a group believes they were created, are those of their direct ancestors.

Although such creationist interpretations of the history of our species seem strange to many scientists, they are shared by a substantial number of non-indigenous people. For example, a recent survey found that about 20% of the people in the United States shared the Christian belief derived from a literal interpretation of the Bible that God created the cosmos about 5000 to 10,000 years ago (Goldhaber, 1996).

Some archaeologists argue that the utilitarian approach of attempting to balance scientific value against descendant rights is an ethnocentric attempt to frame the problem within the “Eurowestern” system of cultural values that emphasizes finding solutions to problems that maximize benefits and minimize costs (Klesert and Powell, 1993). We can all agree that we will never find a culture-free metric for weighing the value of knowing what actually happened in the past against the concerns descendants have about ancestral remains. However, even if we agree that the benefit of giving control over ancestral remains to people who identify themselves as descendants always outweighs their value as a source of scientific information, we still face the problem of determining who should be able to claim standing as a descendant and what is the ethical thing to do when there are competing claims.

When dealing with close relatives, where the genealogical link between ancestor and descendant is known, allocating descendant rights over the remains of their relatives is fairly straightforward. For example, we might establish a hierarchy that gives a person’s spouse, children, parents, and siblings the authority to control the disposition of their remains. Even such a simple scheme as this is open to charges of ethnocentrism because it reifies a western kinship system that emphasizes the importance of genetic relatedness as a criteria for moral authority and invests the rights to make such decisions in a person’s nuclear family. Other societies might give greater authority to elder members of a person’s patriline or matriline or disregard the modern Western preoccupation with genetic relatedness altogether in favor of another culture-dependent conception of relatedness.
Such cultural differences in ways of conceiving the ancestor–descendant relationship can even transcend the species boundary. For example, I know people who claim the moral authority to remove the bones of dinosaurs from museum collections because they believe, based on their creation myths, that these remains are those of their ancestors before they were transformed into human form. What are we to do with people with sincerely held beliefs about an ancestor–descendant relationships such as this when those beliefs conflict with our own?

Even if we are willing to recognize the validity of such claims and agree that the moral authority of belief in a close ancestor–descendant relationship always outweighs any scientific value skeletal collections might have, we are still faced with the dilemma of deciding what to do when there are conflicting claims for the same skeletal collections. This problem is vividly illustrated by several recent cases in the United States in which people with different beliefs about the past have disputed each other’s assertions of moral authority to control archaeological collections. In Hawaii, 15 federally recognized native groups became involved in a dispute over the disposition of ancestral remains from Mokapu on the island of Oahu (NAGPRA, 1994). One of these groups insisted that scientific research be conducted on the remains of these ancient individuals to determine their ancestral relationships, whereas others viewed such work as a deep insult to the spirits of the their ancestors. In a similar case, Stanford University acceded to the reburial demands of one group of Ohlone Indians without scientific analysis over the objections of other Ohlone people who, from the Western genealogical perspective, were equally related to those remains (Gross, 1989; Workman, 1990). Another acrimonious fight over descendant rights has arisen in the American Southwest between the Navajo and Zuni Indians as part of a government-instigated land deal that prohibits the Navajo from burying their dead in certain traditional burial areas and requires them to renounce claims on sacred sites (Benedek, 1992; Cockburn, 1997). Both tribes have publicly asserted their ancestral rights to the remains of what archaeologists call the Anasazi culture. In other disputes, people who have documentary evidence that they are descendants of the indigenous people from an area have objected to the descendant rights claimed by people who lack such documentation (Erlandson et al., 1998; Haley, et al., 1997; Haley 2006; Kelley, 1997).

One option for dealing with the conflicts that arise when several groups of people assert the moral authority that comes with belief in descendancy from distant ancestors is to take refuge in the legal system where lawyers, politicians, government functionaries, and politically astute special interest groups can wrestle with each other to find a solution to the vexing question of who should have legal standing as a descendant. Although they appear to envisage possible exceptions in cases of “extraordinary scientific value,” this is in essence what Klesert and Powell (1993) suggest when they argue that “we must abide by the preferences of the legally recognized descendants” in disputes concerning the excavation and analysis of ancient burials. For those who view our legal system as a distillation of the moral principles of the people that laws govern, turning the ethical problem of defining “real” descendants over to the courts is very appealing. This political strategy, of course, has the added practical advantage of not eliciting legal sanctions. The moral problem of relying on laws to decide which groups have the right to determine the disposition of human remains has its basis in the faulty assumption that we all live in just societies. Laws have, after all, in the recent past been used as the mechanisms through which groups have been defined by democratically elected governments for purposes of apartheid, slavery, and genocide.

The difficulties associated with legislative solutions to the ethical problem of determining the disposition of skeletal collections are illustrated by the problems that have arisen in
Israel and the United States through legislative attempts to resolve disputes over the control of skeletal collections. Ultraorthodox Jewish organizations in Israel, such as the Atra Kadisha, who regard all academic study involving human remains a violation of Jewish law, have long been at loggerheads with physical anthropologists over the excavation and the handling of human remains, including skeletons of extreme antiquity such as those of Neanderthals (Watzman, 1996a, 1996b, 1996c). Because of the compromises necessary for coalitions of political parties to maintain control of the Israeli government, court rulings have been issued that make the study of unearthed human remains impossible.

In the United States, the Native American Graves Protection and Repatriation Act institutionalizes long-standing inequities in the treatment of federally recognized and non-federally recognized descendants (Walker, 1998). Particularly troubling from an ethical standpoint is its failure to acknowledge the existence of authentic descendant groups that, for one reason or another, have either failed to receive or rejected federal tribal recognition. This omission is especially unfortunate for the many federally unrecognized descendants in California and the eastern United States where the vagaries of the colonial process allowed the government to avoid giving Indian tribes the rights of self-determination that go along with federal recognition. Even if such federally unrecognized groups were given legal standing as descendants, the law would still present ethical problems because, with the minor exception of granting rights to people who can show a direct genealogical connection to the remains of known individual, it fails to recognize the rights of the many people of Native-American descent who lack any tribal affiliation.

RESOLVING CONFLICTS

If we cannot rely on our legal systems to make difficult ethical decisions concerning who descendants are and under what conditions their rights should take precedence over the preservation ethic, what basis is there for finding equitable solutions that balance these potentially conflicting ethical principles? First, it is important to recognize that there is no inherent conflict between the maintenance of skeletal collections for scientific research and respect for the dead. As I mentioned, in many countries research on and the public display of ancestral remains are matters of national pride. In other situations, arrangements can often be made that satisfy the religious and symbolic concerns of modern descendants while allowing scientific research on ancestral remains to continue. At St. Bride’s Church, London, the skeletons of people with known descendants whose burials were disturbed during the German bombings of World War II are respectfully maintained in a special room where they are available for scientific research (Huda and Bowman, 1995; Scheuer and Bowman, 1995). In this way, the religious and symbolic concerns of descendants are respected, while at the same time making it possible for these remains to continue to yield important insights into the lives of eighteenth- and nineteenth-century Londoners that are not adequately documented in written records (Walker, 1997, 2004).

In all societies, cultural understandings of sacredness and ethical behavior are constantly being reshaped in response to changing social realities. This is especially true for the issues surrounding the treatment of ancient human remains because the social context of bioarchaeological research is a modern one not confronted by earlier generations. For many indigenous societies the curation of ancestral remains and their study is a new phenomenon that presents practical problems requiring the development of new rituals, new conceptions of sacredness, and new beliefs concerning what is respectful and disrespectful behavior. In other societies, especially sedentary ones accustomed to maintaining large, intensively used cemeteries, a long history of facing the practical and symbolic problems posed by the
disturbance and handling of ancestral remains has resulted in traditional solutions. For example, the Chumash Indians of southern California, with whom I have worked for the past 30 years, had specialists called liwimpshit, which means “custodian of the algebra,” who were familiar with the human skeleton and the art of arranging bones. These medical practitioners not only could set bones, but they could also arrange all the bones of the human skeleton properly, and determine whether those ancestral bones had once belonged to a man or a woman (Walker and Hudson, 1993:46, 48). The need for someone qualified to deal with human bone derived from Chumash burial practices, which emphasize the importance of having the remains of the dead near to the living. Cemeteries were, therefore, located adjacent to or within villages. As the size of Chumash settlements grew, so did the size of their cemeteries and this frequently necessitated the excavation and disturbance of ancestral remains (King, 1969).

Although the social context of the issues surrounding the treatment of the dead that the modern Chumash face are very different from those they confronted in the past, traditional beliefs about the treatment of the dead have served as a basis for creating a situation in which bioarchaeological research can continue while ensuring that due respect is shown for their dead. Through working with tribal members over the years, my colleagues and I have developed a cooperative arrangement through which Chumash ancestral remains and associated burial objects are being repatriated from other universities and museums to a safe keeping place at my campus. This is highly desirable from the perspective of descendants because of our location near the center of the area historically occupied by the tribe. We have constructed a specially designed subterranean ossuary to receive these remains as part of the construction of our new social sciences and humanities building. This ossuary was designed through consultation with both federally and non-federally recognized tribal members to ensure that it meets their spiritual needs, and also solves the practical problem of providing security against future disturbance that would be unavailable in an unguarded reburial area. The ossuary also makes it possible for scientific research on these collections to continue under the supervision of descendants so that future generations can gain a deeper understanding of the history and accomplishments of the tribe.

Mutually acceptable solutions such as this, which balance spiritual and practical concerns of descendants against the important historical information skeletal research can provide, are the outcome of personal relationships, mutual trust and respect, and the recognition of common interests. Such relationships require time to nurture. My academic colleagues and I have spent our entire professional careers working with Chumash descendants to protect and learn from the archaeological record left by their ancestors. This has involved assisting descendants and local law enforcement authorities in the apprehension and prosecution of grave robbers and looters and actively working to minimize the threats urban development poses for their sacred sites and archaeological resources. At the request of descendants we periodically give seminars and workshops on archaeology, osteology, and the intricacies of the laws that governing the management and protection of archaeological resources. Whenever possible, we have actively involved descendants in our research projects. Such collaborations are enormously rewarding, not only on a personal level, but also professionally, because of the important insights descendants can provide into the history of their culture.

Not all groups have religious traditions that can be easily built on to allow scientific research conducted on the remains of the dead. The strong objections ultra-orthodox Jews have to any skeletal studies already have been mentioned (Watzman, 1996c). As the claims of Hopi and the Navajo to archaeological remains from the ancient Anasazi culture show, it is easy for the control of bones and burial sites to become enmeshed in larger
battles over unrelated economic and social issues concerning the control of land and natural resources, environmental preservation, and so on. This of course greatly complicates the problem of finding a basis for compromise. Sometimes collaboration with descendants may be difficult or impossible because of antagonism toward Western science and strong traditional beliefs about the retention of a person’s spirit within their bones. Some native Hawaiians, for example, believe that people possess mana, which after death resides in the bones and have argued in court that the publication of information about skeletal collections is offensive and will steal the mana of their ancestors (Kanahele, 1993). Many Plains Indian tribes also have strong beliefs about the residence of souls in their ancestral remains. This, along with animosity stemming from racism, genocidal attacks by the U.S. military, cultural suppression in boarding schools, and economic marginalization on reservations, makes the prospects for the preservation of skeletal collections from most of the Plains area bleak (Ubelaker, 1994:395).

In situations such as these it may be impossible to obtain a compromise that allows skeletal research to continue. However, from the personal experiences I have had in working with many different groups of indigenous people, once the shroud of mystery associated with what osteologists actually do is removed through direct contacts between people, it is often possible to find a foundation on which mutual understanding and cooperation can be built. The most obvious basis for developing such collaborations is in the identification and analysis of ancient human remains that are inadvertently disturbed through erosion, for example, or during construction projects. In such situations, the value of close collaboration between osteologists and descendants is obvious. After it has been decided that remains are indeed human, the issue of whether they are modern (and thus possibly relevant to a forensic investigation) needs to be resolved. If they are indeed ancient, the question of which modern group of people they are affiliated with needs to be considered. This issue is especially important to some indigenous people who have strong religious sanctions against the burial of non-group members in their cemeteries. The value of osteological research is also self-evident in forensic investigations relating to the prosecution of grave robbers. I have collaborated with Native Americans in several of these cases. In one, we matched a fragment of a mandible confiscated from a suspect’s home with another piece of the same mandible that tribal members had recovered from the area of an ancient grave disturbed by looters. This incontrovertible evidence connecting the defendant with the crime scene resulted in a guilty plea. In another case, we used skeletal evidence to successfully refute a grave robber’s attempt to exonerate himself by claiming that the Native-American remains he excavated were from a person of European ancestry and thus not protected by the state’s Native-American graves protection law. Through the process of working on such cases, I have seen the views of people who once saw little value in skeletal research change dramatically as they increasingly become aware of many important insights skeletal studies can give us into the lives of those who have gone before us.

When skeletal collections are lost because of our inability to find equitable solutions that balance the concerns of modern descendants against the need to preserve collections so that future generations will have substantive information about the past, it is perhaps of some solace to remember that we live in an entropic world in which the natural processes of decay and disintegration and the economic and social realities of modern life continuously conspire to destroy the faint traces our ancestors have left for us in the archaeological record. We cannot turn this tide. All we can do is work to preserve as much of the physical evidence of our common heritage as possible. Those ancestral remains and the facts about the history of our species that they reveal will be our legacy to future generations.
REFERENCES


Augustine D. 1994. Minutes from the Native American Graves Protection and Repatriation Review Committee seventh meeting, Rapid City, South Dakota.


REFERENCES


Cockburn A. 1997. Alexander Cockburn’s America: the Navajo Indians are resisting pressure to capture their lands. New Statesman 126:32.


REFERENCES
Morton SG. 1839. Crania Americana: or, a comparative view of the skulls of various aboriginal nations of North and South America: to which is prefixed an essay on the varieties of the human species. Philadelphia: J. Dobson.


Navran F. 1997. 12 steps to building a best-practices ethics program. Workforce 76:117–120.


Forensic anthropology involves the application of our knowledge and techniques of human skeletal biology to medico-legal issues, especially the study of recovered human remains from the recent past. This subfield of physical anthropology can include the examination of soft tissues, but most practitioners use their skills as physical anthropologists to examine skeletal remains. The aim of this work is basically twofold: to assist in the identification of human remains and in the interpretation of what happened to them. To stay within the parameters of this volume, this chapter focuses primarily on the examination of skeletal remains in this forensic context.

HISTORICAL DEVELOPMENT

To demonstrate the methodological advances in forensic anthropology, one needs to examine the historical foundation on which progress has been made. The beginnings of modern forensic anthropology date back to the eighteenth and nineteenth centuries with European scholarly interests in anatomy and anthropology. One of the earliest contributions to forensic anthropology comes from the Paris art instructor Jean-Joseph Sue (1710–1792), whose measurements of body dimensions at progressive developmental stages led to methods of stature calculation. Improvement in this methodology came from Paul Topinard (1830–1911), Etienne Rollet (1862–1937), and Leonce Manouvrier (1850–1927), although the foundation for modern statistical approaches to stature calculation comes from Karl Pearson (1857–1936) and colleagues, who introduced regression equations and greatly improved statistical methodology. Other important contributions include Matthieu-Joseph Bonaventure Orfila’s (1787–1853) medico-legal textbooks, the founding of the Societe d’Anthropologie de Paris by Paul Broca (1824–1880), and Broca’s development of instrumentation for body measurement.

The academic ancestry of forensic anthropology extends into the nineteenth century when anatomists and early physical anthropologists occasionally were asked to bring their academic skills to focus on problems of human identification. Stewart (1979a) credits Thomas Dwight (1843–1911) for writing the essay in 1878 “The Identification of the Human Skeleton: A Medico-Legal Study” and for launching professional anthropological
interest in this area of physical anthropology. Although an anatomist, he recognized the research need to develop procedures to estimate age at death, sex, and living stature from skeletonized human remains (Dwight, 1878, 1881, 1890a, 1890b, 1894a, 1894b, 1905).

Other notable early pioneers included George A. Dorsey (1869–1931) (Stewart, 1978; Ubelaker, 1999a) and Harris Hawthorne Wilder (1864–1928), (Stewart, 1977, 1979a, 1979b, 1982). Following Dwight’s lead, Dorsey published on the medico-legal applications of knowledge of skeletal anatomy (1897, 1899). In 1897 and 1898, Dorsey testified in a high-profile murder trial in Chicago. A local sausage manufacturer was accused of the murder of his wife and the disposal of her remains in a vat at the sausage factory. Dorsey’s testimony regarding small fragments found within the vat was debated. Although challenged by other experts, his testimony received extensive favorable media coverage in this high-profile trial.

A German anatomist, H. H. Wilder, contributed to two major areas of forensic science. He enlarged on F. Galton’s dermatoglyphic studies to include observations of prints on the palm and sole and advocated the use of dermatoglyphics over the Bertillon system of anthropometric traits for positive identification. Wilder also conducted research on the restoration of soft tissue and facial reproduction (Stewart, 1982).

Recent research has also demonstrated that the early leader in physical anthropology, Aleš Hrdlička (1869–1943), had broad forensic interests, reported on skeletal cases, and initiated a relationship of forensic service provided by the Smithsonian Institution to the Federal Bureau of Investigation (Ubelaker, 1998, 1999b, 1999c, 1999d) that was carried on by subsequent curators, including Angel and Ubelaker. Hrdlička is well known for his seminal role in the development of American physical anthropology. In a career largely spent at the Smithsonian Institution (1903–1943), he founded the *American Journal of Physical Anthropology* in 1918 and the American Association of Physical Anthropologists (first meeting in 1930). His works on the evidence for an early human presence in the New World, anthropometry, and many other areas of physical anthropology are well known. Less recognized are Hrdlička’s training and work in legal medicine. Over his long career, Hrdlička researched broad medico-legal issues, including insanity, and the possible contributions of biological attributes to criminal and other abnormal behavior. He also presented opinions on cases involving insanity, skeletal identification, and ancestry of living peoples. His skeletal casework included trauma interpretation and an early example of photographic superimposition. Hrdlička’s expertise in the identification of skeletonized human remains was recognized as early as 1918 by the Federal Bureau of Investigation (FBI). By the time of his death, Hrdlička consulted regularly with the FBI on skeletal issues and was regarded highly by them as a resource in investigation (Ubelaker, 1999d).

The modern era of forensic anthropology in the United States began with Wilton M. Krogman’s (1903–1987) contributions. He published a “Guide to the identification of human skeletal material” in 1939, which quickly became the authoritative work in forensic anthropology (Stewart, 1979a). Krogman’s main research on growth and development did not hinder his forensic contributions (1943, 1946, 1949); the most significant work was his key text, “The Human Skeleton in Forensic Medicine” (1962). Krogman’s major forensic works were used widely by those involved in identifying remains for the military, such as H.L. Shapiro (1902–1990) in Europe and Charles E. Snow (1910–1967), Mildred Trotter (1899–1991), and T. Dale Stewart (1901–1997) in Hawaii. The work of these individuals and others involved in identification efforts demonstrated the shortcomings of Krogman’s forensic texts and led to important improvements in methodology, such as Trotter’s research on stature (Trotter, 1970; Trotter and Gleser, 1952) and Stewart’s approach
to age changes (McKern and Stewart, 1957). Those researchers involved with military identification gained extensive experience, and this led to important methodological improvements in forensic anthropology, such as Stewart’s edited volume on mass disaster investigation (1970).

In 1972, growth of activity in forensic anthropology led to the formation of a physical anthropology section of the American Academy of Forensic Sciences with 14 founding members. Membership in the section increased steadily to 334 in 2006. Professionalism within forensic anthropology received another boost in 1977 when a certification process was developed. Board membership of 22 in 1978 grew to about 74 in 2006. Currently, to become a board-certified Diplomate of the American Board of Forensic Anthropology, a forensic anthropologist must hold a Ph.D. degree in physical anthropology with a specialty in skeletal biology, human anatomy, and dental anthropology. In addition, applicants for certification must have at least three years of full-time professional experience at least partially devoted to forensic anthropology. Applicants must also have experience doing forensic casework and must pass a written and practical examination usually given in association with the annual meeting of the American Academy of Forensic Sciences. Certification (Diplomate status) provides the recipient with a useful credential for work within forensic anthropology.

Forensic societies including anthropologists are increasingly common throughout the globe. The Canadian Society for Forensic Science (CSFS) includes forensic anthropologists under the section “Anthropology, Medical and Odontology.” With a founding membership of 17 individuals, CSFS membership has grown steadily to include over 450 members to date. A similar organization, the Forensic Anthropology Society of Europe, founded in 2004 in association with the International Academy of Legal Medicine, promises similar stimulus for the development of forensic anthropology in Europe.

Interest in forensic anthropology is growing in Europe, although much of this work continues to be done by nonanthropologists, mostly physicians. The Smithsonian week-long workshop in forensic anthropology has been presented in France in the years 1992, 1994, 1996, 1998, 2000, and 2002 and in Italy in 2004. The first three European courses were sponsored jointly by the Smithsonian Institution and the Université de Bretagne Occidentale, Brest. The fourth and fifth courses were sponsored by the Smithsonian and the Université Montpellier. The 2002 course was presented in Montpellier in conjunction with the International Association of Forensic Sciences. The courses have been well attended by interested parties from many countries, mostly from Europe. Graduates of these workshops recently have organized the Forensic Anthropology Society of Europe in association with the International Academy of Legal Medicine.

Increasingly, forensic anthropologists have published results of their work in forensic journals, especially the Journal of Forensic Sciences (Ubelaker, 1996), edited volumes (Clement and Ranson, 1998; Dupras et al., 2006; Haglund and Sorg, 1997, 2002; Reichs, 1998a), and textbook case studies (Hunter and Cox, 2005; Steadman, 2003). Such research has focused on specific forensic applications and has increased greatly the capability of forensic anthropologists.

Through the involvement of the Diplomates and other forensic anthropologists in the medico-legal system, each year over 1400 cases are reported on in North America. These cases represent mostly skeletonized remains, but they also include fresh, mummified, decomposed, burned remains, and other conditions. Typically, cases originate from civil disputes forwarded from agencies (i.e., local law enforcement, state police, military, coroners, medical examiners, or sheriff’s departments), most of which are sent by medical examiners’ or coroners’ offices, followed by the military. The most common type of anthropological case is skeletal remains, followed by decomposed and recent
cases, respectively. Forensic anthropologists frequently are requested to take on cases involving burned remains, nonhuman material, and remains of archaeological origin. Each case requires different techniques and methodology; some of the more common include dental aging from microscopic structure, age determination from microscopic examination of cortical bone, facial reproduction, photographic superimposition, and photographic comparison.

RELATIONSHIP OF FORENSIC ANTHROPOLOGY TO SKELETAL BIOLOGY

As indicated, the roots of forensic anthropology reach back to early anatomists and pioneer physical anthropologists who responded positively to requests from the law enforcement community to apply their skills. Prominent past forensic anthropologists such as Hrdlička, Stewart, Krogman, and J. Lawrence Angel (1915–1986) represented physical anthropologists who shared their forensic interests with broader issues in anthropology. Hrdlička did some forensic work, including issues of insanity and ancestry of then-living individuals but also pursued research problems in all other areas of physical anthropology and some areas outside of physical anthropology (Ubelaker, 1999d). Stewart considered forensic anthropology only one of three broad interests, the other two being anthropometry and paleoanthropology (Ubelaker, 2000). Angel was regarded highly for his forensic contributions, but he initiated them late in his career, focusing first on skeletal biology of the Near East and anatomy issues (Buikstra and Hoshower, 1990), research that he continued along with his forensic work. Krogman pioneered the modern era of forensic anthropology in North America and was an expert in growth and development (Johnston, 1997). These early workers likely were drawn into forensic applications by their reputations and expertise in the relevant areas of physical anthropology. It is also likely that they welcomed these applications because of the new information the forensic experience provided them. The forensic work represents not only an opportunity for public service and to demonstrate the relevance of our science but also a unique chance to acquire information about contemporary populations and problems. This information and the experience gained through case analysis and court testimony frequently sharpens the skills of the skeletal biologist and improves analysis of remains recovered from archaeological contexts.

THEORETICAL ISSUES

At first glance, it might seem that forensic anthropology differs from the more general field of skeletal biology in that the former concentrates on the individual, whereas the latter addresses larger population issues. On closer examination, this simplistic dichotomy breaks down. Many studies within skeletal biology, especially paleopathology, concentrate on the individual, if not the individual specimen. As with forensic anthropology, the skeletal biologist/paleopathologist uses scientific knowledge from the literature and experience to diagnose a disease or interpret a cultural modification.

As in skeletal biology, the forensic anthropologist considers the individual specimen in the context of total human variation and uses information about the individual to improve techniques and gain insight into broader issues. The immediate goals of forensic anthropology are specific: primarily to help identify the person and figure out what happened to them. The secondary goals are to gather biological/skeletal information about contemporary populations and additional understanding of human variation for several skeletal variables.

Broad anthropological knowledge is almost always needed to properly interpret a forensic case. The seemingly simple determination of human versus nonhuman requires experience
not only in human variation, the human disease process, and the taphonomic effects on human bone, but it also requires awareness of the diverse materials that can resemble human bones and teeth. The more fragmentary and altered the materials, the more difficult this process becomes. A submitted bone fragment may not resemble normal human bone, but could it represent a pathological condition, a fragment from a very young or very old individual, or human bone altered by heat or exposure to postmortem influences? A small fragment may seem to fall within the range of variation of human bone, especially in consideration of the diverse factors outlined above, but can non-human sources be eliminated completely? These interpretations call for broad knowledge of skeletal biology and even other, related fields that ultimately shape the language used within the final report or testimony on the witness stand. Similarly, a broad perspective is needed to address issues of age at death, sex, ancestry, living stature, time since death, facial reproduction, evidence of foul play, and other identifying features frequently involved in forensic anthropology.

The theoretical approach employed in forensic anthropology basically involves a broad anthropological, population perspective applied to the individual. Forensic issues and goals are addressed using anthropological data, techniques, and perspectives. Some foci of forensic anthropological analysis such as interpretation of gunshot wounds and identification of foul play would seem to be unique to the forensic anthropology experience. However, even these areas call for broader anthropological knowledge of bone biomechanics and structure variation for proper interpretation.

As anthropologists, forensic anthropologists also recognize the unique opportunities presented by the casework to improve procedures and to gain knowledge about key issues within contemporary society. The cases examined nearly always present a special problem or reflect materials from unusual contexts that broaden our knowledge. Many of the cases represent known individuals at the time or those who are later identified, offering an opportunity to expand our knowledge about various aspects of human variation. The practice of forensic anthropology clearly expands the experience and expertise of the individual skeletal biologist conducting the work. Through publications and presentations at professional meetings, this special knowledge improves the science.

**THE FORENSIC DATA BANK**

Angel’s work (1974, 1976) offers examples of how forensic anthropologists have used the opportunity offered by casework to gather information about the skeletal biology of contemporary people (Ubelaker, 1990). In 1962, J. Lawrence Angel joined the staff of the Smithsonian Institution and succeeded T. Dale Stewart in consultation with the nearby FBI Headquarters in forensic anthropology. In about 1978, the author assumed responsibility for the FBI work, but Angel continued to report on cases for others. By 1986, his case experience had risen to about 565. He not only learned a lot from doing all of the cases, but he also attempted to apply the data from them to larger anthropological questions. These data became the “modern” sample that added perspective to Angel’s studies of long-term temporal change in patterns of fractures (1974) and other variables (1976). In an internal 1977 Smithsonian document, Angel described his forensic work as both public service and research. Angel noted: “Much time at present goes to careful study of human skeletons brought for identification... since almost all of these are research data. It has taken years to collect adequate samples of modern, and Colonial American adults. The time invested is just beginning to pay off...Applied science is not a dirty word.” (Ubelaker, 1990:197).

As increasing numbers of skeletal biologists became involved in forensic applications, the
need for improved techniques and knowledge of the contemporary population was recognized. Many procedures available for estimating sex, age, ancestry, stature, etc. were based on existing museum collections of human remains, especially the Terry collection at the Smithsonian Institution in Washington, D.C. and the Hamann–Todd collection in Cleveland, Ohio.

The Hamann–Todd collection consists of approximately 3157 skeletons and associated records housed at the Cleveland Museum of Natural History (Lovejoy et al., 1985). The Department of Anatomy of Western Reserve University assembled this collection between 1912 and 1938. According to Lovejoy and colleagues, cause of death information is available for about 9% of the adults and only about 512 are of relatively known age at death. Others show discrepancies between the stated age of the individual and observations on age made by anatomists at the time.

The Terry collection consists of the remains of over 1600 individuals who mostly were dissected at the Washington University School of Medicine in St. Louis, Missouri, during the twentieth century. This collection of predominantly older individuals was assembled mostly by Robert J. Terry (1871–1966) and Mildred Trotter (1899–1991) after dissection. The collection, with associated records, hair samples, and some death masks, is located in the Department of Anthropology of the Smithsonian’s National Museum of Natural History in Washington, D.C.

Both the Terry and the Hamann-Todd collections have been used extensively in skeletal biology research, and both represent individuals whose deaths date from the early part of the twentieth century. As valuable as these collections are, concerns grew about the extent to which they represent human populations from other gene pools, geographic areas, and time periods. Such concern from the forensic anthropologists led to the formation of a forensic data bank in 1984, sponsored by the physical anthropology section of the American Academy of Forensic Sciences. The concept was that as forensic anthropologists throughout North America worked on contemporary forensic cases, they would gather noninvasive data from them that would facilitate the research process. A prominent group of forensic anthropologists met to agree on the data to be collected, and standardized forms were made available to all involved with casework (Jantz and Moore-Jansen, 1988).

By June 1998, the number of individuals in the data bank had increased to about 1550, a size comparable with the older Terry and Todd collections. Emerging from this effort is a large pool of measurements and observations on the skeletal remains of known individuals who were examined as forensic cases. These data not only have obvious utility in supplying anthropological information about the contemporary population but also offer an opportunity to develop improved techniques directly applicable to forensic work.

In 1993, a major product of the database effort emerged, FORDISC 1.0 (Jantz and Ousley, 1993), which is a DOS computer program that enables the classification of an adult cranium by sex and ancestry. The advantage of this system over previously published discriminant function equations was 1) FORDISC was based on the large, diverse, contemporary Forensic Data Bank, of obvious direct applicability to modern forensic cases, and 2) the interactive program allowed the formation of custom functions when not all measurements were available. With FORDISC 1.0, a powerful new tool was added that resulted directly from data generated from the forensic casework.

Three years later, the FORDISC product was upgraded (2.0) (Ousley and Jantz, 1996) using the enlarged data bank and adding additional on-line assistance, including a pictorial guide to measurements and improved graphics. Other additions include post-cranial measurements to improve classification of sex and ancestry and to allow estimation of living stature, mandibular measurements, and worldwide data collected by W. W. Howells (1973, 1989).

FORDISC 3.0 (Ousley and Jantz, 2005) is now available offering an updated database
that includes Guatemalan Mayans, an expanded set of measurements for the Howells database, an opportunity for users to import their own data and enhanced flexibility for stature calculations. The new system represents a substantial improvement and demonstrates the interplay between skeletal biology and forensic anthropology. FORDISC 3.0 can be ordered from the University of Tennessee website (http://web.dii.utk.edu/fordisc/).

Comparison of data in the Forensic Data Bank with those generated from the Terry and Hamann-Todd collections documents the temporal variation involved and the need to maintain the Data Bank. Such a comparison was made by Ousley and Jantz (1998) using measurements derived from both types of collections and by applying discriminant functions in the literature developed from older collections to cases in the Data Bank. This exercise demonstrated differences between these samples, and it documented the importance of modern data in forensic applications.

Although the FORDISC system represents a significant methodological advance, caution is still appropriate in applications to individuals who likely differ substantially from those within the database. As a case in point, Ubelaker et al. (2002a) applied the FORDISC 2.0 system to a sample of Spanish crania dating to the sixteenth and seventeenth centuries. In regard to ancestry, the individual crania within the sample were classified in diverse ways, reflecting the absence of Spanish samples at that time within the FORDISC database.

EVIDENCE RECOVERY

Proper recovery of human remains constitutes an increasingly important aspect of forensic analysis. Most discoveries of human remains that lead to forensic investigation are not made by professionals. Construction workers, hikers, hunters, berry pickers, and family dogs usually are the first to notice human remains in these contexts. Most cases studied by forensic anthropologists are skeletonized and unidentified. Usually this means that they originate in areas not frequently traveled by humans, sites where they can remain undetected while soft-tissue decomposition progresses. Frequently such sites are isolated, rural, wooded areas where ground cover discourages immediate discovery. In many areas of North America, discovery must wait until seasonal activities (hiking, hunting, etc.) bring a potential human discoverer into visual contact with the remains. In areas with cold, winter seasons, low temperatures bring reduced foliage and increased ground visibility. This, coupled with winter treks to the woods by hunters and hikers, leads to discovery.

Once discovery of human remains is made, it is important to recover the remaining evidence as thoroughly and carefully as possible. Although much progress in this area has been made in recent years, much more can be done. If human remains are known or thought to be involved, recovery by a forensic anthropologist can be a great help. Shovels, trowels, and other sharp tools may be necessary to properly excavate the materials, but in the hands of nonprofessionals, these tools may produce alterations on the remains that would complicate analysis. Increasingly, forensic anthropologists are involved in recovery or those involved have had some exposure to principles of excavation through courses or seminars. Publications are available to advise those involved on the principles of recovery of human remains (e.g., Dupras et al., 2006; Haglund and Sorg, 1997; Hunter and Cox, 2005; Killam, 2004; Pickering and Bachman, 1997; Ubelaker, 1999e).

Most recovery missions remain an exercise in common sense using careful archaeological techniques. Advanced technology increasingly is available to supplement the traditional approaches. Ground-penetrating radar, soil resistivity equipment, metal detectors, and other sophisticated equipment can assist, particularly in cases with other evidence suggesting buried remains in a general area. In some circumstances, cadaver dogs can help as well. Use of these approaches, along with surface topography, aerial photography,
vegetation patterns, and other indicators can facilitate difficult decisions on where labor-intensive excavation should be employed. Such a complex search clearly calls for a team approach, since no single expert likely will be knowledgeable in all the specialized procedures and equipment involved.

The major issues involved in evidence recovery are similar to those of traditional archaeology or the recovery of ancient human remains. Decisions must be made regarding the amount of time and effort directed toward recovery. Ideally, all resources available should be used and directed toward the maximum recovery of information. Just as with an archaeological excavation, the recovery operation is destructive and offers a mostly one-time opportunity to learn as much as possible. Practical decisions must be made regarding the use of available resources and time. Frequently these decisions must be made in consideration of other priorities. At the time of recovery, it is difficult to predict what information might prove important later in the investigation. Thus, it is important to have the decisions guided by the most experienced and trained personnel available.

NONHUMAN VERSUS HUMAN REMAINS

Although most forensic anthropologists can easily distinguish intact, well-preserved human remains from those of nonhumans, fragmentary or otherwise altered material can be challenging. Detailed knowledge of human skeletal anatomy is usually sufficient to recognize that evidence is consistent with a human origin. The more precise opinion that remains are of human origin (could not be anything else) requires some recognition of the many other materials that can mimic the human condition. With the reality that molecular approaches potentially offer positive identification from minute bits of evidence, forensic anthropologists can expect to see increasing frequencies of submission of such evidence. Since the molecular procedures are costly and time-consuming, an initial determination by the anthropologist of human vs. nonhuman can be important.

Such determinations usually rely on the experience of the anthropologist, supplemented by comparative collections. Other experts, especially zooarchaeologists or other naturalists, may need to be consulted. In my experience, microscopic examination can provide important diagnostic information. Examination of small evidence with a high-quality dissecting-type microscope (I currently use a Carl Zeiss SV 11 Stereomicroscope with photographic capability) allows clear viewing of morphological surface details that facilitate diagnosis. With fragmentary evidence, the examination of the visible internal surface may reveal helpful aspects of structure.

For especially difficult fragments, it may be desirable to prepare thin sections for even more detailed microscopic study (see Robling and Stout, this volume). In particular, the osteon organization may present a nonhuman pattern (Ubelaker, 1999e). Although human/nonhuman differences in bone histology are recognized in the literature, more research is needed to clarify the distribution of these differences throughout the skeleton and among different species (Enlow, 1962, 1963; Enlow and Brown, 1956, 1957, 1958; Foote, 1916; Mulhern and Ubelaker, 2001, 2003; Ubelaker, 1999e).

If soft tissues are sufficiently preserved, they may provide additional clues on human/nonhuman status. Hair analysis or serological techniques may be able to clarify whether remains are of human origin and perhaps identify species, genus, and so on. Both of these approaches have proven to be useful in determining that some forensic cases of submitted cranial remains represented calves with hydrocephaly rather than human infants (Ubelaker et al., 1991). In this forensic application, extracts were prepared from desiccated soft tissue associated with the specimens. Double diffusion test methods were employed using the extract and antibodies of bovine, deer, horse, sheep, swine, and human. A precipitin line formed in the presence of antibovine serum,
which with additional testing contributed to the final diagnosis that the remains represented calves with hydrocephaly. Hair examination and cranial morphology also contributed diagnostic information.

For extremely fragmentary and/or environmentally compromised material, even recognition if bone or tooth is present can be difficult. For such cases, scanning electron microscopy/energy dispersive spectroscopy (SEM/EDS) can produce diagnostic results. The spectra revealed in such analyses allow identification of constituent elements and their relative proportions. This information can facilitate identification if bone or tooth is present when compared with a relevant spectra database of known materials (Ubelaker et al., 2002b). Such analysis does not allow differentiation of human fragments from those of other animals, however. For such cases, protein radio immunoassay (pRIA) analysis will distinguish between humans and nonhumans and even identify nonhuman fragments to the family represented (Ubelaker et al., 2004).

AGE AT DEATH

Scientific progress in the estimation of age at death involves greater awareness of the variability involved, employing reliable demographic models and understanding new techniques available for different structures, and greater appreciation of the importance of regional and temporal variation in the aging process.

Forensic anthropology strives for an accurate estimation of age at death. Such “accuracy” involves coming as close as possible to actual chronological age at death and realistically conveying the probabilities associated with the estimate. The process is complicated by the fact that many individual techniques are available that have been developed by different researchers working with diverse samples. On the positive side, many techniques are now available reflecting age changes in different populations around the world at varying periods in history. The downside is that comparison of the accuracy of techniques can be difficult if they were developed from different samples. The situation also creates some uncertainty when a method is applied to an individual originating from a population that differs from the one that contributed to the development of the method. Increasingly, the scientific literature respects these problems through utilization of diverse samples and testing of methods on different samples.

Increasingly, biological anthropologists are applying mathematical models to the assessment of demographic distributions, such as Bayes’ theorem (Hoppa and Vaupel, 2002; Milner et al., this volume). Proponents of this methodology suggest that one must first estimate the age distribution of a population before estimating individual age-at-death. To do so, researchers rely on reference collections of populations with known age distributions. A method sometimes employed in estimating paleodemographic profiles is transition analysis, in which a suite of traits from different anatomical areas (e.g., the pubic symphysis, auricular surface, and cranial sutures) are examined (Boldsen et al., 2002). This method provides researchers with probability statistics derived from the reference samples that enhance interpretation.

Decisions regarding which methods to use in estimation of age at death result not just from the relative accuracy of the methods but also from the experience of the investigator, available equipment, and the preservation of the remains. For example, adult age estimation through microscopic examination of long bone cortical microstructure offers one of the most accurate single approaches according to the literature, but it requires time-consuming specimen preparation and equipment that is not universally available. This technique also requires assessment of complex histological structures, which is a process not familiar to all forensic anthropologists (see Robling and Stout, this volume). A technique that may have been accurate when developed by a specialist in a particular process may not be as accurate in the hands of a less-experienced researcher.
Although some techniques for age estimation seem more accurate than others, usually it is advisable to consult as many age indicators as possible in assessing age at death, especially for adults (see Meindl et al., this volume). For immature skeletons, the extent of dental development provides the most accurate age indicator. If teeth are fragmented, absent, or otherwise difficult to recognize, then the size and overall morphology of bones (including epiphyses) or bone fragments can be useful (see Saunders, this volume).

For adults, age assessment is more difficult because the increased number of years have allowed potentially greater internal differences to develop in the various age indicators. A skeleton may present a relatively youthful-appearing pubic symphysis and sternal end of the fourth rib and yet show premature arthritic development and extensive tooth loss. All of these aging systems are variable and not entirely linked. For these reasons, an assessment of all available data is important.

The application of different techniques to the same sample of individuals of known age at death offers some insight into the relative accuracy and reliability of the methods. With such an approach, Virginia Galera, Lee-Ann Hayek, and I (Galera et al., 1995) applied various techniques of estimation of age at death that are popular in both Europe and North America to 963 skeletons of known age at death in the Terry Collection of the Smithsonian Institution. The evaluation included four approaches to cranial suture closure (Acsádi and Nemeskéri, 1970; Baker, 1984; Masset, 1982; Meindl and Lovejoy, 1985), antemortem tooth loss, vertebral osteophytosis (Stewart, 1958), sternal rib morphology (Işcan et al., 1984a, 1984b, 1985, Işcan and Loth, 1986a, 1986b), the pubic symphysis (Acsádi and Nemeskéri, 1970; Gilbert and McKern, 1973; McKern and Stewart, 1957; Suchey and Katz, 1986; Suchey et al., 1986, 1988; Todd, 1920, 1921), and auricular surface changes (Bedford et al., 1989; Lovejoy et al., 1985).

Statistical testing revealed no significant differences in the scoring of these techniques by the three investigators, which suggests that the techniques are not difficult to interpret and apply even with investigators with minimal anthropological preparation (Galera et al., 1995). Application of the cranial methods suggested that those relying on endocranial closure were more accurate than the ectocranial ones (Galera et al., 1998). The different methods demonstrated different strengths when applied to different subsets of the sample, but overall the European methods were more accurate, at least when applied to the crania of the Terry collection.

In another test, various methods of estimating age at death of adults were applied to a sample of 19 French autopsy individuals of known age at death (Ubelaker et al., 1998). Individual techniques tested consisted of the Lamendin et al. (1992) procedure of assessing single-rooted teeth, the morphology of the sternal end of the fourth rib (Işcan et al., 1984a, 1984b, 1985; Işcan and Loth, 1986a, 1986b), the Suchey–Brooks system for the pubic symphysis (Brooks and Suchey, 1990), and the Kerley method (Kerley, 1965, 1970; Kerley and Ubelaker, 1978) of assessing femoral cortical remodeling. In addition, three comprehensive approaches were employed: 1) a mathematical average of the ages derived from the individual approaches, 2) a two-step procedure (Baccino and Zerilli, 1997) that combines aspects of the pubic symphysis assessment with the dental technique, and 3) an overall assessment by each investigator of all available information.

Statistical analysis of the results indicated that the comprehensive approaches were more accurate than any individual approach. Of the single approaches, the dental technique offered the best results. The relative success of the seven methods employed reflected not only the nature of the methods but also the experience of the investigators and the complexity of the structures requiring interpretation, which was most apparent with the complex Kerley method that yielded much more accurate results when applied by an experienced investigator in comparison with those produced.
by a professional with less familiarity with the technique.

In addition to our own studies, anthropologists are increasingly raising methodological standards by conducting research on samples where age-at-death is known. The skeletal analysis and historical demography studies conducted on the St. Thomas Anglican Church sample, a nineteenth-century cemetery in Belleville, Ontario, Canada, by Saunders and colleagues (Saunders, 1992; Saunders et al., 1992, 1993, 1995, 2002) represents an excellent resource for testing methodology for growth and aging studies, as well as ancient DNA research. Research by Cunha and colleagues (Albanese, 2003; Bruzek, 2002; Cunha, 1995; Rissech et al., 2006; Schmitt et al., 2002) on the Coimbra Identified Skeletal Collection at Coimbra University in Portugal also provides important insight into population variation in growth, development, and aging.

Estimation of age at death in a forensic context concentrates on the individual, using experience and data from the literature on age changes within populations throughout the world. Such estimates offered in reports or court testimony should use appropriate language to convey the proper associated probabilities. For example, in the analysis of a forensic case, a particular technique might suggest an age at death of about 34 years. It is appropriate to call attention to this result in a forensic report but equally important to convey the possible range around this estimate. Without this perspective, the authorities may limit their search of missing persons to those of age 34 years and not discover the actual missing person whose age varied somewhat from that estimate. When age at death is known, the cases offer an opportunity to augment our knowledge of skeletal age changes and the influencing factors involved.

SEX (NOT GENDER [Walker and Cook, 1998])

In forensic anthropology, as in the more general field of skeletal biology, the estimation of sex of the adult is generated most accurately from observations of the pelvis. (For a discussion of sex determination of the subadult skeleton, see Saunders, this volume.) In the absence of the pelvis, in incomplete remains, or those with excessive deterioration of the pelvis, the morphology, especially the size, of bones can provide important information. Increasing advances in DNA extraction and identification techniques have led some scientists to identify sex genetically rather than morphologically (Schmidt et al., 2003); however, this is not always the most cost-effective or reliable method of determining sex (Sivagami et al., 2000). A forensic anthropologist is typically called on by medico-legal examiners when remains are fragmentary, decomposed, skeletonized, or otherwise unrecognizable. In these instances, DNA may not be easy to extract or may be cost prohibitive. DNA evidence can also be contaminated if protocol is not strictly followed, and results often take a long time to be processed. It is important to remember that morphological identification of sex is still an invaluable resource for the forensic anthropologist.

Most recent research in this area has been directed toward making data available on sex differences in parts of the skeletons that may be recovered in unusual circumstances. Such research is reflected by recent work on bones of the feet. These data are relevant in the forensic context because bones of the foot can be protected within shoes and stockings from taphonomic forces that otherwise destroy or separate the rest of the skeleton. Boots recovered from water contexts may contain the bones of the foot as the only skeletal representation of the individual, otherwise lost or distributed by decomposition and fluvial factors. Footwear containing bones of the foot also may be separated from the other remains by animals, forcing a specialized analysis of just those remains if only the foot bones are recovered.

To help meet this forensic problem, Introna et al. (1997) examined 80 right calcanei from 40 men and women from a contemporary
southern Italian collection. Eight measurements were taken leading to both univariate and multivariate analyses that compliment earlier work in estimating sex from this bone. The Italian study is especially useful in a comparative sense to help document not only sex differences but how the expression of these differences may vary in populations around the world.

Smith (1997) and Robling and Ubelaker (1997) turned to the Smithsonian’s Terry collection for data on sex differences in the metatarsals (Robling and Ubelaker, 1997; Smith, 1997) and foot phalanges (Smith, 1997). Discriminant function equations were generated to facilitate sex identification of these bones.

**ANCESTRY**

In studying remains of unknown identity, forensic anthropologists usually attempt to estimate the ancestry of the individual or the group with which this person would likely have been identified in their community (the so-called “ethnic identity” of the individual). This information can be useful to law enforcement in attempting to narrow the search for the identity. Knowledge of the ancestry is also useful in other aspects of analysis, especially stature estimation in which different regression equations are available for blacks, whites, and so on, since group variability occurs in stature and body proportions. It is important, however, to recognize the social dimensions of these categories and not to confuse this effort with past typological classifications that suggested greater biological grouping than actually exists.

In forensic anthropology, the estimation of ancestry is accomplished through observation of characteristics such as dental morphology and features of the facial skeleton and through measurements, usually with discriminant function analysis. The former approach has been augmented by increased attention on the variety of anatomical features that display variation (see Gill and Rhine, 1990). The latter approach has grown primarily through the acquisition of measurements from larger, more diverse samples, with special effort aimed at the needs of the forensic community.

In recent years, the estimation of ancestry has been strengthened considerably in forensic anthropology through the forensic data bank discussed earlier in this chapter. The data bank and the computer program FORDISC 2.0 and 3.0 have improved the ability of forensic anthropologists to predict ancestry (Ousley and Jantz, 1996). Like any other discriminant function program, FORDISC forces a classification into the groups defined within the database. Intelligent use of this program, and any other such functions, must interpret these classifications using all available information.

As an example of the application of this system, measurements from a recent forensic case of known identity were entered into the program. The individual was an adult woman of known living stature of about 63 inches (160 cm) and was known in the community as being white. Once the data were entered on a standard IBM-compatible computer, only a few seconds were required to classify the measurements as originating from a white woman with a posterior probability of 0.969 and a typicality of 0.842 (Ubelaker, 1997). Posterior probability relates the likelihood that the unknown originates within the group it is classified into, assuming it originates from one of the groups used in the analysis. Typicality probability presents the likelihood that the unknown originates within any of the groups used in the analysis. For additional detail on these statistics, see Ousley and Jantz, 1996.

A second classification using the database of W.W. Howells (1973, 1989) required nine minutes to classify the “unknown” as a Zalavar woman from western Hungary, ninth to tenth century. The other close groupings using the Howells database were Egyptian and Taiwanese (women). Like the forensic database, the Howells system had classified correctly the measurements as originating from a woman, but the closest population classifications were disparate geographically. Intelligent use of this system must regard the classifications within
the context of other information. It obviously would be incorrect to conclude from the Howells classification that the individual originated from the Zalavar group of western Hungary, ninth to tenth century, as suggested by the face value of the analysis. Rather, an individual classified in that manner by the Howells database likely would have been classified socially as white in contemporary North American society. The Howells database is mostly composed of archaeologically recovered remains. Although such ancient samples may not be represented directly by a modern forensic case, they can provide a useful perspective to assist in the overall evaluation.

The estimation of ancestry calls on our skills as physical anthropologists to assess the biological information displayed by the forensic remains in the context of worldwide human variation. This aspect of forensic analysis also requires broader anthropological interpretation of how an individual with certain skeletal characteristics likely would be regarded (social classification or “ethnic identity”) by the community in which he or she lived. Once again, language on this issue must be selected carefully to convey the proper levels of probability involved.

LIVING STATURE

Recent progress in the estimation of living stature has been registered on three fronts: 1) the detection of errors and clarification of past formulas for estimating stature; 2) improvements in methods with the new forensic database, and 3) recognition of the potential error involved in “known” living stature.

In recent years, most skeletal biologists in North America have relied heavily on the published stature regression equations of Trotter (1970). These formulas were generated from multiple databases of measurements of long bones from individuals of known stature, many of military origin whose statures likely were measured and thus accurately documented. Of all the long bones, the tibia is especially problematic to measure because it presents extensions from the articular surfaces on both the proximal (intercondylar eminence) and the distal (medial malleolus) ends. Research with Trotter’s original measurements raised questions about the manner in which she measured the tibia. Her published definition of the measurement used in the stature calculation from the tibia excluded the intercondylar eminence but included the medial malleolus. Since Trotter had measured individuals in the Terry collection, those tibiae were available and could be remeasured for comparison with her original measurements (Jantz et al., 1995). This procedure suggested that, at least in the Terry sample, the medial malleolus was excluded, but questions remain on how the measurement was taken on other skeletons used in Trotter’s research. The example is illustrative of the importance of careful definition of measurements and the need to critically evaluate such information in the existing literature.

Improvements in stature estimation also result from the incorporation of new information from additional skeletal samples. Most relevant to forensic anthropology are revisions resulting from the Forensic Data Bank. Using the new contemporary information in the Data Bank, Jantz (1992) offered new regression equations for both sexes and the various population groups. Comparison of these formulas with those previously published revealed differences, but it was difficult to determine whether they represented secular trends in the groups represented or relative error in the reporting of living stature.

The error involved in so-called known living stature remains a persistent problem, not only in developing new equations from the Forensic Data Bank but also in applying estimates to cases. Research has shown a human tendency to erroneously report living stature, especially in men (Himes and Roche, 1982; Willey and Falsetti, 1991). Thus, even if stature can be estimated accurately from skeletal remains, the “known” statures of missing persons available for comparison may be sufficiently different from actual stature to complicate the identification process. Estimates of stature from
remains can be useful not only to help identify unknown remains but also for exclusion and to assist in the sorting of commingled remains.

**FACIAL REPRODUCTION**

A central goal of forensic anthropology is the accumulation of information about an individual from their remains to facilitate identification. At times, even with individuals thought to be deceased only short periods of time, information on sex, age at death, ancestry, living stature, and other information still does not lead to an identification. There are many reasons for this lack of identification but among them is the possibility that the lifestyle of the individual is such that the community has not considered the possibility of death, or for other reasons, the individual is not registered officially as a missing person. In such cases, authorities may request a facial reproduction. The goal here is to estimate the living facial appearance of the individual so that the image can be presented to the public through the media. Hopefully, such visual public exposure will trigger the memory of an acquaintance and lead to identification. The facial reproduction is not used directly for the positive identification; however, it may help generate the recognition and documents needed for such identification.

Techniques of facial reproduction can consist of three-dimensional reproduction using clay or a similar material deposited on the skull or a cast of the skull, an artist’s sketch prepared from visual examination of the skull, or various computer approaches (Clement and Ranson, 1998; Taylor, 2001; Ubelaker, 1993; Ubelaker and O’Donnell, 1992; Wilkinson, 2004). Although these approaches have varying procedures and advantages, they all require considerable experience and rely on estimates of the depth of the soft tissue at various points on the skull. Current research seeks to improve data on soft tissue depth through studies of ultrasound (Manhein et al., 1998) and magnetic resonance imaging (MRI) (Marks et al., 1998).

**PHOTOGRAPHIC SUPERIMPOSITION**

The closely related procedure of comparing a recovered skull with a photograph taken during life has been augmented by computer technology (Ubelaker et al., 1992). This technique is termed photographic superimposition and is employed in cases when identification of remains is suspected but not confirmed by other evidence. The technique is used largely for exclusion, to demonstrate that the recovered skull could not have originated from the individual depicted in the photograph. Although it is possible that the technique also could be used for positive identification, usually when a “match” is made, it can only establish that the remains could have originated from the person shown in the photograph.

An early example of a variant of this technique is provided by the 1935 analysis of John Glaister and colleagues who compared photographs of two missing persons with recovered remains (Glaister, 1953; Glaister and Brash, 1937). They compared the photographs of the living with photographs taken of the remains, arranged to approximate the positions of the heads in the living photographs. The comparisons contributed to the identification of the skeletons.

The method of comparison was enhanced by the introduction of the use of video (Brown, 1982, 1983; Helmer and Gruner 1976a, 1976b). Use of two video cameras, an electronic mixing device, and a viewing screen offered a more dynamic method of comparing the two images.

The computer-assisted approach involves the use of a video camera and a computer. Images of both the remains and the photograph of the living individual are captured with the video camera, digitized, and stored within the computer. Software then allows both images to be brought up on the screen simultaneously and manipulated for detailed comparison (Ubelaker et al., 1992). Variations of the above techniques can be used to compare any two objects to evaluate their similarity of shape. They have proven most useful in forensic anthropology to
compare recovered skulls with antemortem photographs, but applications are not limited to these materials. Since the technique is usually not used for positive identification, it can be used for exclusion or in support of other evidence for identification. The technique merely facilitates comparison. The success of the evaluation of the comparison depends on the experience of the investigator. For example, the technique can be used to establish that the two images are consistent. The investigator must then use knowledge of variation to determine whether the features examined are sufficiently unique to warrant identification.

Fenton and Sauer (2006) use a similar system to interpret photograph/face comparisons. This approach is employed to assess whether a particular person is represented on a photographic image. Analysis may involve photographing the suspected person to replicate the correct size and position depicted in the image. Comparison involves a broad selection of individual traits on the aspects of the body visible in the photograph.

**TIME SINCE DEATH**

Recent research in the interpretation of time since death consists of improvements in recognizing the taphonomy behind postmortem alterations (Bassett and Manheim, 2002; Carson et al., 2000; Durić et al., 2004; Haglund and Sorg, 2002; Mandojana et al., 2001; Pickering, 2001; Pickering and Carlson, 2004; Puskas and Runnem, 2003; Thompson, 2005; Warren and Schultz, 2002) and specific dating techniques that can document whether recovered remains can be dated to a specific time range or period (Carter and Tibbett, 2003; Courtin and Fairgrieve, 2004; Davis and Goff, 2000; Hobischak and Anderson, 2002; Megyesi et al., 2005; Tibbett and Carter, 2003; Tibbett et al., 2004; Vass et al., 2002, 2004; Weitzel, 2005; Yan et al., 2001). Research in taphonomy indicates that climate, soil composition, seasonality, burial depth, burial container type, body size, body condition, microenvironment, preparation of body prior to burial, and association with vegetation all may affect tissue preservation and consequently the interpretation of postmortem interval. Any single variable can assist in interpretation, but variation will be considerable and multiple factors should be assessed.

Determination of the postmortem interval is often a task that can only be solved by forensic investigators. Research using controlled specimens (especially such as that conducted at the facility at the University of Tennessee) has contributed a great deal to improving our understanding of the complex factors operating on postmortem alteration. The Anthropological Research Facility at the University of Tennessee was created in 1972 by William Bass. This outdoor field laboratory enables forensic scientists to scientifically document postmortem change. The research is made possible by donated remains that are studied in differing environmental conditions; the remains also provide a modern osteological teaching collection. The success of the Anthropological Research Facility in Tennessee has spurred other important contributions to the study of postmortem interval and decomposition, including research on pigs (Anderson and Hobischak, 2004; Archer and Elgar, 2003; Banaschak et al., 2005; Prez et al., 2005; Tomberlin et al., 2005) and humans (Kovarik et al., 2005; Megyesi et al., 2005; Moriya and Hashimoto, 2004; Prieto et al., 2004).

For relatively fresh remains, entomological evidence is valuable; however, it is often difficult to interpret the postmortem interval in skeletonized remains and even those that may display soft-tissue preservation. In addition to entomological data, volatile fatty acids in the soil (Vass et al., 1992), plant growth (Willey and Heilman, 1987), mechanical trampling (Ubelaker and Adams, 1995), and many other factors can influence the process. Haglund and Sorg have edited two volumes (1997, 2002) that summarize aspects of this recent research. Much of this information limits specific anthropological interpretation since patterns can vary greatly and frequently not
all of the influencing factors are known in a case.

Important information on time since death also can be gained from analysis of radiocarbon, especially in regard to the “bomb curve.” Atmospheric testing of thermonuclear devices between 1950 and 1968 produced artificially high levels of radiocarbon that through the food chain have been incorporated into organic material worldwide. The increased levels peaked in about 1964 and have subsequently declined; today, however, they remain above the pre-1950 values. If radiocarbon analysis of human remains recovered in a possible forensic context does not reveal the “bomb-curve” increased levels, the investigator knows that they originate from an individual who died before the bomb curve or that average values of radiocarbon in the tissue sampled are below those suggested by the bomb-curve. If radiocarbon levels in bone are increased, this means that they likely are of medico-legal interest (Ubelaker, 2001; Ubelaker and Houck, 2002).

Interpreting the postmortem interval using the radiocarbon approach requires an understanding of the dynamics of tissue formation and remodeling (Ubelaker and Buchholz, 2006). Since dental enamel does not remodel, radiocarbon analysis of this tissue reveals the levels of artificial radiocarbon at the individual age of formation of that tissue, during the period of growth. When radiocarbon contents of dental enamel reach increased levels, an approximate birth date of the individual can be estimated based on the age of enamel formation and the specific radiocarbon value (Ubelaker and Buchholz, 2006). Because of the remodeling process, certain factors must be considered, such as age at death, medical treatment, and type of bone tissue sampled. The selection of multiple tissues, such as trabecular and cortical bone, will assist in proper placement of the values on the higher or lower range of the curve, thus facilitating the interpretation of date of death (Ubelaker and Buchholz, 2006; Ubelaker et al., 2006).

**POSITIVE IDENTIFICATION**

Positive identification is achieved when unique characteristics known to exist within an individual are found in recovered remains. Traditionally, such evidence originates from dental records. These records are most commonly radiographs, which can be compared with radiographs of unidentified jaws and teeth. Other useful dental characteristics include restorations, fillings, and crowns. Unique aspects of dental or skeletal morphology may also be useful. Increasingly, positive identifications are made using molecular techniques, frequently from minute evidence.

Research in positive identification has been stimulated by legal debate, as outlined in *Daubert v. Merrell Dow Pharmaceutical, Inc.* (1993), primarily in North America, where questions regarding the objectivity of methodology have been raised. The *Daubert* ruling requires that the validity of a particular methodology used by an expert witness be assessed by a trial judge. The criteria for an acceptable methodology include its testability, subject to peer review and publication, known error rates, existing standards, and widespread acceptance in the scientific community (Keierleber and Bohan, 2005). Thus, current research in positive identification has focused on creating larger databases that support methodology as well as on objective testing of the accuracy of these methods.

Usually, positive identifications from dental work or molecular evidence are made by specialists in those areas (forensic odontologists and geneticists, respectively). Forensic anthropologists are involved in the analysis that leads to identification by narrowing the range of possibilities and are involved directly in the interpretation of skeletal and dental morphology from photographs, radiographs, and other evidence. This interpretative endeavor calls for experience and judgment. In some areas of the world where dental and/or medical records may not be available and the costs of molecular analysis are prohibitive, traditional anthropological techniques are
employed to assist in identification. In such cases, identifications may be based on clothing or personal belongings, pathology, family oral history, the biological profile of the decedent, and morphological anatomical variants such as nose shape (Prokopec and Ubelaker, 2002).

In my experience, positive identifications from forensic anthropology result primarily from comparison of skeletal anatomical details with antemortem radiographs, such as frontal sinus patterns (Ubelaker, 1984). If identical features are found, attention focuses on the uniqueness of those features. Are the points observed unusual enough that positive identification can be established? On the other hand, if differences are found, are they of the nature and magnitude to preclude identification? Differences may reflect only techniques of preparing the radiographs, the length of time between death and the date the antemortem radiographs were taken, taphonomic factors, or others. Experience and knowledge of human skeletal anatomy is needed to sort all this out and to draw a reasonable conclusion.

Investigators have demonstrated that positive identifications can be made from dental evidence within cremated remains (Delattre, 2000), evidence of recent dental extractions (Carmichael, 2002), and orthopedic device identification (Ubelaker and Jacobs, 1995). Moreover, the use of molecular techniques to identify disease organisms also shows potential for future positive identifications (Donoghue et al., 1999; Ubelaker et al., 2000).

**MOLECULAR APPROACHES**

Recent years have witnessed a surge of activity in molecular approaches to forensic science. Aspects of this research have an impact and complement efforts in forensic anthropology. The areas of greatest impact are positive identification, sex determination, and potentially, ancestry evaluation.

Techniques of forensic DNA extraction and amplification from human bone have increased dramatically (Alonso et al., 2003; Arismendi et al., 2004; Coble and Butler, 2005; Rennick et al., 2005; von Wurmb-Schwark et al., 2003, Yang et al., 2003). Such techniques have proven useful in establishing positive identification, even in such cases as skeletal remains submerged in water for three years (Crainic et al., 2002) and severely burned bone (Calacal et al., 2003; Ye et al., 2004). Increasingly, DNA evidence is used to identify victims from mass disasters when only bone fragments are present (Alonso et al., 2005), such as with the identification efforts after the World Trade Center attack in 2001 (Brenner and Weir, 2003; Budimlija et al., 2003; Holland et al., 2003). In cases in which forensic anthropological analysis does not clearly indicate sex, molecular techniques may be able to provide the necessary information (Schmidt et al., 2003; Sivagami et al., 2000).

Considerable information is also compiled on molecular patterns in worldwide populations (Bell et al., 1997; Blaxter, 2004; Brandon et al., 2005; Budowle et al., 2004; Gill et al., 2006; Imaizumi et al., 2002; McCartney, 2004; Polanskey and Budowle, 2005; Sullivan et al., 2004). Many laboratories claim to provide detailed information about an individual’s genetic ancestry for a fee (e.g., www.genelex.com). These companies use DNA sequencing methods to identify haplogroups that have been linked to particular populations worldwide. Some purport to provide percentage values of your ethnic background (e.g., 35% European, 50% Asian, 15% African), but these services are controversial (see www.gene-watch.org for summary articles).

Despite the controversy, future advances may allow estimates of ancestry with increased precision. Even with these developments the role of the forensic anthropologist is still important in providing rapid initial descriptions and in narrowing the field of identification in cases where investigators have no preliminary suspicions of who the person might be.

**EVIDENCE OF FOUL PLAY**

Forensic anthropologists are in a unique position to offer opinions about some types of evidence relating to foul play. A careful eye
(sometimes aided by a microscope) is needed to spot some evidence for trauma. Knowledge of the reaction of bone to a variety of stimuli is then required to determine the nature of the alterations and whether they represent antemortem, perimortem, or postmortem conditions. A single forensic case may present evidence of all types. Evidence of bone response, remodeling (or lack of it), coloration patterns, and related observations all contribute to the solution of the puzzle. Since anthropologists routinely work with human remains from many different contexts (archaeologically recovered skeletons, museum collections, autopsy specimens, etc.), they, more than any other professionals, have the necessary knowledge.

The exposure to evidence of foul play in modern cases also facilitates interpretation of trauma in the archaeologically recovered remains encountered in skeletal biology. This area of forensic anthropology represents a clear example of the desirable “two-way street” interplay between forensic work and the more general field of skeletal biology. Knowledge of diverse taphonomic processes operating on human bone gained through work with archaeological samples is needed to interpret examples of trauma in forensic cases. Conversely, the knowledge gained about perimortem trauma through forensic work is critical to proper interpretation of such evidence in skeletal biology.

Such involvement has lead to significant research by anthropologists in trauma interpretation. Examples are especially noteworthy in interpretation of gunshot wounds and sharp-force trauma. Ross (1996) used her research approach as a forensic anthropologist to address the long-standing, recognized difficulty in estimating bullet caliber from characteristics of cranial alterations. She examined the cranial alterations in 73 individuals with cranial gunshot trauma of known bullet caliber. Measurements were recorded on the alterations, and the data were analyzed using sophisticated statistics. The study revealed that the size of the alterations produced in gunshot trauma results not only from the caliber of the bullet but also from the thickness of the bone at the site of impact. Information was generated by the study that assists in the interpretation of gunshot skeletal trauma.

In another innovative research approach to better understanding skeletal trauma, Houck (1998) devised a “cutting machine” to examine aspects of sharp-force trauma in bone. Bovine tibial diaphyses were cut with three different knives using the machine. All resulting cutmarks were examined for class and individual characteristics. As with Ross’s study cited above, the data were analyzed statistically by Houck producing information of great importance in the forensic interpretation of sharp-force trauma. Additional discussion of anthropological interpretation of skeletal trauma is provided by Sauer (1998), Berryman and Symes (1998), Reichs (1998b), and Symes et al. (1998).

FUTURE PROSPECTS

Interest and participation in forensic anthropology has grown steadily since the casual, occasional encounters of a few pioneers early in the history of American physical anthropology. The sustained growth of this area of physical anthropology can be measured in student interest, the numbers of anthropologists involved in casework, the increase in the membership of the physical anthropology section of the American Academy of Forensic Science and the Diplomates of the American Board of Forensic Anthropology, and the increase in the number of research publications focusing on topics in forensic anthropology. This growth shows no sign of diminishing and has led to the recognition of forensic anthropology as a vigorous subfield of physical anthropology. Slowly, job opportunities have expanded from the traditional anthropological employment sites of universities and museums to specific research and teaching in forensic anthropology, as well as employment in crime laboratories, medical examiner’s offices, and the military. Increasingly, forensic anthropologists are integrated into evidence recovery teams and the
medico-legal investigation of death. Forensic anthropologists have provided an important perspective in the international investigation of possible war crimes and human rights issues.

The field of forensic anthropology has progressed so far that one might think that few research problems remain. Such is not the case. In fact, progress has highlighted the tremendous need for new data and an anthropological perspective on most areas of forensic anthropology. Major questions remain regarding population variation in many problems routinely assessed by forensic anthropologists. Much more work needs to be done in the areas of assessment of time since death, animal versus human recognition, taphonomic change, environmental factors, foul play, and positive identification. Even more traditional areas of scholarship such as the estimate of sex, age at death, stature, and ancestral origins would benefit greatly from new research and perspective.

Most training in forensic anthropology is centered in university departments of anthropology with an emphasis on human skeletal biology and its forensic applications. For students seeking education in this area, I recommend consulting the “Guide” to departments of anthropology published by the American Anthropological Association (e.g., American Anthropological Association, 2005; www.aaanet.org/pubs/guide.htm) along with the list of current diplomates of the American Board of Forensic Anthropology, Inc. (website http://www.csuchico.edu/anth/ABFA). University degree programs can be supplemented with various courses and seminars in forensic anthropology that are offered periodically. Students seeking experience should also consider internships with professional forensic anthropologists who are active with casework.

Training remains firmly entrenched within skeletal biology, but increasingly students are acquiring skills in additional areas, including law and various technical specialties. Although this additional perspective is undoubtedly valuable, I feel it should supplement and not displace the traditional anthropological education.

Anthropologists working in the forensic arena need to be aware of the increasingly specialized contributions made in medico-legal cases. However, it is primarily the anthropological training with archaeological techniques and samples and other experiences within anthropology that separate the contributions of forensic anthropologists from those of others and make us unique. The future is indeed bright for those skeletal biologists willing to accept the challenge of applying their skills to the resolution of forensic problems.

CASE STUDY

One morning in Spring 2006, detectives received a report from the local prison warden that a correction officer had overheard a conversation between inmates indicating that one of them had been involved in the disposal of a murder victim 31 years earlier in 1965. Subsequent interviews with the inmates confirmed the report and further indicated that burial had taken place in a rural field and that the victim was a 20 year old man of European ancestry who had been reported missing at about that time. The missing person report indicated that living stature was about 6 feet 2 inches, the person had suffered major facial trauma in a car accident four years prior to death, and the person had three missing teeth and extensive dental work. Since the burial had taken place at night, the inmate could not recall details of the burial site, only its general location.

Armed with this information, the authorities visited the scene accompanied by a forensic anthropologist, employed at the local university. They had located the anthropologist by checking the website of the American Board of Forensic Anthropology and discovered that an experienced board-certified forensic anthropologist was available to work with them.

At the scene, they noted several surface features that might indicate a burial had taken place. Unusual vegetation was present in one area. In another place, a slight depression...
was found; in another, a slight mound was found. Careful examination of the ground surface discovered a concentration of possible bone fragments in yet another area. The anthropologist declared that two of the larger fragments clearly presented morphology indicating nonhuman, probably deer. Since the anthropologist was not sure about the origin of the other fragments, a decision was made to postpone additional site investigation until the other smaller fragments had been identified.

Back in the laboratory, the anthropologist carefully examined the remaining fragments with the aid of a high-quality dissecting microscope. Of the six fragments, five were clearly bone, but morphological markers were not sufficient to distinguish human from nonhuman. One very small fragment was so eroded that it was not clear whether it represented bone or some other material.

The very small fragment of uncertain origin was subjected to analysis by SEM/EDS. A very small sample from this fragment revealed spectra suggesting an elemental composition inconsistent with bone. Comparison with the spectra database suggested the fragment likely originated from a shell, probably a land snail.

The remaining five fragments that represented bone were analyzed using the pRIA technique. Small samples taken from these fragments revealed an antibody response also consistent with deer.

Armed with the information that none of the fragments recovered was of human origin, authorities returned to the scene for additional testing. Ground-penetrating radar equipment was brought in to examine the subsoil for evidence of variation in compaction perhaps suggesting the location of the burial. A proton magnetometer also was employed to examine variation in magnetic patterns within the soil. Both approaches located two areas in the field that looked promising; both associated with unusual surface features as well.

The anthropologist then directed careful excavation of both areas. The first area revealed a clear pit outline indicating that indeed they had discovered a hole that had been dug previously and then filled in. At the bottom of the pit, they found an articulated skeleton (bones in anatomical order) but that of a dog. Apparently someone had chosen this site for the burial of their deceased pet.

The second excavation also revealed a clear pit outline but larger than that of the dog burial. At the bottom of this pit, excavators found a fully articulated adult skeleton with some clothing remnants as well. After exposure of the entire skeleton in situ and careful documentation, the remains were removed for analysis.

Back in the laboratory, the anthropologist carefully laid out the remains on the examination table. Detailed inventory documented the skeleton was complete and in excellent condition. The extent of root formation on the third molars, epiphyseal closure, morphology of the pubic symphysis and sternal end of the ribs, lack of root translucency in the anterior teeth, and other indicators suggested an age at death likely between 18 and 24 years.

Morphology of the skeleton, especially general robusticity and features of the pubis, strongly indicated male sex.

Measurement of the long bones, especially the femur and fibula, indicated a living stature of about 6 feet 1 inch.

Observations of the cranium, especially the bones of the facial area, indicated likely European ancestry. Such ancestry also was suggested by detailed cranial measurements analyzed using FORDISC 3.0.

The pattern of missing teeth and dental restorations were consistent with those of the missing person. Detailed examination of the restorations by a forensic odontologist indicated they matched those of the missing person (dental radiographs of the missing person had been obtained from a local dentist). In addition, mitochondrial DNA analysis of bone from the skeleton matched a sample donated from a maternal relative.

Examination of the cranium, including radiography, detected evidence of gunshot injury...
with an entrance site in the left temporal and an exit in the right parietal. The perforation in the left temporal was smaller than that of the parietal and displayed interior beveling and numerous associated radiodense particles.

The final report concluded that the skeleton was identified positively as that of the missing person, and the gunshot injury was documented. The inmate’s story had been verified, and another old case had been solved.

REFERENCES


Baccino E, Zerilli A. 1997. The two step strategy (TSS) or the right way to combine a dental (Lamendin) and an anthropological (Suchey-Brooks System) method for age determination. Proceedings. Am Acad For Sci 3: 150 (abstract).


Dwight T. 1905. The size of the articular surfaces of the long bones as characteristics of sex; an anthropological study. Am J Anat 4:19–32.


Hoppa RD, Vaupel JW. 2002. The Rostock Manifesto for paleodemography: the way from...


Krogman WM. 1946. The reconstruction of the living head from the skull. FBI Law Enf Bull 15:11–18.


Lovejoy CO, Meindl RS, Pryzbeck TR, Mensforth RP. 1985. Chronological metamorphosis of the


Renick SL, Fenton TW, Foran DR. 2005. The effects of skeletal preparation techniques on


INTRODUCTION

This chapter is a discussion of the numerous factors that affect the composition and condition of human bone assemblages recovered from archaeological contexts. Circumstances of death and mortuary programs have an impact on the human remains that we eventually study; excavation strategies and research paradigms determine what is recovered in the field; museum practices and research trends impact the treatment and use of curated collections, and repatriation laws determine access to (and continued existence of) collections. Thus, an assemblage of bone reveals the biological life history of the individuals represented, but it also embodies the history of the assemblage as a culturally created entity. The kind and quality of information available to the researcher varies considerably between assemblages, but data recovery and the understanding of prehistoric people and their lives can be maximized by the broadest possible understanding of the depositional context and history of the skeletal assemblage.

Formally defined as, “the study of the transition of organics from the biosphere into the lithosphere or geological record” (Lyman, 1994:1), taphonomy originated as a subfield of paleontology (Efremov, 1940). In friendlier terms, taphonomy is the study of “the physical and chemical processes (induced by human, animal, or natural agents) that modify an organism after its death and through which it is incorporated into geological deposits” (LaMotta and Schiffer, 2005:122). Archaeological assemblages are never perfectly preserved or perfectly complete. Taphonomy provides the framework in which we can investigate the multiple processes and events that cumulatively determine the content and condition of skeletal assemblages from archaeological sites. In practice, the study of bone modification in human skeletal assemblages is based on methods developed in zooarchaeology: recording weathering and fragmentation patterns, evidence of animal scavenging, thermal alteration, and the location and orientation of tool marks on skeletal elements. Through the study of assemblage history taphonomy restores the critical link between field recovery and laboratory analysis (Nawrocki, 1995). This connection is often lost in the postexcavation stage of projects when human remains are sent to the osteologist, often with little contextual information, whereas animal bones and artifacts are sent to others,
minimizing the opportunity for holistic analysis (Jones, 2002; Outram et al., 2005).

Since the 1970s, taphonomy has been an increasingly important component of zooarchaeology and paleontology (Lyman, 1994), as a perusal of the literature pertaining to the antiquity of hunting and mortuary ritual in hominins and the analysis of faunal remains from archaeological sites will show. Forensic taphonomy incorporating archaeological and zooarchaeological methods is an integral part of the medico-legal study of human remains, where the distinction between premortem and postmortem modification of human remains is critical (Boddington et al., 1987; Dupras et al., 2006; Haglund and Sorg, 1997, 2002; Hunter et al., 1996; Micozzi, 1991). The importance of taphonomy and the adoption of zooarchaeological methods as part of the study of human remains from archaeological sites have been formalized gradually over the past several decades. More recent human osteology texts include chapters on postmortem modification by natural and human agents (e.g., White and Folkens, 2005), and taphonomic observations are included in the widely used Standards for Data Collection From Human Skeletal Remains (Buikstra and Ubelaker, 1994) and the British counterpart Guidelines to the Standards for Recording Human Remains (Brickley and McKinley, 2004). The taphonomic approach is especially important in the study of fragmentary and commingled skeletal assemblages (mass graves, ossuaries, cremations) where we need to distinguish between human and nonhuman agents of modification. And where human actions are the source of modification, the combination of taphonomy, biological analysis, and fine-grained contextual analysis is essential if we are to discern the intentions behind those actions. The most visible application of taphonomy in bioarchaeological research has been in the highly publicized studies of Neanderthal and Anasazi cannibalism. But involvement of biological anthropologists with taphonomy and forensics is not new (Buikstra et al., 2003), and it is not restricted to the kinds of stories that make it into Discover magazine.

**TAPHONOMY AS ASSEMBLAGE HISTORY**

In the study of human remains, the scope of taphonomy encompasses the natural and the cultural events, processes, and agents that modify human remains from the time of death until the time of analysis. Human taphonomy is quintessentially interdisciplinary as it involves consideration of archaeological site formation processes, burial, decomposition, physical weathering and chemical degradation (diagenesis), modification of bone by animals, and by the intentional and unintentional activities of humans in the past and in the present. Among all of these, mortuary practice is the most important determinant of the condition of human remains (Henderson, 1987:49). Skeletal assemblages—whether we find them as intact primary burials or as fragmentary and commingled bone deposits—are not randomly created entities. Mortuary ritual is remarkably variable across time and human cultures (Carr, 1995; Chapman et al., 1981; Metcalf and Huntington, 1991; Parker Pearson, 2002; Rakita et al., 2005), and appreciation of this fact is crucial to our understanding of archaeological assemblages and the agents and intentions that created them. But modification of bone takes place at all stages of taphonomic time. These stages include the pre-depositional (antemortem) stage; the depositional (perimortem) stage when, except for the case of secondary burials, human remains represent the newly dead; the postdepositional (postmortem) stage; and the postrecovery stage (Sorg and Haglund, 2002).

We can map the events and processes in the history of human skeletal assemblages in a manner similar to that used in artifact life histories (Gosden and Marshall, 1999; LaMotta and Schiffer, 2005), by extending the osteobiography of individuals and assemblages beyond death and into the present. As listed in Table 3.1 the stages in the history of a human skeletal assemblage include death, mortuary processing and deposition, soft-tissue decomposition, skeletonization and hard-tissue
decomposition in the burial context; discovery and archaeological excavation; documentation, collection and transport of remains from the site; curation and study; and data archiving, and for some assemblages, repatriation.

**MORTUARY PROGRAMS AND THE ARCHAEOLOGICAL RECORD**

If you learn skeletal biology in a climate-controlled laboratory populated by well-preserved, complete skeletons, each found in a discrete grave feature with the normative gender-specific grave goods, then the need to understand ancient mortuary practice can seem far removed. But most assemblages of human remains are not complete skeletons from discrete graves, and biological sex assigned to the skeleton does not always match the gender associated with the grave goods. The link between the social, ritual, and political aspects of funerary practice and the work of the skeletal biologist might be stronger than you think. Mortuary programs determine what is buried: cremated remains (often referred to as “cremains”) or the entire skeleton, or the skeleton minus the skull or other elements that may be retained for spiritual or economic purposes—as a relic or as material for weapon manufacture. And mortuary programs dictate where people are buried. Consider the range of cemeteries in use within the past century: pet cemeteries, military cemeteries, poorhouse and asylum cemeteries, church parish cemeteries, town cemeteries, family cemeteries, cemeteries of leprosaria, reliquaries containing the finger bones of saints, memorial parks where you can watch video-eulogies of the deceased, mausoleums with cremains in marble niches,

<table>
<thead>
<tr>
<th>Taphonomic Events and Processes</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>— Death —</td>
<td><em>Deceased portion of population</em></td>
</tr>
<tr>
<td>Perimortem modification: trauma, exposure and weathering, animal scavenging, skull, or other element curation</td>
<td></td>
</tr>
<tr>
<td>— Mortuary program —</td>
<td><em>Interred bodies</em></td>
</tr>
<tr>
<td>Processing, deposition (interment), secondary processing activities</td>
<td><em>Preserved skeletal assemblage</em></td>
</tr>
<tr>
<td>— Decomposition, weathering, diagenesis —</td>
<td><em>Recovered assemblage</em></td>
</tr>
<tr>
<td>Rate determined by wrapping, container/coffin type; presence of organic or metal artifacts; type of grave furniture and facility; soil/grave surface chemistry; temperature fluctuation; wetting and drying cycles; vegetation and microorganisms</td>
<td><em>Curated assemblage</em></td>
</tr>
<tr>
<td>— Discovery and excavation —</td>
<td><em>Potential study assemblage</em></td>
</tr>
<tr>
<td>Excavation design and rationale; damage from mechanized equipment or hand excavation tools</td>
<td><em>Cumulative knowledge base created</em></td>
</tr>
<tr>
<td>— Documentation, collection, transport —</td>
<td><em>Conserved potential study assemblage</em></td>
</tr>
<tr>
<td>Provenience and condition documentation; screening of burial soils; field discard; reinterment; field exposure to weather; cleaning application of consolidants packaging; transport method</td>
<td><em>Curation and study —</em></td>
</tr>
<tr>
<td>Conservation methods, storage conditions; access restrictions; wear damage</td>
<td><em>Data archiving —</em></td>
</tr>
<tr>
<td>Curation and continued accessibility of paper and electronic field and laboratory records</td>
<td><em>Sources: Galloway, 1997; Garland and Janaway, 1989; Gutierrez, 2001; Haglund and Sorg, 1997; Henderson, 1987; Merbs, 1997; Mays, 1998; McCartney, 2002; Micozzi, 1991; Millard, 2001; Nawrocki, 1995; Waldron, 1987.</em></td>
</tr>
</tbody>
</table>
catacombs beneath cathedrals, and so on. The list is long because it reflects the various ways in which we humans identify and organize ourselves: by kin group, religion, social class, and occupation. Consider also that many people in the past did not use what we would recognize as a cemetery as the place of disposal for the dead. A funeral might take decades, involve several stages of ritual feasting and manipulation of the corpse and then the bones, and may not ever include inhumation burial of a complete or even nearly complete skeleton.

It is also important to recognize the performance aspects of funerary ritual. Death provides an array of powerful symbols that are harnessed for various purposes. Funerals invoke displays of economic and political power as well as control of ritual knowledge. Monumental tombs may be built for insignificant people to legitimate the power of the builder (Metcalf and Huntington, 1991:150), and ancestral tombs may be co-opted for use by a conquering regime as a means of legitimating a new power regime and discrediting the old (Arnold, 2002). Death promotes the creation of memories at many scales: the personal, the collective, the level of the grave, and the level of ritual landscape (Cannon, 2002). Mortuary ritual plays many roles in private and public lives, from the household to the global scale. All of these social meanings reside in the contextualized skeletal assemblage, intimately connected to the biological history of individuals and social groups also held therein.

It is not only where people are buried that creates variability and meaning in the mortuary and bioarchaeological record, but it is also how. Many social, physical, religious, and economic factors influence mortuary practice (Carr, 1995) and what remains as evidence of those practices hundreds or thousands of years later. The cause, location, and circumstances of death all influence treatment of the deceased and the initiation and rate of soft-tissue decomposition and skeletonization. Massive injuries and active infections may hasten soft-tissue destruction by microbial agents, as will exposure to the elements of an unburied body on a battlefield. Climate and weather affect decomposition and the timing of funeral operations. Decomposition is managed deliberately in mortuary programs that involve prolonged corpse visitation periods or multiple stages of body processing. The tissues that decompose soonest are the organs of the digestive system; then the heart and circulatory organs; then the lungs, kidney, and bladder; brain and nervous tissue; skeletal muscles; and lastly, the collagen-bearing connective tissue of the skeletal system (Gill-King, 1997:98). Clothing or burial shrouds, coffin material and hardware, and the presence of organic or metal grave goods also influence the rate of soft-tissue decay, as do the shape and depth of the grave, as detailed by Garland and Janaway (1989). All of these factors affect the chemical environment, temperature, and humidity in the grave, and they control the actions of microbial agents of decomposition. Coffins act (at least temporarily) as barriers to soil, vegetation invasions, and scavengers, but they may also retain water. General climate is also a factor: extremes in temperature and humidity may hasten or retard degradation (Galloway, 1997).

The range of variation in the archaeological record of mortuary practice should be evident in Table 3.2, which is drawn from Burial Terminology, a manual for archaeologists intended to promote consistency in burial recording (Sprague, 2005) and from Carr’s (1995) extensive ethnographic review of mortuary behavior. Many versions of both interment and cremation exist: temporary interment and disinterment, exposure and then interment, burial in water, nonburial as in the Tibetan sky burial, and cremation in one or two stages that may be followed by interment. And these versions may be preceded by embalming, mumification, fermentation, disarticulation, and processing of the body for various purposes and with various implements.

Mortuary features may contain one or many individuals, and they may represent one or several burial events. Bodies may be arranged in very particular ways or sprawled in a haphazard manner. However, the position of the skeleton may change both during and after
soft-tissue decomposition as a result of muscle and ligament decomposition; movement by water, forces of gravity, and overburden; and depending on how much space there is in the grave (Roksandic, 2002). For example, the position of the skull and the mandible in relation to the cervical vertebrae may shift in the grave and erroneously suggest premortem decapitation (Nawrocki, 1995). Burial position cannot be automatically assumed to reflect the original placement in the grave. In these cases, an understanding of the anatomy and sequence of soft-tissue decomposition is essential, and a search for cutmarks or perimortem trauma to the skull, cervical vertebrae, and mandible would be essential to test hypotheses about the origins of the burial position (Table 3.2).

The location of a specific grave relevant to other graves, to landscape and astrological features, and to other site components are all relevant to understanding assemblage history and the particular biases that may be present in the archaeological assemblage. Choice of burial location is a critical factor in preservation. A cemetery on a flood plain is likely to suffer repeated cycles of inundation. Cemeteries in desirable locations will be subject to disturbance as the spot is reused for various purposes by people over time. Burials inside structures are disturbed during architectural remodeling.

<table>
<thead>
<tr>
<th>TABLE 3.2 Variable Lists Used in Recording Mortuary Features</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Body Characteristics</strong></td>
</tr>
<tr>
<td>Body treatment</td>
</tr>
<tr>
<td>Mutilation; embalming; autopsy; cremation; Defleshing, disarticulation; excarnation; packaging bones or ashes</td>
</tr>
<tr>
<td>Complexity of disposal</td>
</tr>
<tr>
<td>Simple (one event), compound (two or more events)</td>
</tr>
<tr>
<td>Form(s) of disposal</td>
</tr>
<tr>
<td>Interment; temporary interment; water burial; exposure; cremation; fermentation</td>
</tr>
<tr>
<td>Individuality</td>
</tr>
<tr>
<td>Single, double, multiple, mass (implies disarticulation)</td>
</tr>
<tr>
<td>Articulation</td>
</tr>
<tr>
<td>Articulated; semi-articulated; disarticulated; disturbed</td>
</tr>
<tr>
<td>Position (flexure)</td>
</tr>
<tr>
<td>Body/trunk, knee, arm position; arranged or haphazard</td>
</tr>
<tr>
<td>Body orientation</td>
</tr>
<tr>
<td>Re: cardinal directions, landscape, structure, or spiritual location</td>
</tr>
<tr>
<td>Deposition</td>
</tr>
<tr>
<td>On back, side, face, sitting, standing</td>
</tr>
<tr>
<td><strong>Grave Characteristics</strong></td>
</tr>
<tr>
<td>Form</td>
</tr>
<tr>
<td>Size, shape building materials, decoration</td>
</tr>
<tr>
<td>Orientation</td>
</tr>
<tr>
<td>Cardinal directions; re: built or cosmological features</td>
</tr>
<tr>
<td>Location</td>
</tr>
<tr>
<td>Extra or intramural; Cemetery, mound, house, midden</td>
</tr>
<tr>
<td><strong>Grave Furniture</strong></td>
</tr>
<tr>
<td>Kinds</td>
</tr>
<tr>
<td>Functional/ornamental; local/exotic; broken/intact</td>
</tr>
<tr>
<td>Arrangement of</td>
</tr>
<tr>
<td>Spatial arrangement of grave goods in the grave</td>
</tr>
<tr>
<td>Energy expenditure</td>
</tr>
<tr>
<td>On grave, and on grave furniture</td>
</tr>
<tr>
<td><strong>Disposal Area</strong></td>
</tr>
<tr>
<td>Local location</td>
</tr>
<tr>
<td>Re: domestic structures, religious structures, economic areas</td>
</tr>
<tr>
<td>Regional location</td>
</tr>
<tr>
<td>Re: topography, ritual landscape</td>
</tr>
<tr>
<td>Demarcation</td>
</tr>
<tr>
<td>Marked cemetery? Marked or unmarked graves</td>
</tr>
<tr>
<td><strong>Cemetery Organization</strong></td>
</tr>
<tr>
<td>N of burials (graves)</td>
</tr>
<tr>
<td>Primary, secondary, single, multiple, mass, cremation</td>
</tr>
<tr>
<td>Types of burials</td>
</tr>
<tr>
<td>Distance from center of Chief’s ancestor’s grave</td>
</tr>
<tr>
<td>Location re: central grave</td>
</tr>
<tr>
<td>Location re: other graves</td>
</tr>
<tr>
<td>Family groups; age or gender groups</td>
</tr>
<tr>
<td>Arrangement of graves</td>
</tr>
<tr>
<td>Linear arrangement; spokes; irregular</td>
</tr>
</tbody>
</table>

Sources: Sprague, 2005 (body characteristics); Carr, 1995.
ARCHAEOLOGICAL RECOVERY OF HUMAN REMAINS

The excavation methods, purpose, and strategy of an archaeological project are also important determinants of the composition, condition, and representativeness of the recovered skeletal assemblage. Bones are damaged on discovery during backhoe trenching, grading, or other earth-moving work, and additional damage may result from handheld excavation tools. Some institutions use burial recording forms that ask the excavator to specify the specific event(s) leading to discovery of a skeleton, what equipment was in use, and what tools were used in exposing and removing the skeleton. These forms aid the osteologist in distinguishing new from ancient damage.

Newly exposed bones may dry and crack very rapidly during the course of documentation and removal (Nawrocki, 1995), and some skeletal remains are so fragile that they simply cannot be removed from the ground without completely falling apart. Sometimes consolidants are applied to fragile remains in the field. In some instances, the best procedure is removal of the entire soil block containing the burial for excavation in the laboratory. In some contexts the bone is dissolved completely and all that remain are bone outlines, stains, “sand bodies” as at Sutton Hoo (Bethell and Carver, 1987), or “pseudomorphs” or casts as in some Upper Paleolithic burials (Arias and Álvarez-Fernández, 2006).

Archaeological mitigation projects that are circumscribed by construction plans result in the incomplete excavation of burial features that fall partially outside the project boundaries. Even long term, research-driven archaeological projects cannot be assured to recover all or even a representative assemblage of the human remains present at a given site. It is extremely rare for an entire site to be completely excavated, so even a cemetery project will generally not recover all the burials—just those in the area to be affected by development or renovation. And equally important, conservation is a guiding principal of field archaeology. Since excavation literally destroys a site, it is standard to leave portions of a site unexcavated and undisturbed.

Specific research goals determine the excavation focus on certain areas of the site and de-emphasize others. For example, a sampling program designed to recover food refuse for subsistence strategy and diet reconstruction might focus on cooking hearths and other food processing areas, storage features, trash middens, and agricultural facilities. These areas may or may not be the same ones that were used for mortuary purposes. Some prehistoric communities interred the deceased beneath house floors or in subsurface features of outdoor use areas at residential sites rather than in a demarcated cemetery. The extent to which subfloor and subsurface features are investigated would thus determine the rate of recovery of human remains from a community with this kind of mortuary program. And increasingly, archaeologists in the United States operate with the explicit objective to avoid disturbing human remains when their locations are known or surmised. This situation is exemplified in the excavation history of Grasshopper Pueblo in the mountains of central Arizona, which was the location of the University of Arizona archaeological field school from 1963 until 1992. This village of Mogollon (and later also Anasazi) horticulturalists, hunters, and foragers was occupied between 1275 and 1400 A.D. Excavation of human burials ceased after 1979, and the human remains—representing 674 individuals—ultimately will be repatriated to the White Mountain Apache Tribe. The Grasshopper Pueblo skeletal assemblage has been studied by many researchers (recently summarized by Whittlesey and Reid, 2001) and is notable for the number of subadult skeletons recovered: 156 infants under age 1, and 230 children aged 1 through 16 years old at death (Hinkes, 1983:15). High infant mortality and immigration are partial explanations for the age distribution of the Grasshopper burials, but the archaeological program is also major factor. The excavation focused on the rooms of the
pueblo, where 90% of the subadults were found, whereas extramural areas like plazas, where more of the adults were buried, were excavated less intensively (Whittlesey and Reid, 2001:73).

In addition to the specific portions of a site that are investigated and their potential intersection with mortuary areas, collection methods also influence the nature of the skeletal assemblage recovered. The manner in which bones are removed from archaeological contexts and the extent to which soils from the grave feature and surrounding areas are screened is a major determinant in the recovery of small elements like carpals, infant bones, loose teeth, and bone fragments (Baker et al., 2005; Mays, 1998). Screening burial soils, inventory and analysis of fragments and of bone from non-mortuary contexts, and inspection of bone collected as nonhuman faunal material can all have a significant impact, as evident in the skeletal assemblage recovered from the beachside Hyatt Hotel Site in Tumon Bay, Guam (See Table 3.3).

This effort was a large-scale (approximately 11 acres) mitigation project; human remains representing 482 individuals were recovered from mortuary features and adjacent disturbed areas. The site dates to the Latte Phase (1000–1521) of the Chamorro culture, named for the distinctive Latte stone architecture: alignments of 6 to 12 paired stone foundation pillars with hemispherical capstones, which supported a wood and thatch superstructure. People were buried beneath the coral paving between the Latte stones, oriented toward the sea (Thompson, 1932), and sometimes between or adjacent to Latte structures. Latte stones are not always preserved in situ, but the presence and alignment of the burial clusters signifies their former presence. Earlier analyses of burial data from Hans Hornbostel’s excavation of 14 Tumon Bay Latte structures in the 1920s led researchers to conclude that, “infants and children were largely excluded in the formal mortuary areas adjacent to or within Latte sets” (Graves, 1986:147), which was taken as evidence of a ranked society with an ascribed status system focused on adult males (Graves, 1991:182–183). This early (and significantly smaller) excavation resulted in a substantially different kind of assemblage than the one recovered in the mitigation project conducted by Paul H. Rosendahl, PhD Inc., in 1990–1991.

The excavation and analysis of skeletal remains from the Hyatt site were conducted with careful attention to the multiple taphonomic factors that affect archaeological assemblages in the Marianas Archipelago: storm surges that displace and mix beach deposits, trees that grow in the middle of features resulting in movement and displacement of body parts, burial pits intruding into earlier graves, secondary burials, prehistoric disturbance of graves for the purpose of long bone acquisition for tool and weapon manufacture, and modern disturbance, including trenching during World War II and occasional ordinance removal by the local bomb squad (Douglas et al., 1997; Hanson and Butler, 1997; Stodder, 1997). All bulk soil samples and fill from burial areas were screened in the laboratory, and all faunal remains were checked for human material, which clearly altered the recovery of subadult remains. The anatomical and provenience-based inventory of every skeletal and dental element allowed us to reunite dispersed elements from within and around the Latte set assemblages with their original burials, some of which included multiple individuals.

As shown in Table 3.4, the skeletal assemblages from the five largest burial clusters include substantial numbers of subadults, from about one fourth to more than half of the total number of individuals represented in each group, compared with the 15% subadult

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Subadults*/Total</th>
<th>% Subadults</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>32/66</td>
<td>48</td>
</tr>
<tr>
<td>2</td>
<td>24/57</td>
<td>42</td>
</tr>
<tr>
<td>4</td>
<td>18/55</td>
<td>33</td>
</tr>
<tr>
<td>5</td>
<td>19/40</td>
<td>47</td>
</tr>
<tr>
<td>6</td>
<td>21/89</td>
<td>24</td>
</tr>
</tbody>
</table>

component reported by Graves (1991:182) in the collection made by Hornbostel in the 1920s. Clearly, infants and children were not excluded from these burial contexts at the Hyatt site. It is possible that these two sets of data represent two different time periods, and two different burial strategies, but there is not sufficient data to refine the dates. Additional revision of earlier reconstructions of prehistoric Chamorro social organization is afforded by the gender distribution of adults in the Hyatt site burial clusters. Contrary to the model of the adult male-dominated latte structure burial program, the Hyatt site burial groups show no significant gender bias and instead suggest kinship rather than gender as a basis for the burial groupings (Ryan, 1998).

Selective collection of human remains in the field is well known from nineteenth-century archaeological expeditions. Well-preserved adult (preferably male) crania were crated and sent back to the museums for craniological studies addressing racial taxonomy. Sometimes pathological specimens were kept, but there was less interest in the infracranial skeleton in the early years of physical anthropology (see Walker, this volume). The quality of documentation—provenience, condition, associations—of human remains also varies considerably between archaeological programs. Limitations may be the result of skill or time constraints on excavation programs. Human bones are not always found in what the archaeologist perceives to be a mortuary feature or an otherwise meaningful context. Bone found in contexts other than a typical burial feature may be dismissed as the product of rodent disturbance and may not be carefully recorded, and perhaps not even saved. Forensic taphonomists refer to this expectation-based aspect of investigator bias as “the search image” bias (Haglund and Sorg, 1997:20).

Much of taphonomy addresses alterations to the skeleton and assemblage that prevent us from seeing what the skeletal biologist typically wants to see: whole bones with well-preserved cortical surfaces, and complete skeletons of all ages and sexes in the same proportions as in the living population. The taphonomic approach allows us to systematically document and understand the sources of missing data: incomplete burials, poorly preserved bone, skewed age, and sex composition. In this sense, taphonomy deals with information to strip away, biases to account for and accommodate, and noise to screen out of our analyses. But taphonomic data are not just negative data: they allow us to explore the role of human vs. nonhuman agents in the condition of an assemblage, and to examine important realms of human behavior in the past (Lyman, 2002; Micozzi, 1991:2; Sorg and Haglund, 2002:7).

**PRESERVATION AND DECOMPOSITION OF HUMAN BONE**

In addition to mortuary practice, the preservation of the skeleton is mediated by factors
intrinsic to the bones, such as bone density, size, and shape, and by extrinsic factors, such as characteristics of the soil and water in the burial microenvironment. It is important to distinguish between two dimensions of preservation: completeness—the degree of fragmentation of bone—and condition—the degree of destruction of bone (Marean, 1991). There are several methods for documenting and characterizing both of these dimensions of preservation. White’s human taphonomy protocol (1992:110–117) includes these data points: percentage of intact outer cortex (0–100%); degree of fragmentation (50% or more of the element present; less than 50% present); presence of striae on the bone surface; degree of weathering; preparation damage; element portion present; and percent of shaft circumference present for appendicular bones. Standards for Data Collection (Buikstra and Ubelaker, 1994) includes this measure of completeness for long bones: the diaphysis should be divided into proximal, middle, and distal thirds, and each of these sections and the epiphyses (or articular surfaces) should be recorded separately. Values to record are similar to White’s: blank denotes a missing element or diaphyseal section; a score of 1 is assigned if more than 75% is present, 2 is assigned when 25% to 75% is present; and a score of 3 denotes that less than 25% of the element is present (Buikstra and Ubelaker, 1994:7). Element completeness is included in the Center for Archaeology report guidelines for reports on human skeletal remains issued by English Heritage (Mays et al., 2004). Discussion of inventory methods for disarticulated and commingled remains and for cremated remains is included in the British Association for Biological Anthropology and Osteoarchaeology’s Guidelines to the Standards for Recording Human Remains (Brickley and McKinley, 2004).

Marean’s (1991) “Completeness Index” is an assemblage-wide composite measure of element representation meant to characterize the average completeness of specimens of each skeletal element in an assemblage. This measure enables comparisons between elements and between assemblages. A version of these measures and characterizations of preservation specifically for human skeletal assemblages is proposed by Bello et al. (2006): the Anatomical Preservation Index, based on percentage of the element preserved; the Bone Representation Index: the number of times a specific element is represented in the assemblage compared with the expected number of those elements; and the Qualitative Bone Index, which captures the degree of preservation of the cortical surface in six classes.

Methods used in recording which portion of an element is present vary according to the kind of bone—long bones, vertebrae, flat bones, and with the purpose of the analysis (see Buikstra and Ubelaker, 1994; Knu¨sel and Outram, 2004; Outram et al., 2005). Related to element portion survival is the overall degree of fragmentation in an assemblage. Fragment size directly affects the identification of an incomplete skeletal element and thus the interpretation of the entire assemblage. The smaller the fragment size, the less likely it is that anatomically (and taxonomically) diagnostic landmarks will be preserved (Darwent and Lyman, 2002:356). Fragmentation affects the number of identifiable specimens (NISP) in an assemblage, the count of elements represented (minimum number of elements or MNE), and the number of individuals represented (minimum number of individuals or MNI). Unless all fragments are present and each skeletal element can be reconstructed through conjoining, assemblages of broken bones produce higher NISP counts than assemblages of unbroken bones (Darwent and Lyman, 2002).

Several different methods are used to estimate MNE and MNI in fragmentary and commingled skeletal assemblages: visual matching of elements by size, age, sex, and robusticity; sorting based on osteometric methods (Byrd and Adams, 2003); the Lincoln Index, which uses pair matching (Ubelaker, 2002); and the Most Likely Number of Individuals (Adams and Konigsberg, 2004), which also is based on pair-matching. Although the basic MNI goal is...
to determine the minimum number of individuals represented in a recovered skeletal assemblage, the LI and MLNI are designed to estimate the number of individuals in the original assemblage (Adams and Konigsberg, 2004). Thus the LI and MLNI are better suited to the purposes of the paleodemographer. For nondemographic studies, assemblage-based MNE and MNI are more appropriate since one cannot observe the presence or absence of traits or modifications of skeletal elements that are not present in the assemblage.

These data form the basis of the biological reconstruction of the assemblage represented, but element representation, completeness, and preservation are crucial for quantification of the number of observable cases for most types of pathology and for metric and nonmetric data collected in paleopidemiological and biodistance research (Waldron, 1987, 1994), and for perimortem and postmortem modification of bone.

Extrinsic Factors in Bone Preservation: Physical and Chemical Weathering

Degradation of the organic (about 90% collagen) and mineral (bone apatite) components of bone is the result of contact with microbial agents, plants, soil, and ground water in the depositional environment. Events prior to burial will affect preservation, as an exposed body may be subject to wetting and drying, freezing and thawing, animal trampling and chewing, plant disturbance, fluvial transport, and other events. And, as discussed, numerous specific aspects of the burial program, the burial environment, weathering, and other natural processes are mediated by accompanying artifacts (e.g., copper, which can enhance tissue preservation), grave type, and coffin type.

Weathering—the degree of physical destruction visible on the bone—is most often recorded on a six-point scale developed by Behrensmeyer (1978) based on actualistic studies of ungulate taphonomy in the Amboseli Reserve in Kenya. The scale starts at zero with no evidence of weathering. Stages 1 through 5 track the progression of deterioration starting with development of longitudinal cracks, flaking off of thin layers, and then flakes of bone leaving a patchy surface of fibrous bone and eventually deeper cracking and splitting and disintegration. These changes are illustrated in Behrensmeyer (1978) and in Buikstra and Ubelaker (1994:98–99).

Although the general sequence of degradation is applicable to any assemblage and provides a standard for description, the stages cannot be equated with time elapsed since death. Even within a single assemblage, bone of the same age can exhibit markedly different states of preservation, including the weathering stage (Lyman, 2002; Nielsen-Marsh et al., 2000). Weathering can also mimic damage from thermal alteration, so where there is no color change, it can be difficult to distinguish between these two sources of bone surface deterioration (White, 1992). Field studies since Behrensmeyer’s have shown that the timing of weathering stages is markedly distinct in different environments. Elephant bones exposed in the Ituri tropical rainforest in Zaire survived much longer without longitudinal cracking and splitting than the ungulate bones on the Amboseli savannah (Tappen, 1994). A study of chimpanzee bone taphonomy in the Kibale Forest of Uganda also documents delayed weathering (Kerbis-Peterhans et al., 1993). In the rainforest, the bones are protected from extreme temperature variation and the action of ultraviolet radiation, which damages collagen fibrils and the mineral crystals embedded therein (Nielsen-Marsh et al., 2000:440; Tappen, 1994). Repeated inundation and exposure of burial sites by flooding cycles accelerates both disarticulation and weathering in burials (Littleton, 2000).

Another critical aspect of weathering for taphonomy studies is that climate has a major impact on the length of time a bone remains in its “green” or vital state (with intact collagen and lipids) before becoming “dry bone” from loss of organic component and increased porosity and brittleness (Nielsen-Marsh et al., 2000). This aspect complicates the
interpretation of perimortem and postmortem damage to human bone. Dry and green bone fracture in distinctive ways, but if bone remains in a green state long after death, then we might be misinterpreting perimortem modification as postmortem modification (Nielsen-Marsh et al., 2000; Sorg and Haglund, 2002).

Bone degradation takes place by two processes: the destruction of the collagen component by bacterial collagenase and the chemical demineralization of bone apatite (Nielsen-Marsh et al., 2000:441). It has been thought that the collagen phase of bone was destroyed first and then the destruction of the mineral phase began: the mineral crystals are embedded in collagen fibrils, so collagen degradation exposes the mineral crystals to dissolution. But these processes must be contemporaneous to some extent. Some demineralization and increase in pore size have to occur to allow the large-size collagenase molecules into the bone (Child 1995:168; Collins et al., 2002; Nielsen–Marsh et al., 2000). As more and more bone mineral is dissolved, pore size increases, and the bone collagen is increasingly available to microbial action. Soil microbiology (population of bacteria and fungi) is a component in differential preservation, but this is largely a matter of soil chemistry, which is a far more important determinant in bone preservation.

The most important factor in dissolution of bone mineral is ground water, which acts as the medium for mineral ion exchange between the bone and the immediately surrounding soil, rock, or grave surface (Gill-King, 1997:105; Henderson, 1987; Millard 2001; Nielsen–Marsh et al., 2000:442). The chemistry of the ground water relative to bone and the degree of fluctuation in ground water contact with bone determine the rate of dissolution of the bone mineral. Ions are exchanged between the bone and the soil solution or water until the bone and the soil or water are in equilibrium. Wetting and drying cycles thus have a more dramatic impact as this process is repeated. The exchange of ions between bones and their environment, known as diagenesis, has important implications for research involving bone chemistry for biomolecular archaeology, dating of bone, and isotope-based dietary reconstruction (see the chapters by Burton, Katzenberg, and Stone, this volume). Diagenesis studies measure the degree of specific types of bone alteration at the histological level using a variety of methods, including optical microscopy, and mercury intrusion porosimetry (Millard, 2001; Nielsen-Marsh et al., 2000).

The impact of grave form and microenvironment on differential bone preservation in burials within a site is demonstrated in Merbs’s study of Eskimo graves at the Kamavik and Silumiut sites in the Northwest Territories of Canada. Comparing only adult graves, lower skeletal inventory scores (presence of a skeletal element even if incomplete) were found in graves with bedrock floors compared with those where the floor was soil, gravel, or another type of rock (Merbs, 1997:259). This is attributed to the acidity of the bedrock and poor drainage in the bedrock-floored graves. Of the remains on bedrock-floored graves, it was observed that, “bone at the contact zone had simply disintegrated, giving the appearance of having dissolved” (Merbs, 1997:251).

**Intrinsic Factors in Bone Preservation: Size, Shape, and Density**

Size, shape, surface area, and bone density are all intrinsic aspects of bone that affect their survival. Building on zooarchaeological research on the accumulation and condition of animal bone deposits associated with contemporary hunter-gatherer settlements (Binford, 1981; Brain, 1981), Waldron’s (1987) study of the survival of different bones in burials from a Romano-British cemetery in London demonstrated the importance of size and anatomical position in bone survival. He calculated the number of each skeletal element expected in the assemblage given the number of discrete graves. He found that the best represented bones were the “dense, relatively heavy bones such as the petrous temporal bone and mastoid
in the skull, mandible, acetabulum, sciatic notch and proximal ulna” (1987:62). He observed under-representation of bones in anterior positions: the sternum, coracoid, acromion, pubis, and patella (Waldron, 1987:62). The least well-represented bones were phalanges, carpals, the coccyx, scapula, and small tarsals. Waldron cautions that this particular set of findings might not be typical at another site, but these general patterns are typical of nonhuman remains assemblages as well.

**Bone Density and Element Survival.** Differential survival of skeletal elements is also well documented for the human remains from the Crow Creek (South Dakota) massacre assemblage (Willey, 1990; Willey et al., 1997). This assemblage representing a minimum of 486 individuals (based on MNE of the right temporal bone) was discovered eroding out of a fortification ditch in the Crow Creek village site in 1978. The remains were studied in five months and reburied in 1981. The two bone beds, the larger a deposit of bones four and a half feet thick, date to the Initial Coalescent Period, ca. 1325 A.D. The bodies underwent an array of taphonomic processes: “[F]irst there was murder and mutilation of the villagers, followed by scavenging and decomposition, stabilized by collecting and burying the body parts, then exposure by erosion, later looting and excavation, and completed with cleaning, analysis, and reburial” (Willey, 1990:xxiv).

Element survival in this assemblage was affected by dismemberment during the attack, damage by scavengers, loss of body parts that may have been overlooked when the bodies were buried by the survivors, differential preservation after burial, and loss during excavation, processing, or inspection (Willey, 1990:20). The impact of disarticulation and transport is reflected in the element survival: larger, denser bones and those closer to the torso were more frequently present than the smaller, lighter, more distal ones (Willey, 1990:14). Comparison of the element survival patterns in the Crow Creek assemblage with bone density studies of modern skeletons (Galloway et al., 1997) shows that the survival of elements and element portions in the Crow Creek assemblage correlate closely with differential bone mineral density (Willey et al., 1997).

Galloway and colleagues (1997) generated bone mineral density data for specific anatomical portions of skeletal elements in modern human adult skeletal remains to provide “baseline data” for anthropological research. Scans of modern long bones were made with a single photon absorptiometer at major anatomical landmarks and at sites 20%, 30%, 50%, 65%, and 80% of the diaphyseal length from the distal end. Their data for the humerus and tibia are shown in Fig. 3.1. Sexual dimorphism in bone density was documented at most scan sites as well as age-dependent bone density loss (Galloway et al., 1997). The midshaft densities of the longbones are ranked (highest to lowest) in the same order as their frequency of recovery in forensic cases involving exposure and scavenging: femur, tibia, humerus, radius, ulna, and fibula (Galloway et al., 1997:314).

The relative mineral density of specific portions of the long bones is important for methodological reasons. Willey used certain parts of the skeleton to calculate MNE (and thus MNI) for the Crow Creek assemblage: the external auditory meatus and petrous of the temporal bone, deltoid tuberosity of the humerus, base of the radial tuberosity, base of the coronoid process of the ulna, anterior crest of the tibia at the nutrient foramen, base of the lesser trochanter of the femur, and lateral aspect of the distal end of the fibula shaft (Willey, 1990:10–11). These osteological features are recognizable and are easy to side, and they were generally well preserved in the Crow Creek assemblage. The modern bone density data showed that the points used for MNE estimates of the radius, ulna, tibia, and femur are among the bone portions with the highest density, but that other points on the humerus and fibula with higher mineral density would probably have been represented more frequently in the assemblage. Use of these areas would have increased the MNE.
counts (Willey et al., 1997:524). “Denser elements and denser element portions are more likely to survive and provide better estimation of the “maximum” minimum number of individuals than less dense parts” (Willey et al., 1997:527).

The differential density within skeletal elements, and the correlation between intra-element density and element portion preservation in the Crow Creek assemblage, demonstrates the value of bone density data in predicting element representation. “Absence of certain segments, where bone mineral density is low, should be expected as the postmortem interval increases or where environmental conditions are particularly destructive. The absence of bone that is more densely constructed however, suggests that some form of selection, perhaps by opportunistic scavengers or deliberate human modification of remains, has occurred” (Willey et al., 1997:527).

**Bone Density and Paleodemography.**

Bone mineral density is low in infants and in the elderly, the two age classes typically under-represented in skeletal assemblages. But bone density varies throughout life and is vulnerable to reduction during normal growth, pregnancy, and illness or periods of nutritional stress. Bone density at birth is determined by many aspects of maternal health, including diabetes, perhaps because of decreased transplacental mineral transfer, and vitamin D deficiency (Namgung and Tsang, 2003). Small-for-gestational-age infants have decreased bone formation *in utero* and lower than normal bone density. Season of birth is an important factor in maternal vitamin D status, as biosynthesis of vitamin D is lowered by reduced exposure to sunlight during winter months. Bone mineral content in the forearm and lumbar spine of adults differs according to season by an average of less than 2%, but in newborns, the difference is between 8% and 12% (Namgung and Tsang, 2000:58). Even in full-term neonates, bone mineral density is very low and remains so in the first year of life, but then it increases rapidly throughout childhood with influences from weight-bearing, activity level, and nutrition (Rauch and Schoenau, 2001). Boys tend to have higher bone density than girls, but this is a
factor of body size (Gilsanz, 1998). Bone density in childhood is affected negatively by osteoporosis (loss of bone mineral and matrix) as in juvenile idiopathic osteoporosis or as a result of chronic disease (van der Sluis and de Munick Keizer-Schrama, 2001). Rickets results from loss of bone mineral caused by reduced calcium-phosphate levels in vitamin D deficiency. Childhood bone density is also decreased by osteogenesis imperfecta and congenital or developmental conditions involving mutations in collagen genes (Gilsanz, 1998).

Peak buildup of bone density is reached at the end of adolescence, maintained through young adulthood in healthy individuals, and then slowly depleted depending on the individual’s activity level and health status. During pregnancy bone density fluctuates as bone turnover rates vary between the trimesters and the post-partum period (Black et al., 2000; Shataheri et al., 1999). Both premenopausal and postmenopausal osteoporosis are crucial factors in adult bone density. Although diet and activity levels are important determinants of bone density in adults, it is estimated that as much as 80% of variation in bone mineral density may be controlled genetically by variants in genes related to vitamin D, collagen, and estrogen (Van der Sluis and de Munick Keizer-Schrama, 2001:818). Studies of subadult bone density variation at the populational level are complicated by sampling and methodological issues: methods used in measuring bone density, bone and site used, and the difficulty of controlling for differential development of adolescents (Gilsanz, 1998; Van der Sluis and de Munick Keizer-Schrama, 2001), but bone density is greater in individuals of African ancestry than in Europeans, Asians, or Native Americans (Galloway et al., 1997; Lam and Pearson, 2005), but as Lam and Pearson warn, density should not be used as a univariate proxy for bone survival while other intrinsic and extrinsic factors are ignored (2005:107).

Differential preservation nearly always is mentioned as a factor in the under-representation or absence of infants and subadults in archaeological skeletal assemblages, as exemplified by empty child-sized graves in historic cemeteries known to have been used for children’s burials (Bello et al., 2006:29; Guy et al., 1997:226; Larsen et al., 1995:142). Looking more closely at the age trends in bone mineralization, Guy and colleagues suggest that a threshold at about age 3, when mineralization substantially increases, accounts for the increased representation of older infants in the archaeological record, not “the deadly effects of weaning” (1997:226). Some question exists about the timing of this “threshold” since the number of different methods used to measure and calculate bone density (especially problematic in children) makes the results of many studies incomparable (Gilsanz, 1998; Lam et al., 2003; Petit et al., 2005; van Rijn et al., 2003), and Bello et al. (2006) suggest that the threshold is later and not as abrupt as portrayed by Guy and colleagues.

But contrary to the notion that most infants and children disappear from the archaeological record, there are archaeological skeletal assemblages with substantial numbers of infants and children. Grasshopper Pueblo, mentioned above, is one. The Honokohua site on the island of Maui is another: 39% of the skeletal sample ($N = 712$) were subadults under age 15 (Pietrusewsky et al., 1991). In the skeletal assemblage from Kok Phanom Di in Thailand ($N = 107$), 36% are under age 15 (Tayles, 1996). There are assemblages of all children such as those from pre-Roman Geto-Dacian contexts where children and infants were buried in settlements, at shrines, and in natural rock crevices near settlements (Sirbu, 2005). The Eastern Necropolis at Deir el Medina (New Kingdom Egypt) had an area reserved for the youngest children: infants, fetuses,
still-born babies, and even placentas (Meskell, 1994). And there are skeletal assemblages with no infants: Mokrin, an Early Bronze Age cemetery (Rega, 1997), and Peqi’in, a Chalcolithic burial cave in Galilee (Nagar and Eshed, 2001).

It is unusual to find infant and child remains treated in the same manner as adults in the archaeological record, and these are the assemblages that stand out (Parker Pearson, 2002:103), not the ones where children are treated differently, or buried in other, special places. Interpretation of age-patterns in a burial assemblage must be based on knowledge of mortuary practice at the household, community, and regional levels, since there may be several kinds of mortuary programs and locations in use at once. If there are no infants in an assemblage, then yes, they might have been there and disappeared, especially if preservation is very poor overall. But they might also be somewhere else. The absence of infants should be interpreted as a reflection of cultural practice, not as an indication of taphonomic erasure of all infant bones. The age at which children appear in a cemetery is likely to reflect a culturally defined threshold: dental eruption in ancient Rome or baptism in some Christian cultures (Arnold, 1991:116, citing Häusler, 1968).

Part of the difficulty in understanding culturally constructed skeletal assemblages lies in our confusion of biological categories with social categories (Sofaer, 2006). As osteologists we group skeletons by age and sex, and most of our research follows biologically based lines of inquiry. In this process we may conflate biologically defined groups with social groups and inappropriately assume certain social roles for groups of people, especially children, in the past (Scott, 1997). Childhood is constructed culturally, and mortuary treatment of infants and children reflect specific, culturally defined thresholds in personhood. Skeletal assemblages are modified by numerous factors before and after deposition, but they are culturally created entities, and the particular age and sex composition has cultural meaning.

**Taphonomy and Paleopathology**

Poor preservation of skeletal assemblages limits the ability to observe pathological conditions (Waldron, 1994) and obliterates evidence of pathological lesions even when the bone is still present. New periosteal bone on a subadult long bone can be lost (Fig. 3.2), enamel can be chipped off teeth, etc., which can happen in the burial environment; during excavation, transport, or cleaning; and in the laboratory and museum. Skeletal remains are damaged cumulatively by handling by researchers and students (Caffell et al., 2001; Janaway et al., 2001).

Postmortem damage to skeletal remains may also include changes to the shape or surface of the bone that mimic pathology. In the Eskimo

---

**Figure 3.2** Periosteal new bone chipping off an infant tibia (NMNH 314,341).
skeletons mentioned, Merbs noted lichens on the bones that looked like neoplasias and local areas of bone destruction (from contact with acidic rock) resembling lytic foci of carcinoma (Merbs, 1997:151). Eroded holes in Greenlandic Eskimo crania resulting from the same process could be mistaken for trephination (Merbs, 1997:251, citing Pales et al., 1952). Bones that have been warped by dampness or overburden may suggest rickets and other bowing deformities (Glencross and Stuart-Macadam 2000; Merbs, 1997). Taphonomic changes in the skeleton can also mimic traumatic injury: postmortem warping, separation of sutures, and other damage to (especially subadult) cranial bones may be interpreted as evidence of antemortem or perimortem cranial trauma (Crist et al., 1997). Burials that were wrapped tightly in textiles that are not preserved at the time of excavation can suggest congenital deformation or traumatic injury to the hands or feet when their contorted positions actually only reflect their compression in the shroud (Valentín et al., 2001).

Cremation poses special challenges to the collection of osteological and paleopathology data as bones and teeth are fragmented, warped, discolored, and subject to highly variable rates of shrinkage. Cremation is typically a multistage mortuary ritual. Once sufficiently cooled, the remains are collected and placed in a pit feature, pot, basket, or other receptacle, but it is rare that all of the remains are collected, and not all of that is identifiable (McKinley, 2001:284; Reinhard and Fink, 1994). Some variations include secondary reburning of remains and division of the remains for burial in more than one place, practices that reduce further recoverable biological information (Beck, 2005). Because of the fragmentation and shrinkage, metric data from cremated remains are generally considered to be unreliable (but see Correia, 1997 and McKinley, 2000 for debates on this). Age sometimes can be assigned to broad categories based on cranial suture closure, but auricular surfaces of ilia are represented by small fragments if at all. McKinley reports that only about 4% of 4000 British cremation burials had pubic symphyses preserved (McKinley, 2000:410). Sex can sometimes be assessed visually (McKinley, 2001). The position of the body on the pyre affects differential preservation of different parts of the body, but generally denser bones and those with the most tissue covering are more commonly preserved (Correia, 1997), including cranial vault portions with suture lines, mandible fragments, head of the femur and humerus (Correia, 1997), unerupted tooth crowns, and the petrous portion of the temporal, which can be used in age estimation. Preserved cranial features can be used for nonmetric trait analyses (Merbs, 1967). Specific pathological changes that are reported from cremations include trepanation, dental disease affecting tooth roots, diffuse idiopathic skeletal hyperostosis (McKinley, 2001), porotic hyperostosis, antemortem tooth loss, vertebral osteophytosis and osteoarthritis in appendicular joints (Reinhard and Fink, 1994), neoplasias, traumatic injuries, vertebral compression fractures, congenital anomalies in the vertebra and sacrum, nonspecific infection, cribra orbitalia, possible treponemal infection, or leprosy (Blau, 2001). Heat altered bone is often very well preserved, so representation of age and sex diagnostic portions of the skeleton is a more critical limiting factor than degree of fragmentation (Merbs, 1967:595).

**ANIMAL AGENTS IN HUMAN BONE MODIFICATION**

Animals are a significant source of modification to human remains before and after burial. Animals trample, transport, collect, and redeposit bones, and they chew, fracture, and consume body parts. These activities result in disarticulation of the body and the scattering and destruction of skeletal elements, and they leave a variety of marks on bone surfaces: gnawing marks, bite marks, fractures, scratches and carrying marks, and surface modification by digestive acids. Carnivores are especially destructive; dogs, coyotes, wolves, hyenas, and leopards
break open bones to get at the marrow, and they can completely consume soft trabecular regions of bone (White and Folkens, 2005:55). The scavenging behavior of canids (Haglund 1997a; Haglund et al., 1989), pigs (Berryman, 2002; Greenfield, 1988; Spennemann, 1994), rodents (Haglund, 1997b), and even polar bears (Merbs, 1997) have been studied by forensic anthropologists. The zooarchaeological literature addresses scavenging behavior in other species (e.g., bears, lions, hyenas, and wolves in Haynes, 1983) in the several decades of research papers addressing the question of hunting versus scavenging by early hominids. Carnivore damage to ungulate remains (and other prey species) at hominid fossil sites like Swartkrans in South Africa (Pickering et al., 2000) and the FLK Zinjanthropus site at Olduvai (Dominguez-Rodrigo and Barba, 2006; Selvaggio and Wilder, 2001, and many more) is studied to determine the origins and sequence of tooth marks and tool marks on animal bones, based on species-specific tooth morphology and behavior, fracture patterns, and cutmark distribution.

These methods are used in studying human remains to discern (when possible) the sequence of events in postmortem modification of bone assemblages, and to distinguish between animal and human agents of modification. Rodent gnaw marks are perhaps the most commonly observed animal-modification on human bones. These marks are generally distinctive, appearing as rows of fine scratches or parallel sets of scratches. They may appear almost as beveling on crests and ridges (Fig. 3.3). Punctures in bone made by the canines or carnassials are also distinctive and leave holes on bones and scalloped or crenulated edges on bone ends (Binford, 1981:44). But some chewing marks are not so obvious. Furrows and pits (incomplete punctures) result from the bone being carried in an animal’s mouth or from a tooth traveling across but not penetrating the bone (Binford, 1981; Milner and Smith, 1989). Pits and furrows may be smooth or textured depending on the morphology of the teeth involved and the “freshness” of the bone. The bone fragments in Fig. 3.4 were chewed by pigs and probably also by dogs. The striations in the bite marks (Fig. 3.4c and 3.4d) result from the topography of the pig molars (which have many accessory cusps) and the transverse motion in pig chewing (Herring, 1976). Tooth pit size depends on the size of the tooth and

Figure 3.3 Rodent gnawing, mandible.
its function, specifically crushing, grinding, or piercing, but also on the density of the bone being chewed. Cancellous bone near the metaphyses is less resistant than cortical bone in the diaphyses and midshafts, and so larger tooth pits may be present at bone ends (Selvaggio and Wilder, 2001). Bone chewing proceeds from the softer, cancellous bone at the ends to the harder shaft of the bone, which the animal might not be strong enough to puncture, so marks on long bone shafts are more likely to be pits and furrows (Binford, 1981:46).

Systematic study of scavenging behavior by canids (dogs and coyotes) indicates a sequence of five relatively consistent stages of scavenging: 0 = soft-tissue scavenging with no removal of body parts or modification to the skeleton; 1 = destruction of the ventral thorax and removal of upper extremities, including clavicles and scapulae; 2 = complete or partial removal of lower extremities; 3 = disarticulation of all skeletal elements except vertebral column; and 4 = total disarticulation, only cranium and miscellaneous elements and fragments remaining (Haglund, 1997a:368; Haglund et al., 1989). In forensic cases these stages correspond to time interval since death starting with a few hours at stage 1 to 22 months for stage 4.

Scavenging patterns depend on the condition of the body, cause of death, and degree of disarticulation, but in general bone ends are consumed by pigs (Greenfield, 1988), and they are targeted by carnivores as well (Milner and Smith, 1989). In the Norris Farms (Illinois) skeletal assemblage, long bone ends as well as portions of flat bones were consumed by carnivores (Milner and Smith, 1989). The remains of 30 individuals in an Oneonta cemetery (dated to ca. 1300) apparently were exposed to carnivore scavengers before being buried. In this assemblage, the most frequent damage was to the ends of long bones, face, and bones in the abdominal and gluteal regions. This is very similar to the patterns of scavenging in the Crow Creek assemblage: chewing attributed to dogs, wolves, and coyotes was most frequent on the cancellous bones, which have projections (os coxae, sacra, vertebrae) and on the ends of long bones (Willey, 1990:151).

It can be difficult to distinguish between less extreme forms of animal modification on bone. Shallow scoring or gnaw marks have been interpreted as cutmarks and as simply

Figure 3.4 Chewing marks on long bone fragments: (a) gnawed edge, (b) a tooth pit, (c) a shallow depressed scored area, and (d) a deep furrow with scoring.
wear from the burial substrate. Animal chewing damage may obliterate evidence of antemortem trauma (of particular importance in massacre assemblages), mortuary processing, and other biological evidence in human remains. This is another instance when familiarity with the prehistoric environment, mortuary practice, and the faunal remains recovered from an archaeological site is extremely important in the interpretation of human skeletal assemblage histories.

**HUMAN AGENTS AND HUMAN INTENTIONS IN BONE MODIFICATION**

Human activity produces three major categories of modification to human remains: (1) tool marks on the bone surface, (2) fractures and fracture products, and (3) thermal alteration. Identifying ancient human activity as the agent of modification is predicated on systematic evaluation of the range of possible sources of modification: animal chewing, contact with sharp rocks in the sediment, excavation damage, preparation, and curation modification (e.g., Pickering et al., 2000). In human-modified assemblages, taphonomy studies try to distinguish among three broad categories of activities: violence resulting in traumatic injury and death; perimortem or postmortem body processing for secondary burial or a particular stage of extended mortuary ritual; and body processing for consumption (cannibalism). These activities would seem to be distinct behaviors, but they are not mutually exclusive, nor do they result in entirely distinctive patterns of bone modification. Quantitative and qualitative taphonomic data do not in themselves answer anthropological questions about the motivations and emotions of people in the past. But the combination of systematic taphonomy data and contextualized analysis does enable us to explore the human intentions at work in creating skeletal assemblages.

### Toolmarks on Human Bone

Categories of toolmarks include cutmarks made by a sharp tool held perpendicularly to the bone surface; scrape marks—sets of several shallow, narrow, closely spaced marks across a bone surface; chop marks—shorter and broader than cuts; and percussion marks from hammerstones or other heavy tools or weapons used to fracture a bone (White, 1992; White and Folkens, 2005:61–62). Observation of these modifications is dependent on the condition of the bone surface, and interpretation is dependent in large part on identification of a bone or bone fragment as to element and anatomical portion. The number, size, and anatomical location (and for cut and scrape marks, the orientation) of tool marks are all important, as is the exploration of what kind of tool was used: a knife, a scraper, a hammerstone, a sharp pointed weapon, or a blunt object. The material is also important; was the implement made from obsidian, flint, metal, bamboo, shell, or bone? Cutmarks made by stone tools are typically v-shaped in profile, with some internal striations, whereas those made by softer material like shell (DeGusta, 1999; Toth and Woods, 1989) and bamboo (Spennemann, 1990) are shallower and more u-shaped (and more closely resemble animal tooth furrows or scratches). Cutmark morphology is analyzed by casting the cut in latex and studying the cast with the scanning electron microscope.

Cuts and other toolmarks are interpreted as being purposeful in their location and orientation (Lyman, 1994:298), and some basic patterns can be associated with certain activities. Toolmarks at joints indicate disarticulation or dismemberment; toolmarks at areas of muscle attachment are interpreted as evidence of defleshing or meat removal; cutmarks around the face suggest defleshing; cutmarks on the cervical vertebrae and basicranium suggest decapitation.

### Postcranial Modification Patterns

As shown in Fig. 3.5a, defleshing the body for secondary burial results in many series of cutmarks where soft tissue, including ligaments and periosteum, has been removed (Olsen and Shipman, 1994:380). The consistent orientation of the marks in this case suggest systematic processing of this individual, an adult male,
prior to disarticulation and deposition of the bones in the burial mound at Boundary Mound, a Middle Woodland site. Secondary burial may involve disarticulation and associated marks, but if the body is exposed after death or buried in a temporary grave for initial decomposition first, then extensive cutting may not be needed and no toolmarks will record this disarticulation procedure (Stodder and Reith, 2003). The individual shown in Fig. 3.5b exhibits only a few cutmarks on the distal femur and cranium but also has perimortem fractures on the right femur and cranial trauma. These remains are from a large assemblage of mixed human and animal bone from Velim Skalka, a fortified Bronze Age site in the Czech Republic (Outram et al., 2005). Typical mortuary practice for this culture was inhumation in tumuli, so one question addressed in the analysis was whether the people found in pits, partially disarticulated, and in haphazard positions had been butchered or showed any evidence of postmortem processing. The modification pattern does not suggest defleshing or intentional widespread breakage of the bones.

Figure 3.5 (a) Defleshed secondary burial, Boundary Mound Burial 6. (Adapted from Olsen and Shipman, 1994:380, fig. 5.) (b) Cutmarks, cranial trauma, perimortem fractures (no defleshing), a Bronze Age trauma victim. (Adapted from Outram et al., 2005:1706, fig. 6). (c) Cutmarks on defleshed skeleton, Alfred Packer-historic cannibalism assemblage. No disarticulation. (Adapted from Rautman and Fenton, 2006:331, fig. 6.)
and most likely reflects injuries sustained during battle, rather than butchering or other postmortem processing (Outram et al., 2005). The distribution of cutmarks in Fig. 3.5c clearly contrasts with the other two examples; the cuts are located at areas of large muscle attachment near the proximal and distal ends of the long bones and were produced during meat removal; the body was not disarticulated, and no evidence exists of any other processing (Rautman and Fenton, 2005). This is one of the individuals from the historically documented incident of cannibalism by Alfred Packer, stranded in the Rocky Mountains in the winter of 1874.

Cranial Modification Patterns. Examples of cranial modification resulting from different activities are shown in Fig. 3.6. Scalping generally produces linear marks on the frontal and parietals (Fig. 3.6a) as in Burial 62B from the massacre assemblage at the Larson Site (Olsen and Shipman, 1994:383). Although we associate scalping with violence, it is important to note that skeletal remains exhibiting signs of scalping are also found in formal burials with grave goods and without indication of traumatic injury or violent death (Owsley, 1994:338). Defleshing the skull for secondary burial in Plains groups results in many more cutmarks on the cranial vault, face, and mandible, as shown in the drawing of Burial 8 from Split Rock Creek (Fig. 3.6b), a Late Woodland Site (Olsen and Shipman, 1994: 381). Defleshing marks are also evident on the skulls of Aztec sacrificial victims as shown in Fig. 3.6c. Some of these exhibit perforations of one or both temporals so that the skulls could be threaded in rows on a tzompantli, or skull rack. The skulls with perforation of only one temporal were placed at the left or right end of a row of skulls.
Figure 3.6 (a) Scalping, Larson Site Burial 62B. (Adapted from Olsen and Shipman, 1994:383, fig. 8.) (b) Defleshing for secondary burial, Split Rock Creek Burial 8. (Adapted from Olsen and Shipman, 1994:381, fig. 6.) (c) Cutmarks typically found on Aztec sacrifice victims and temporal perforations on a skull prepared for mounting on a skull rack (*tzompantli*). (Adapted from Pijoan Aguade and Lory, 1997:231, fig. 8.8.) (d) Curated skull from Ibunda, Lower Sepik, Papua New Guinea. Incised design on parietais, pitch and clay on frontal, holes drilled for attaching mandible and for hanging skull, defleshing marks, cuts from fiber ties, orbit and nasal damage from pith eyes and nose (FMNH 43486).

(Pijoan Aguade and Lory, 1997:229). Figure 3.6d shows cutmarks associated with defleshing, disarticulation, and reattachment of the mandible on nineteenth-century curated skulls from New Guinea. Holes were drilled for fiber ties to hold the mandible and to hang the skull from a skull rack. Other decoration is applied as incised designs, paint, clay, and artificial eyes and noses made of wood pith. The wood leaves scratches and small fractures
along the eye orbits and nasal margins, and the fibers leave cutmarks on the mandible and zygomatics (Stodder, 2006).

**Skeletal Modification in Mortuary Ritual**

Secondary mortuary rites, rather expansively defined by Metcalf and Huntington (1991:97) as “the regular, and socially sanctioned removal of the relics of some or all deceased persons from a place of temporary storage to a permanent resting place” encompasses a wide range of treatments of the dead which are practiced in many societies past and present. Because there may be such a range of activities associated with manipulation of the dead, the recognition of secondary mortuary ritual and resulting bone modification is in large measure dependent on context. Olsen and Shipman provide a systematic comparison between the modification of skeletal remains resulting from conflict and those resulting from secondary burial practices in the Southern Great Plains (Table 3.4). This comparison is extremely useful because although the same patterns might also be seen in other archaeological cultures, the interpretation must be culturally specific. The mere presence of toolmarks—even when their specific purpose, removal of the cranium for instance, is evident—does not distinguish violence from mortuary ritual as the source of bone modification. Bone modification patterns can help to distinguish between communal interments resulting from periodic secondary mortuary rituals involving large numbers of people, such as the Huron Feast of the Dead ritual, which was held every 8–10 years (Kidd, 1953), and mass graves resulting from warfare or epidemics (Hutchinson and Aragon, 2002).

Intensive processing of the body can represent perimortem mutilation, torture, and violent death, or it can be the result of veneration and honorific procedures. Energy expenditure in mortuary ritual is considered a reflection of an individual’s importance (Tainter, 1975). In this model, more modification of the skeleton can be interpreted as a sign of higher prestige. Speal (2006) suggests that postmortem processing (including drilled holes in long bones and removal of bone disks from crania) in Younge Mortuary Complex burials from the Riviere au Vase and other Central Great Lakes Late Woodland sites can be seen as a sort of proxy for status in the same sense as grave goods. Processing can reflect positive or hostile intent; it may reflect treatment of an extremely prestigious or an extremely dangerous individual. One interpretation of the extreme processing in several Southwestern U.S. skeletal assemblages suggests that this was a feature of witch execution (Darling, 1999; Margolis, 2000). The line between veneration and violation is culturally drawn and may not be readily interpretable by archaeologists (Duncan, 2005).

Mortuary ritual in Melanesia often involves multistage mortuary programs and curation of skulls and other skeletal elements. These ethnographically documented practices are reflected in the archaeological record of secondary burials, missing elements, and toolmarks indicating disarticulation (Stodder, 2005; Stodder, 2006). Historically, skulls of some individuals in the Papuan Gulf and Sepik regions of Papua New Guinea were cleaned, decorated, and kept in various contexts. Taphonomic analysis of curated skulls collected in the early 1900s reveals the numerous additive and reductive modifications to the skulls (see Fig. 3.6d). Some of these modifications—cutmarks, and enlargement of the foramen magnum—might be attributed to cannibalism if the particular context was not considered. Ethnographic evidence also indicates that some procedures, like detaching the cranium and cleaning the skull, could have been done without use of any tools and without leaving any taphonomic record since temporary interment or exposure in a tropical setting were used as means of skeletonization before additional processing (Stodder and Rieth, 2003). Forensic investigations in tropical settings indicate that decomposition of ligaments sufficient for manipulation and secondary...
interment could be complete in about three years (Spennemann and Franke, 1995:71).

Patterns of modification in the cranial and the postcranial skeleton all suggest certain behaviors (and sometimes motivations) on the part of the people who handled the deceased, but it is critical to interpret these patterns in their specific archaeological and anthropological context, against the backdrop of local and regional mortuary practice, comparison between treatment of animal and human remains, and through assessment of alternative possible explanations for observed conditions—not only alternative agents but alternative behavioral scenarios. As the above examples show, defleshing is not always an act of desecration; not all curated skulls are trophy skulls taken in warfare or head-hunting; not all scalps were taken in violence; and not all postcranial fragmentation and cutmarks are the result of cannibalism. This is equifinality: more than one process can lead to the same result. The two mandibles in Fig. 3.7 exhibit cutmarks in the same location on the anterior edge of the ramus, which suggests they were made during removal of the masseter or temporalis muscles. The mandible in Fig. 3.7a is part of the cannibalized assemblage from Site 5MTUMR 2346 in Mancos Canyon Colorado described by White (1992). The mandible in Fig. 3.7b is from a decorated skull (Fig. 3.7c) collected between 1910 and 1913 in Kerau, a village on the Sepik River.

![Figure 3.7](a) Cutmarks on mandible from cannibalized assemblage, Southwestern Colorado (fig. 7.39, White, Tim D. Prehistoric Cannibalism at Mancos 5MTUMR2346. © 1992 Princeton University Press. Reprinted by Permission of Princeton University Press). (b) Cutmarks on mandible, curated skull (FMNH 43366). (c) Curated skull with incised design, pith eyes and nose, mandible attachments at nasal, and zygomatic, decorative carrying frame. Kerau, Lower Sepik River, Papua New Guinea (FMNH 43366).
Distinguishing Perimortem and Postmortem Trauma: Fractures and Fracture Products

Fracture types are important in distinguishing between ancient and modern fractures. Bone in a “green” or vital state with high moisture content and intact collagen tends to fracture in a helical or curvilinear manner, whereas “dry” bone with degraded collagen or complete destruction of the organic component has more angular fracture patterns. Although this information does not tell us what broke the bone, fracture types help to distinguish between bone broken before or shortly after death—antemortem and perimortem fractures—and bone broken later—postmortem fractures (see Lovell, this volume).

In the longbone shaft fracture classification scheme shown in Fig. 3.8 (from Marshall, 1989:14), the spiral, v-shaped, sawtoothed (oblique angles on the fracture edges) are characteristic of vital bone, as are flaking fractures and the depressed fracture with adhering flakes (Galloway, 1999; Lovell, 1997; Lyman, 1994; Marshall, 1989; White, 1992). In the cranium, depressed fractures, concentric circular fractures, and radiating fracture lines occur in vital bone. The presence of healing—some degree of callus formation, new woven bone at the injury site—is the only reliable indication of an antemortem fracture. The time to initial healing, at least 3 weeks for initial callus formation (Lovell, 1997:145), depends on the type of fracture, location, and age of the injured person. Fractures generally heal more quickly in children than in adults (Galloway, 1999) and more quickly in trabecular bone than in cortical bone (Lovell, 1997). Stepped, longitudinal, and perpendicular fractures (acute angles and squared fracture edges) indicate postmortem breaks in nonvital bone.

The perimortem versus postmortem distinction is not as straightforward as it seems because the transition between vital and nonvital bone is gradual and many bone fragments will have more than one type of fracture, or some intermediate version of one of these types (Villa and Mahieu, 1991:34). However, fracture outlines do provide valuable information: a long bone with numerous perpendicular fractures as shown in Fig. 3.9 indicates postmortem breakage due to overburden. A bone reduced to long bone splinters with spiral fractures and irregular ends suggests perimortem fracture as a result of indirect trauma—as for marrow extraction. Cutmarks and

**Figure 3.8** Fracture types in long bone shafts. (Adapted from Marshall, 1989:14.)
perimortem fractures, or toolmarks and a combination of perimortem and postmortem fractures (Fig. 3.10) are clearly suggestive of processing.

Fracture products also characterize the bone status when the force was applied, as well as the kind of implement used in the fracturing process. These products include conchoidal scars, areas of crushing, incipient fracture cracks that propagate outward from a point of contact, percussion pits and associated striae, adhering flakes, and peeling in which small splinters of bone adhere the fracture edge (like a green twig). Fracture products are often associated with animal chewing damage and other percussive damage to bone. An example of crushing damage associated with perimortem damage to the posterior maxilla is shown in Fig. 3.11. This perimortem damage contrasts with the recent postmortem damage to the anterior maxilla, which is identifiable by the markedly lighter color of the newly exposed bone.

Fracture types and the timing of traumatic injury are essential aspects of the reconstruction of bone modification in an archaeological assemblage, but more importantly, they constitute physical evidence of the existence and nature of conflict in earlier human societies in a manner distinct from stories and texts (Walker, 2001). Perimortem injuries in mass-grave assemblages like Crow Creek tell the story of the violent deaths and postmortem mutilation and dismemberment of the victims: 40% had depressed cranial fractures, 90% had

Figure 3.9 Reconstructed tibia with postmortem fractures.

Figure 3.10 Pig tibia diaphysis with a cutmark, perimortem fractures, and chewing damage. (Copyright by the Field Museum, neg. #A113870.17, photographer John Weinstein.)
evidence of scalping, and 25% had evidence of decapitation as cutmarks on the occipital bone or the first or second cervical vertebrae (Willey, 1990:149). The record of antemortem trauma, lethal injuries, and postmortem treatment in the skeletons from Plazas A and C at the Pyramid of the Moon provides physical evidence of the treatment of sacrifice victims previously only known from Moche iconography (Verano, 2005). The victims exhibit old fractures as well as fractures in the initial stages of healing, documenting traumatic injuries suffered earlier in life and at the time of capture. Cutmarks on the cervical vertebrae indicate cause of death, and cutmarks at areas of muscle attachment indicate defleshing and dismemberment of Moche victims. Other bioarchaeological studies that combine paleopathology with forensic taphonomy have identified probable victims of domestic violence or have exemplified the existence of a social underclass as in the group of Ancestral Puebloan women from New Mexico’s La Plata Valley with markedly high rates of healed traumas and haphazard interment (Martin and Akins, 2001).

Thermal Alteration of Human Bone
Thermal alteration of bone results from post-mortem processing, including cremation, cooking for consumption, heating to facilitate dismemberment, and incidental contact with fire before or after death. This type of bone modification is of interest to archaeologists, forensic anthropologists, and zooarchaeologists as it relates to mortuary practices, medico-legal investigation, and animal food processing and discard patterns. Heat changes the structure of bone mineral crystals; cooked bone is fragmented more readily and is more difficult to identify (Lyman, 1994; Roberts et al., 2002; Stiner et al., 1995).

Thermal alteration is recorded as change in color and texture of bone. Color ranges from ivory or yellow to brown, reddish brown, black, gray-blue, and white (calcined) (Lyman, 1994:385). The color of thermally altered bone generally corresponds to the temperature of the heat source (e.g., Shipman et al., 1985:62), but color change is also dependent on the state of the bone when burned (Correia, 1997:279). Textural changes range from exfoliation and longitudinal cracking that mimics weathering (White, 1992) to warping and cracking in a checkerboard pattern (checking). Studies of the differential response of bone to thermal alteration when fleshed, defleshed but still vital, and nonvital (dry), seem to result in inconsistent findings (Correia, 1997:279), but experimental studies suggest that bone burned

Figure 3.11  Postmortem (recent) damage to the alveolar bone with recent loss of several anterior teeth, and perimortem damage to the posterior alveolar bone with adhering flakes (FMNH 43112).
when dry is not subject to the same kind of warping and longitudinal splitting characteristic of vital bone (Buikstra and Swegle, 1989, Lyman, 1994).

The pattern of burning on an articulated skeleton or a specific element is determined by the nature of heating (direct contact, steaming, boiling, wrapped, unwrapped, etc.), by the degree of articulation in the remains, and by the thickness of soft tissue covering of the bone (Buikstra and Swegle, 1989; Buikstra and Ubelaker, 1994). Determining whether bones were burned accidentally or intentionally, before or after other modifications, is dependent on the association of burning and butchery marks, the recovery context, and the part of the bone that is burned—the whole bone or high points only (Lyman, 1994). Color difference between the external surface and broken edges can reveal whether the bone was whole or already broken when burned.

Cannibalism in the Archaeological Record: Conceptual and Analytical Frameworks

Cannibalism is a term that covers a range of behaviors known in a wide range of species. Human cannibalism is typically divided into three main types: 1) starvation cannibalism—the Alfred Packer party and the stranded soccer players in the Andes, 2) gustatory or nutritional cannibalism—consumption of human flesh not as a starvation avoidance, and 3) ritual cannibalism—consumption of specific body parts not as a food source but as part of mortuary ritual. Gustatory and ritual cannibalism can each be motivated by positive or negative emotions about the deceased: consumption of a loved one or a powerful warrior can be a means of incorporating their essence into the living, whereas consumption of a feared or hated individual may constitute the ultimate desecration. Hypotheses about cannibalism in hominin remains from palentological and archaeological assemblages are not new (see White, 1992 and Turner and Turner, 1999 for extensive reviews), but today assemblages of human bone that are fragmentary, and exhibit toolmarks, burning, or non-normative mortuary treatment are hypothesized routinely to represent the victims of cannibalism. An assemblage of fragmentary human remains mixed with faunal remains is regarded as particularly suggestive of cannibalism.

The foundation of taphonomy studies of potentially cannibalized remains is the premise that if humans were eaten like animals, then they were likely to have been butchered and cooked the same way as animals, and disposed of the same way as refuse from game animals (Villa et al., 1986). “Cannibalism can be differentiated from all other forms of bone damage and mortuary practice by a distinctive signature which matches that seen in the bone refuse of large and small game animals” (Turner and Turner, 1999:2). Quantitative comparison of taphonomic profiles (percentage of bones with thermal alteration, toolmarks, etc.) in human and animal bone assemblages is the dominant analytical approach (e.g. Degusta, 1999; Defleur et al., 1999; Edgar and Sciulli, 2006; Turner and Turner, 1999; White, 1992). One practical limitation to this approach is that not all assemblages of human and faunal remains are comparable. You cannot compare butchery methods of humans and fish or shellfish or human and nonhuman butchery patterns when the nonhuman remains are primarily skulls or other non-economically relevant parts.

The second analytical approach used here is the comparison of the taphonomic profile of a potentially cannibalized assemblage with that created by local mortuary practice through analysis of modification to bones from primary or secondary burials or other contexts. This process involves taphonomic analysis of all human remains from a site, not just those from apparently nonfunerary contexts, and is particularly important in the study of prehistoric cultures with a broader range of mortuary practice. It is important to consider cannibalism as part of the bigger picture of ritual interaction with the dead.

It is suggested that cannibalized remains mixed with faunal remains represent gustatory
(nutritional) cannibalism, whereas human-only assemblages more likely represent ritual cannibalism (Fernandez-Jalvo et al., 1999). Presumably a mixed assemblage would not represent starvation cannibalism since other food resources were obviously present, but the distinction between gustatory and ritual cannibalism is not as clear—either behaviorally or archaeologically. Does such a distinction imply a lack of ritual in the butchery of game animals? It seems reasonable to assume that the processing and consumption of a human being would have been attended by some sort of ritual, even if the entire corpse were consumed delightedly. Goodale’s description of the preparation of the skeleton in the mortuary rites of the Kualong of West New Britain reveals the ritual component common to both human and pig processing in Melanesia:

A specialist is called in to exhume and prepare the bones. The skull and jaw bones are certainly removed and sometimes also the long bones of the arm, the scapula and collarbones, and their breastbones. (I believe it to be significant that these bones are the same as those in the butchered head portion of a pig). (Goodale, 1995:236, parentheses in original).

Macintyre writes about Tubetube Island, Milne Bay Province, Papua New Guinea, as follows:

The people of this region were cannibals and their definitions of humanity varied according to the political relationships prevailing. The substitution of pigs for people can also be seen as a reflection of the fact that in certain social contexts people were meat for ceremonial distribution (1984:114). . . . Pigs are living creatures that are like human beings, but are really human food. Enemies are living creatures that are like pigs in that they provide food, but are really human beings.” (Macintyre, 1984:117).

It may be argued that Melanesians’ relationships with pigs are not relevant to the interpretation of fragmentary skeletal assemblages, but this should serve to highlight the importance of human–animal relationships such as those signified in animal totems. We should not discount the ritual aspect of subsistence activities. Nor should we expect consumption of a person to have the same cultural meaning as consumption of an animal (Rautman and Fenton, 2005:333). The question of ritual is gratuitous if all we want to know is whether somebody ate somebody else. But is that all we want to know?

Cannibalism is the last topic to be addressed in this chapter because in the archaeological record these assemblages are part of the continuum of modified bone assemblages just as the several forms of cannibalism form a continuum. The suggested dichotomy between ritual and nonritual cannibalism and the implicit suggestion that this was an act potentially devoid of meaning seems to stem from our adoption of the explicitly economic, zoological, analytical framework without also employing a conceptual framework grounded in the scope of mortuary ritual. We need the methods of zooarchaeology and forensic taphonomy to conduct systematic analyses of these kinds of assemblages, but we should look beyond subsistence behavior for meaning and motive.

The limit of this economic analytical approach is revealed in the archaeological skeletal assemblages where there is convincing evidence that humans were processed, cooked, maybe eaten, and disposed of in a nonconsiderate manner, but where the human remains stand out in one very significant dimension: extreme fragmentation. Some of the sites attributed to cannibalism are argued to represent perimortem mutilation without consumption (Kantner, 1999), whereas other assemblages may be the result of both processes. The people that created these assemblages may or may not have used efficient means of butchery perfected in the processing of deer, but they proceeded with additional processing (more properly considered destruction), such that the human remains exhibit substantially more modification than comparable assemblages of animal remains (Billman et al., 2000; Edgar
and Sciulli, 2006; Lambert et al., 2000) These assemblages signify that something other than eating was on the prehistoric social agenda—revenge, terrorism, persecution, and mass executions. And because mortuary ritual has public, political, and performance aspects, Rautman and Fenton (2005:336) ask “what message extensive body fragmentation and commingling of remains might have been intended to convey to viewers?” The social and political climate is an important aspect of archaeological context since cannibalism does co-occur with interpersonal and institutionalized violence (Turner and Turner, 1999).

Cannibalism in Context

Perimortem modification of human and animal bone is not the only data that have to be considered in the analysis and interpretation of possibly cannibalized assemblages. As with other assemblages, the archaeological context is crucial in interpreting the modification patterns. First is the distinction between a burial context and nonburial context, which requires knowledge of the nature of primary and secondary (and possibly other) burial forms. Assemblages that include human and animal remains could represent sequential rather than simultaneous deposition, or they could be the result of postdepositional disturbance and mixing of distinct deposits. Associated animal and human remains also represent burials with animal offerings or remains of mortuary feasts. Do the faunal remains reflect a “typical” assemblage of food debris, or does the species or anatomical distribution of the faunal remains suggest something else? For example, a man might be buried with trophy skulls representing his prowess as a hunter.

Where were the bones? Were they found in an outdoor trash heap, an intramural pit feature filled with kitchen refuse, a hearth? Or were they found in a ritual context such as an Anasazi kiva floor (e.g., Malville, 1989) or a cave with bone deposits but no evidence of habitation or artifacts associated with daily life, like Ana Manuku in the Cook islands (Steadman et al., 2000)? These domestic versus ritual contexts have distinctive implications for the nature of the assemblage. What is the temporal relation of the deposit to the history of the structure or the site? Features or bone deposits associated with site abandonment have different implications. The mutilated and burned human remains at the Cowboy Wash site are related clearly to settlement abandonment; this was not an ordinary (ritual-free) event after which daily life of the community continued (Billman et al., 2000).

The “Taphonomic Signature”

Over the course of three decades of research on perimortem damage to human remains assemblages from the Four Corners region of the U.S. Southwest, Christy Turner (Flinn et al., 1976; Turner, 1961, 1983, 1989, 1993; Turner and Morris, 1970; Turner and Turner, 1990, 1995, 1999) developed a list of assemblage characteristics referred to as the taphonomic signature of cannibalism. The “minimal taphonomic signature” includes breakage, cutmarks, anvil abrasions, burning, many missing vertebrae, and pot polishing (as defined by White, 1992), which results from cooking in a ceramic pot (Turner and Turner, 1999:24). Additional characteristics listed by Turner are as follows: a single, short-term depositional episode is represented; good-to-excellent bone preservation; large numbers of bones and bone fragments (between 400 and 3500) per depositional episode; a high rate of missing elements including vertebrae; disarticulation of all or most body segments; massive perimortem breakage in 40–100% of elements; butchering marks on 1–5% of all elements; animal chewing marks on fewer than 5%; burning on about 20%; and cutmarks on about 3% (Turner and Turner, 1999:22–23). Assemblages with evidence of violence and trauma but not cannibalism exhibit fewer cutmarks, perimortem damage is focused on the head and face, and embedded projectile points or other weapons are more likely to be found with the bones (Turner and Turner, 1999:50–51). White’s study of the Mancos assemblage includes pot polish and other variables
especially useful in documenting fracture products and quantifying fragmentation and surface damage.

Although there are disagreements with Turner’s approach and with his interpretation of several specific assemblages (e.g., Ferguson et al., 2001; Kantner, 1999; Martin, 2000), most researchers use some version of Turner’s (1995, 1999) attribute list, White’s more detailed and exhaustively illustrated published protocols (1992), and the “processed like animals” model (Villa et al., 1986) to guide standardized data collection for the analysis of fragmentary skeletal assemblages and assessment of possible cannibalism. The data categories (listed in Table 3.5) include the ones discussed throughout this chapter and others beyond the current scope.

As more assemblages are studied with these methods, it becomes clear that substantial variability exists between supposedly cannibalized assemblages around the world, in the Southwest, and even between different assemblages from a single site (Hurlbut, 2000; Lambert et al., 2000). Within the general set of attributes, there is no exact quantitative threshold of cutmarks, burning, or other damage that can define cannibalized assemblages from all times and places, and there is a wide variation in the anatomical distribution of elements present in these assemblages. This is not surprising given the myriad behavioral and taphonomic variables at work in the creation of every assemblage. Although there are a limited number of ways in which we expect a human body to be processed (e.g., cutting muscles to detach the mandible), there is not one sequence of events that defines cannibalism. The extremes are represented by the systematic recovery of meat portions from the table.

### TABLE 3.5 Example of Variables in Taphonomy Data Collection Protocol Based on White (1992)

<table>
<thead>
<tr>
<th>Identifier</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identification</td>
<td>Chopmarks: number, location, orientation</td>
</tr>
<tr>
<td>Unique ID#</td>
<td>Chopmarks: number, location, orientation</td>
</tr>
<tr>
<td>Conjoin set#</td>
<td>Scrapemarks: number, location, orientation</td>
</tr>
<tr>
<td>Vertical provenience</td>
<td>Percussion striae: number, location, orientation</td>
</tr>
<tr>
<td>Horizontal provenience</td>
<td></td>
</tr>
<tr>
<td>Species</td>
<td></td>
</tr>
<tr>
<td>Element</td>
<td></td>
</tr>
<tr>
<td>Element portion</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td></td>
</tr>
<tr>
<td>Fragmentation/surface condition</td>
<td>Maximum dimension of bone/fragment</td>
</tr>
<tr>
<td></td>
<td>Maximum dimension of conjoined set</td>
</tr>
<tr>
<td></td>
<td>% of shaft circumference (score 0–5)</td>
</tr>
<tr>
<td></td>
<td>% of cortex remaining (score 0–3)</td>
</tr>
<tr>
<td>Weathering stage (score 0–6)</td>
<td></td>
</tr>
<tr>
<td>Rolling</td>
<td></td>
</tr>
<tr>
<td>Trampling/random striae</td>
<td></td>
</tr>
<tr>
<td>End polish/pot polish</td>
<td></td>
</tr>
<tr>
<td>Beveling</td>
<td></td>
</tr>
<tr>
<td>Animal chewing damage</td>
<td></td>
</tr>
<tr>
<td>Chewing damage</td>
<td></td>
</tr>
<tr>
<td>Rodent gnawing</td>
<td></td>
</tr>
<tr>
<td>Tooth punctures</td>
<td></td>
</tr>
<tr>
<td>Toolmarks</td>
<td></td>
</tr>
<tr>
<td>Cutmarks: number, location, orientation</td>
<td></td>
</tr>
</tbody>
</table>
bodies at the Alfred Packer site (see Fig. 3.5c) to the Anasazi assemblages where damage to bodies and body parts far exceeds anything related to meat and marrow acquisition, and the incidents clearly precipitated abandonment of habitation sites (Billman et al., 2000; Kuckleman et al., 2002; Lambert et al., 2000).

Inter-assemblage variation is also technological. Different cooking methods will have different taphonomic impacts. Obviously there will not be pot polish on bones from aceramic cultural contexts. Use of an earth oven to cook by steaming is likely to result in a different range of macroscopic thermal alteration than roasting above ground, and the degree of color change in bone may be variable depending on the degree of defleshing, nature of wrapping around the meat, and duration of cooking. Some earth oven-cooked remains, like those from Ana Manuku Cave in the Cook Islands, exhibit widespread thermal alteration (Steadman et al., 2000). But prolonged heating of well-wrapped bone that never comes in contact with coals or fire might result in minimal thermal alteration, as observed in human remains found in an earth oven in Quarańc Cave, Fiji (Piétrusewsky et al., 2003) or in thermal alteration that mimics weathering (White, 1992). The long cooking time required by earth ovens may result in the kinds of diagenetic changes produced in actualistic studies of long-term boiling of bone (Roberts et al., 2002), so evidence of this cooking that did not result in obvious color and texture changes could perhaps be identified histologically.

Applying Quantitative Taphonomy

Quantitative taphonomy provides a valuable tool for describing and understanding the history of assemblage, and for comparing the taphonomic profiles of assemblages created by different sets of agents and mechanisms. Table 3.6 lists some categories of taphonomic data for several assemblages of human and animal bone from sites in Fiji: Navatu (A.D. 50–1900) and Vunda (A.D. 800–1600) (DeGusta, 1999, 2000) and the North Coast of New Guinea-New Guinea Research Program Sites 23 and 46 (A.D. 700) (Stodder, 2005; Stodder and Rieth, 2003; Terrell and Welsch, 1997). Primary burials from Vunda, secondary burials from NGRP 23, and cannibalized human bone from the Navatu midden are listed in the first three columns from the left and animal bone in the other three columns. Assemblage sizes (NISP) range from 56 pig bone fragments from NGRP 46 to 1159 human bone fragments from the secondary burials at NGRP 23.

Looking at the three Navatu assemblages, note that the primary burials are unmodified except for cutmarks on one femur (DeGusta, 1999:228). Comparing the rates of modification between the cannibalized human bone and the mammal bone from Navatu indicates more burning in the animal bone and a higher incidence of peeled fractures in the human bone but similar or identical rates of crushing, cutmarks, and percussion pits. Anatomical representation of the human bone in the Navatu midden assemblage is different from the burials; the midden assemblage includes cranial vaults and arm, hand, and leg bones but no fragments of the innominates or sacrum or scapula (DeGusta, 1999:221). Comparing human and nonhuman assemblages from the New Guinea sites, the different taphonomic profiles in the small assemblage of pig bone and the human bone clearly distinguishes intensity of the modification between food remains and mortuary processing (Table 3.6).

Comparing the three human assemblages, burning, crushing, and cutmarks are most abundant in the cannibalized bone from the Navatu midden, but the secondary burials from Site 23 have the highest rate of peeling and percussion pits of the human assemblages. This is an assemblage of extremely fragmentary bone with under-representation of crania and upper limb bones, suggestive of bone curation similar to the practice documented in ethnographic literature for the Sepik Coast (Stodder, 2005).

Fragmentation of tubular bones (long bones, metacarpals, metatarsals, phalanges) in these
assemblages is shown in Table 3.7. The primary burials from Vunda have the most elements or fragments with complete shaft circumferences (cylinders) preserved. Shaft circumference preservation of the spatially isolated and more completely represented Burial 6 from Site 23 is substantially better than the mass secondary burial at this site and the animal bone from the three sites. This Site 23 sub-assemblage excluding Burial 6 was subject to substantial damage by pig chewing and trampling (Figs 3.4 and 3.9). Fragmentation is essentially equivalent in the cannibalized human and midden mammal assemblages from Navatu (as per the taphonomic signature).

Intrasite variation in preservation at NGRP 23 is documented further in the observations of weathering, rolling, and trampling (Table 3.8). Compared with other bone from this site, the Burial 6 bones exhibit less physical weathering, wear from sediment and rocks, and less damage from animal trampling. Several factors are thought to contribute to the better preservation of this set of remains. One scenario is that small assemblages of bone from the other six individuals (mixed ages and sexes) were

<table>
<thead>
<tr>
<th>Table 3.6 Taphonomic Alterations in Burials, Secondary Burials, Cannibalized Remains, and Mammal Remains from Fiji and New Guinea</th>
</tr>
</thead>
<tbody>
<tr>
<td>Site</td>
</tr>
<tr>
<td>Assemblage</td>
</tr>
<tr>
<td>NISP</td>
</tr>
<tr>
<td>Burning %</td>
</tr>
<tr>
<td>Crushing %</td>
</tr>
<tr>
<td>Cut marks %</td>
</tr>
<tr>
<td>Peeling %</td>
</tr>
<tr>
<td>Percussion pits %</td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th>Table 3.7 Preservation of Tubular Bone Shaft Circumference in Human and Nonhuman Bone Assemblages from Fiji and New Guinea</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assemblage</td>
</tr>
<tr>
<td>Vunda primary burials</td>
</tr>
<tr>
<td>NGRP 23 secondary burial Burial 6</td>
</tr>
<tr>
<td>Navatu mammal</td>
</tr>
<tr>
<td>Navatu midden human*</td>
</tr>
<tr>
<td>Vunda mammal</td>
</tr>
<tr>
<td>NGRP 23 mass secondary burial excludes Burial 6</td>
</tr>
<tr>
<td>NGRP 46 pig</td>
</tr>
</tbody>
</table>

brought to this hilltop rock crevice after Burial 6 (an adult male) had been placed there. Ethnographic and archaeological evidence suggest that primary inhumation is not likely to ever have been practiced on the North Sepik Coast before the 1900s (Stodder and Rieth, 2003). The remains from this site, mixed with shell, fish bones, and animal bone, were thought to have been cannibalized, but the taphonomic data do not support such an interpretation. Although there is evidence of processing in the human bones, much of the fragmentation is from animal-induced damage, and the animal bones in the deposit turned out to be virtually all pig mandibles rather than meat-bearing parts. It is suggested that these remains represent two stages of multistage mortuary program. Without the quantitative information generated from the systematic taphonomy analysis, this would have been an uninterpretable assemblage of fragmentary bone.

### STUDYING MUSEUM COLLECTIONS

There are some special considerations to be aware of when studying museum collections of human remains from archaeological sites. The life history of an assemblage continues in the museum or laboratory, and one hopes that the documentation of its previous life is preserved there as well. Some researchers use information provided in the museum catalog for dating, cultural affiliation, and other background even if that information is 100 years old. This can be a problem if the site has been re-excavated, re-dated, reinterpreted as to cultural affiliation, or renamed since the materials were accessioned and cataloged. The validity of museum-based research will be undermined if the biological anthropologist does not take the time to learn the correct site name or cultural affiliation or confuses archaeological dating notations like B.C. and B.P. It seems obvious that one should read everything already written about a skeletal assemblage and about the site it came from, but unfortunately this is not always done, and thus, some publications list sites with blatantly incorrect information.

It may also be important to examine excavation notes and museum registration records. Trades between museums and excavations at the same site by different institutions can mean that one is examining only part of the material recovered from a site. Field notes may reveal that a collection includes only a portion of the human remains encountered in excavation. Some remains are left in situ, and since for most of the history of physical anthropology interest was in well-preserved adult crania and examples of pathological specimens, many skeletal remains were not collected. Excavation records can provide a truer picture of the age distribution of people buried at a site. As demonstrated in Fig. 3.12, field records from excavations at San Cristobal Pueblo, a late prehistoric Pueblo in New Mexico, reveals a different age distribution—specifically, more subadults—than the age distribution derived from laboratory analysis of the curated collection (Stodder, 1990).

Familiarity with the history of conservation of a collection is advantageous as well, since consolidants have different effects on bone over time and may interfere with radiocarbon dating, stable isotope and trace element analysis, DNA, retrieval, and macroscopic and SEM

### TABLE 3.8 Intra–site Differences in Assemblage Weathering and Wear at Site NGRP 23

<table>
<thead>
<tr>
<th></th>
<th>Burial 6</th>
<th>All Other Human Bone</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean weathering score</td>
<td>0.54</td>
<td>0.87</td>
</tr>
<tr>
<td>(Behrensmeyer, 1981)</td>
<td>N = 121 SD = 0.85</td>
<td>N = 374 SD = 1.01</td>
</tr>
<tr>
<td>Rolling, rounding</td>
<td>45/116 8.79%</td>
<td>227/343 60.86%</td>
</tr>
<tr>
<td>Trampling, striae</td>
<td>7/131 5.34%</td>
<td>65/373 17.43%</td>
</tr>
</tbody>
</table>

*Difference is significant at 0.05 or better.*
analysis of surface modifications (Johnson, 1994:228; Millard 2001). It is common to find teeth glued into the wrong sockets, sand embedded in consolidants, and other well-intentioned repairs that can impede your research, but some modification can be misinterpreted as prehistoric cultural modification like the scratches on the Engis skull from the sharp tips on the craniometer (White and Toth, 1989). Conservation may aid or interfere with research objectives (Johnson, 1994; Kres and Lovell, 1995; Panagiaris, 2001), so it is hoped that conservators work closely with analysts in planning conservation for skeletal assemblages.

It can be difficult to find out what collections exist at museums, as portrayed in Whittington’s (2001) description of the task of locating human remains from Maya sites in the Guatemalan Highlands. Depending on the level of detail in museum records and the kind of information they are willing to provide, it can also be very difficult to obtain sufficient information on the completeness and preservation of a collection to plan a research project. There are several logistical and political reasons for these difficulties, but it is essential to meet with museum staff and to carefully examine a collection in person, at the earliest possible point in a research program. Although funding is always an issue, access may be an even bigger hurdle since some museums now require researchers to obtain permission from federally recognized Tribes or other groups with repatriation claims even when the collections remain in the museum.

Museums routinely request that visiting researchers send a copy of their published results, which can be expensive if you have to buy another copy of a book, but it really is important since your research is part of the assemblage life history. The possibility of repatriation must be acknowledged for all human skeletal remains, and this brings a new responsibility to each researcher to use standard methods and to document original innovations in methods so that when data are all that remain, others will be able to build on the work you have done.

CONCLUSION

The goal of this chapter has been to demonstrate that each archaeological assemblage of human remains is a unique, invaluable resource but not one to be studied in isolation. Context is critical—minimally to the point that you locate the assemblage in space and time! Skeletal assemblages are created and modified in many ways from the time of their deposition until
the day they are studied, and that history is recorded in the assemblage composition and the taphonomic profile. Taphonomy provides the framework for documenting these changes and for using these data as information about the past rather than simply something to work around. The unusual age distribution of a cemetery may very well be the result of deliberate mortuary behavior, not just the results of sampling error to remedy with palaeodemographic modeling. Implicit in this is the importance of understanding past cultures and the archaeological site as a record of some fraction of human behavior. Taphonomy restores the link between the site formation processes and the laboratory analysis of the human skeleton. We will certainly maximize our understanding of the past if we study the entire history of skeletal assemblages.

ACKNOWLEDGMENTS

The assistance of these mentors and colleagues is gratefully acknowledged: John Edward Terrell, Diane Gifford-Gonzales, Susan Antón, Bettina Arnold, Michelle Bonogofsky, Michele T. Douglas, Heather Edgar, David DeGusta, Michael Pietrusewsky, and Timothy Rieth. The drawings in Figs 3.4 and 3.9 are by Eric Wert, and other illustrations were improved greatly by the work of Christine Sulok, Sulok Graphics, Inc.

REFERENCES

REFERENCES


Roksandic M. 2002. Position of skeletal remains as a key to understanding mortuary behavior.


Stodder ALW. 2006. The taphonomy of cranial modification in Papua New Guinea: implications for the archaeology of mortuary ritual British archaeological reports, international series. In: Bonogofsky M, editors. Skull Collection,


PART II

MORPHOLOGICAL AND DEVELOPMENTAL ANALYSES
Until the twentieth century, juvenile mortality was high in most human populations, but juvenile skeletons are often hard to find in archaeological excavations of cemeteries. Estimating sex in juvenile skeletons is also a difficult task, but estimation of age at death can be very specific because of the nature of developmental changes in skeletal and dental hard tissues. This chapter addresses some of the apparent problems and contradictions that bioarchaeological researchers encounter when studying juvenile mortality samples.

**INTRODUCTION**

In 1968, Johnston was justified in claiming that physical anthropologists studying the skeletal biology of earlier human populations concentrated on adults and excluded infants and children from their research (Johnston, 1968). This neglect is even more surprising since typical mortality curves for nonindustrial and pre-twentieth-century groups graphically portray the precariousness of the growth period with its great nutritional requirements and susceptibility to disease (Lewis, 2007; Saunders and Barrans, 1999). Since the 1960s, and in large part influenced by Johnston’s early efforts on the Indian Knoll skeletal sample from Kentucky (Johnston, 1961, 1962, 1969), there have been several attempts to redress the oversight and look for prehistoric population differences in juvenile mortality, child growth, and child development.

The assessment of general population health has also become a fundamental part of the interpretation of the life ways of extinct or past populations with a shifting focus on multiple rather than single indicators of physiologic stress (Buikstra and Cook, 1980; Goodman and Martin, 2002; Goodman et al., 1988; Larsen, 1997; Mensforth et al., 1978; Ubelaker et al., 1995; Verano and Ubelaker, 1992). General confidence exists that it is now possible to test hypotheses and assess the relationship between a population and its environment from the study of cemetery skeletal remains (Goodman and Martin, 2002; Johnston and Zimmerman, 1989). But the question remains, is this confidence justified, especially when trying to study juvenile mortality?

This chapter examines in some detail the practical problems of analyzing the skeletons of immature individuals from archaeological samples and whether these problems are surmountable. It addresses questions of sampling, the estimation of sex, and the estimation of age.
The balance of the chapter focuses on the assessment of growth and development using population comparisons, studies of growth rates, and the influence of environmental factors on growth. The possible detection of health stress and/or cause of death of juveniles are broad fields that are not included here because of space limitations. Readers wishing more information should consult relevant chapters in this book (Lovell, and Milner et al., this volume) and the following sources: (Larsen, 1997; Lewis, 2007; Ortner, 1998; Ortner and Mays, 1998; Saunders and Barrans, 1999; Schultz, 1989). It is clear that certain theoretical and methodological difficulties are still intractable. Thus, what is possible? This chapter concludes with recommendations for future research.

**SAMPLING**

First in any study of skeletal samples is the evaluation of bone preservation. Counts of juvenile skeletons from prehistoric and historic cemeteries are often very low, introducing a bias into skeletal samples (Guy et al., 1997; Jackes, 1992; Lewis, 2002; Saunders and Katzenberg, 1992). Several authors have identified several factors as responsible for the under-representation of immature individuals. These factors are cultural beliefs about infants and children influencing mortuary behavior, the effects of biological (intrinsic) and environmental (extrinsic) processes causing differential preservation of immature bones, and incomplete archaeological recovery because of biased excavation techniques.1

Considerable evidence exists for differential burial practices that will frequently bias against infants and children or alter the proportions of subadult and adult skeletons from a cemetery. The practice of infanticide has been and still is widespread and relatively common among many human cultures (Lewis, 2007; Scrimshaw, 1984). Both the deliberate killing of babies and “passive infanticide” in the form of neglect, by increasing the risks to infant survival, decrease the likelihood that some deceased infants will receive formal cemetery burial. In addition, the definition of life after birth is usually dependent on a cultural definition of when life begins. Some groups acknowledge infant life several days after birth, whereas others do not consider children fully human for several years (Saunders and Barrans, 1999). Lewis (2007) has recently summarized several examples where biological anthropologists have interpreted archaeological discoveries of concentrations of neonatal burials as cases of infanticide. She points out that the evidence is equivocal since results do not meet expectations. Ancient DNA studies do not confirm the prevalence of female deaths at some of these sites, which would be expected if, as is common, males are more valued (although mortality samples may have more males because of their physiological susceptibility to mortality). In addition, observations of unusually high concentrations of newborn skeletons may be affected by the aging methods used, and the interpretation of the absence of the neonatal line in tooth enamel as indicating infanticide shortly after birth ignores the fact that burials may be normal stillbirths. The archaeological context of burials is still important when making inferences about infanticide (Smith and Kahila, 1992).

Infants who die of natural causes are frequently buried far away from cemeteries, in house floors, and entryways or in other contexts. Taran is a well-known term in Gaelic referring to the ghosts of unbaptized babies buried outside of normal cemeteries. Ethnohistoric accounts by the Jesuits of the burial practices of seventeenth-century Iroquoians of southern

---

1 Others divide all influences on preservation into just two categories, intrinsic and extrinsic. Intrinsic factors refer to the resistance of the element (e.g., bone) to postmortem damage based on its biomechanical properties, and extrinsic factors refer to the geographic and geologic environment of a site, the nature of the local flora and fauna, and the activities of humans (which includes human cultural choices) (Henderson, 1987).
Ontario refer to the burial of infants along pathways so that their souls might reenter the womb of a passing woman (Thwaites, 1896). Modern excavations of Iroquoian village sites have uncovered high proportions of newborn infants buried in long house floors (Saunders and Spence, 1986; Spence, 1986), and it seems that the choice of the central corridor of the long house often represented the “path” along which an appropriate woman might walk (Saunders and Fitzgerald, 1988).

It is also believed widely that the bones of infants and children, because they are small and fragile, do not preserve and therefore are often lost to the excavator (Johnston and Zimmerman, 1989). Some years ago Gordon and Buikstra (1981) found that soil acidity levels are correlated significantly with bone preservation in both adult and children’s skeletons, with juvenile bone durability declining rapidly with decreasing soil pH. Certainly intrinsically, the bones of young individuals have high organic and low mineral content and are less dense than the bones of adults, which makes them more susceptible to decay (Currey and Butler, 1975; Specker et al., 1987). Research on living individuals has shown that the bone mineral density or BMD (the mass of inorganic mineral matter per unit volume) in a normal infant skeleton decreases in the first year after birth and then increases through childhood and adolescence (Rauch and Schoenau, 2001). However, lower than normal bone density can occur because of problems of maternal health as well as of several other pathological factors after birth (see Stodd, this volume). The question is whether the bioarchaeologist can devise investigations that could detect such phenomena in skeletal mortality samples.

Some researchers have compared observed skeletal samples with expectations from documentary data to judge relative preservation of different sex and age groups. Walker et al. (1988) examined mortality profiles derived from an analysis of burial records and skeletal collections from a nineteenth-century Franciscan mission cemetery in California. The baptismal and death records of the mission indicated that most people buried in the cemetery were either infants or elderly adults, which fits the typical U-shaped distribution of mortality observed in earlier populations. However, the skeletal sample contains mainly young adults. Even though the skeletal sample size is small (only 2% of all people buried in the cemetery), a random sample of burials should not deviate so much from the known age distribution of people buried in the cemetery. The authors concluded that age-specific differences in preservation account for the missing children and elderly adults. Later, Guy et al. (1997) claimed that the proportion of infants (under one year) in archaeological cemeteries generally fluctuates around 5% or 6%, whereas the proportion of infants dying in pre-vaccination populations recorded by historical demographers should never fall below 25% of live births. They identify taphonomic variability (susceptible bones in poor burial environments) as the main cause of low percentages of juvenile bones in skeletal samples, arguing therefore that paleodemographic reconstruction is restricted uniformly. The same explanation was offered by Jones and Ubelaker (2001) who compared an excavated skeletal sample from a nineteenth-century Swiss–German church cemetery in Pennsylvania to parish registers listing the burials. Of those burials listed in the registers as infants or children less than five years of age, only 50% preserved as skeletal remains.

On the other hand, others have argued (Lewis, 2007; Sundick, 1978) that even the smallest growth centers and the most fragile of the flat bones of the vault of the skull or the scapula of subadult skeletons can be as well preserved as the bones of the robust adult skeleton. Some attribute incompleteness to the degree of skill of the excavators and the need to be able to recognize juvenile bones. Few would care to admit that excavation methods at a site had been less than perfect, but obviously deficient
training may contribute to inadequate recovery of skeletal remains.

More recently, Bello et al. (2006) have proposed a new method for evaluating bone preservation that takes into account the different aspects of what represents “well-preserved,” “present,” and “complete.” Skeletons from an excavation may be missing whole elements, some of the elements that are present may be broken, and the cortical surfaces of intact or broken bones may be damaged. These authors evaluated preservation in three different ways. They recorded an anatomical preservation index (the percentage of bones preserved out of the total anatomical number of bones of the skeleton per individual) and a bone representation index (the ratio of the actual number of bones recovered at excavation to the total number of elements of the skeleton expected). Lastly, they calculated a qualitative bone index or the ratio between sound cortical surfaces and damaged cortical surfaces for each bone. They evaluated skeletal samples from three historical cemeteries for these indices. They found juvenile bones to be more poorly preserved and less well represented than adult bones, especially for the 0–4-year group. They also observed that female immature skeletons did not survive as well as male immature skeletons in the Christ Church, Spitalfields sample (London, U.K.). An evaluation of extrinsic factors such as burial location within a site and burial depth (graves of infants and children tend to be shallower) found no evidence that these factors strongly influenced the state of preservation of the bones from these three sites. However, intrinsic anatomical properties of the bones themselves (i.e., bone mineral density) do seem to be important (as observed by the negative correlation between preservation and age), and extrinsic factors, in some situations, may serve to exacerbate this general phenomenon. But these authors did not present general demographic statistics to evaluate the overall proportions of juveniles in their samples compared with adults. This comparison can help to detect human influences such as differential mortuary practices for infants and children.

Thus, we should expect that explanations for juvenile sample bias in skeletal collections are diverse, multifactorial, and often unique to the cultural and archaeological context of the site (see, for example, Hoppa and Gruspier, 1996; Hoppa and Saunders, 1998; Nawrocki, 1995; Paine and Harpending, 1998; Saunders, 1992). On the other hand, Guy et al. (1997) reject cultural explanations for the lack of infant skeletons, citing an example from an historic French cemetery where the proportion of infants represents a fifth of the recorded proportion in the parish registers. They note that this population punished infanticide and practiced the baptism of neonates, presuming then that most infant skeletons should have been found. Yet, several examples exist from historic cemeteries where it is clear that spatial locations of burials are age-specific (Lilley et al., 1994; McWhirr et al., 1982; Molleson and Cox, 1993) or where substantial proportions of immature individuals, even infants, are present (Cook and Buikstra, 1979; Farwell and Molleson, 1993; Hutchins, 1998; Lewis, 2007; Owsley and Jantz, 1985; Saunders et al., 1993; Saunders et al., 1995a; Sperduti et al., 1997).
Almost ten years of research conducted on a large nineteenth-century church cemetery sample from Canada, where skeletal observations could be compared with a complete set of burial records as well as other documents, has produced some interesting data on juvenile skeletal preservation (Saunders et al., 1995a; Saunders et al., 1995b) (Fig. 4.1). This work has shown that a complex of factors explains the relative proportions of juveniles and infants in the skeletal sample when compared with documentary information. In 1989, archaeologists excavated 597 burials at St. Thomas’ Anglican Church cemetery in Belleville, Ontario (Herring et al., 1991; McKillop et al., 1989). Burials took place between 1821 and 1874, and the 1989 excavation recovered 40% of all interments as recorded in parish registers covering the 53-year cemetery period. The estimated proportion of all sub-adults to adults in the skeletal sample matches closely with adult and sub-adult burials recorded in the registers (Herring et al., 1991). However, the proportion of infants (under one year, the appropriate demographic definition) in the skeletal sample significantly exceeds that recorded in the registers over the full period of cemetery use, contrary to widely held expectations. Initially, we interpreted this difference as a temporal bias in the archaeological excavation of the site. During the 1820s and 1830s, many of the region’s pioneer inhabitants avoided the distance required to travel to town and buried deceased babies in family plots. The proportions of infant burials increased over the decades as the population grew (a clear illustration of how increasing fertility rates can increase the number of infants in a cemetery; see Milner et al., this volume). It seemed possible that the cemetery excavation had disproportionately recovered later period burials containing more infants although the excavated burials were those closest to the church structure.2

Additional work on the St. Thomas’ sample has shown that although the quality of record keeping in the parish registers is very good, some infant deaths may be under-reported, which is a common phenomenon in other past documents (Lynch et al., 1985), since estimates of infant mortality rates are lower than what is typically reported for the time period. However, other cultural reasons may account for the presence of surreptitious infant burials whose interment was hidden from the minister. The local newspaper provides accounts of probable infanticide cases (Saunders et al., 1995b), and other historical documents of the region refer to concealed infant burials in church cemeteries by those avoiding the payment for a plot (Daechsel, personal communication).

Were the juvenile skeletons in this sample more poorly preserved than the adult skeletons, as might be expected? One way to evaluate this is to compare the proportion of preserved age indicators on nonadult and adult skeletons since other skeletal markers of disease or stress cannot be evaluated until age at death is estimated. Table 4.1 shows that in fact, a similar or even better proportion of the juveniles (individuals aged under 16 years although most are under 6 years) in the St. Thomas’ sample preserved usable age indicators. In 98% of the juveniles, dental, diaphyseal, or both indicators were available for age estimation. Overall conditions of most bones from the site were “good” to “excellent” because of the well-drained, sandy soil that improved the situation for smaller, less-mineralized bones. Excavators took special care to search for and to recover developing tooth germs and bone growth centers. In another example, Spence (1986) anticipated the discovery of fetal skeletons from within an early Iroquoian village site and arranged to have all maxillae and mandibles prewrapped in the field to preserve fetal tooth germs and, thus, estimate fetal ages more accurately. These examples suggest that factors such as differential burial practices and excavation techniques can sometimes prove just as important to juvenile skeletal preservation as differential tissue survival.

2No map of the burials was found in the original church documents.
SEX ESTIMATION

One universal problem presents itself with the study of subadult skeletons, the apparent difficulty of assigning sex (Workshop of European Anthropologists, 1980). Knowing the sex of the individual has a bearing on estimation of age because of sex variability in growth. Generally, females grow faster and mature earlier than males, meaning that non-sex-specific age estimates of immature skeletons will be much broader than they might be. Yet, knowledge of the sex of the individual is important to anthropological inquiry since it provides the opportunity of studying gender-biased care and status differences among children as well as sex differences in disease susceptibility, sex biases in morbidity and mortality, and a host of cultural factors associated with sex and the individual’s role in society (Lewis, 2007; Stini, 1985).

The differentiation between the sexes is in large part determined by the testes, because if cells are not masculinized by the presence of androgen, they will develop along ovarian lines (Haseltine and Ohno, 1981). In male infants, circulating levels of luteinizing hormone (LH), follicle-stimulating hormone (FSH), and the secretion of testosterone begin to rise during early prenatal life and decline again just before birth. After birth, total testosterone levels rise again and reach peak values that approach the low normal adult male range between 1 and 3 months of age. They then fall to normal juvenile levels by 6 to 8 months of age and remain low until puberty (Mann and Fraser, 1996). Probably as a consequence of these early hormonal differences, male newborns have greater muscle mass and higher average birth weights than females (Malina and Bouchard, 1991). It has been proposed, therefore, that sexual dimorphism in infant bones and teeth may arise in response to sex-specific hormones detected by receptors at susceptible skeletal sites (Loth and Henneberg, 2001).

---

#### Figure 4.1
Example of an adult burial from the St. Thomas’ Anglican Church Cemetery. This case is the only one out of almost 300 adults in which a postmortem autopsy was conducted.

#### TABLE 4.1 Frequency of Preserved Age Estimation Indicators for Subadult Skeletons and Adult Skeletons from the St. Thomas’ Anglican Church Sample

<table>
<thead>
<tr>
<th>Age</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subadults</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dental and/or Diaphyseal data</td>
<td>275/281</td>
<td>97.9</td>
</tr>
<tr>
<td>Dental formation data</td>
<td>240/281</td>
<td>85.4</td>
</tr>
<tr>
<td>Diaphyseal data</td>
<td>242/281</td>
<td>86.1</td>
</tr>
<tr>
<td>No information for age</td>
<td>6/281</td>
<td>2.1</td>
</tr>
<tr>
<td>Adults</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Auricular surface data</td>
<td>238/278</td>
<td>85.6</td>
</tr>
<tr>
<td>Pubic symphysis data</td>
<td>171/278</td>
<td>61.5</td>
</tr>
<tr>
<td>Intact cranium (analysable)</td>
<td>232/278</td>
<td>83.5</td>
</tr>
<tr>
<td>Infra-cranial measurements</td>
<td>255/278</td>
<td>91.7</td>
</tr>
</tbody>
</table>

---

Multiple loci are likely involved in the ultimate genetic control of sex differentiation. Endocrine function then contributes to producing human sexual dimorphism. Before the onset of puberty, male infants and children are, on average, larger than females for such features as head for weight index (Ounsted et al., 1981), bone thickness, and bone density (Mazess and Cameron, 1972; Specker et al., 1987). But, in addition, from the twentieth week of life in utero, the female fetus is approximately 10% more mature than the male and this growth difference will persist until the attainment of full maturity (Stini, 1985). It is likely that a complex of factors influences this difference in maturation rate between the sexes.

Because of the sex difference in maturation, the sex of subadolescent human skeletons might be inferred by comparing dental development with bone development in the same individual, since males mature more slowly skeletally than females, whereas the rate of dental formation is more similar (Hunt and Gleiser, 1955). The procedure is to estimate age independently from dental and skeletal remains using both male and female standards. If the dental and skeletal age estimates from male standards were closer to one another than for female standards, then the unknown individual would be male. A test of this method using dental and skeletal radiographs from living children obtained accuracy levels of 73–81% (Gleiser and Hunt, 1955). In this study, skeletal ages were based on hand bone maturation, which is an impractical method for excavated skeletons, but the authors suggest using bone formation at the knee joint. Sundick (1977) reported success with this approach in skeletons of individuals over 12 years where sex was confirmed by pelvic morphology and skeletal age was estimated from diaphyseal lengths, but accuracy rates were not reported. Later, Goode-Null (1996) achieved high accuracy by comparing skeletal development at the knee with a generalized (not sex specific) estimate of dental age in a documented sample of fetal and infant remains. Although questions arise about the influences of population variation and pathological stress on the effectiveness of this method, the reports of success suggest it warrants additional testing.

Applying another approach, Choi and Trotter (1970) used long bone weight and length ratios to produce a sex classification accuracy of 72% for fetal skeletons, but the use of bone weights makes this method currently inapplicable for exhumed bones. Furthermore, the classification accuracy is low; a minimum criterion for classification accuracy should be 50% better than chance or at least 75% (DeVito and Saunders, 1990).

Since the pelvis is the most sexually dimorphic part of the adult skeleton, it would make sense to look for dimorphism in the subadult pelvis and there is a long literature on this subject (Boucher, 1955, 1957; Kelly and Reynolds, 1947; Morton, 1942; Morton and Hayden, 1941; Reynolds, 1945; Thomson, 1899). These authors say that infant males have longer ilia, ischia, and femoral necks, and females have longer pubic bones and wider greater sciatic notches. Weaver (1980) reevaluated some of the earlier metric methods using a large fetal dry bone sample of known sex but found almost no significant sex differences for measurement indices. Schutkowski (1987) used the raw data from hip and femur dimensions of a sample of fetuses and neonates collected by Fazekas and Kósa (1978) to examine sex differences. He derived discriminant functions from these measurements, but the maximum classification accuracy was only 70%. Another application of his calculated functions to two samples of known age and sex was unsuccessful (Majó et al., 1993).

Weaver (1980, 1986) proposed that subjectively recorded variation in the height of the iliac auricular surface is sex dependent in fetal and infant skeletons. He tested this nonmetric trait on a known sample and obtained a classification accuracy of 43% to 75% for females and 73% to 92% for males. In an attempt to test the reliability of the trait, Hunt (1990) compared the ratio of raised with nonraised auricular surfaces with a 1 : 1 expected sex distribution in a large
sample of subadult ilia from several archaeological sites (where sex was unknown). Although burial practices and other factors could alter an expected 1:1 infant sex ratio in archaeological samples, Hunt found severely unrealistic sex ratios, including an age-related shift from newborns to young adolescents. A later study of a small known sex sample (Mittler and Sheridan, 1992) also identified an age influence on this trait. In addition, Holcomb and Konigsberg (1995) explored sciatic notch shape in a large fetal sample using an objective morphometric approach. Although they did find significant sexual dimorphism in the shape of the notch, overlap between males and females was too great for the notch to be useful as a sex indicator at early ages.

Recently, eight previously proposed, morphological traits, said to be sex-related in the juvenile skeleton, were tested on a sample of known-sex mummies from Chile (Sutter, 2003). This author found that all traits showed a significant relationship to sex, but only four of them produced minimally acceptable levels of accuracy for estimating sex, whereas others only performed slightly better than chance. Of course, it is not clear whether the useful traits would work equally as well in another sample from another population, but they offer some prospect for those wishing to estimate sex ratios of juvenile subgroups in archaeological samples without too much technical effort (especially since this was a blind test conducted by someone who did not evaluate the original reference samples). Nevertheless, readers need to beware. As Klepinger (2006) has pointed out, a major source of problems is that many workers have difficulties perceiving and scoring the standards for traits established by others, even when pictures are provided! Researchers need to conduct more blind tests of the so-called useful traits by studying samples of known sex, which unfortunately, are hard to come by.

The fact that significant sexual dimorphism occurs in the permanent dentition has prompted the claim that for children, the teeth might represent the only factor useful for sex diagnosis (Workshop of European Anthropologists, 1980). However, the magnitudes of dimorphism are small and most data come from permanent teeth that emerge only later in childhood. Several studies have determined that a small but significant dimorphism does exist in the deciduous dentition, but only two studies have employed classificatory procedures for separating the sexes. One of these studies (Black, 1978), concluded that deciduous teeth show much less dimorphism than permanent teeth and that discriminant functions calculated from the diameters of deciduous teeth are much less accurate for sex classification. DeVito and Saunders (1990), using three to five measurements of deciduous teeth as well as combinations of deciduous and permanent measurements, produced discriminant functions in which 76% to 90% of holdout samples were classified correctly by sex, which means that the level of classification accuracy of the deciduous teeth at least approaches the levels achieved using the permanent teeth. However, the pattern and degree of sexual dimorphism reported for various groups shows considerable population variation.

DeVito and Saunders (1990), using three to five measurements of deciduous teeth as well as combinations of deciduous and permanent measurements, produced discriminant functions in which 76% to 90% of holdout samples were classified correctly by sex, which means that the level of classification accuracy of the deciduous teeth at least approaches the levels achieved using the permanent teeth. However, the pattern and degree of sexual dimorphism reported for various groups shows considerable population variation.

The author conducted a test of these discriminant function equations using a small sample of personally identified skeletons of children from two historic nineteenth-century pioneer cemeteries (Saunders, 1992). Of the seven females represented, six were correctly classified, but of the eight males, only three were correctly classified. The teeth of the individuals from the archaeological samples are smaller than the living Canadian reference sample so that males are identified as females, which suggests an interesting illustration of biased male mortality in this sample (Wood et al., 1992).

4The tested traits included those previously proposed by Weaver (1980) and Schutkowski (1993).

5Recognition of this limitation in modern populations raises questions about paleoanthropologists’ ability to estimate the sex of juvenile fossil hominins as was done recently for A. afarensis from Ethiopia (Alemseged et al., 2006).
Recently, we have examined sexual dimorphism in permanent mandibular canines and premolars of the St. Thomas’ skeletal sample using area measurements of the enamel and dentine-pulp core taken from longitudinal thin sections (Saunders et al., 2007). The relative dentine-pulp area of the teeth displayed high levels of sexual dimorphism as well as statistically significant mean differences between the sexes. The male canines and premolars have significantly more dentine than their female counterparts as well as relatively more dentine with respect to overall crown size. The female canines and premolars have significantly more enamel relative to overall crown area than those of the males. These results suggest that relative area measures of crown tissues are more predictable measures of sexual dimorphism than absolute measures, and tissue proportions may remain constant despite intra-sex variation in overall tooth crown size.

Hopes have sometimes been expressed for methods of chemical or elemental identification of sex from the skeleton (Beattie, 1982; Dennison, 1979; Gibbs, 1991; Lengyel, 1968), but sex differences in bone chemical composition usually are dependent on post-pubertal physiologic differences and can be altered dramatically by the burial environment. The end of the long arm of the Y chromosome is visible with special staining in living cells undergoing division. This method of sex determination has been attempted in forensic cases (Mudd, 1984; Sundick, 1985), but human cells do not remain patent for long and investigators may identify artifacts as stained chromosomes.

With the advent of molecular techniques for sequencing human cellular DNA, sex can be defined at the molecular level because the X and Y chromosomes have their own distinctive DNA sequences. Unlike direct microscopic observations of the chromosomes, molecular analysis does not require that the DNA be in a viable condition as long as DNA molecules are chemically available. Now that several researchers have successfully determined sex in archaeological skeletons by extracting DNA (Faerman et al., 1995; Hummel and Herrmann, 1994; Stone et al., 1996; Yang et al., 1998, see also Stone, this volume), the problem of juvenile sex classification seems solved. However, aDNA (ancient DNA) analysis is far from foolproof, and the methodology is not readily available to bioarchaeologists. The cost and time involved in testing large samples is still prohibitive, and problems with extraction and contamination of archaeological specimens plague researchers. Significant proportions of archaeological samples do not yield any amplifiable DNA, and for those samples that do, several independent tests are required to avoid the high likelihood of false negatives and false positives. For juvenile skeletons, the special problem is that there are no reliable morphological or metric methods to compare with DNA results, as with adults. Nevertheless, we should be optimistic. Sex estimation by DNA analysis offers some of the greatest potential for future research because laboratory techniques in molecular biology continue to make great strides.

AGE ESTIMATION

Most critical, of course, to the identification of immature individuals from skeletal samples is the problem of age estimation. Juvenile age at death estimations are more accurate than adult age estimations because of the telescoped time span of human growth relative to the total life span over which age variability is assessed. However, difficulties with estimating sex in juveniles increase the range of error. Age estimation of the skeleton involves establishing physiologic age (developmental changes in the tissues) and then attempting to correlate this with chronological age at death. Additional sources of error besides the sex difference contribute to the discrepancy between physiologic and chronological age. These sources include 1) random individual

6Although see FitzGerald and Rose (this volume) for an explanation of the odontochronological method of age estimation.
variation in maturation and 2) the systematic effects of environmental and genetic factors on growth. Tooth emergence, a piercing of the gum or alveolar bone by the developing tooth, has been studied extensively and used widely in archaeological and forensic efforts to estimate age at death of unknown skeletons. However, many local factors can affect tooth emergence such as infection or premature extraction of the deciduous predecessor (cf. Demirjian, 1978; El-Nofely and Işcan, 1989), and published data from living populations usually refers to emergence through the gums, not to the bony alveolus that skeletal biologists observe (for a detailed review of tooth eruption in humans, see Liversidge et al. (2003) and Scheuer and Black, (2000)).

Because of greater variation in the timing of tooth eruption, dental formation is a better measure of physiologic maturity. The formation of tooth crowns and roots is much less affected by hormonal influences, local and general environmental factors, and nutrition and social factors than tooth emergence, skeletal development, weight, or height (Demirjian, 1978; El-Nofely and Işcan, 1989; Smith, 1991). Developing teeth show morphologically distinct stages of formation and mineralization that can be identified radiographically and microscopically. Human biologists studying growth in living children started in the middle of the 20th century to use dental formation (or sometimes calcification, which microscopically evaluates the mineralization process) rather than emergence as a maturity indicator. There are two major advantages to this approach. First of all, dental formation is independent of skeletal maturity and most closely approximates chronological age (Demirjian et al., 1973; Garn et al., 1959; Garn et al., 1960; Moorrees et al., 1963a; Nolla, 1960). In addition, the dental formation system is the only system that is uniformly applicable for estimating age from prenatal stages to late adolescence since formation is a continuous process (Demirjian, 1978).

What methods should skeletal biologists use to record dental maturation, and what standards exist for estimating chronological age from dental age? Since skeletal biologists studying archaeological samples now seem to recognize that dental formation is the preferred method, it is important that careful and consistent methodologies are used. Research in recent decades has shown that chronological age can be determined most accurately from microscopic assessments of incremental structures in tooth tissues, the method of odontochronology (see the chapter by FitzGerald and Rose, this volume). However, since odontochronology is a destructive method that requires considerable technical expertise, currently, the simplest and least expensive method of assessing large archaeological samples comes from X-rays. Obtaining good radiographs of all juvenile dentitions from cemetery samples should be common practice. Many researchers now have their own X-ray facilities, and most have ready access to medical or dental X-ray services. Even small, portable X-ray units with independent power sources are becoming more widely available for those having to work in difficult field situations. Besides their value for comparing tooth formation with existing standards, which are all based on X-rays, the films also serve as primary data sources if skeletons must be reburied.

During tooth maturation, a series of morphological stages is recognizable beginning with actual formation of the tooth crypt and ending with closure of the apex of the fully formed root. Every tooth follows the same sequence, but to study the process, some system of measurement or evaluation is required. Most researchers have chosen ordinal or ranked systems of observation, but the numbers of tooth formation stages in the different systems have ranged from 3 to more than 20 (Table 4.2, and see Demirjian, 1978 for a detailed discussion of these systems; Liversidge, 2003; Liversidge and Molleson, 2001). The difficulties with this type of observation include problems of defining stages and subjectivity in identifying stages such as the difference between one-quarter and one-half formation of root length (Macho and Wood,
**TABLE 4.2 Comparative Table of the Stages of Dental Formation According to Different authors (taken from Demirjian, 1978)**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Presence of crypt</td>
<td>1</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>1</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Initial cusp formation</td>
<td>2</td>
<td>1</td>
<td>—</td>
<td>—</td>
<td>2</td>
<td>1(A)</td>
<td>1</td>
</tr>
<tr>
<td>Coalescence of cusps</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Occlusal surface completed</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>—</td>
<td>2(B)</td>
<td>—</td>
</tr>
<tr>
<td>Crown formation completed</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>3</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Crown</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>—</td>
<td>3(C)</td>
<td>—</td>
</tr>
<tr>
<td>Crown</td>
<td>6</td>
<td>—</td>
<td>4</td>
<td>—</td>
<td>4</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Crown</td>
<td>—</td>
<td>5</td>
<td>—</td>
<td>3</td>
<td>5</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Crown formation completed</td>
<td>7</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>6</td>
<td>4(D)</td>
<td>—</td>
</tr>
<tr>
<td>Initial radicular formation</td>
<td>8</td>
<td>7</td>
<td>6</td>
<td>5(½)</td>
<td>—</td>
<td>2</td>
<td>—</td>
</tr>
<tr>
<td>Initial radicular bifurcation</td>
<td>8 A, B</td>
<td>8</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Root 1/4</td>
<td>9</td>
<td>9</td>
<td>7</td>
<td>6</td>
<td>—</td>
<td>5(E)</td>
<td>—</td>
</tr>
<tr>
<td>Root 1/4</td>
<td>10</td>
<td>—</td>
<td>8</td>
<td>7(½)</td>
<td>7</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Root 1/4</td>
<td>11</td>
<td>10</td>
<td>9</td>
<td>8</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Root 1/4</td>
<td>12</td>
<td>—</td>
<td>10</td>
<td>9(½)</td>
<td>8</td>
<td>6(F)</td>
<td>—</td>
</tr>
<tr>
<td>Root 1/4</td>
<td>13</td>
<td>11</td>
<td>11</td>
<td>10</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Root completed</td>
<td>14</td>
<td>12</td>
<td>12</td>
<td>11(½)</td>
<td>9</td>
<td>7(G)</td>
<td>—</td>
</tr>
<tr>
<td>Apex 1/4 closed</td>
<td>—</td>
<td>13</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Apex closed</td>
<td>15</td>
<td>14</td>
<td>13</td>
<td>12</td>
<td>10</td>
<td>8(H)</td>
<td>3</td>
</tr>
</tbody>
</table>

*Reproduced from Demirjian (1978), with permission of the publisher.

*Apex closure: 1/4, 1/4, 1/4.*
Collapsing of categories reduces interobserver and intraobserver error. Radiographic standards provided by Moorees et al. (Moorrees et al., 1963a; Moorrees et al., 1963b) (MFH) are often used by bioarchaeologists because they are based on a large sample of children observed longitudinally and include standards for both deciduous and permanent teeth from the same sample. However, the timing of tooth formation in this sample is earlier than in other documented samples (Goodman and Song, 1999; Liversidge, 2003), and Liversidge notes that this may be attributable to the large number of formation stages defined for this system. Some researchers have dealt with this problem by adjusting the MFH age predictions to the Demirjian eight-stage scoring system (see below) (Cardoso, 2007b; McVeigh, 1999).

In the 1970s, Demirjian et al. (1973) developed a system for evaluating dental maturation in a large sample of French Canadian children. They used eight carefully defined stages of formation based on X-ray pictures, diagrams, and written criteria. Each tooth was assigned a score depending on its state of development, and weighted scores for all available teeth were added to produce a total maturity score. Because of its methods, this system has proven difficult for bioarchaeologists to employ. In addition, several subsequent studies on living children have shown advancement in dental maturation compared with the Demirjian standard, particularly between the ages of 5 and 8 years (Liversidge, 2003). Recently, Liversidge and colleagues (Liversidge et al., 2006) reevaluated the timing of seven of Demirjian’s tooth formation stages7 in many children (>9000) from eight different countries. They found no evidence of population differences (thus, failing to explain the previous findings of differences using Demirjian’s dental maturity method). Significant for bioarchaeologists, though, is that they produced a table of mean ages of children “in a tooth formation stage” for seven permanent teeth for all combined groups (see Table 4.3).

Another method by which bioarchaeologists might evaluate tooth formation is to measure the absolute lengths of developing teeth. Such an approach measures developing crown-root lengths both directly from loose teeth and from radiographs (Carels et al., 1991; Israel and Lewis, 1971; Mörnstad et al., 1994). The test of the method on a documented archaeological sample of infants and young children of known age from the eighteenth/nineteenth-century Spitalfields church crypt in England is the most comprehensive and easy to apply (Liversidge et al., 1993). Advantages of the developing tooth length approach are that it is quantifiable on a continuous scale, does not seem to be variable by sex, and was designed specifically to generate prediction equations for both deciduous and permanent teeth of cases of unknown age. Results of the Spitalfields studies showed a clear relationship between developing root length and the age of the individuals. Later, Liversidge et al. (2003) tested the Spitalfields standards on permanent mandibular teeth of a modern sample of 145 children (two other standards, one from Sweden and one from the Netherlands, were also tested) but only for children between 8 and 14 years. Results showed that age was underestimated with all three methods and that accuracy decreased with increasing age of the child (Spitalfields standards produced a mean difference between estimated and chronological ages of −0.79 (S.D. 0.93) for boys and −0.63 (S.D. 0.92) for girls). The Swedish standards were the most accurate for the 8 years + group.

Most recently, Cardoso has tested the Spitalfields standards against a sample of juvenile skeletons of documented age from Lisbon, Portugal, separately for deciduous and permanent teeth (Cardoso, 2007a). Results were highly accurate with tests of deciduous teeth showing an average difference between estimated and chronological age of 0.17 and −0.14 years for single teeth, and 0.04 years

7The end stage (mature root apex) is omitted since it is not possible to predict how much time has passed since a child entered this stage.
TABLE 4.3 Summary of Mean Ages of Children “In a Tooth Formation Stage” from a Large Sample of Combined Groups from Australia, Belgium, Canada, England, Finland, France, South Korea and Sweden

<table>
<thead>
<tr>
<th>Tooth</th>
<th>Stage</th>
<th>Girls Mean</th>
<th>SE</th>
<th>SD</th>
<th>n</th>
<th>Boys Mean</th>
<th>SE</th>
<th>SD</th>
<th>n</th>
<th>Girls and boys Mean</th>
<th>SE</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>I1</td>
<td>D</td>
<td>4.06</td>
<td>0.08</td>
<td>1.01</td>
<td>163</td>
<td>4.21</td>
<td>0.08</td>
<td>1.19</td>
<td>225</td>
<td>4.15</td>
<td>0.06</td>
<td>1.12</td>
</tr>
<tr>
<td></td>
<td>E</td>
<td>5.14</td>
<td>0.04</td>
<td>0.74</td>
<td>298</td>
<td>5.38</td>
<td>0.05</td>
<td>1.00</td>
<td>427</td>
<td>5.28</td>
<td>0.04</td>
<td>0.91</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>6.36</td>
<td>0.05</td>
<td>0.87</td>
<td>270</td>
<td>6.61</td>
<td>0.05</td>
<td>0.82</td>
<td>309</td>
<td>6.49</td>
<td>0.04</td>
<td>0.85</td>
</tr>
<tr>
<td></td>
<td>G</td>
<td>7.66</td>
<td>0.04</td>
<td>0.89</td>
<td>462</td>
<td>8.00</td>
<td>0.05</td>
<td>1.06</td>
<td>562</td>
<td>7.85</td>
<td>0.03</td>
<td>1.00</td>
</tr>
<tr>
<td>I2</td>
<td>C</td>
<td>3.38</td>
<td>0.13</td>
<td>0.70</td>
<td>29</td>
<td>3.82</td>
<td>0.11</td>
<td>0.73</td>
<td>42</td>
<td>3.64</td>
<td>0.09</td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>4.42</td>
<td>0.06</td>
<td>0.96</td>
<td>254</td>
<td>4.68</td>
<td>0.06</td>
<td>1.08</td>
<td>389</td>
<td>4.58</td>
<td>0.04</td>
<td>1.04</td>
</tr>
<tr>
<td></td>
<td>E</td>
<td>5.71</td>
<td>0.05</td>
<td>0.87</td>
<td>337</td>
<td>6.03</td>
<td>0.05</td>
<td>1.05</td>
<td>437</td>
<td>5.89</td>
<td>0.04</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>7.06</td>
<td>0.04</td>
<td>0.80</td>
<td>353</td>
<td>7.49</td>
<td>0.05</td>
<td>0.95</td>
<td>432</td>
<td>7.30</td>
<td>0.03</td>
<td>0.91</td>
</tr>
<tr>
<td></td>
<td>G</td>
<td>8.31</td>
<td>0.04</td>
<td>0.98</td>
<td>645</td>
<td>8.83</td>
<td>0.04</td>
<td>1.10</td>
<td>688</td>
<td>8.58</td>
<td>0.03</td>
<td>1.08</td>
</tr>
<tr>
<td>C</td>
<td>C</td>
<td>4.15</td>
<td>0.08</td>
<td>1.02</td>
<td>174</td>
<td>4.54</td>
<td>0.07</td>
<td>1.19</td>
<td>330</td>
<td>4.40</td>
<td>0.05</td>
<td>1.15</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>5.35</td>
<td>0.05</td>
<td>1.03</td>
<td>372</td>
<td>5.82</td>
<td>0.05</td>
<td>1.14</td>
<td>514</td>
<td>5.63</td>
<td>0.04</td>
<td>1.12</td>
</tr>
<tr>
<td></td>
<td>E</td>
<td>7.08</td>
<td>0.04</td>
<td>0.97</td>
<td>544</td>
<td>7.74</td>
<td>0.04</td>
<td>1.07</td>
<td>749</td>
<td>7.46</td>
<td>0.03</td>
<td>1.08</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>8.81</td>
<td>0.03</td>
<td>1.08</td>
<td>996</td>
<td>9.78</td>
<td>0.04</td>
<td>1.22</td>
<td>1069</td>
<td>9.31</td>
<td>0.03</td>
<td>1.25</td>
</tr>
<tr>
<td></td>
<td>G</td>
<td>10.85</td>
<td>0.04</td>
<td>1.28</td>
<td>821</td>
<td>12.02</td>
<td>0.05</td>
<td>1.33</td>
<td>819</td>
<td>11.43</td>
<td>0.04</td>
<td>1.43</td>
</tr>
<tr>
<td>P1</td>
<td>B</td>
<td>3.57</td>
<td>0.18</td>
<td>1.31</td>
<td>56</td>
<td>3.78</td>
<td>0.08</td>
<td>0.69</td>
<td>69</td>
<td>3.68</td>
<td>0.09</td>
<td>1.02</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>4.73</td>
<td>0.04</td>
<td>0.78</td>
<td>308</td>
<td>4.99</td>
<td>0.05</td>
<td>1.10</td>
<td>436</td>
<td>4.88</td>
<td>0.04</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>6.26</td>
<td>0.05</td>
<td>0.96</td>
<td>408</td>
<td>6.64</td>
<td>0.04</td>
<td>1.04</td>
<td>559</td>
<td>6.48</td>
<td>0.03</td>
<td>1.03</td>
</tr>
<tr>
<td></td>
<td>E</td>
<td>7.92</td>
<td>0.03</td>
<td>0.93</td>
<td>798</td>
<td>8.35</td>
<td>0.04</td>
<td>1.05</td>
<td>808</td>
<td>8.14</td>
<td>0.03</td>
<td>1.02</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>9.77</td>
<td>0.04</td>
<td>1.14</td>
<td>893</td>
<td>10.29</td>
<td>0.04</td>
<td>1.24</td>
<td>930</td>
<td>10.04</td>
<td>0.03</td>
<td>1.22</td>
</tr>
<tr>
<td></td>
<td>G</td>
<td>11.46</td>
<td>0.05</td>
<td>1.18</td>
<td>659</td>
<td>12.14</td>
<td>0.05</td>
<td>1.23</td>
<td>591</td>
<td>11.78</td>
<td>0.04</td>
<td>1.25</td>
</tr>
<tr>
<td>P2</td>
<td>A</td>
<td>4.14</td>
<td>0.16</td>
<td>1.43</td>
<td>78</td>
<td>4.32</td>
<td>0.14</td>
<td>1.37</td>
<td>90</td>
<td>4.24</td>
<td>0.11</td>
<td>1.39</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>4.79</td>
<td>0.07</td>
<td>0.88</td>
<td>160</td>
<td>4.86</td>
<td>0.07</td>
<td>1.07</td>
<td>223</td>
<td>4.83</td>
<td>0.05</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>5.90</td>
<td>0.07</td>
<td>1.24</td>
<td>340</td>
<td>6.02</td>
<td>0.06</td>
<td>1.22</td>
<td>448</td>
<td>5.97</td>
<td>0.04</td>
<td>1.23</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>7.29</td>
<td>0.05</td>
<td>1.09</td>
<td>496</td>
<td>7.46</td>
<td>0.05</td>
<td>1.16</td>
<td>577</td>
<td>7.38</td>
<td>0.03</td>
<td>1.13</td>
</tr>
<tr>
<td></td>
<td>E</td>
<td>8.54</td>
<td>0.04</td>
<td>1.14</td>
<td>745</td>
<td>9.06</td>
<td>0.04</td>
<td>1.23</td>
<td>777</td>
<td>8.81</td>
<td>0.03</td>
<td>1.21</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>10.51</td>
<td>0.04</td>
<td>1.29</td>
<td>975</td>
<td>10.98</td>
<td>0.05</td>
<td>1.43</td>
<td>958</td>
<td>10.74</td>
<td>0.03</td>
<td>1.38</td>
</tr>
<tr>
<td></td>
<td>G</td>
<td>12.33</td>
<td>0.05</td>
<td>1.32</td>
<td>701</td>
<td>12.94</td>
<td>0.06</td>
<td>1.42</td>
<td>633</td>
<td>12.62</td>
<td>0.04</td>
<td>1.40</td>
</tr>
<tr>
<td>M1</td>
<td>D</td>
<td>3.74</td>
<td>0.08</td>
<td>0.77</td>
<td>101</td>
<td>4.02</td>
<td>0.12</td>
<td>1.40</td>
<td>138</td>
<td>3.90</td>
<td>0.08</td>
<td>1.18</td>
</tr>
</tbody>
</table>

(Continued)
when using all available teeth. Tests of permanent teeth showed an average difference between estimated and chronological age of 0.56 and 0.05 years when using single teeth and 0.16 years when using all available teeth, although there was a tendency to overestimate age in individuals less than 6 years. This is perhaps more surprising since Cardoso has also shown that age estimates from maxillary deciduous teeth tend to overestimate true age versus their mandibular deciduous teeth pairs (Cardoso, 2007a) and the Spitalfields standards combine data from both maxillary and mandibular jaws. Partial explanation for the poor performance of the Spitalfields standards for the modern children (Liversidge et al., 2003) versus the comparisons made by Cardoso may be from sample ages, sample sizes, and methods employed. However, it is also possible that the Spitalfields standards, reflecting a biased mortality sample, will more closely estimate age of another mortality sample (i.e., the Lisbon collection) than a sample of healthy living children.

The accuracy of age prediction is dependent on the age range that is studied and the number of teeth observed. Variation in accuracy increases at older ages as sources of variability in dental development increase, but the more teeth that are examined, the more accurate the estimates. Smith’s (1991) comparison of the chronological age of several children with the stages of permanent mandibular tooth formation derived from the data of Moorrees et al. (1963a) suggest that dental age can be estimated to within two months for young children (the children ranged in age from 4 to 10 years).

Several others have tested age estimates for small samples of personally identified individuals (Bowman et al., 1992; Liversidge, 1994; Saunders and Hoppa, 1993), but they have used different methods and achieved variable results. Saunders et al. (1993) also compared the overall distribution of age estimations from tooth formation in a large nineteenth-century Canadian sample to documented ages from burial records for the cemetery to determine the representativeness of the skeletal sample.
Trying out several combinations of published dental formation standards, they observed that the combination of permanent and deciduous mandibular tooth standards based on the sample of children studied by Moorrees et al. (1963a; 1963b) produced the best comparisons of the skeletal sample to the age distributions recorded in the parish registers. They argue, mainly for methodological reasons, that the mandibular standards of Moorrees et al. work best, at least for North American or European samples. Although other workers might prefer other standards, these authors point out that care must be taken not to apply truncated reference standards to cases of young children. For example, the standards of Anderson et al. (1976) begin at 3 years of age, making them inappropriate for estimating the ages of infants and toddlers. As has been shown with age estimation reference standards for adults, the distributions of juvenile age estimates for archaeological samples may be biased by and thus reflect the samples that were used to assign age. Depending on which standard is used, there are tendencies to under-age or over-age the individuals (Saunders and Spence, 1986).

Age evaluation for the prenatal and perinatal periods is generally based on fewer standards produced from much smaller samples. But Kraus and Jordan (1965) published data on very early crown coalescence from a sample of 737 human fetuses examined between 1954 and 1963. These standards have been applied to archaeological specimens in several cases known to this author (for examples, see Saunders and Spence, 1986; Spence, 1986). Deutsch et al. (1985) have also provided data on deciduous anterior tooth crown length measurements and weights from a sample of 50 anatomically normal fetuses and infants aged 0–46 weeks, based on their total body size (actual gestational age is almost never known for fetal samples). Obviously, only the developing crown length standards can be used on archaeological samples. Skinner and Goodman (1992) have noted that there is still a distinct shortage of detailed standards for the early formation of deciduous tooth crowns that can be applied to late fetal and neonatal skeletons.

Physiologic age of the immature skeleton, minus the teeth, must be assessed from either the appearance and union of bony epiphyses or bone size. Diaphyseal length measurements are the common sources of skeletal age estimates from before birth to mid-teens because long bone epiphyses deteriorate or are frequently lost at excavation. Even if they are recovered, trying to attach them to separate diaphyses would simply introduce more error. This limits the bone size technique of skeletal age estimation to a shorter portion of the total growth period, approximately late fetal to 12 years.

Since epiphyseal appearance is not readily applicable to excavated skeletons because of recovery problems, epiphyseal union becomes the favored method in the mid-teens when the process of gradual union of the epiphyses begins. Ubelaker (1989) provides a thorough discussion of the various standards available for age estimation from epiphyseal union. Important points to bear in mind are that sex differences also exist in the timing and sequence of epiphyseal union, many published studies fail to report full ranges of variation for the timing of union of epiphyses, and methodological problems exist with observation of the process and duration of union. Several years can elapse between the beginning and the final closure of an epiphysis (McKern and Stewart, 1957) so that various “stages” of union can be defined. However, others have shown that inter-observer error increases with the number of ranked stages of union defined for any one epiphysis (Owings Webb and Suchey, 1985). Some studies have proposed using specific

---

8The Moorrees et al. standard still suffers from problems. The high number of formation stages creates problems with precision of observations, and most seriously, the charts of results of age variability for each stage are poorly reproduced and hard to convert to numerical values.
examples of growth center appearance, size increase, and early union for estimating skeletal age of infants and young children. These examples include the mandibular symphysis (Becker, 1986), the tympanic plate of the temporal (Curran and Weaver, 1982; Weaver, 1979), and development of the occipital bone (Redfield, 1970), but the age ranges appropriate to these criteria are often limited. More recently, careful examinations of the development of individual bones in the fetal, infant, and child skeleton have begun to produce more useful standards (Scheuer and MacLaughlin-Black, 1994).

Diaphyseal length is used as an estimate of skeletal age when calcifying teeth are missing. It must be remembered, though, that bone development studies using skeletal samples require both dental and skeletal data so that dental age may be treated as the closest approximation to chronological age, whereas skeletal age serves as a marker for alterations and defects of growth. In human osteological research, we assume that dental development is less influenced by environmental insults than skeletal development. Cardoso (2007a) recently tested this assumption in the well-documented Lisbon sample and confirmed it, but his results also show that children of low socioeconomic status are delayed severely in skeletal growth and frequently also delayed in dental development.

Sources of diaphyseal length data include the radiographic samples mentioned above and dry bone measurements taken from archaeological samples (Armelagos et al., 1972; Hoffman, 1979; Hummert and Van Gerven, 1983; Johnston, 1962; Mensforth, 1985; Merchant and Ubelaker, 1977; Sundick, 1972; y’Edynak, 1976). The purpose of recording diaphyseal measurements from archaeological samples using dental development as an approximation of chronological age (true chronological age is unknown) is to try to account for population variation in bone size as well as to attempt assessments of population variation in growth. Ubelaker (1989) has shown that the calculated age at death from femur diaphyses using several different population standards varies widely. The total number of years within the range of estimates he calculated varies from 3.5 to 8.5 years and increases with increasing size of the bone. Consequently, it is important to choose skeletal age standards appropriate for the known or suspected population affiliation of the individual or sample because this will greatly increase accuracy.

Specific bone size standards also exist for estimating fetal and perinatal age (Fazekas and Kósa, 1978; Malinowski and Młodziejowski, 1978; Olivier and Pineau, 1960; Palkama et al., 1962; Scheuer et al., 1980). Here again, population variation is a factor (Scheuer et al., 1980). Ubelaker (1989) showed that a considerable range of variation exists in fetal age estimations, exceeding one-half lunar month, when using the Fazekas and Kósa (1978) regression equations. One other potential source of data for fetal and perinatal age estimation is fetal femur length measured using ultrasound (O’Brien and Queenan, 1981). An examination of mean ultrasound femur length at term is very close to mean length based on cadaver samples (Fazekas and Kósa, 1978; Olivier and Pineau, 1960).

The sample reported by Fazekas and Kósa (1978) is one of the most important because of its large size although skeletal size is compared with body size in this sample and not to true gestational age. Fortunately for skeletal researchers, meticulous analysis of the Fazekas and Kósa data as well other samples by Sellier et al. (Brzez et al., 1997; Sellier et al., 1997) has produced new equations for estimating body length (and therefore, age) from diaphyseal length in the skeletons of stillborns and neonates, which provides opportunities for

---

9Malinowski and Młodziejowski (1978) provide a useful review of early studies of the size of limb bones of fetuses and infants.
researchers to concentrate on fetal and infant mortality as indicators of population health in past populations when these bones are well preserved in skeletal samples (Saunders and Barrans, 1999).

Other age estimation methods

Although the dental and skeletal methods described above are the main means of estimating age, the possibility also exists of using histological methods with subadult bones. Several important studies of cortical bone histology have included assessments of infants and children (Amprino and Bairati, 1936; Jowsey, 1960; Kerley, 1965). Bone turnover in the growing skeleton is usually observed as too complex to be used in age estimation because of the superimposition of bone modeling over remodeling of cellular-based structures (Stout, 1992). Nevertheless, there may be some useful applications in the future (Saunders and DeVito, 1990).

One other possibility is to use a modification of the method of functional dental wear developed by A.E.W. Miles (1962, 1963, 1978). The advantage of this method is that it is specifically tailored to each population to which it is applied. One begins with dental age estimates from formation and emergence of preadult dentitions seriated from youngest to oldest. The degrees of occlusal molar wear observed on permanent molars of age-estimated children and adolescents is then used to assign ages to the adults in a sample. The method is based on the premise that the first permanent molar will average six more years of wear than the second permanent molar, and so on. “Functional molar age” is then defined as the number of years for which the molar had been in use. Functional dental age estimates based on wear might be applied to older children and adolescents, whereas deciduous tooth wear could also be examined as a check on dental age variability among individuals. Jackes (1992) indicates that wear assessments for estimating age should be supplemented with dental measurements.

GROWTH STUDIES OF ARCHAEOLOGICAL SKELETAL SAMPLES

Although he was one of the first to attempt to study growth using archaeological skeletal samples, Johnston (1962) was also the first to caution that these samples do not represent the normal, healthy children in the population who lived but those who died to become part of a biased, mortality sample. On the other hand, Lovejoy et al. (1990) have argued that most infant deaths among earlier groups were the result of acute, rather than chronic, diseases and, therefore, should not drastically alter dental or osteological maturation. They claim, as did Sundick (1978), that skeletal samples compare favorably with their counterparts who survived to adulthood. Other researchers are more conservative, arguing that deceased juveniles represent the minima rather than the modes of those who lived (Buikstra and Cook, 1980). To some, differential selection at death and the heterogeneity of mortality samples is a given (Whittington, 1991). Saunders and Hoppa (1993) addressed this issue by examining the literature on survivors and nonsurvivors in living populations. They observed that nonsurvivors have higher morbidity per age and higher levels of stress resulting in shorter height-for-age than survivors. After modeling the potential magnitude of biological mortality bias on linear growth by generating survivor and nonsurvivor distributions of height-for-age, they found that the difference in stature between survivors and nonsurvivors is significant, but the actual measurable difference in total femoral length is probably never more than several millimeters. These findings suggest that the effect of mortality bias on long bone lengths of juvenile skeletons from archaeological samples is minimal. Other factors such as sample sizes, aging
methodologies, and preservation status may have a greater impact on efforts to investigate growth retardation resulting from increased stress in earlier populations.¹⁰

Konigsberg and Holman (1999) have addressed the “estimation of age at death” problem pointing out that age estimation from one or more skeletal/dental traits requires producing the distribution of possible chronological ages for each skeleton. Ignoring the distributional information around the mean or median of estimates of age generates a false sense of security about the statistical power of any additional analyses, such as generating growth trajectories. Few researchers have taken this approach with growth-related studies of archaeological samples.

One of the difficulties with the earliest growth-related studies is that chronological age was often estimated from combinations of dental and osseous criteria or from dental emergence alone (Armelagos et al., 1972; Johnston, 1962; Lallo, 1973; y’Edynak, 1976) (see Table 4.4). Some early researchers also attempted to separate male and female juveniles without any methodological justification. A turning point came with the use of dental development to estimate chronological age at death (Merchant and Ubelaker, 1977; Sundick, 1972). Nevertheless, because of constant efforts to “tinker” with age estimation methodologies, few existing publications are comparable even if we ignore the problem of mortality bias.

Measures of diaphyseal length, when compared with estimated dental ages in archaeological samples, are necessarily cross-sectional but with the superimposed problems of sex and age estimation. Longitudinal growth studies of living children, which follow a series of individuals continuously over substantial periods, can examine individual growth patterns, the timing of significant growth events, and the relative velocities of growth. Only longitudinal growth studies can examine adequately individual variability in growth rates and patterns. Many who have studied juvenile long bone size in archaeological samples persist in referring to their results as growth curves. It must be remembered that these are not true growth curves in the sense that they are used in living growth studies, either cross-sectional or longitudinal, since, of course, the dimensions represent deceased individuals who never reached maturity.

Bearing in mind the above limitations, until recently, studies of growth-related change in archaeological samples have been restricted to examining the adequacy of growth of children as an index of overall community health or the adaptation of the population to its environment (Johnston and Zimmerman, 1989). No one has discovered any major differences in the direction or apparent pattern of skeletal size in the past. It could be argued that this means there have not been any major changes in human growth patterns over time, but such changes would have to be drastic to be detected in archaeological skeletal samples. Most or all skeletal samples of past groups appear shorter for age than modern groups. Might this represent genetic differences (see y’Edynak, 1976) or the effects of a harsh environment on the growth of disadvantaged children (Johnston and Zimmerman, 1989)? Given what is known about the effects of the environment on the growth of the skeleton, we would expect that most cases represent environmental effects, that is, populations suffering from nutritional and disease-related stress. But our study of the historic-era skeletal sample from St. Thomas’ Anglican Church, Belleville, Ontario (Saunders et al., 1993), suggests that these children followed a pattern of growth not unlike that of modern American children, up to at least 12 years of age, with perhaps the exception of those under 2 years who are slightly smaller than modern standards. This finding is contrasted with a sample of skeletons from the Raunds Site, a tenth-century A.D. Anglo-Saxon cemetery (Hoppa, 1992) and a sample of skeletons from the Roman period cemeteries at Poundbury in England (Farwell

¹⁰But note my earlier comments about the possibility of mortality bias being reflected in tooth size as well as the work by Cardoso (2007b).
## TABLE 4.4 Some Previous Growth Studies of Archaeological Samples

<table>
<thead>
<tr>
<th>Source (Year)</th>
<th>Sample and Temporal Context</th>
<th>Sample Size</th>
<th>Age Range</th>
<th>Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Johnston, 1962)</td>
<td>Indian Knoll, Kentucky, Archaic Period. 3000 B.C.</td>
<td>3000 B.C.</td>
<td>165</td>
<td>fetal–5.5 yrs.</td>
</tr>
<tr>
<td>(see Armelagos et al., 1972; Mahler, 1968)</td>
<td>Nubians, Wadi Halfa 350 B.C.–1350 A.D.</td>
<td>1350 A.D.</td>
<td>115</td>
<td>6 mos.–31 yrs.</td>
</tr>
<tr>
<td>(Walker, 1969)</td>
<td>Late Woodland, Illinois</td>
<td>1700–1750 A.D.</td>
<td>43</td>
<td>nb*–12 yrs.</td>
</tr>
<tr>
<td>(Sundick, 1972)</td>
<td>Indian Knoll, Kentucky, Archaic Period, 3000 B.C.</td>
<td>1700–1750 A.D.</td>
<td>128</td>
<td>nb–18 yrs.</td>
</tr>
<tr>
<td>(see Goodman et al., 1984; Lallo, 1973)</td>
<td>Dickson Mounds, Woodland–Mississippian</td>
<td>1700–1750 A.D.</td>
<td>557</td>
<td>nb–35 yrs.</td>
</tr>
<tr>
<td>(Sundick, 1978)</td>
<td>Altenerding, W., Germany, 6th–7th centuries A.D.</td>
<td>1700–1750 A.D.</td>
<td>82</td>
<td>nb–18 yrs.</td>
</tr>
<tr>
<td>(Jantz and Owsley, 1984a; Jantz and Owsley, 1984b; Jantz and Owsley, 1985)</td>
<td>Arikara, seven samples</td>
<td>1700–1750 A.D.</td>
<td>.500</td>
<td>nb–12 yrs.</td>
</tr>
<tr>
<td>(Cook, 1984)</td>
<td>Illinois Valley, Woodland and Mississippian, ca. 244</td>
<td>244</td>
<td>180</td>
<td>nb–6 yrs.</td>
</tr>
<tr>
<td>(Mensforth, 1985)</td>
<td>Libben Late Woodland &amp; Bt-5 Archaic</td>
<td>1615 A.D.</td>
<td>85</td>
<td>nb–10 yrs.</td>
</tr>
<tr>
<td>(Saunders and Melbye, 1990)</td>
<td>Late Ontario Iroquois, 146</td>
<td>1615 A.D.</td>
<td>147</td>
<td>nb–15 yrs.</td>
</tr>
<tr>
<td>(Lovejoy et al., 1990)</td>
<td>Libben Site, Ohio, Late Woodland, 800–1100 A.D.</td>
<td>1615 A.D.</td>
<td>152</td>
<td>nb–12 yrs.</td>
</tr>
<tr>
<td>(Hoppa, 1992)</td>
<td>Raunds, Berinsfield and Exeter Anglo-Saxon medieval</td>
<td>1615 A.D.</td>
<td>90</td>
<td>nb–16 yrs.</td>
</tr>
<tr>
<td>(Saunders et al., 1993)</td>
<td>St. Thomas’ Anglican Church cemetery, 19th cent. 1821–1874</td>
<td>1821–1874</td>
<td>281</td>
<td>fetal–15 yrs.</td>
</tr>
<tr>
<td>(Farwell and Molleson, 1993)</td>
<td>Poundbury Camp, Dorchester, U.K. Late Roman</td>
<td>1821–1874</td>
<td>&gt;400</td>
<td>&lt;20 yrs.</td>
</tr>
</tbody>
</table>

(Continued)
and Molleson, 1993) as well as several other medieval and later period cemeteries from Britain (Molleson and Cox, 1993; Ribot and Roberts, 1996; Wiggins and Rogers, 1995), where individuals of comparable dental age had much shorter diaphyses.

More recently, Humphrey (2003) included the St. Thomas’ sample data in a comparison of growth of the femoral diaphysis for several archaeological samples from around the world. She compared the data in terms of percentage of adult size attained by dividing femoral diaphyseal length by estimated adult femur length taken from the adult skeletons in each sample. This approach places the emphasis on the rate of progress toward adult size rather than toward actual diaphyseal size attained at any given age. A decline in growth attainment in infancy and childhood was an almost universal feature of every sample. One might argue that this decline simply reflects universal mortality bias in skeletal samples (the children died before reaching their adult potential). But results for the St. Thomas’ sample were different. This group is very similar to the modern reference sample of living children from Denver and then even rises far above it at later ages. This result might be explained in part by low values for mean adult femoral length. The period of use of the cemetery was relatively short in archaeological terms (53 years), and it is known that the proportion of juveniles increased over that time (Saunders et al., 2002). Perhaps adult skeletons are over-represented by recent migrants from Britain in the earlier part of the century whose completed stature was relatively short. However, as noted, absolute femur lengths of the children from St. Thomas’ are similar to those from Denver supporting the interpretation that most deaths were from acute causes rather than from chronic ones that would have compromised growth.

### TABLE 4.4

<table>
<thead>
<tr>
<th>Source</th>
<th>Sample and Temporal Context</th>
<th>Sample Size</th>
<th>Age Range</th>
<th>Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Miles and Bulman, 1994)</td>
<td>Ensay, Scotland 1500–1850 A.D.</td>
<td>120</td>
<td>Fetal–20 yrs.</td>
<td>dental development</td>
</tr>
<tr>
<td>(Dillingham, 1996)</td>
<td>Toqua Site, Tennessee Late Mississippian Dalla Phase 1200–1650 A.D.</td>
<td>145 subadults 33 femora</td>
<td>prenatal–2.5 yrs.</td>
<td>dental development</td>
</tr>
<tr>
<td>(Ribot and Roberts, 1996)</td>
<td>Raunds, Chicester 8th–12th centuries A.D.</td>
<td>180</td>
<td>&lt;18–20 yrs.</td>
<td>dental and osseous development</td>
</tr>
<tr>
<td>(Hutchins, 1998)</td>
<td>Milwaukee County Almshouse late 19th–early 20th centuries</td>
<td>138</td>
<td>fetal–6 mos.</td>
<td>dental development</td>
</tr>
<tr>
<td>(Lewis, 2002)</td>
<td>Raunds Furnells, Wharram Percy, St. Helen-on-the-Walls, Christ Church Spitalfields</td>
<td>831</td>
<td>fetal–17 yrs.</td>
<td>dental development</td>
</tr>
<tr>
<td>(Humphrey, 2003)</td>
<td>Published data from 11 populations from 2000 B.C.–19th cent.</td>
<td>1000</td>
<td>birth–12 yrs.</td>
<td>dental development</td>
</tr>
</tbody>
</table>

*nb=newborn*
Several ways exist of assessing alterations to the growth process in skeletal samples. One can look for temporal changes that reflect changes in environmental quality within groups or possibly by comparison with modern samples with adequate control of population factors (age estimation, etc.). As long as we know that population migration was minimal, rapid changes over relatively short time periods are easily attributable to environmental changes since genetic change takes place much more slowly. Several studies have identified temporal changes in growth-related bone size. Laloo (1973, and see Goodman et al., 1984) found a decrease in the attained size of diaphyseal lengths in the Missippian period (A.D. 1200–1300) in the central Illinois River valley and associated this with a dietary change from mixed foraging and farming to intensive maize agriculture.

Jantz and Owsley in their series of papers on the Arikara (Jantz and Owsley, 1984a, 1984b; Owsley and Jantz, 1985) showed that the latest group (in the late eighteenth and early nineteenth century) experienced the lowest rates of diaphyseal increase in size, particularly in late childhood and in the perinatal group (late fetal/early neonatal) of the archaeological samples. They attribute this to under-nutrition, introduction of epidemic diseases, depopulation and intertribal conflict, and especially stress effects on the mother, later reflected in the perinatals. One might ask here how epidemic diseases would influence altered growth characteristics in the later samples since most contact period epidemics were acute and not chronic. They suggest that the native groups were undergoing declines in health status as European influence and encroachment from other tribes increased over time.

An association between smaller individuals in specific growth periods and other skeletal indicators of bone pathology is also possible. Hummert and associates (Hummert, 1983a; 1983b; Hummert and Van Gerven, 1983) have examined the relationship between diaphyseal lengths and cortical bone volume, Harris lines, and histomorphology in two medieval period samples from Sudanese Nubia. They identified differences between the early and late samples. However, although increases in bone lengths seem to have been well maintained, the percent cortical area revealed excessive endosteal resorption. This resorption was attributed to nutritionally related stress. However, a later study of the geometric properties of these bones (Van Gerven et al., 1985; see also Ruff, this volume) suggested that the reduction in the percent cortical area could just as easily be interpreted as a response to increased bending strength (developed as the children grew normally) and that no evidence of environmental stress on growth might be found. On the other hand, definite evidence exists that bone growth in length will be maintained at the expense of growth in cortical thickness in the face of nutritional and disease stress (Himes, 1978; Husselsmore, 1981). These kinds of comparisons require additional study.

Several examinations of temporal and within-site variations in growth-related features have also been conducted by Mensforth (1985) and Lovejoy et al. (1990) on prehistoric sites in Kentucky and Ohio. Mensforth (1985) identified early long bone growth retardation in the Late Woodland Libben, Ohio sample (A.D. 800–1100) when compared with Late Archaic period foragers from the Carlston Annis Bt-5 site in Kentucky (2655–3992 b.c.), specifically in the 6-month to 4-year age range and suggested high levels of infectious disease in the first years of life as the implicating factor. This suggestion was based on the identification of a high prevalence of periosteal reactive bone in the subadults, which indicated infection. Lovejoy et al. (1990) have provided documentation of this situation in their study of the normalized values for the Libben sample compared with healthy Euro-American children. Both studies note that evidence shows that the Libben people ate a nutritionally adequate diet. They attribute increased disease loads to higher population density and to a greater degree of sedentism compared with seasonally mobile hunter-gatherers. Although this finding is difficult to demonstrate for past
populations, it is a compelling argument that has been used in other investigations (Katzenberg, 1992; Larsen, 1997; Ribot and Roberts, 1996; Saunders et al., 1992).

More recently, Lewis (2002) examined femoral size increase and morbidity profiles in four sites from medieval and post-medieval England. She observed that the urban children from the industrial period (Spitalfields sample) were up to 3 cm shorter than their late medieval counterparts, that skeletal indicators of stress appeared in very young infants from this site, and that metabolic stress (particularly rickets) was very common. She argues that rather than differences in rural and urban environments, it was industrialization that had the greatest impact on child health.

SUGGESTIONS FOR FUTURE RESEARCH

Sex Estimation

Logic says that sufficient dimorphism should exist for sex separation in fetal and early infant skeletons because of the presence of high levels of testosterone. Dimorphism should increase again at adolescence as pubertal changes begin to occur. But the percentages of observed subadult skeletal dimorphism are believed to be low compared with levels observed in the adult pelvis. In Weaver’s (1980) study of hip bones, two indices, although not significantly different between the sexes but which fit expected patterns, showed percent dimorphism ranging from 0.2% to 9.9%. Specker et al. (1987), in their analysis of a bone mineral index in 5–7-year-old children, observed 17.7% (surprisingly high) dimorphism. The ranges of percent dimorphism for a variety of adult pelvic indices (Kelley, 1979; MacLaughlin and Bruce, 1986; Schulter-Ellis et al., 1983) are between 10% and 26%, with the highest levels deriving from Kelley’s sciatic notch/acetabular index.

The teeth, because of their constancy of size after development, should be good indicators of subadult sexual dimorphism. However, the magnitude of tooth dimorphism is also fairly low, as indicated. We now know that it is the relative amount of dentine that contributes to sex differences, at least for permanent teeth. This phenomenon requires additional investigation.

It would help if in the future skeletal researchers could have access to large samples of subadult skeletons of known sex and age, be they accumulations of data from forensic cases (which are, fortunately, still rare) or, more likely, from identified individuals excavated from historically documented cemeteries. Sex determination of archaeological samples using DNA extraction and analysis of sex-specific gene fragments is of great promise but not yet ready to solve all our problems. Issues of consistency, control, false negatives, and false positives remain in DNA analysis as well as the current high cost of testing large samples from which substantial proportions may never yield amplifiable DNA. However, in the future, we should see controlled studies of the various morphological and metric methods of juvenile sex estimation compared with DNA methods and documented sex from archaeological samples.

Age Estimation

A need exists for more examination of dental development in a variety of population groups. In particular, we need information on deciduous tooth development (quantified at macroscopic and microscopic levels) gathered from living individuals. Although there is not much in the way of longitudinal tooth development data since the widespread use of research X-rays has been curtailed, the possibility exists of amassing large samples of cross-sectional data from clinical, radiographic databases (Trodden, 1982). Yet, archaeological samples themselves may be the best sources of data for devising and improving on methods of skeletal and even dental age estimation. Identified cases from historic archaeological sites can, again, act as a database for exploring deciduous
crown and root development, dental wear as a subadult age estimation technique, and histological age changes in cortical and possibly even trabecular bone.

If population differences in the rates and timing of dental development exist (but see Liversidge et al., 2006), then the problem of how to compare two samples using consistent estimates of dental age and thereby search for meaningful differences in skeletal growth still must be solved. This work is only beginning, and analyses of microscopic enamel and dentine formation will help to determine the range of variability in dental development.

Growth Related Research

To continue to pursue growth-related research on archaeological skeletal samples, we need some solutions to the problems identified previously. Pragmatically, the most useful current research would be to use ranges established for population variation in growth and development in the kinds of living groups that are comparable with the archaeological samples we are studying. We will never conduct longitudinal studies of growth in the past, but we can refine our cross-sectional comparisons between archaeological samples and modern group data.

Since little likelihood exists of examining the skeletons of willed bodies of juveniles, occasional forensic cases of immature individuals serve as very important sources of data for improving our methods of identification. One other source that has considerable potential is the recovery of identified individuals from historic cemeteries with associated documentation. Stored databases of clinical radiographs might also prove useful in this regard.

This survey has been cautionary, but it is not pessimistic. Considerable potential exists for the analysis of growth-related phenomena in archaeological skeletal samples. Studies of historical samples, especially those with associated documentation, may allow us to test some of our assumptions about the nature of human mortality samples so that we can reach confident conclusions about prehistoric peoples.

ACKNOWLEDGEMENTS

Work conducted on the St. Thomas’, Belleville, sample was made possible by collaboration with Dr. Heather McKillop, and I am thankful for the permission for study provided by St. Thomas’ Anglican Church and its congregation. The skeletal research on the St. Thomas’ sample has been supported by grants from the Ontario Heritage Foundation, the Bridge St. Foundation, the Arts Research Board of McMaster University, and the Social Sciences and Humanities Research Council of Canada. This chapter could not have been written without the collaboration and support of the SSHRC grant team: Gerry Boyce, Ann Herring, and Larry Sawchuk. I would also like to thank Rob Hoppa for stimulating research discussions and for his association as a scientific colleague. Thanks are also due to the Canada Research Chairs program.

REFERENCES


REFERENCES


REFERENCES


Miles AE. 1978. Teeth as an indicator of age in man. In: Butler PM, Joysey KA, editors. Development,


Saunders SR, Barrans L. 1999. What can be done about the infant category in skeletal samples?


Whittington SL. 1991. Detection of significant demographic differences between subpopulations of prehispanic Maya from Copan, Honduras, by
CHAPTER 5

HISTOMORPHOMETRY OF HUMAN CORTICAL BONE: APPLICATIONS TO AGE ESTIMATION

ALEXANDER G. ROBLING
SAM D. STOUT

INTRODUCTION

Quantitative bone histology (histomorphometry) has been used to estimate age at death for nearly a century, the first published report of which appeared in 1911 (Balthazard and Lebrun, 1911). Bone histomorphometry offers a powerful tool to the skeletal biologist, and its application to age estimation in modern, historic, and prehistoric populations has been met with encouraging results. These successes can be enhanced by incorporating several recently elucidated factors affecting reliability and accuracy into histological age estimates. Such factors include sex and population variability, adequate sampling techniques—including reference sample composition, choice of skeletal element, and topographic sampling procedure—and the effects of pathological conditions and the biomechanical loading environment. Indeed, proper use of histomorphometric age estimation techniques requires an understanding of the effect of these intrinsic and extrinsic factors on derived age estimates. This chapter reviews the physiologic basis for histomorphometric age-estimation techniques, the histomorphology of cortical bone and its relation to age estimation, some factors known to affect derived estimates, and finally, considers future directions in the field. Appendix A provides working examples of two techniques commonly used for histomorphometric age estimation and includes labeled schematic diagrams for both examples. Appendix B profiles selected age estimation techniques.

THE PHYSIOLOGIC BASIS FOR HISTOMORPHOMETRIC AGE ESTIMATION TECHNIQUES: BONE MODELING AND REMODELING

Development of the adult skeleton is achieved by growth, modeling, and remodeling. Growth and modeling are two processes that work in concert in the normal growing individual and thus will be considered together. In long bones, growth increases bone length and diameter (both internal and external) as specified by the genetic program of the organism. This baseline architecture is modified by the modeling process, which sculpts the bone’s size, shape,
and curvature to optimally sustain the mechanical loads typically borne by that bone. The separate effects of growth and modeling are apparent in the limb bones of paralyzed, growing children and animals, which usually lack significant bone curvature, develop subnormal cortical thickness, and exhibit a roughly circular cross section. Modeling adjusts bone architecture and mass via modeling drifts, which add bone to some surfaces and remove it from others. Modeling drifts move bone through tissue space (Fig. 5.1) and can simultaneously increase or decrease the cross section’s size by selectively inhibiting or promoting cellular activity at the resorptive and appositional surfaces accordingly. In the normal developing skeleton, growth and modeling result in the production of organized, parallel sheets of primary lamellar bone—circumferential and endosteal lamellae—typically visible in diaphyseal cross sections (Fig. 5.2a). Some circumferential and endosteal lamellae deposited early in the developmental period are removed or “modeled out” as the bone drifts. Thus, the adult cortex comprises a collection of lamellae exhibiting an array of different ages. Their mean age, however, is always less than the individual’s chronological age. The circumferential and endosteal lamellae deposited during modeling provide the canvas on which discrete units of cortical remodeling leave their mark (see below).

Once skeletal maturity is reached, modeling reduces to a trivial level compared with that

![Figure 5.1](image)

**Figure 5.1** (A): During development, modeling drifts in the middle third of the sixth rib remove bone from the internally facing surfaces and deposit bone on the externally facing surfaces. As the rib drifts laterally, the formation surfaces outpace the resorptive surfaces, and the rib gains cross-sectional area. Note that none of the tissue present in the young rib (a) is present in the adult rib (c); it has been completely “modeled out.” (B): Enlarged view of the drifting rib cortex (trabeculae have been removed for clarity). The younger bone (lamellae) is darker, and the older bone is lighter. The cross section thus comprises a mosaic of different aged lamellae. (C): The same drifting rib cortex illustrated in B. The stippled region represents the size and position of the rib cortex at time 1 (arbitrarily designated as age 6 in the figure). A year later (shaded silhouette), some cortex that was present at age 6 is still present (region exhibiting both stippling and shading). However, nearly half of the cortical bone present at age 7 did not exist in the same rib just one year earlier. Bone modeling is a dynamic process in the growing skeleton that regularly and rapidly alters the size, shape, relative position, and age of bone tissue.
which occurs during development. Renewed modeling in the adult skeleton can occur, however, in some disease states and in cases where the mechanical loading environment has been altered radically. These observations are relevant to those using histomorphometric techniques to estimate age because renewed adult modeling, in the absence of a concurrent increase in remodeling, will decrease the mean tissue age and ultimately result in age estimates lower than actual age.

Unlike modeling, which involves either resorption or formation (but not both) at a locus, bone remodeling always follows an activation → resorption → formation sequence at a locus (Fig. 5.3). Remodeling removes and replaces discrete, measurable “packets” of bone. These packets, or bone structural units (BSUs), form the basis for most histomorphometric age estimation techniques. Within the cortex of bone, BSUs comprise secondary osteons.
Bone is remodeled by a complex arrangement of cells, collectively called the basic multicellular unit (BMU). Intracortical BMUs tunnel through long bone diaphyses in a nearly longitudinal orientation (Hert et al. 1994). The leading region of the BMU is lined with osteoclasts—specialized cells capable of bone resorption. The diameter of the tunnel excavated by osteoclasts, which typically reaches roughly 250–300 μm, defines the cross-sectional size of the osteon that will form in its wake. A histological section that transects the resorptive phase of a BMU (the cutting cone) will exhibit a cavity with rough, scalloped edges (Howship’s lacunae), known as a resorptive bay (Figs. 5.2a and 5.3). Howship’s lacunae are characteristic of all actively resorbing bone surfaces.

Following closely behind the osteoclasts is a group of mononuclear cells. The exact function of these cells remains unclear (Eriksen and Langdahl, 1995). Mononuclear cells line the resorptive bay during the reversal phase (the period between resorption and formation). It is likely that they smooth off the scalloped periphery of the resorptive bay in preparation for the deposition of a reversal line—a thin, mineral-deficient, sulfur-rich layer of matrix that separates an osteon from surrounding interstitial lamellae (Schaffler et al., 1987). An appreciation of reversal lines is particularly important in applying histomorphometric age estimation techniques because their presence can be used to differentiate secondary osteons (a product of remodeling) from primary osteons (a product of modeling).

Behind the mononuclear cells, rows of osteoblasts adhere to the reversal zone and deposit layers of osteoid (unmineralized bone matrix) centripetally. The size of the remodeling space constricts as more concentric osteonal lamellae are deposited and mineralized. At a specified point, deposition ceases leaving a Haversian canal in the center.

### CORTICAL BONE HISTOMORPHOLOGY AND AGE ESTIMATION

Several types of osteons exist in compact bone. Many methods employ different osteon types as variables; thus, an understanding of their morphologies and the ability to distinguish among different osteon types are necessary in applying such methods.

During modeling, some blood vessels in the periosteum become incorporated into the circumferential lamellae being deposited, producing primary vascular channels in the cortex. In some primary vascular channels
(non-Haversian canals), a few rudimentary concentric lamellae are deposited around the periphery of the channel, thus producing a primary osteon with a central non-Haversian canal. Others do not deposit lamellae, and they are simply non-Haversian canals (Fig. 5.2a). Note that primary channels with and without concentric lamellae exhibit non-Haversian canals. As bone remodels, fewer non-Haversian canals are present in the cortex. Several researchers have made use of the association between age and the prevalence of non-Haversian canals/primary osteons by incorporating their frequency into age-predicting equations (e.g., Ericksen, 1991; Fangwu, 1983; Kerley, 1965).

Several varieties of secondary osteons exist in cortical bone. As mentioned, secondary osteons (including all subtypes) can be distinguished from primary osteons by the existence of a reversal line at the periphery of the former. Common secondary osteons (also called type I) are formed by BMUs as described in the previous section (Figs. 5.2 and 5.3). Their number tends to increase with age, and many existing methods employ the number of type I osteons in a specified area or per unit area as an age-predicting variable (see Appendix B). One must pay particular attention to the definition of secondary osteons (and of all other structures) provided in the method intended for use. Table 5.1 provides examples of five different sets of criteria for counting a structure as a secondary osteon. Proper use of a method requires strict adherence to the definitions.

As age increases, the cortex becomes crowded with secondary osteons. As a result, new osteons remove portions of older ones, thus creating osteon fragments. The number of osteon fragments is also correlated positively with age, which is an observation many researchers have incorporated into their age estimation methods. Some methods (e.g., Stout and Paine, 1992) combine the density (number per millimeter squared) of intact and fragmentary osteons to form a new variable, the osteon population density (OPD). OPD increases with advancing age until an asymptote is reached—a point at which subsequent osteon creations remove all evidence of previous ones. When the cortex reaches asymptote, OPD cannot increase anymore. The asymptote imposes obvious limitations on histological methods, the ramifications of which are discussed below.

Other age estimation variables derived from type I osteons include mean number of lamellae per osteon (Singh and Gunberg, 1970); mean osteon area (Iwamoto and Konishi, 1982; Narasaki, 1990; Thompson, 1979; Yoshino et al., 1994) and perimeter (Narasaki, 1990; Thompson, 1979); and mean Haversian canal area (Singh and Gunberg, 1970; Thompson, 1979; Yoshino et al., 1994), diameter (Balthazard and Lebrun, 1911; Iwamoto and Konishi, 1982; Rother et al., 1978; Samson and Branigan, 1987), perimeter (Narasaki, 1990; Thompson, 1979), and population density (Samson and Branigan, 1987). Reported associations with age exhibited by some of these variables (and others described below) vary and are still open to question (Table 5.2).

Type II (embedded) osteons are smaller versions of type I osteons that form by radial remodeling of a preexisting Haversian canal (Jaworski et al., 1972; Richman et al., 1979).

---

1Thompson (1979), for example, includes primary osteons in his Haversian canal counts.
The completed type II osteon appears as a small Haversian system—complete with a reversal line and concentric lamellae—embedded completely within a larger, parent osteon (Fig. 5.2a). Some authors (Ericksen, 1991; Yoshino et al., 1994) report that the number of type II osteons/mm² increases with age, a factor they have incorporated into their age-predicting equations. Richman et al. (1979), however, report no significant association with age.

Double-zonal osteons exhibit a hypercalcified ring within their concentric lamellae, demarcating a point during the formation phase where matrix elaboration temporarily ceased (Fig. 5.2b). Double-zonal osteons typically exhibit two "zones" of dissimilar radiopacity on microradiographs, separated by the hypercalcified ring (arrest line) (Pankovich et al., 1974; Stout and Simmons, 1979). They can be distinguished from type II osteons by their lack of an internal reversal line and by the parallel contours of lamellae and osteocyte lacunae between the inner and the outer zones. The observation by Yoshino et al. (1994) that the number of double-zonal osteons/mm² squared decreases with age has been incorporated into their predicting equations. Pankovich et al. (1974), however, report an increase with age, and Stout and Simmons (1979) report no change.

Drifting osteons form from BMUs that simultaneously travel longitudinally and transversely through the cortex, a process that results in a transversely elongated osteon exhibiting a hemicyclic lamellar "tail" (Robling and Stout, 1999) (Fig. 5.2a). Despite their documented association with age (Coutelier, 1976; Sedlin et al., 1963), drifting osteons have not been incorporated in published age estimation techniques. They are, however, prominent features employed in the subadult aging method proposed by Streeter (2005) described below.

An alternative method that avoids distinguishing among the different types of osteons and their frequency was first proposed by Ahlqvist and Damsten (1969), who measured the percent of the microscopic field occupied by remodeled bone, i.e., by bone of any secondary osteon subtype. This approach purports to reduce interobserver error associated with counting the number of specific microstructures in a given field (see Lynnerup et al., 1998). Several subsequent methods have used percent remodeled bone as a variable in their age estimation equations (e.g., Uytterschaut 1985, 1993). Ericksen (1991) and Cool et al. (1995) use a variation of this approach by quantifying

### TABLE 5.2 Variation in Reports of Association Between Histomorphological Features and Age

<table>
<thead>
<tr>
<th>Microstructural Variable</th>
<th>Increases with Age</th>
<th>Decreases with Age</th>
<th>No Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type II osteon population density</td>
<td>Yoshino et al. (1994)</td>
<td>None</td>
<td>Richman et al. (1979)</td>
</tr>
</tbody>
</table>

*Other, less equivocal relations (e.g. OPD and age) are not included. Superscripts denote bone(s) examined: **mandible, *femur, #humerus, *rib, *tibia.*
the percent of bone occupied by the different microstructures (% osteonal bone, % fragmental bone, % unremodeled bone).

Several methods have incorporated other, nonhistological variables, such as sternal rib end morphology (Stout et al., 1994), cortical thickness, volume, weight, and density (Thompson, 1979), and cortical cross-sectional area (Cho et al., 2002). The prevalence of these and other variables discussed above is tabulated in Appendix B. Dudar et al. (1993) report that averaging morphological and histological age estimates provides a higher correlation between estimated and known age and a lower standard error of estimate.

EFFECTS OF INTRINSIC AND EXTRINSIC VARIABILITY ON HISTOMORPHOMETRIC AGE ESTIMATES

Several factors besides age are known to affect cortical bone remodeling and, consequently, histological age estimates (Frost, 1987). Those factors most relevant to bioanthropology and related fields are discussed below.

Sex-Related Variation

Differences in the rate at which cortical bone is remodeled in postmenopausal women and in men of the same age have been known for some time (e.g., Heaney et al., 1978a, 1978b; Oursler et al., 1996). An increase in bone remodeling occurs in women immediately after the onset of menopause (Parfitt, 1979). What is less clear, however, is 1) whether this geriatric sex difference has an appreciable effect on the overall accuracy of age estimates derived from pooled-sex equations and 2) whether differences exist in males and females of younger (premenopausal) ages. The data addressing these issues are equivocal. Kerley (1965) found no significant sex difference in his sample of 126 lower limb bones, nor did Stout and colleagues (Stout and Paine, 1992; Stout et al., 1994, 1996) in their samples of ribs and clavicles. Ahlqvist and Damsten (1969) and, later, Uytterschaut (1985, 1993) did not investigate sex differences, citing Kerley’s conclusion that they do not exist. Singh and Gunberg (1970) began their study with a sample of 52 males and 7 females, but subsequently dropped the females from the reference sample for use as an independent test sample. The male-based predicting equations produced errors “within experimental and measurement error and [were] not significant” when tested on the 7 females (1970:376). Yoshino et al. (1994), Iwamoto and Konishi (1982), and Samson and Branigan (1987) also used reference samples composed solely of males.

Thompson (1979) seems to have been the first author to present sex-specific predicting equations. He later (Thompson, 1981) tested their use on a series of 54 autopsy specimens, and despite the lower standard errors obtained for males in his original article, he obtained better results for females in the autopsy series. The pooled-sex equations were not tested. Ericksen (1991) found that sex-specific equations gave better results than those based on pooled sexes. She notes that females accumulate intact osteons until the sixth decade of life, compared with the tenth decade in males. Sex differences in fragmentary osteon accumulation were also noted.

In addition to the rate at which osteons are created, male–female differences in osteon size could also have an effect on age estimates. For example, female Pecos Indians exhibit greater mean femoral osteon area (0.041 mm²) than their male counterparts (0.034 mm²) (Burr et al., 1990). This difference potentially could produce greater age estimates in males than in females of the same age, because an average millimeter squared of cortex can accommodate fewer large osteons than small osteons. The packing arrangement, however, would also have an effect on both the observed OPD and the percent remodeled bone (Frost, 1987). Additionally, Burr et al. note that osteons seem to increase in size with advancing age among females, whereas in males, they decrease. Exactly the opposite conclusion was
reached by Broulik et al. (1982) in a modern sample. The apparently conflicting results could be from population differences or from the sampling location (Broulik et al. sampled cortex from the ilium rather than from a long bone). Recent data on osteon size in samples from twentieth-century South Africans, nineteenth-century Canadian settlers, and eighteenth-century English suggest that no significant sex differences exist in osteon size in the femoral midshaft or in the sternal one third of the sixth rib (Pfeiffer, 1998).

Skeletal modeling differences between males and females could also affect age estimates derived from pooled-sex equations, particularly at the younger (18–22 years) age range. The well-documented differences in maturation rates and in the timing of growth cessation (e.g., epiphyseal union) between males and females (e.g., Tanner, 1978) could reflect differing ages at which the adult cortex is completed by modeling. If females complete their adult cortex at an earlier age, i.e., they have a greater effective age of adult compacta, then the mean tissue age—and consequently the amount of time osteons would have had time to accumulate in the cortex—would differ between a young man and woman of the same chronological age. Currently, no data are available that address sex differences in effective age of adult compacta, but future research could elucidate them, and their effect, if any, on histological age estimates.

Physical Activity

Studies of animal models have demonstrated that intracortical bone remodeling can accelerate as a result of increased mechanical loading (Bouvier and Hylander, 1981; Hert et al., 1972; Lieberman, 1997). It is likely that skeletal strain levels, distributions, and frequencies experienced by past populations were variable from population to population and were different from those experienced by industrialized, modern populations.

Technological developments, landscape topography, and food procurement strategies all contribute to the strain milieu experienced by the skeleton. Ruff (this volume) presents data indicating that past populations practicing different subsistence economies often exhibit differences in the gross geometry of their long bones that may be strain related. These differences in gross geometry are also manifested intracortically (Burr et al., 1990). Age estimates in highly physically active populations that are based on regression equations from more sedentary (modern) individuals, therefore, could exhibit significant bias.

For example, in an analysis of two Late Woodland ossuaries from southern Maryland, Ubelaker (1974) found markedly different age estimates using Ahlqvist and Damsten’s (1969) method and methods based on the gross metamorphosis of the pubis. He notes that individuals comprising Ahlqvist and Damsten’s reference sample (modern Finns) may have been less physically active than individuals from the ossuaries and that this difference may account for the greater histological age estimates in the Indian samples. Additionally, Walker et al. (1994) have shown that OPD in the midshaft femur correlates more strongly with the cross section’s bending rigidity (a hypothesized indicator of physical activity) than with age among older (50+) individuals.

Severe disuse also can have profound effects on bone remodeling. Animals with artificially induced limb impairments exhibit a dramatic increase in bone turnover immediately after onset of disuse (Uhthoff and Jaworski, 1978). Stout (1982) reports histomorphometric differences between totally immobilized limbs and those retaining partial use in a 51-year-old quadriplegic female. However, severe disuse states often can be recognized in skeletal remains by grossly visible atrophy of the cortex.

In addition to its effect on remodeling, physical activity levels also seem to affect modeling. Lieberman (1996) has shown that in pigs, increased exercise promotes modeling throughout the skeleton, not just in those bones being loaded. Consequently, in physically active individuals or populations, mean tissue age
could be less than age-matched individuals from a more sedentary population. Histomorphometric age estimates in the former population could be lower than actual age because secondary osteons would have had less time to accumulate in the cortex and because older osteons would be removed by subsequent modeling drifts. The interaction between lower mean tissue age and increased remodeling resulting from a more physically active lifestyle is currently unclear.

Population Variation

Many organs in the human body exhibit substantial interpopulational variation in metabolism (e.g., melanin production by the integument), and the skeleton seems to be among them. It is this attribute, coupled with the fact that bone microstructure provides a record of past metabolic events, that makes histological investigations of past and present skeletal populations so fruitful. However, population-level differences in bone remodeling dynamics potentially can lead to complications in age estimates when equations based on one population are applied to others. Some examples of population differences in bone remodeling follow:

Weinstein and Bell (1988) found that African Americans exhibit lower trabecular bone turnover rates than U.S. whites, an observation later confirmed in cortical bone by Cho et al. (2002) in a series of historic African-American skeletons from Missouri. Ironically, black Africans exhibit bone turnover rates more similar to whites than to African Americans (Schnitzler, 1993). However, both African Americans and black Africans show decreased (compared with U.S. whites) susceptibility to osteoporotic fracture (Farmer et al., 1984; Solomon, 1979), which is a puzzling finding that could reflect greater bone quality in blacks than in whites.

Ericksen and Stix (1991) tested Ericksen’s wedge technique on a series of interred nineteenth-century African-American skeletons from Philadelphia. Their results compared favorably with macroscopic age estimates. Interestingly, most individuals whose age estimates fell outside one standard error were overaged by the histological method, which is a surprising result given the slower turnover noted for African Americans. The inclusion of a significant number of blacks (from the Dominican Republic) in Ericksen’s reference sample may partially account for the overestimates, whereas the comparison of bone remodeling rates for samples of African-American and European-American descent by Cho et al. (2006) could explain the resulting underestimates. It should also be noted that Cho et al. (2002) had exact age at death data from burial markers, whereas Ericksen’s histological estimates could be compared only with macroscopically estimated ages (see Aiello and Molleson, 1993; Pfeiffer, 1980).

Eskimos, on the other hand, seem to exhibit greater bone turnover rates than U.S. whites. Thompson and Gunness-Hey (1981) reported that their sample of femoral cores from a series of nineteenth-century Eskimo femora exhibited greater osteon densities than cores from modern (autopsy and cadaver) U.S. whites. When Thompson and Cowen (1984) estimated age in two Eskimo mummies using pubic morphology and Thompson’s (1979) histological technique, they found that histological age estimates exceeded mean macroscopic estimates in both individuals. Eskimos also seem to exhibit greater turnover than other Native American populations, including Arikara and Pueblo Indians (Ericksen, 1973; Richman et al., 1979). Thompson and Gunness-Hey conclude that population-specific equations are required to overcome genetic differences in bone remodeling, which is a recommendation supported by the poor age estimates they obtained using Thompson’s (U.S. white-based) technique on the large Eskimo sample.

Despite Kerley’s success in applying his method (Kerley, 1965) to skeletal remains from different populations (interred Indian skeletons from Kentucky, Florida, Virginia, and the Aleutian and Philippine Islands), Ubelaker (1977, 1986) found that Kerley’s
method produced errors averaging 11 years when applied to a modern sample of black and mulatto femora from the Dominican Republic. Fangwu (1983) applied Kerley’s method with a sample of 35 modern Chinese femora. He found that the 42% of the individuals’ ages were estimated correctly to within ±5 years, 25% were between ±5 and ±10 years, and the remaining 33% produced estimates with errors greater than ±10 years. Watanabe et al. (1998) undertook a study to estimate age from the femur of Japanese. Their sample from cadavers included femora from 72 males and 26 females with age ranges of 43 days to 92 years (mean age = 50.38 ± 21.53) and 2 to 88 years (mean age = 48.85 ± 28.54), respectively. Predicting variables included the area, maximum and minimum diameter, and perimeter of complete osteons and Haversian canals, the number of type II and fragmentary osteons, and the “area of a triangle” (the area of a triangle formed by lines drawn from three adjacent Haversian canals). The authors report a multiple correlation coefficient of 0.9484 and differences between known estimated ages ranging from 0.5 to 7.2 years. Aiello and Molleson (1993) tested Kerley’s method on a series of eighteenth- and nineteenth-century skeletons of known age from London. Although Kerley’s method was biased, underestimating true age in most (70%) individuals, 65% of the estimates were reasonably accurate, producing age estimates within ±10 years of actual age.

Given the evidence for population-level variation in bone remodeling dynamics, population-specific histological age estimation methods should provide greater accuracy. Most histological age-predicting methods currently available are based on samples derived primarily from individuals of European biological ancestry. Pratte and Pfeiffer (1999) found that histological age predicting equations developed on predominantly European descent Midwestern North American cadaver samples performed very poorly when applied to samples from South Africa representing diverse biological ancestry. The results of this study, however, were complicated by the fact that the South African cadaver sample had a relatively high mean age (63.5 years), which can adversely affect histological age estimation. However, a dearth of histological methods are applicable to skeletal remains of different biological ancestry. Cho et al. (2002) offer a population-specific histological age estimating method for the rib that can be applied when biological ancestry is known to be African or European American. Their equation is derived from a sample of 154 ribs (103 African American and 51 European American) obtained from an African-American cemetery, forensic cases, and autopsies. This method is modeled after Stout and Paine (1992), with the addition of mean osteonal area (On.Ar) and relative cortical area (Ct.Ar/Tt.Ar) as predicting variables. The equation includes an indicator variable for biological ancestry. A separate age predicting equation is also provided for use when biological ancestry is unknown. Appendix B provides some examples of age estimating methods for samples of various biological ancestries.

Estimating Age in Subadults and the Elderly. As with any method, those developed for histological age estimation are limited to age ranges included in the sample from which they are derived. In addition, no histological methods are applicable to the entire human age span; age estimation for the very young and the elderly are particularly problematic. Because of the affects of growth and modeling discussed above, the pattern for age-associated cortical bone histomorphometry is unique for subadults. In addition, remodeling rates are higher and more variable for subadults than for adults. For example, the range for activation frequencies (birth rate of osteons) for sixth ribs between the ages of 9 and 23 years of age is reported to be 14.06/mm²/yr to 3.74/mm²/yr, whereas the range for activation frequencies between ages 32 and 59 years is 1.34/mm²/yr to 2.14/mm²/yr (Pirok et al., 1966). Most of the histological age estimation methods available today have been
developed on samples that include very few, if any, subadults. These methods, therefore, are not applicable to most subadult bones.

Streeter (2005) addresses this limitation for histological age estimation. Using a sample of 72 subadult rib autopsy samples with a mean age of 14.1 ± 5.2 years and a range of 2–21 years, she describes the systematic changes in the microstructure of the subadult rib cortex and provides a method for age at death estimation. Her histological method is qualitative in nature and provides four distinct age categories based on the amount and location of woven bone, circumferential lamellae, secondary osteons, drifting osteons, primary vascular canals, and primary lamellar bone (Table 5.3). Streeter also reports that the association between age and OPD for the rib is not significant for ages younger than approximately 15 years. It is noteworthy that this is similar to the effective age of adult compacta suggested by Wu et al. (1970). A sex difference was observed only after roughly 16 years of age, probably reflecting the effects of puberty on bone growth and development.

Agnew (2006) tested Streeter’s (2005) age categories on a sample of subadult ribs from a medieval (eleventh to twelfth century) Polish cemetery. All four phases were identifiable, but when the phases determined for the Polish sample were compared with age estimates from standard osteological methods, including dental eruption, epiphyseal closure, and diaphyseal length, there was only 38% agreement between these age estimates and the original age ranges assigned to the histomorphological phases by Streeter. When the Polish skeletal samples were seriated using diaphyseal length, however, the assigned histomorphological phases corresponded well to increasing diaphyseal length. These results suggest that different rates of growth and development occurred in the medieval Polish sample compared with the modern autopsy population sample used to develop the method. This finding suggests the interesting possibility of using this method to describe and compare patterns of growth and development in archaeological and paleontological skeletal remains. The effects of population variability on this method need additional study.

When dealing with older skeletal remains, our ability to estimate age histologically is limited by the age at which a given bone reaches its OPD asymptote as discussed above. When confronted with skeletal remains that are older than this age, histological methods can only tell us that the individual in question is the predicted age or older. However, it is possible to use additional histomorphological evidence to estimate whether an individual was close to the asymptotic age or older. Stout and Streeter (2004) provide an example of this approach. The age at death for Janaab’Pakal, the Mayan ruler from the temple of inscriptions at the site of Palenque, is uncertain. His estimated age at death based on standard osteological methods is reported to be 40–50 years. However, epigraphic analysis of inscriptions from the site indicates that he was 80 years old at the time of death. A histological analysis of a rib sample was undertaken to resolve the issue of age at death for Pakal. The method of Cho et al. (2002) produced an age estimate of 52 years. This age is near the age at which OPD asymptote occurs for the rib; therefore, it is consistent with an age of 52 years or older. To ascertain whether Pakal was indeed in his sixth decade of life, or older at the time of death, his osteon size (On.Ar), absolute (Ct.Ar), and relative (Ct.Ar/Tt.Ar) cortical cross-sectional area were compared with values established for a small sample (N = 7) of ribs from the site of Palenque. The absolute and relative cortical areas of 12.7 mm² and 0.34 mm², respectively, for Pakal’s rib are both more than two standard deviations below mean values for other rib samples from Palenque. The mean size of Pakal’s osteons (On.Ar) is 0.014 mm², which is also well below the means for the other Palenque samples. These results suggest an elderly age. When compared with modern values for the rib (Pirok et al., 1966; Wu et al., 1970), Pakal’s histomorphometry is consistent with that of an individual in their eighth or ninth decades of life. These histological results,
therefore, support the age at death provided by the inscriptions.

Pathological Conditions

Several pathological conditions can affect bone modeling and remodeling, and consequently, histological age estimates can be affected. Metabolic disturbances are manifest in the skeleton via several pathways. For example, osteogenesis imperfecta (OGIM) is associated with an increase in the number of osteons created per year (activation frequency). This increase would have the effect of overestimating true age in individuals with OGIM. However, OGIM is also associated with sustained modeling in adulthood. Consequently, mean tissue age in adults with OGIM is less than that in age-matched normals. The lower mean tissue age has a profound effect on observed osteon populations: despite an accelerated activation frequency, individuals with OGIM exhibit subnormal OPDs (Wu et al., 1970). Thus, histological age estimates in individuals with OGIM could be biased toward a younger age.

Hyperparathyroidism (HP), on the other hand, is associated with an increase in activation

| TABLE 5.3 Subadult Age Categories Based on Cortical Bone Histomorphology* |
|-----------------------------|----------------|----------------|----------------|
| Phase I 5-9yo               | Remodeling     | Woven bone     | Cutaneous cortex |
| Primary Lamellar bone       | Rare           | Most of both cortices | Thinner, mostly woven bone, many primary vascular canals |
| Rare, small areas on pleural Endosteal and cutaneous periosteal surfaces initially | Rare | | Thicker, some woven bone, primary lamellae form initially endosteally |
| New primary lamellar bone on pleural cortex | Large drifting osteons on pleural cortex originating at periosteum | Some areas on cutaneous cortex, rare woven bone on pleural cortex | Thinner, mostly intracortical woven bone, many primary vascular canals |
| Phase III 10-17yo           | Drifting osteons on both cortices | Thin rind on cutaneous periosteal surface | Thinner, mostly lamellar bone with some remodeling, periosteal woven bone, large resorptive bays (drifting osteons) |
| Both cortices intracortically | Thinner, mostly lamellar bone on pleural cortex | Thicker, mostly primary lamellar bone, few drifting osteons, many Volkmann’s canals |
| Phase IV 18-21yo            | Both cortices several rows thick, fewer drifting osteons, more Type I osteons | None | Thinner, denser osteons 3 to 4 rows thick, primary lamellar bone periosteally |
| Both cortices periosteally  | Thinner, denser osteons 4 to 5 rows thick, occasional areas of primary lamellar bone |

*Table provided by Margaret Streeter.
frequency, with no known affect on bone modeling. Individuals suffering from HP usually exhibit greater osteon counts than age-matched normals; thus, histological age estimates can be affected. HP (and especially secondary HP) is a particularly relevant condition when studying bone remodeling in extinct populations because many groups are known to have suffered from dietary stresses. An inadequate supply of dietary calcium or a reduction in the body's ability to absorb calcium from the gut (e.g., vitamin D deficiency) can result in low levels of serum calcium—a factor known to trigger excessive release of parathyroid hormone and subsequently increase osteoclast differentiation, i.e., increase the number of active BMUs. Cook et al. (1988) have presented data supporting a diagnosis of hyperparathyroidism in a 50–60-year-old female excavated from a Roman Period site in central Egypt.

Several local noxious stimuli can affect age estimates by increasing the remodeling rate locally. Frost (1983) has called this factor the “regional acceleratory phenomenon” (RAP). Fractures, bone infections, local circulatory complications, and many other factors can result in RAP. The effects of RAP are local; thus, they may not be detectable from biopsies at distant sites. The effect of RAP on histomorphometrics—and on derived age estimates—can go unchecked if a small tissue sample from a RAP-affected area is analyzed. This risk can be reduced by reading a large number of fields distributed throughout a cortical cross section. Some age-estimation techniques that employ a small sample of tissue (e.g., core or wedge) have been criticized by Stout (1989, 1992) on exactly these grounds.

Some ways to evaluate whether a pathological condition may be affecting an age estimate follow:

1. Scan the cross section for any signs of abnormal histomorphology, including abnormal modeling drifts, and the presence of woven bone. The latter condition can be pathognomonic of several metabolic disorders, including Paget's disease, renal osteodystrophy, or OGIM.

2. Consider the cortical thickness from which the thin section comes. High turnover rates on the endosteum can result in reduced cortical thickness, which is a common symptom of osteoporosis-related disorders.

3. If possible, analyze a sample of trabecular bone (preferably from the ilium) from the same individual. Because of its much greater surface-area-to-volume ratio, trabecular bone usually exhibits the effects of systemic metabolic bone disease more rapidly and to a greater extent than cortical bone. Abnormal values of trabecular bone volume, surface density, mean BSU wall thickness, and total resorptive surface can be diagnostic of certain disease states. These static parameters are all measurable in well-preserved archaeological remains (Cook et al., 1988), and reference values for age-matched normal individuals are available in the literature (Lips et al., 1978; Merz and Schenk, 1970). Moreover, needle biopsy can be used to remove small bone cores (e.g., from the ilium) with very little damage to the specimen (Weinstein et al., 1981). Age estimates should be suspect when trabecular parameters deviate substantially from age-matched normal values.

Undoubtedly, individuals affected with conditions known to affect histomorphometry have made their way into many samples from which predicting equations have been generated. Ericksen (1991) purposefully included individuals in her reference sample that were found to have pathological conditions known to affect bone remodeling, such as diabetes and chronic renal disease. Her aim was to produce age-estimation equations that included conditions that are likely to be encountered in unidentified individuals. Thompson (1979)
provides specific equations for a pathological subset of his sample. However, given that pathological conditions can either accelerate (e.g., hyperparathyroidism) or retard (e.g., diabetes) remodeling rates, their utility is questionable.

A study by Paine and Brenton (2006) illustrates the possible effects of a particular pathological condition on histological age estimation. They evaluated the use of the Stout and Paine (1992) histological age estimating formula on a sample of ribs from 26 black South Africans of known age at death for whom records indicate cause of death to have been from malnutrition or the niacin deficiency disease pellagra. Age estimates produced by the histological method averaged 29.2 years below known age. These results illustrate the potential impact of pathology on the accuracy of histological-based methods based on healthy reference populations. Since the study did not include healthy individuals from the same population, however, the study could not address the extent to which the poor age estimation results were from population variation. It would be interesting to see how the population-specific method developed by Cho et al. (2002) performed on this sample.

Research Methods

The importance of reading an adequate amount of cross-sectional area has been known for at least 30 years: “Two serial cross sections of rib with only 10 mm² cortical area each may exhibit several hundred percent difference in bone formation, whereas two serial cross sections of the femur of the same patient, with more than 300 mm² cortical area each, will differ by less than 5%” (Wu et al., 1970: 216). Thus, the amount of cortical area read per individual is a paramount concern in accurately quantifying the remodeling history from a region of bone.

Sampling area is a difficult issue to deal with in histological investigations of archaeological or fossil bone because permission to remove complete cross sections rarely is granted. These considerations have lead to the development of techniques that do not require complete cross sections and that leave the original specimen largely intact. Thompson’s (1979) method requires only a small (0.4 cm in diameter) core of bone. Ericksen’s (1991) method involves removal of a small wedge (1 cm in transverse width) of bone from the anterior cortex, providing a larger cross-sectional area than that obtained by Thompson’s technique. The sampling technique that provides the greatest amount of cross-sectional area yet leaves the original specimen intact is described in Aiello and Molleson (1993). At the femoral midshaft, they made two parallel transverse cuts, starting from the anterior surface, extending about three quarters of the way through the bone. Using a chisel, they removed a c-shaped block, from which sections—exhibiting all but the posterior quadrant—were prepared. The original shape, length, and most of the diaphysis of the femur were thus retained for future morphological study. This sampling technique is recommended when complete cross sections cannot be removed.

The distribution of histological structures in a typical long-bone cross section is not uniform, a fact that has been known for some time. Jowsey (1966) showed that the regions of bone near the marrow cavity are remodeled more heavily than more peripheral (subperiosteal) regions. Drusini (1987) reported on the great variability in osteon density around the periphery of the femoral cortex (see also Pfeiffer et al., 1995). Thus, it is imperative that the exact topographic sampling location described in a particular method is followed precisely. This caveat can impose severe limitations on a method in archaeological investigations, if for example the region or quadrant of bone necessary to apply a particular method is missing. More commonly, significant wear on the periosteal surface caused by taphonomic processes removes subperiosteal bone (of an unknowable depth) from the remains. Most age estimation methods are based on the subperiosteal cortex (see Appendix B), so the application of methods based on the subperiosteal cortex to remains exhibiting significant wear can produce erroneously high age estimates. Pfeiffer (1992) and Ericksen (1997)
encountered this very problem when attempting to estimate age at death in samples of interred nineteenth-century Canadians and Preceramic Chileans, respectively.

To circumvent the influence of significant periosteal wear on age estimates, Stout and colleagues (Stout, 1986; Stout and Paine, 1992; Stout et al., 1994, 1996) have incorporated into their methods a topographic sampling procedure where every other field in the entire section is read. This procedure results in a “checkerboard” sampling pattern, which has the advantage of not relying on any one region of the cross section for histomorphometrics. Although this procedure works well for bones with minimal cross-sectional area (e.g., rib or clavicle), it may be impractical in larger bones (e.g., femur or tibia).

For larger bones, age estimation methods that sample the endosteal region of the cortex would be of particular use in archaeological contexts. Such a strategy avoids complications introduced by significant periosteal wear—a factor that seems to be the rule, rather than the exception, in many archaeological collections. Hauser et al. (1980) presented age estimation equations based on the juxtamedulary (endosteal) region and on the subperiosteal region of the femur and tibia. Correlation between the juxtamedulary histomorphometry and age was greater than that between the subperiosteal histomorphometry and age—a result challenging the observation that the deeper cortex does not reflect age changes as accurately as the peripheral cortex (Aiello and Molleson, 1993; Kerley, 1965). Additional support for the use of the endosteal region in age estimation methods is presented in Pfeiffer et al. (1995), who found that fields at or near the endosteum were similar to those located subperiosteally, but both locations were significantly different from fields in between the subperiosteal and the endosteal fields. Given their utility in archaeological material, age estimation techniques that employ endosteal regions of the cortex warrant additional investigation.

Additionally, statistical methods used in producing age-predicting equations can have a significant effect on age estimates. Most methods present equations derived from the “classic” regression model, in which the quantified histological variable (e.g., OPD) is the dependent variable and is regressed onto age at death, the independent variable. Rogers and Stout (1998) report that classic regression can reduce bias and provide more accurate age estimates for histological methods derived from samples for which correlation coefficients are moderate; i.e., $r \approx 0.82$. Inverse regression was found to provide better results when correlation is high ($r \geq 0.82$) or low ($r \leq 0.65$). These observations are based on a limited set of data, and additional research is needed to confirm their utility in other existing and new methods.

### EVALUATION OF HISTOLOGICAL AGE ESTIMATION METHODS

Unlike macroscopic methods, histological age estimation methods have not been subjected to testing of their accuracy and precision, particularly with reference to their application outside their reference samples. For histological methods, the literature is confusing and often conflicting regarding recommended methods and reported rates of precision and accuracy (Crowder, 2005). Considering the requirements for acceptance of evidence in courts of law such as the Frye and more recent Daubert standards, this situation is of significant concern.

Willows (1991) undertook a study to assess and compare the accuracies of the femoral methods of Kerley and Ubelaker (1978) and Thompson (1979). She found that the two methods produced different age at death estimates. The Kerley and Ubelaker method provided the best age-estimating accuracy for the complete sample, four subsamples composed of males and females, and individuals either younger or older than 45 years of age.

Several studies compare the Kerley method (Kerley and Ubelaker, 1978) with other methods. Bouvier and Ubelaker (1977) compared the precision and accuracy of the Kerley and by Ahlqvist and Damsten (1969) methods for age estimation for the femur. Both methods were found to be equally precise,
based on comparable intra- and interobserver errors, but the Kerley method was found to be more accurate. However, this comparison of accuracy is problematic for two reasons. First, only Kerley’s formula for complete osteons was employed. Second, the test sample used in the comparison was composed of bone sections obtained from the original sample used by Kerley to develop his method and, therefore, does not represent an independent sample. Bekaert (2004) also compared the performance of the Kerley (Kerley and Ubelaker, 1978) and Ahlqvist and Damsten (1969) methods for estimating age from the femur and concluded that the modified Kerley method was more accurate and easier to perform. Stout and Gehlert (1980) compared the relative accuracy and reliability of the Kerley (Kerley and Ubelaker, 1978) methods for the femur, tibia, and fibula and that of the Ahlqvist and Damsten (1969) for the femur using a small independent sample of 13 individuals of known age at death. Kerley’s femoral intact osteon formula produced the greatest accuracy for individuals 51 years of age and younger, whereas the fibula osteon fragment formula was most accurate for older ages (≥60 years). Age estimates determined by averaging the estimates for all Kerley predicting formulas produced the greatest accuracy and reliability for all age classes.

Crowder (2005) reviews the literature relating to the evaluation of histological methods and provides a model for the evaluation of methods. He evaluates multiple histological methods to determine accuracy, precision, and ease of application for six histological age estimation methods, three for the rib (Cho et al., 2002; Stout, 1986; Stout and Paine, 1992) and three for the femur (Ericksen, 1991; Singh and Gunberg, 1970; Thompson, 1979). The study sample is composed of 258 individuals from the early nineteenth-century cemetery at Christ Church, Spitalfields from London. It concludes that the Cho et al. (2002) method is the most accurate rib-based method, whereas for the femur, both Thompson (1979) and Ericksen (1991) methods performed similarly. The Thompson method, however, is recommended because of its lower interobserver error and applicability from its requirement of only a 0.4-cm-diameter core of bone. The latter advantage is exemplified by the study by Streeter et al. (2001) in which the formula of Thompson and Galvin (1983) was used to estimate the age at death for the Middle Pleistocene hominin fossil remains known as Boxgrove to the fourth decade of life using only a medial mid-shaft fragment of the tibia. Crowder’s study also found that the rib methods provide the best accuracy and reliability for those with minimal experience in histomorphometry, and similar to the findings of Stout and Gehlert (1980), the method of Singh and Gunberg (1970) was found to be exceptionally inaccurate. Finally, he reports that the femur may be more accurate than the rib when the age of the individual is over 50 years, and that coupling the age estimates from the rib and femur does not improve accuracy because the two bones do not perform equally for all age cohorts. These findings by Crowder are in disagreement with Stout and Gehlert (1980), who report that combining estimates from different bones of the lower extremity provides better accuracy and reliability.

CONCLUSIONS: FUTURE DIRECTIONS AND CONSIDERATIONS IN HISTOLOGICAL AGE ESTIMATION

It is clear from the preceding discussion that several factors—both physiological and methodological—are germane to the proper use of histomorphometric age estimation techniques. Advances in bone biology and histology made over the past 30 years have elucidated many factors that have known effects on bone remodeling dynamics. Knowledge of their effects was, for the most part, not available to Kerley and other early pioneers of histological age estimation methods. Today, efforts should focus on finding ways to account for some variation in age estimates resulting from these influences, i.e., differences in physical activity levels, sex, genetic differences, and states of health. We suggest some guidelines and
potentially fruitful areas of investigation for those endeavoring to develop new age-estimation methods and some issues to consider for consumers of existing and forthcoming methods.

1. The composition of the reference sample used in any new method should be reported in as much detail as possible, including total \( n \), the \( n \) for each sex, age ranges and mean ages for each sex, and the ethnic composition of the sample. Given that the age range and distribution of the reference sample can bias considerably the age estimates produced for other samples (Lazenby, 1984; Masset, 1976), a complete description of the reference sample should be provided so those using the method are aware of potential biases they may encounter. Additionally, given the histomorphometric variability in sampling location (Drusini, 1987, 1996; Iwaniec, 1997; Pfeiffer et al., 1995), the exact topographic sampling technique should be described and illustrated. A brief review of Appendix B reveals some sample-reporting deficiencies among several methods currently in use.

2. If ethnically heterogeneous reference samples are used, some indication of the influence of the subgroups on the method should be reported. For example, reporting mean residual scores for different subgroups will give the reader some indication of potential biases and their expected direction, if the method is applied to individuals from a particular population.

3. Examining sections from multiple skeletal elements is recommended for several reasons:

   A. Biomechanical factors affecting bone remodeling seem to be local, usually affecting only those bones being strained (Bouvier and Hylander, 1981; Tommerup et al., 1993). By sampling more than one bone—preferably including bones from both the axial and the appendicular regions—the risk of arriving at an age estimate substantially altered by excessively vigorous or trivial physical activity is reduced.

   B. RAP is also a locally acting phenomenon. Multiple-site sampling also reduces the risk of RAP having a severe effect on an age estimate.

   C. Several authors (e.g., Lazenby, 1984; Martin et al., 1981; Walker et al., 1994; Willows, 1991) have discussed the inability of histological methods to produce accurate age estimates among older (50+) individuals. Most deficiencies can be attributed to the fact that the cortex reaches asymptote in later years. That bones vary in their baseline remodeling rates (Marotti, 1976) could be used to circumvent the problem associated with the remodeling asymptote. Some slow-turnover bones (e.g., tibia or metatarsals) may reach asymptote at a later age than other, more rapid-turnover bones (e.g., femur or rib), although this remains to be confirmed. Using bones with low and high turnover rates in conjunction should allow more accurate age estimates in samples that include individuals of young and advanced ages.

4. In addition to devising age-estimation methods that employ microstructural data from multiple skeletal sites, we also recommend exploring techniques that incorporate both microscopic and macroscopic variables into one method (e.g., Stout et al., 1994). Studies that have compared microscopic and macroscopic age estimates show that their combined use is usually better than either alone (Aiello and Molleson, 1993; Dudar et al., 1993).

5. Studies of topographic variation in histomorphometry have the potential to overcome curator-imposed limitations on tissue sampling (Pfeiffer et al., 1995). For example, Iwaniec (1997) determined that
95% of the variation in OPD in the anterior quadrant of the femur is accounted for by two 1-mm-wide columns that span from periosteum to endosteum. This discovery significantly reduces the number of fields required to quantify accurately bone turnover in the anterior cortex. However, these results apply only to the anterior quadrant; studies on the remaining three quadrants are needed. Thus, it might be possible to remove three or four bone cores, from the appropriate cortical locations and to quantify accurately bone turnover in the entire cross section (most of which could remain in the intact bone). Iwaniec (1997) reports that in the pig, only 50% of the midshaft femoral cross section need be read to account for over 90% of the variation in the entire cross section. This area of research warrants more investigation, particularly in humans.

6. As discussed, it has become clear that mechanical loading has profound effects on cortical bone microstructure. To augment some differences in age estimates attributable to physical activity, it may be necessary to include in the analysis other variables, such as cortical thickness or normalized second moments of area, that can potentially account for some variability associated with physical activity differences.

7. Both inverse and classic regression should be explored in the development of new age estimation methods. This step has the potential to substantially reduce bias and to increase the accuracy of estimates derived from the resulting prediction equations.

8. Differences in mean tissue age can introduce substantial error into histological age estimates. Investigations of sex and population differences in skeletal maturation and mean tissue age would be particularly enlightening for the bone palaeohistologist, in both age estimations and other investigations (e.g., remodeling rates). Research in this direction has already begun and seems promising (Stout and Lueck, 1995).

9. Age-estimating methods need to be evaluated using independent, known age samples to determine their accuracy and precision, especially if they are to be employed in forensic analyses that ultimately may be presented as evidence in a court of law.

10. Lastly, the histological methods available today have two important limitations; they are invasive and rely on the observation and quantification of histological features from two-dimensional cross-sections. It has been demonstrated that three-dimensional images can be created for histomorphological study (Stout et al., 1999). Cooper (2005) employed high-resolution microcomputed tomography (micro-CT) to visualize and quantify the dynamics of the cortical bone canal network and to evaluate its potential for estimating age at death. Although this study did not improve precision of histological age estimation, it provides groundwork for the development of histological methods that use this technology and paves the way for methods that employ three-dimensional histomorphometry.

REFERENCES


APPENDIX A: WORKED EXAMPLES OF TWO AGE ESTIMATION METHODS

I. Kerley (1965) and Kerley & Ubelaker (1978) method for the femur.

The black ring represents the field of view at 100×. The visible area has been expanded to allow structures traversing the periphery of the field—which were included in Kerley’s counts—to be seen and counted. This is accomplished in practice by moving the stage controller back and forth slightly at each field. Four microscopic fields, one at each of the four subperiosteal fields along the anatomical axes, are read per section. Results for the lateral field illustrated in the photomicrograph (above) are highlighted in the table below. Results from the remaining three fields (not illustrated) are also tabulated below.

The field diameter = 1.68 mm, therefore the field size = 2.22 mm². To correct for differences in field size between Kerley’s microscope and yours, divide your field size into Kerley’s original field size of 2.06 mm². The quotient is a correction factor by which the sums of each variable (e.g., osteons, osteon fragments) are multiplied. In this example, the correction factor = 2.06 mm² / 2.22 mm² = .93. To calculate age at death, insert the corrected sums (Σcorr.) into their respective predicting formulae (see equations below) and apply to the profile chart provided in Kerley (1965), or simply average the two estimates as recommended in Stout & Gebler (1980).

<table>
<thead>
<tr>
<th>Cortex</th>
<th>Osteons</th>
<th>Fragments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anterior</td>
<td>22</td>
<td>9</td>
</tr>
<tr>
<td>Medial</td>
<td>39</td>
<td>15</td>
</tr>
<tr>
<td>Lateral</td>
<td>38</td>
<td>22</td>
</tr>
<tr>
<td>Posterior</td>
<td>27</td>
<td>13</td>
</tr>
<tr>
<td>Σ</td>
<td>126</td>
<td>59</td>
</tr>
<tr>
<td>Σcorr.</td>
<td>117.18</td>
<td>54.87</td>
</tr>
</tbody>
</table>

\[
\text{Osteons: } \text{Age} = 2.28 + 0.187X + 0.00226X^2 = 55.23
\]

\[
\text{Fragments: } \text{Age} = 5.241 + 0.509X + 0.017X^2 - 0.00015X^2 = 59.57
\]

*Mean estimated age = 57.40

*Actual age at death from death certificate = 55 years.

The superimposed grid is a Merz counting reticule containing 36 intersections or "hits." Structures exhibiting ~1/2 or more of their area within the grid are included in the count. Every other field in the section is read, so the total number of fields read will depend on the cross-sectional size of the rib. The field illustrated in the photomicrograph is one of eleven fields read for this particular section, and is recorded as field 7 in the table below. Results from the remaining ten fields (not illustrated) are also tabulated below.

The distance between intersections (hits) = 170 μm (or 0.17 mm); therefore one hit represents (0.17 mm)^2 or 0.0289 mm^2 of bone area. The area read equals 252 hits × 0.0289 mm^2 per hit = 7.28 mm^2. Osteon population density (OPD) is the population of intact osteons (P_I) plus the population of osteon fragments (P_f) divided by the area read.

In this example, OPD = 178 ÷ 7.28 mm^2 = 24.45/mm^2.

To calculate age at death, insert rib OPD into Stout and Paine (1992) rib formula and solve for age:

\[ \ln(\text{age}) = \frac{2.343 + 0.050877 \times \text{OPD}_{\text{rib}}}{\text{OPD}_{\text{rib}}} - 3.59 \]

\[ e^{1.59} = 36.1 \]

Estimated age \( \hat{\text{age}} \) = 36.1 years.

*Actual age at death from death certificate = 38 years.
III. Schematic illustrations of counted structures (shaded) from the previous two examples. Counted intact osteons are labelled with numbers, osteon fragments are labelled with letters.
<table>
<thead>
<tr>
<th>Bone</th>
<th>Method</th>
<th>Sample Parameters: (N); Age Range ({x} ) Sex ({c^x; q^x; ?}) and Ethnic Composition</th>
<th>Specimen Sampling Technique</th>
<th>Variables Used in Analysis</th>
<th>Reported Accuracy</th>
<th>Comments; [references for test of method]</th>
</tr>
</thead>
<tbody>
<tr>
<td>OCCIPITAL</td>
<td>Cool et al., 1995</td>
<td>(17); 21–70 {46.8} [17:0:0] Caucasian</td>
<td></td>
<td>fractional volume of: 1. primary osteons 2. secondary osteons 3. secondary osteon fragments 4. unremodeled lamellar bone</td>
<td>(r^2 = 0.10) (secondary osteons) (r^2 = 0.44) (osteon fragments) (r^2 = 0.32) (lamellar bone)</td>
<td>Both outer and inner tables of cortex were investigated. Authors concluded that the method is not recommended for age estimation because of low correlation of variables with age, even when subjected to multivariate analysis.</td>
</tr>
<tr>
<td>2nd METACARPAL</td>
<td>Kimura, 1992</td>
<td>(227); 30–98 {68} [114:113:0] Modern Japanese</td>
<td></td>
<td>1. cortical bone density 2. cortical thickness 3. total number of osteons and osteon fragments in section 4. total osteon density [OPD]</td>
<td>S.E.E. = 9.99–14.82 (for total sample); 5.51–8.64 (for 30–65 year old group); 5.64–7.68 (for 66–98 year old group)</td>
<td>Microradiographs used. Complete cross section is sampled. Excellent sample size. Author found much better results when sample was analyzed separately in two groups (ages 30–65 and 66–98).</td>
</tr>
<tr>
<td>MANDIBLE</td>
<td>Singh and Gunberg, 1970</td>
<td>(52); 39–87 {64} [52:0:0] Cadavers from US; no ethnic data provided</td>
<td></td>
<td>1. total number of osteons in two microscopic fields 2. mean number of lamellae per osteon 3. mean Haversian canal diameter</td>
<td>S.E.E. = 2.55–3.83 95% of sample within ± 5 yrs. 67% of sample within ± 3 yrs.</td>
<td>Authors found the mandible provided the greatest accuracy of the three bones they investigated. The sample of sections analyzed included both decalcified (33) and undecalcified (19) preparations. Equations based on males only.</td>
</tr>
<tr>
<td></td>
<td>Druzisini and Businaro, 1990</td>
<td>(50) 18–97 {35.3} [32:18:0] Modern Italians?</td>
<td></td>
<td>1. mean number of secondary ostons per (\text{mm}^2) 2. mean number of secondary osteon fragments per (\text{mm}^2)</td>
<td>S.E.E. = 6.42–11.45 6.84–13.19 for pooled sex 5.14–7.09 for females</td>
<td>Singh and Gunberg’s sampling technique is followed, but only undecalcified sections are used. Much better topographical sampling of the section than Singh and Gunberg.</td>
</tr>
</tbody>
</table>
### APPENDIX B  Continued

<table>
<thead>
<tr>
<th>Bone</th>
<th>Method</th>
<th>Sample Parameters: (N); Age Range {X}, Sex [♂:♀:?], Ethnic Composition</th>
<th>Specimen Sampling Technique</th>
<th>Variables Used in Analysis</th>
<th>Reported Accuracy</th>
<th>Comments: [references for test of method]</th>
</tr>
</thead>
<tbody>
<tr>
<td>FIBULA</td>
<td>Kerley, 1965 &amp; Kerley and Ubelaker, 1978</td>
<td>(25); 0–83 {34.5} Caucasoid and Negroid; distribution in sample not provided</td>
<td>Number of 1) intact osteons, 2) osteon fragments, 3) non-Haversian canals, and 4) % unremodeled bone in a 2.06-mm$^2$ field</td>
<td>S.E.E. = 5.27–10.85 for single variables. Accuracy is increased when overlap of age range from all variables is used.</td>
<td>Of the three lower-limb bones Kerley analyzed, the fibula provided the most accurate age estimates. In osteopenic fibula (e.g., from the elderly), the diameter of the microscopic field can exceed the thickness of the cortex (Stout and Gehlert, 1980; Cera and Drusini, 1985).</td>
<td></td>
</tr>
<tr>
<td>ULNA</td>
<td>Thompson, 1979</td>
<td>(31); 7–? {68.8} New England whites</td>
<td>19 variables analyzed. See original article for description</td>
<td>S.E.E. = 7.9–10.6</td>
<td>Core technique minimizes destruction to specimen. Size of area read may increase risk of sampling an atypical field.</td>
<td></td>
</tr>
<tr>
<td>FEMUR</td>
<td>Kerley, 1965 &amp; Kerley and Ubelaker, 1978</td>
<td>(67); 0–95 {41.6} Caucasoid and Negroid; distribution in sample not provided</td>
<td>Number of 1) intact osteons, 2) osteon fragments, 3) non-Haversian canals, and 4) % unremodeled bone in a 2.06-mm$^2$ field</td>
<td>S.E.E. = 9.4–13.9 for single variables. Accuracy is increased when overlap of age range from all variables is used.</td>
<td>Widely used technique. Several authors have expressed difficulty in consistently identifying microstructures (Bouvier and Ubelaker, 1977; Stout and Gehlert, 1980; Cera and Drusini, 1985; Ubelaker, 1986; Wilkows, 1991; Aiello and Molleson, 1993; Walker et al., 1994).</td>
<td></td>
</tr>
<tr>
<td>Ahlqvist and Damsten, 1969</td>
<td></td>
<td>(20); 7–? {55} Autopsy specimens; no ethnic data provided</td>
<td>% remodeled bone in 1-mm$^2$ square field</td>
<td>S.E.E. = 6.71</td>
<td>Small sample size. Variable employed appears less subjective than others (e.g., counting different microstructures) (Bouvier and Ubelaker, 1977; Stout and Gehlert, 1980; Uytterschaut, 1993).</td>
<td></td>
</tr>
</tbody>
</table>
Singh and Gunberg, 1970 (33); 39–87 [62.3]
Cadavers from US; no ethnic data provided

1) total number of osteons in two microscopic fields
2) mean number of lamellae per osteon
3) mean Haversian canal diameter

S.E.E. = 3.24–5.01
Decalcified sections were used in study sample (possible shrinkage). Authors caution that results for femur are preliminary and should be interpreted conservatively (due to small sample size). Equations based on males only (no sex differences reported).

Thompson, 1979 (116); 30–97 [71.7]
New England whites

19 variables analyzed. See original article for description

S.E.E. = 7.1–8.6 (pooled sex); 6.4–8.3 (males) 7.2–9.4 (females)
Core technique minimizes destruction to specimen. Size of area read may increase risk of sampling an atypical field (Thompson, 1981; Cera and Drusini, 1985; Willows, 1991; Pfeiffer, 1992).

Ericksen, 1991 (328) 14–97 [62.8]
251 US whites, 1 US oriental, 12 US blacks, 6 Chilean hispanics, 58 Domin. Repub. blacks

Density of 1) osteons, 2) type II osteons, 3) fragments, 4) resorption spaces, and 5) non-Haversian canals. Also, mean % 6) unremodeled, 7) osteonal, and 8) fragmental bone.

S.E.E. = 10.1–12.2 (pooled sex); 10.1–12.0 (males) 10.0–11.6 (females)
Excellent sample size; ethnic mixture of sample is favorable for applying predicting equations to individuals or populations of unknown origin. Sampling technique is a good compromise between destruction of original specimen and adequate size of tissue sample (Ericksen and Stix, 1991; Ericksen, 1997).

Fangwu, 1983 (35); 5–86 [39.1]
Modern Chinese

Number of intact osteons, osteon fragments, and non-Haversian canals in 100X field, and mean thickness of outer circumferential lamellae

80% of sample within ±5 yrs. 90% of sample within ±10 yrs.
This technique applies Kerley’s research design to the femora of modern Chinese. Resulting equations are more accurate on modern Chinese than are Kerley’s formulas.

(Continued)
### APPENDIX B  Continued

<table>
<thead>
<tr>
<th>Bone</th>
<th>Method</th>
<th>Sample Parameters: (N); Age Range [x]; Sex [♂:♀:♂♀] and Ethnic Composition</th>
<th>Specimen Sampling Technique</th>
<th>Variables Used in Analysis</th>
<th>Reported Accuracy</th>
<th>Comments: [references for test of method]</th>
</tr>
</thead>
</table>
| Hauser et al., 1980 | Modern Europeans | (96); 21–87 [♀] [♂:♀:♂♀]     | ![Diagram](image1.png) | 1. population density of osteons  
2. mean minimum diameter of Haversian canals  
3. % of surface occupied by Haversian canals | S.E.E. = 10.7–11.4 | Separate formulas provided for the subperiosteal region and endocortical region.  
Endocortical formulas could be useful in archaeological cases where periosteal surface is worn (e.g., Ericksen, 1997; Pfeiffer, 1992). |
| Cera and Drusini, 1985; Drusini, 1987 | Modern Italians | (20); 19–50 [28.8] [♀:♀:♂♀] | ![Diagram](image2.png) | number of secondary osteons per mm² | S.E.E. = 3.92 | Small sample size. Large portion of the cortex sampled. Osteon fragments are not included in this method, as they are in Drusini's later (e.g., Drusini and Businaro, 1990) methods. |
| Samson and Branigan, 1987 | Caucasians | (58); 16–91 [♀:♂:♂♀] | ![Diagram](image3.png) | 1) number of Haversian canals per mm²  
2) mean Haversian canal diameter  
3) mean cortical thickness excluding linea aspera | S.E.E. = 6 for males; 16 for females | A potentially useful approach for poorly preserved bones.  
Method purports to allow for age at death determination, but prediction equations are not provided. Criteria used to accept or reject Haversian canals for measurements and counts were decided on by the authors, but not reported.  
Good topographic sampling (Aiello and Molleson, 1993). |
| Narasaki, 1990 | Modern Japanese | (52); 43–98 [♀:♀:♂♀] | ![Diagram](image4.png) | core 1) thickness and 2) weight; 3) intact osteonal bone area; mean secondary osteon:  
4) density, 5) area, 6) σ of area, 7) perimeter, 8) σ of perimeter | S.E.E. = 9.28 (males); 9.95 (females) | This study applies Thompson's core technique and 6 of his 19 variables (plus 2 new ones) to modern Japanese. |
<table>
<thead>
<tr>
<th>Reference</th>
<th>Sample</th>
<th>Methods</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Watanabe, 1998</td>
<td>Modern Japanese</td>
<td>osteon perimeter, Haversian canal length, number of osteon fragments</td>
<td>Variables are averaged for three 0.707744-mm² fields. A second formula using osteon width, Haversian canal length, number of osteon fragments, and area of triangle (see text) produced an S.E.E. of 4.88.</td>
</tr>
<tr>
<td>Kerley, 1965 &amp; Kerley and Ubelaker, 1978</td>
<td>Caucasoid and Negroid; distribution in sample not provided</td>
<td>Number of 1) intact osteons, 2) osteon fragments, 3) non-Haversian canals, and 4) % unremodeled bone in a 2.06-mm² field</td>
<td>Several authors have expressed difficulty in consistently identifying microstructures (Stout and Gehlert, 1980; Cera and Drusini, 1985).</td>
</tr>
<tr>
<td>Singh and Gunberg, 1970</td>
<td>Cadavers from US; no ethnic data provided</td>
<td>total number of osteons in two microscopic fields, mean number of lamellae per osteon, mean Haversian canal diameter</td>
<td>Decalcified sections were used in study sample (possible shrinkage). Authors caution that results for tibia are preliminary and should be interpreted conservatively (due to small sample size). Equations based on males only (no sex differences reported).</td>
</tr>
<tr>
<td>Thompson and Galvin, 1983</td>
<td>41 US Caucasians, 12 US Blacks</td>
<td>core weight, cortical bone density and thickness, secondary osteon number, area, perimeter, and Haversian canal area and perimeter</td>
<td>Thompson's original (1979) method for the tibia produced poor results among individuals below 55. The new method (based primarily on osteon number instead of size) produces more accurate age estimates.</td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Bone</th>
<th>Method</th>
<th>Sample Parameters: (N); Age Range {x}, Sex [♂;♀;?] and Ethnic Composition</th>
<th>Specimen Sampling Technique</th>
<th>Variables Used in Analysis</th>
<th>Reported Accuracy</th>
<th>Comments; [references for test of method]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hauser et al., 1980</td>
<td>Modern Europeans</td>
<td>(31); 18–88 {♀} [♂:♀:?]</td>
<td>1. population density of osteons</td>
<td>S.E.E. = 13.5–16.0</td>
<td>Small sample size; large standard error. Separate formulas provided for the subperiosteal region and endocortical region. Endocortical formulas could be useful in archaeological cases where periosteal surface is worn (e.g., Ericksen, 1997; Pfeiffer, 1992).</td>
<td></td>
</tr>
<tr>
<td>Uytterschaut, 1985; 1993</td>
<td>Cadavers from the Netherlands; no ethnic data provided</td>
<td>(20); 17–92 {♀} [♂:♀:?]</td>
<td>% remodeled bone in 1 mm² square field</td>
<td>S.E.E. = 6.29</td>
<td>This technique applies Ahlqvist and Damsten’s research design for the femur to the tibia. Small sample size may contribute to the relatively low S.E.E.</td>
<td></td>
</tr>
<tr>
<td>Balthazard and Lebrun, 1911</td>
<td>0–84 {♀} [♂:♀:?]</td>
<td>mean diameter of (100–200) Haversian canals in anterior tibial cortex</td>
<td>67% of sample within ± 5 years</td>
<td></td>
<td>First published method for estimating age from bone microstructure. Deslypere and Baert (1958) found poor results using this method on an independent test sample.</td>
<td></td>
</tr>
<tr>
<td>HUMERUS</td>
<td>Thompson, 1979</td>
<td>(31); ? – ? {♀} [♂:♀:?]</td>
<td>19 variables analyzed. See original article for description</td>
<td>S.E.E. = 6.2–9.5</td>
<td>Core technique minimizes destruction to specimen. Size of area read may increases risk of sampling an atypical field. Thompson later (1981) discourages use of humeral cores for age estimations because of innacuracy.</td>
<td></td>
</tr>
</tbody>
</table>
Yoshino et al., 1994

(40) 23–80 [47.6]
[40:0:0] Modern Japanese
density of: 1) secondary, 2) type II, 3) double-zonal, and 4) low-density osteons, 5) fragments, and 6) resorption spaces. Also, mean (7 & 8) and total (9 & 10) Haversian canal and osteon area

S.E.E. = 6.1–8.8

Microradiographs are necessary to distinguish accurately different types of osteons. Sections can be removed by Ericksen’s wedge technique with minimal damage to the original specimen.

Iwamoto, 1978

(42) 41–102 [69.1]
[42:0:0] Modern Japanese
1) osteons/mm² 2) osteon size 3) Haversian canal diameter 4) interstitial lamellae/mm²

95% of estimates within ±11.7 yrs.

Reference sample comprises only older individuals (none below 40 years). Variables are somewhat confusing (e.g., interstitial lamellae per mm² = number of fragments?; number of circumferential lamellae?)

Rother et al., 1978

(70) 20–81 [7] [42:28:0]
Modern Germans
diameter of 1) Haversian canals and 2) osteons; 3) osteocyte; 4) Haversian canal, and 5) Volkmann’s canal density; 6) % osteonal and interstitial bone

S.E.E. = 8.5–9.7 years

Authors found greater accuracy when macroscopic variables were entered into the regressions. They also present results from a factor analysis on the histologic variables.

6th RIB Stout and Paine, 1992

(40) 13–62 [28.6]
[32:7:1] 32 US whites, 4 US blacks, 4 ?

sum of intact and fragmentary osteons per mm² (OPD)

mean absolute difference between estimates and actual age = 3.9 years

Two sections per bone sampled. Good topographic sampling technique. Better results were obtained using a formula combining clavicle and rib OPD (Dudar et al., 1993).

Cho et al., 2002

(103) 17–82 [50.4]
[??.?] 51 U.S. Whites 103 U.S. Blacks

On.Ar
OPD
On.Ar
Ct.Ar/Tt.Ar
Grouping Variable

On Ar

Most accurate rib-based method. Best accuracy and reliability for those with minimal experience (Crowder, 2005).

(Continued)
## APPENDIX B  Continued

<table>
<thead>
<tr>
<th>Bone</th>
<th>Method</th>
<th>Sample Parameters: (N); Age Range [x], Sex [♂:♀:?] and Ethnic Composition</th>
<th>Specimen Sampling Technique</th>
<th>Variables Used in Analysis</th>
<th>Reported Accuracy</th>
<th>Comments; [references for test of method]</th>
</tr>
</thead>
<tbody>
<tr>
<td>4th RIB</td>
<td>Stout et al., 1994</td>
<td>(59) 11–88 [39.2] [♀:♂:?] US whites (autopsy)</td>
<td>sum of intact and fragmentary osteons per mm² (OPD) (also sternal rib phase)</td>
<td>S.E.E. = 10.43 for rib histology alone; 7.18 for combined rib histology and sternal rib end phase</td>
<td>Reference sample combines both historic (1800s) and modern (autopsy) specimens. This method exploits the available information (gross morphology and microstructure) to increase accuracy of age estimates.</td>
<td></td>
</tr>
<tr>
<td>CLAVICLE</td>
<td>Stout et al., 1996</td>
<td>(123) 13–75 [34.0] [73:49:1] 32 US whites; 4 US blacks; 4? (autopsy); 83 19th cent. Swiss</td>
<td>sum of intact and fragmentary osteons per mm² (OPD)</td>
<td>$r^2 = 0.85$</td>
<td>Supersedes Stout and Paine (1992) method for clavicle. Good sample size and topographic sampling technique. Age estimates using this method are less affected by periosteal wear than others that sample only the subperiosteal cortex.</td>
<td></td>
</tr>
</tbody>
</table>

*Predicting equations are not provided in the table. To use a method properly, the original article in which formulas are presented must be consulted.*
Biomechanics is the application of mechanical principles to biological systems. The potential scope of biomechanical studies is very broad, encompassing everything from analyses of the movements of microorganisms to the design of tree trunks (Vogel, 2003; Wainwright et al., 1982). Biomechanics as applied to skeletal material has a long history stretching back to Galileo (1638). Much of the development of skeletal biomechanics was carried out by anatomists and orthopedists in the 1800s and early 1900s (for historical reviews, see Evans, 1953; Koch, 1917; Roesler, 1987); orthopedic biomechanics is still a very active area of research, given its clinical importance (Bartel et al., 2006; Burstein and Wright, 1994; Frankel and Nordin, 2001; Martin et al., 1998; Mow and Huiskes, 2004). Anthropologists have also long been interested in how mechanical principles can be used to explain skeletal variation among past and present populations (Washburn, 1951). Biomechanics theory is commonly used today to explore a variety of issues in biological anthropology, such as the evolution of human walking (Wang et al., 2003), primate locomotion (Bertram, 2004), and the effects of diet on skull morphology (Ross et al., 2005). The purpose of this chapter is to explain and review how bioarchaeologists have applied biomechanics theory to reconstruct past behavioral patterns among human populations. The emphasis here is on structural analyses of long bone diaphyses, in part because this has been the most common area of interest among anthropologists, and because the techniques involved are relatively simple and straightforward.

**“WOLFF’S LAW” AND BONE FUNCTIONAL ADAPTATION**

At the heart of any attempt to reconstruct behavior from skeletal remains is the concept that bone is adapted to its mechanical environment during life. If bone is not responsive to mechanical loadings (the magnitude and orientation of mechanical forces), then its preserved morphology after death will not reflect accurately the particular loadings that it was subjected to, precluding attempts to infer the behaviors that produced those loadings. The general concept that mechanical loading influences bone structure is often referred to as “Wolff’s Law,” named after a nineteenth-century orthopedic surgeon who popularized the concept (Wolff, 1892). However, Wolff’s Law in its original meaning was somewhat different than how we
understand it today, and it is probably better to substitute the term “bone functional adaptation” instead (Ruff et al., 2006a).

A simple feedback model illustrating bone functional adaptation is shown in Fig. 6.1. It is based on the mechanical deformation, or strain, of bone tissue under mechanical loading. Increased strain, for example, through an increase in body size or muscle activity, stimulates deposition of new bone tissue, which strengthens the bone and reduces strain to its original level. Decreased strain, for example, caused by spaceflight, paralysis, or inactivity, leads to bone resorption, which weakens the bone and again restores original strain levels.

This general model is supported by many experimental and observational studies, although there are also many qualifications and limitations to the model (Lieberman et al., 2004; Pearson and Lieberman, 2004; Ruff et al., 2006a). The target strain or “optimum customary strain level” for bone can vary depending on its anatomical location and such systemic factors as diet, disease, age, hormonal factors, and genetic background. Thus, it is important to consider these factors when interpreting bone structural properties. For example, because bone is normally lost with age in adults, age differences must be taken into account when making comparisons between individuals or populations with different average ages at death (also see below). Both diet and behavior may combine to produce a particular morphological result. For example, compared with other Native American samples, archaeological samples from the Great Basin were found to have relatively less bone in their limbs, which was probably a result of a relatively poor diet, but this bone was remodeled into a form that preserved bone strength, which was likely a result of their very physically demanding lifestyle (Ruff, 1999). Genetic differences between individuals and populations also contribute to skeletal morphological variation; however, such differences are modified greatly by actual mechanical loading during life, especially during the “growth” years, which include early adulthood. Limiting comparisons to similar skeletal locations with similar overall biomechanical environments (e.g., the femoral midshaft in humans) simplifies interpretations.

With suitable caution, inferring differences in mechanical loadings during life from differences in preserved skeletal morphology is well justified. It must always be remembered, though, that skeletal morphology is in some sense a compromise between mechanical and other influences (Ruff et al., 2006a).

METHODS FOR ANALYZING LONG BONE DIAPHYSISAL STRUCTURE

Beam Model and Cross-Sectional Properties

It has been shown that long bone diaphyses behave much like engineering beams when they are mechanically loaded (Huiskes, 1982); thus, they can be analyzed using the same theory used by engineers in designing structures. In a beam model, stresses (pressures inside of a material) resulting from externally applied loadings can be calculated given the cross-sectional geometric properties of the beam, which usually are measured perpendicular to the long axis of the beam. When stresses reach a certain critical point, the structure will
fail (fracture); the ability to resist breaking is referred to as strength. The resistance of a structure to deformation, prior to failure, is referred to as rigidity. Both characteristics are important for bone—remaining rigid for support of the body and not breaking under load. Also, different kinds of loadings exist—axial compression and tension, which are forces that act along the long axis of a structure and compress or pull it apart; bending, which produces both compression and tension (on opposite surfaces of a cross section); and torsion, in which a structure is twisted about its long axis, producing diagonal (shearing) stresses. More information and illustrations of these loading conditions and their mechanical consequences can be found in the recent texts given near the beginning of this chapter or in standard engineering texts (e.g., Gere and Timoshenko, 1990); Larsen (1997:197–201) also reproduces some useful diagrams.

Table 6.1 lists the cross-sectional geometric properties that can be used to evaluate the rigidity and strength of a beam, or bone. Rigidity and strength in pure compression and tension are proportional to the cross-sectional area of material in the beam; this is equivalent to bone cortical area (CA) in long bone diaphyses. However, for a variety of reasons, bones rarely are subjected to pure tension or compression. The more mechanically important loadings are bending and torsion. Bending and torsional rigidity are proportional to cross-sectional properties known as second moments of area (sometimes also referred to as "cross-sectional moments of inertia"). Second moments of area (SMAs) are calculated about either an axis through the section (for bending) or about the central point (centroid) of the section (for torsion). They are the product of small unit areas of material multiplied by the squared distances of these areas to this axis or point; thus, they are given in linear units to the fourth power. SMAs can be calculated about any axis through a section, but they are most commonly measured about the anatomical axes of the bone—mediolateral (M-L) or anteroposterior (A-P)—or as maximum and minimum SMAs. These are proportional to bending rigidities in the A-P and M-L directions, and maximum and minimum bending rigidities, respectively, are designated as I followed by the appropriate subscript (Table 6.1). The orientation of maximum bending rigidity relative to anatomical axes can also be calculated ($\theta$). The SMA calculated about the section centroid is referred

<table>
<thead>
<tr>
<th>Property</th>
<th>Abbreviations</th>
<th>Units</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cortical area</td>
<td>CA</td>
<td>mm$^2$</td>
<td>compressive/tensile strength</td>
</tr>
<tr>
<td>Total subperiosteal area</td>
<td>TA</td>
<td>mm$^2$</td>
<td>area within outer (subperiosteal) surface</td>
</tr>
<tr>
<td>Medullary area</td>
<td>MA</td>
<td>mm$^2$</td>
<td>area within medullary cavity</td>
</tr>
<tr>
<td>Percent cortical area</td>
<td>%CA</td>
<td>%</td>
<td>$(CA/TA) \times 100$</td>
</tr>
<tr>
<td>Second moment of area about M-L (x) axis</td>
<td>$I_x$</td>
<td>mm$^3$</td>
<td>A-P bending rigidity</td>
</tr>
<tr>
<td>Second moment of area about A-P (y) axis</td>
<td>$I_y$</td>
<td>mm$^3$</td>
<td>M-L bending rigidity</td>
</tr>
<tr>
<td>Maximum second moment of area</td>
<td>$I_{\text{max}}$</td>
<td>mm$^3$</td>
<td>maximum bending rigidity</td>
</tr>
<tr>
<td>Minimum second moment of area</td>
<td>$I_{\text{min}}$</td>
<td>mm$^3$</td>
<td>minimum bending rigidity</td>
</tr>
<tr>
<td>Polar second moment of area</td>
<td>$J$</td>
<td>mm$^4$</td>
<td>torsional and (twice) average bending rigidity</td>
</tr>
<tr>
<td>Theta</td>
<td>$\theta$</td>
<td>degrees</td>
<td>orientation of maximum bending rigidity</td>
</tr>
<tr>
<td>Section modulus about M-L (x) axis</td>
<td>$Z_x$</td>
<td>mm$^3$</td>
<td>A-P bending strength</td>
</tr>
<tr>
<td>Section modulus about A-P (y) axis</td>
<td>$Z_y$</td>
<td>mm$^3$</td>
<td>M-L bending strength</td>
</tr>
<tr>
<td>Maximum section modulus</td>
<td>$Z_{\text{max}}$</td>
<td>mm$^3$</td>
<td>maximum bending strength</td>
</tr>
<tr>
<td>Minimum section modulus</td>
<td>$Z_{\text{min}}$</td>
<td>mm$^3$</td>
<td>minimum bending strength</td>
</tr>
<tr>
<td>Polar section modulus</td>
<td>$Z_p$</td>
<td>mm$^3$</td>
<td>torsional and (twice) average bending strength</td>
</tr>
</tbody>
</table>
to as the polar second moment of area \((J)\), and it is proportional to torsional rigidity, as well as (twice) the average bending rigidity in any two perpendicular planes. \(J\) is thus a good index of overall rigidity of a beam or bone.

Bending and torsional strengths rather than rigidities are estimated using related but slightly different cross-sectional properties. Because the maximum stress under bending or torsion occurs on the outermost surface of a cross section, strength-related properties are calculated by dividing SMAs by the distance from this surface to the appropriate axis or the centroid. It results in properties referred to as section moduli, commonly designated \(Z\) with an appropriate subscript, given in linear units to the third power (Table 6.1). In practice, section moduli usually are calculated by dividing SMAs by half the appropriate diameter (external breadth) of the section; they can also be approximated by taking SMAs to the 0.73 power (Ruff, 1995, 2002).

Other cross-sectional properties that are of interest from a more morphological rather than biomechanical perspective are the total subperiosteal area \((TA)\), the area encompassed by the outer perimeter of the section, and the medullary area \((MA)\), the area of the medullary canal. These parameters reflect the relative distribution of bone in the section, including the cortical thickness. The percent cortical area \((%CA)\) is sometimes calculated as \((CA/TA) \times 100\), and it is another measure of relative cortical area or thickness, although this parameter has no direct mechanical significance.

From the description of SMAs and section moduli above, it is apparent that distribution of bone in addition to the amount of bone in a cross section is very important in mechanical analyses: the same bone area distributed further from the center of the section will result in much greater bending and torsional rigidity and strength. This principle is important in interpreting many of the findings presented later in this chapter.

It should be recognized that in addition to geometry, material properties such as bone density also affect rigidity and strength. These properties are difficult if not impossible to determine in archaeological bone, given varying diagenetic (postdeposition compositional) changes. However, fortunately, most bone functional adaptation, both evolutionary (genetic selection) and ontogenetic (during the lifetime of an individual), seems to occur via changes in geometry rather than material properties (Erickson et al., 2002; Robling et al., 2002; Woo et al., 1981). It also means that stress and strain are proportional in bone and in that sense can be considered interchangeable, which is useful in relating beam properties (defined on the basis of stress) to the proximate stimulus for bone functional adaptation, strain (Fig. 6.1).

It is also clear from experimental studies that true mechanical loading situations during life are more complex than those modeled using simple beam analysis (Lieberman et al., 2004; Ruff et al., 2006a). However, the cross-sectional properties described here are still the best available for evaluating mechanical performance, and they are strongly related to at least relative rigidity and strength in vivo (Lieberman et al., 2004; Ruff et al., 2006a).

Case Example 6.1: Cross Sections of the Tyrolean “Iceman”

The Tyrolean “Iceman” is a mummified body from the Neolithic period (about 5200 years ago) that was found in the Tyrolean Alps near the Italian–Austrian border in 1991 (Dickson et al., 2003; Seidler et al., 1992). Figure CE1 shows two images produced by CT (Murphy et al., 2003) of his midshaft femur (left image) and midshaft tibia (right image). Bone is white in these images (the fibula is included with the tibia); the gray material is dessicated soft tissue. Anterior is up, and lateral is to the right. The images were analyzed using a computer program to derive bone cross-sectional geometric properties (Ruff et al., 2006b), some of which are listed below the images. See Table 6.1 for definitions and units for these properties. The Iceman had relatively very thick bone cortices and a
small medullary canal (MA), which usually are indicative of high mechanical loadings during life (Ruff et al., 1993), although see later in the main text for qualifications to this interpretation. Comparisons of the Iceman with other prehistoric European skeletal samples are given in Case Example 6.2.

An example of a long bone diaphyseal cross-sectional analysis is shown in Case Example 6.1, which was part of a recent study of the Neolithic Tyrolean “Iceman” (Ruff et al., 2006b). Although both section moduli and SMAs were analyzed in that study, for simplicity only the SMA properties, together with cross-sectional areas, are given here.

**Derivation of Cross-Sectional Properties**

There are several ways to both obtain and analyze cross-sectional images of long bones. Before section images are obtained, it is important to preorient the bones in standard reference planes, which are defined on the basis of anatomical axes. These planes have been described and illustrated in detail elsewhere (Ruff, 2002). Sections are then obtained at standard locations that usually are defined as a percentage of bone length (e.g., 50% of bone length or midshaft). Bone lengths are defined in specific ways (see Ruff, 2002 for one methodology). Obtaining sections at 15% intervals from 20% to 80% of bone length is adequate for fully describing morphological variation along the diaphysis (Ruff and Hayes, 1983a); many studies have used only one section, usually taken at or near midshaft, to characterize diaphyseal rigidity or strength. For the humerus, a section at either 35% or 40% of bone length from the distal (lower) end has been found to be useful because it avoids a large muscle crest (deltoid tuberosity) that lies near midshaft.

Images can be obtained directly from broken or cut sections. In the case of broken sections, it is important that they be approximately transverse to the long axis of bones, as anatomically defined. Such sections can be photographed and then either manually digitized or scanned in using an optical scanner.

More commonly, images are obtained noninvasively using either radiography or computed tomography (CT). CT is the method of choice if available, since it is very rapid and produces accurate two-dimensional images of inner and outer bone contours, although care must be taken to use the correct display thresholds (i.e., what the scanner recognizes as bone vs. air) (Ruff and Leo, 1986). CT was used to
obtain cross-sectional images of the Iceman (Case Example 6.1).

However, CT can be expensive and/or difficult to arrange, especially for skeletal collections curated in remote areas. In these cases, radiography has been shown to be a viable alternative, particularly if combined with molding of external bone contours (O’Neill and Ruff, 2004). In this technique, a mold of the outer surface of the section is made and traced (or scanned using an optical scanner); biplanar radiographs are taken and used to measure bone cortical thicknesses in the A-P and M-L planes, which are then used to establish the contours of the medullary cavity, modeling that as an ellipse. The technique has been shown to be accurate to within 5% compared with direct measurements of section properties (O’Neill and Ruff, 2004). Also, formulas are available for estimating cross-sectional properties from only radiographic breadths, with varying degrees of sophistication (Biknevicius and Ruff, 1992b; O’Neill and Ruff, 2004). These properties produce greater error, but in some circumstances, they may be the only available method for reconstructing cross sections (e.g., Ruff, 2003). Some formulas for converting properties determined from biplanar radiographs alone to true properties are available; these work fairly well for the femoral midshaft (O’Neill and Ruff, 2004).

Several programs have been developed for deriving cross-sectional geometric properties directly from section images. They all work from basic engineering principles (see Nagurka and Hayes, 1980 for a description of the algorithms). Free software that works with popular (and also free) image analysis packages, including NIH Image and ImageJ, is available at www.hopkinsmedicine.org/iae/CR.htm.

Finally, some more specialized methods or equipment have been used to derive cross-sectional properties from various bone image data. Peripheral quantitative CT (pQCT) uses a small, high-resolution CT machine and has built-in software for calculating section properties as well as bone density (Ferretti et al., 1996; Moisio et al., 2003). Bone absorptiometry (“densitometry”) data, which are commonly used in clinical situations, can also be reanalyzed to extract geometric section properties (Beck et al., 1990, 2000).

Size Standardization of Cross-Sectional Properties

When comparing bone structural properties between individuals or populations, it is important to control for differences in body size, because body size constitutes a mechanical load (i.e., resistance of the skeleton to gravity), and because it is related to other factors that also influence mechanical loading, such as muscle size. “Skeletal robusticity” can be defined as “the strength or rigidity of a structure relative to the mechanically relevant measure of body size” (Ruff et al., 1993). For cross-sectional areas, the best body size measure to standardize against is body mass (weight), whereas for section moduli, the best body size standardizing factor is body mass \( \div \text{bone length}^2 \) (Ruff, 2000). It can be shown that the best body size standardizing factor for SMAs is body mass \( \times \text{bone length}^2 \). Body mass can be estimated from different skeletal dimensions; femoral head breadth and a technique that combines estimated stature and bi-iliac (maximum pelvic) breadth have both been shown to provide reasonable results (Auerbach and Ruff, 2004; Ruff et al., 2005). The body mass for the Tyrolean Iceman (Case Example 6.1) was calculated using both techniques, which yielded an identical estimate of 61 kg. It was used in making comparisons between his femoral and tibial mechanical properties and those of other archaeological samples (Ruff et al., 2006b) (see Case Example 6.2).

If body mass cannot be estimated, for example, because of very fragmentary available material, it is also possible to divide cross-sectional dimensions by powers of bone length (estimated if necessary) to “size” standardize them. The recommended powers are bone
length\textsuperscript{5,33} for SMAs and bone length\textsuperscript{3} for areas (Ruff et al., 1993). However, this procedure assumes equivalent body shapes, which can be problematic (Ruff, 2000). Thus, whenever possible, body mass should be estimated directly.

EVOLUTIONARY TRENDS IN LONG BONE ROBUSTICITY

Examining long-term trends in skeletal biomechanical properties provides not only insights into past behavioral patterns but also valuable context for interpreting modern skeletal variation. It has long been recognized that early humans tend to look more skeletally robust than living humans (Boule, 1911; Weidenreich, 1941). However, a comprehensive analysis of temporal trends in skeletal robusticity among humans was not carried out until relatively recently (Ruff et al., 1993). Results of a more recent analysis of these data plus some additional data are shown in Fig. 6.2.

Figure 6.2 shows changes over the past 2 million years in the midshaft femoral polar section modulus (left) and midshaft femoral cortical area (right) in the genus Homo, both standardized for body size as discussed earlier. The x (temporal) axis is given in logarithmic units because it has been shown that temporal trends are exponential rather than linear; that is, there is an exponentially increasing decline in relative bone strength through time (Ruff et al., 1993). The individual data points represent individual fossil and archaeological specimens from around the Old World; the three stars are sample means of more recent populations, including the archaeological Pecos Pueblo Native American sample, a modern East African sample (Ruff, 2000) (open stars), and a modern U.S. white autopsy sample (Ruff, 1987) (filled star). Two standard deviation ranges are shown for the Pecos and East African samples; only the mean for the autopsy sample is given because only mean body size data were available (see Ruff, 2005 for details).

About a 15% decline occurs on average in relative bone strength from 2 million to 8000 years ago, which is statistically significant for both section properties ($p < 0.05$). The dotted lines in Fig. 6.2 are extrapolations of the regressions through the earlier data points continued through to modern times. All modern human sample means fall on or below these lines, indicating that relative bone strength has continued to decline during the Holocene, by another approximately 15% on average.

![Figure 6.2](image-url)
Thus, a total loss of almost a third in relative bone strength has occurred within our own genus since close to its inception, with half of that occurring over the past few thousand years.

The most likely explanation for this exponential decline is that mechanical loadings of the skeleton have also declined exponentially over this period, because of technological advancements that have progressively shielded the body from physical demands (Ruff et al., 1993). The data point for the modern U.S. white sample is particularly low in Fig. 6.2. It is tempting to attribute this to the more sedentary lifestyle of modern industrial populations (the East African population lived in the early twentieth century and was likely not heavily industrialized). However, more data from relatively recent populations are needed to test this possibility more comprehensively.

Comparisons of these data with those for even earlier human ancestors (Ruff, 1998; Ruff et al., 1999) and with nonhuman primates (Ruff, 2000, 2003) demonstrate even greater declines in relative bone strength among modern humans. Simply put, modern humans represent the endpoint of a long lineage of gradually accelerating reductions in bone strength, probably in parallel with reductions in physical activity and muscle strength. The clinical consequences of this “gracilization” of the modern human skeleton may include an increased risk of fracture in old age, that is, osteoporosis (Agarwal and Grynpas, 1996; Lees et al., 1993; Ruff, 2005; Ruff and Hayes, 1988; and see Agarwal, this volume). Osteoporosis is considered in more detail below, but this comparison highlights the general value of placing modern skeletal variation into deep historical and evolutionary context.

VARIATION WITHIN RECENT HUMAN POPULATIONS

Subsistence Strategy

As discussed, behavioral changes associated with subsistence technology were probably important influences on long-term changes in diaphyseal structure during human evolution. Evidence also exists for the effects of both subsistence strategy and physical environment on long bone structure among more recent populations, although these effects are often complex.

Studies of archaeological material from the Georgia (U.S.) coast demonstrated a decline in relative bone strength of the femur and humerus from preagricultural to agricultural groups, which was interpreted to result from increased sedentism and reduced workload in the agriculturalists (Ruff and Larsen, 2001; Ruff et al., 1984). Examples of typical cross-sectional changes are shown in Fig. 6.3. Interestingly, Native populations that were missionized by the Spanish in the 1600s show a partial reversal of this trend, which may be a result of the heavier workloads imposed on them (Ruff and Larsen, 2001).

Studies of other Native North American samples undergoing subsistence strategy changes have produced variable results (Bridges, 1989; Bridges et al., 2000; Brock and Ruff, 1988). A “meta-analysis” of several different prehistoric North American samples that attempted to control for physical terrain demonstrated no significant consistent effect of subsistence strategy on femoral robusticity (Ruff, 1999). This same finding was reported recently for another set of North American samples, carried out using just external long bone data (Wescott, 2006). One problem with such broad comparisons is that they lump together many different groups with subtly differing behavioral characteristics as well as physical environments. “Hunting and gathering,” for example, may involve relatively long-distance traveling over very rough terrain or much less vigorous activities such as shellfish collecting. Attempts to subdivide subsistence categories more finely are only partially successful (Wescott, 2006). It is likely that changes in subsistence strategy will have varying effects on both activity and mechanical loading of the skeleton, depending on the particular culture as well as on the physical environment.
In contrast, in the same “meta-analysis” of North American samples, a strong effect of geographical terrain on femoral robusticity was found (Ruff, 1999). Results of this comparison are shown in Fig. 6.4. When subsistence strategy and sex are controlled for statistically, population samples from mountainous regions have greater femoral robusticity than those from plains or coastal regions, which do not differ from each other. True body size information was not available for all of these samples, so they were standardized by powers of bone length; the additional error produced by this method was probably not too great, however, since body shape was relatively similar between the samples (Ruff et al., 1993). The greater relative strength of the femur in samples from mountainous regions is consistent with the predicted mechanical consequences of traveling through this more rugged terrain. Intriguingly, although the available data are much more limited, the humerus does not seem to show this same pattern, which is again consistent with the hypothesized cause and effect (since the humerus has only non-locomotor functions in humans) (Ruff, 1999).

A special case of “terrain” effects on long bone relative strength is use of water versus land for transportation. Ocean kayakers such as the Aleut have long been hypothesized to have particularly strong upper limbs because of the increased demands on their arms from this mode of transport (Churchill, 1994; Laughlin et al., 1991). Figure 6.5 shows the results of a study of humeri from several prehistoric and protohistoric Native North American samples, including those who did no rowing or paddling on water (a Southwest population), those who paddled on rivers (two samples from the Georgia coast), and those who paddled on the ocean (from British Columbia and Aleuts from Alaska) (Weiss, 2003). Several cross-sectional variables were combined into an “aggregate” robusticity measure, expressed as a Z score (i.e., standardized to the mean and standard deviation of the entire pooled sample). The results shown in Fig. 6.5 were

**Figure 6.3** Changes in cross-sectional shape of the femoral diaphysis with the transition to agriculture on the Georgia coast. (Reproduced with permission from Ruff et al., 1984.)

**Figure 6.4** Effects of terrain on femoral midshaft robusticity (polar second moment of area, standardized for bone length and controlled for subsistence strategy and sex) in Native North American archaeological samples. Mean ± 1 SE. (Data from Ruff, 1999.)
standardized by body mass and humeral length; since only sample mean body masses could be estimated, the data are mean values only. It was predicted that males, who did most of the paddling, would show a progressive increase in strength from nonrowing to river-rowing to ocean-rowing samples, with ocean rowing considered to be the most physically demanding activity. Females should not have been as affected by this difference, since they did little rowing (Drucker, 1955; Hrdlicka, 1945; Oliver, 1988; Saunders, 2000).

Results shown in Fig. 6.5 largely bear out these predictions: there is a progressive increase in relative humeral strength between rowing categories among males, whereas females show much less patterning. The only partial exception is among female Aleuts, who also show an increase in strength relative to other females despite doing little if any paddling (Hrdlicka, 1945; Oliver, 1988). It is possible that other activities, such as processing of whale carcasses, also produced high mechanical loads on the arms of females in this population (Weiss, 2003). Such complexities illustrate some of the difficulties in isolating the effects of particular behaviors within a total behavioral repertoire.

Another interesting study more directly compared relative strengths of the upper and lower limb bones in populations who practiced primarily maritime versus terrestrial modes of life (Stock and Pfeiffer, 2001). Figure 6.6 compares relative humeral and femoral rigidity in protohistoric Andamanese Islanders, who had limited terrestrial but high marine mobility, and Later Stone Age (2000–11,000 years B.P. [before present]) South Africans, who foraged over rough terrain on land. Relative humeral rigidity is greater in the Andamanese, and relative femoral rigidity is greater among the South Africans, as predicted. Marked behavioral contrasts such as these thus seem to be fairly well represented in skeletal structural dimensions.

Sexual Dimorphism

Some differences in skeletal structure between males and females are likely attributable to differences in mechanical loading. Greater upper body strength in males is reflected in

![Figure 6.5](image1.png) **Figure 6.5** Combined measure of humeral diaphyseal strength standardized for body size (see text) in Native North American nonrowing, river-rowing, and ocean-rowing population samples. (Data from Weiss, 2003.)

![Figure 6.6](image2.png) **Figure 6.6** Relative femoral and humeral midshaft rigidity (polar section modulus standardized for body size) in Andamanese Islanders and South Africans. Mean ± 1 SD. (Data from Stock and Pfeiffer, 2001.)
greater average upper limb bone robusticity in most populations (Fig. 6.5), although this varies depending on specific behavioral patterns. For example, females of the Southwestern Native American sample in Fig. 6.5 have slightly stronger humeri than males relative to body size, which is possibly related to their very arduous lifestyle (Weiss, 2003). The wider hips of females in general, because of obstetric requirements, would be expected to create larger M-L bending loads in the proximal (upper) femur (Ruff, 1995), and in fact, the proximal femur is M-L strengthened in females (Ruff, 1987).

An association among subsistence strategy, sexual division of labor, and shape of the femoral and tibial diaphyses in the region surrounding the knee joint (i.e., from the mid-femur through the mid-tibia) has been shown consistently in several studies (Ruff, 1987, 1999; Ruff and Larsen, 2001). Figure 6.7 is a summary of those studies, plus some additional data from studies of earlier (European Paleolithic) samples (Ruff et al., 1993; Trinkaus and Ruff, 1999; Trinkaus, personal communications). Almost identical results have been reported recently for another series of North American samples using only external metrics (Wescott, 2006).

The results shown in Fig. 6.7 indicate that hunter-gatherers have more sexual dimorphism in A-P/M-L bending rigidity of the femoral midshaft than agriculturalists, who in turn have more than industrial societies. This finding neatly parallels behavioral differences that are broadly characteristic of these types of cultures, wherein hunter-gatherers generally exhibit high sexual dimorphism in mobility (males more mobile), agriculturalists less, and industrial societies very little (Ruff, 1987). Long-distance travel over rough terrain should produce theoretically high A-P bending loads in the region near the knee, and this seems to be the case. It is interesting in this regard that the three highest Holocene data points in Fig. 6.7 are all from the Great Basin, with its very rugged terrain. The position of the two Paleolithic data points—for Neandertals and an Early Upper Paleolithic sample—are interesting, too, in that they imply equivalent sexual division of labor to that of modern hunter-gatherers. This division exists despite the fact that the overall cross-sectional shape of the femur varies greatly between these groups, in part probably because of differences in body shape and its effects on mechanical loading of the femur (Trinkaus and Ruff, 1989; Trinkaus et al., 1998; Weaver, 2003).

The effects of body shape on proximal femoral diaphyseal bending were mentioned earlier in connection to sexual dimorphism in hip structure. The body shape factor is important, and it needs to be considered when attempting to reconstruct behavior from long bone structural analyses, particularly of lower limb remains. Evidence exists that activity-related effects on long bone robusticity and cross-sectional shape are more distinct in the distal elements of both upper and lower limbs (Stock, 2006). In the lower limb, this may be related to the confounding effect of body variability.

Figure 6.7 Sexual dimorphism [(male-female)/female x 100)] in femoral midshaft A-P to M-L bending rigidity (second moments of area) in relation to subsistence strategy. Filled and open squares: Native North Americans; circles: modern Japanese and U.S. whites; filled star: Neandertals; open star: Upper Paleolithic humans. (Reproduced with permission from Ruff, 2005.)
shape on the more proximal element (i.e., the femur), with body shape varying in part from climatic constraints (Ruff, 1994; Ruff, 2005; Stock, 2006; Weaver, 2003). An example of how body shape and activity may interact in producing variation in lower limb bone structure is given in Case Example 6.2, which is a comparative analysis of the Iceman’s femur and tibia (also see Case Example 6.1 above). Such factors may explain why relationships between mobility and femoral robusticity and cross-sectional shape are not always straightforward (Ruff et al., 2006b; Stock, 2006; Wescott, 2006).

Case Example 6.2: Comparative Study of the Tyrolean “Iceman”

The Iceman’s femoral and tibial cross sections (see Case Example 6.1) were compared with those of a large sample of prehistoric European males (Ruff et al., 2006b). These cross sections can be divided into five temporal/cultural periods: Early Upper Paleolithic (EUP, 20–30,000 B.P.), Late Upper Paleolithic (LUP, 10–19,000 B.P.), Mesolithic (MES, 5300–9000 B.P.), Neolithic (NEO, 4200–4800 B.P.), and Bronze Age (BRZ, 3500–4200 B.P.). The Iceman (star in figure) falls within the early Neolithic (5200 B.P.). In terms of overall strength (polar section modulus) relative to body size (body mass × bone length), his femur is about average for Neolithic males (top left graph), whereas his tibia is well above average (top right graph). In terms of A-P/M-L bending strength, his femur is close to average for Neolithics (bottom left graph), but his tibia is again above average (bottom right graph). The greater

Figure for case example 6.2.
strength and more A-P buttressed shape of the Iceman’s tibia compared with most contemporaneous males is probably a result of his very vigorous lifestyle—he was likely involved in frequent long-distance travel over very mountainous regions. In these respects, he more closely resembles earlier, pre-Neolithic (pre-agricultural) males, who were probably also more mobile on average. His more circular femur may be a result of his relatively wide body (Ruff et al., 2006b), which would tend to preferentially increase M-L bending of the femur (but not tibia) (also see the main text). Thus, both body shape and behavior likely contributed to the mechanical loadings, and thus structure, of his lower limb bones. The sharp reductions in both overall robusticity and tibial A-P/M-L bending strength in Neolithic males may be a result of greater sedentism (i.e., reduced mobility) in those samples. The more gradual decrease in femoral A-P/M-L bending rigidity across the entire temporal range may reflect the combined effects of body shape and mobility changes (Holt, 2003; Ruff et al., 2006b). [Lines fit through comparative data using LOWESS, a nonparametric regression technique (Cleveland, 1979).]

**VARIATION WITHIN INDIVIDUALS**

**Ontogeny**

Biomechanics theory can also be used to examine and explain changes in bone structure within the lifetime of an individual, both during the subadult “growth” period and during adulthood, as well as structural variation present at any one time within an individual skeleton, including bilateral asymmetry. Such studies can shed light on many interesting questions, including the general health of a population (i.e., growth disruptions and their causes), age-related behavioral changes, and the relative effects of genetics and environment in determining skeletal form (Larsen, 1997; Ruff et al., 2006a; Trinkaus et al., 1994).

In terms of behavioral reconstruction, one question that can be asked is how early in life are “adult” patterns of behavior established? The answer to this question has important implications regarding social structure and resource acquisition/allocation within past populations (e.g., see Hewlett and Lamb, 2005). As shown above, one of the most consistent adult contrasts in behavior reflected in skeletal morphology is between males and females in preindustrial societies, with males being more mobile on average, which leads to greater A-P/M-L bending strength/rigidity in their distal femur and proximal tibia. The Pecos Pueblo, New Mexico, archaeological sample was the first in which this pattern was demonstrated (Ruff, 1987; Ruff and Hayes, 1983b). The degree of sexual dimorphism in femoral midshaft A-P/M-L bending rigidity among adults from this sample falls near the middle of the distribution for archaeological samples shown in Fig. 6.7 (Pecos Pueblo is the most sexually dimorphic of the agricultural samples, perhaps because of its rugged terrain and the long-distance traveling of males during hunting expeditions; see Ruff and Hayes, 1983b).

Long bone structural properties of a sample of juveniles from Pecos Pueblo were analyzed subsequently (Ruff et al., 1994). Figure 6.8 shows age changes in A-P/M-L bending rigidity in a section 35% of bone length from the distal end of the femur, within the region shown to be the most sexually dimorphic in shape among adults (Ruff, 1987). Juveniles under the age of about 15 years cannot be sexed accurately, so sex designations are given only for older adolescents. Means and standard deviations for young adult (20–24 years) males and females are also given for comparison. The older adolescents, 15–19 years of age, already show evidence of sexual dimorphism in bone shape: although extensive overlap exists between the sexes (as there is in adults), differences between males and females in this age group in A-P/M-L bending rigidity
are statistically significant ($p = 0.05$, t-test between groups), and the mean percentage difference between them (12.7%) is similar to that between Pecos adults as a whole (13.3%) (Ruff, 1987).

Thus, the skeletal evidence implies that “adult” patterns of behavior, at least regarding sex differences in mobility, were already present among older adolescents from Pecos Pueblo. This result is consistent with ethnographic evidence indicating the assumption of at least some adult behavioral roles during childhood and adolescence among many “traditional” non-Western societies (Hewlett and Lamb, 2005).

Relative cortical thickness of long bones has been used as a health marker for children in archaeological samples (Cook, 1984; Hummert, 1983). However, it is likely that an interaction exists between dietary and mechanical effects on cortical geometry during growth (Specker and Binkley, 2003), and the mechanical consequences of changes in geometry may be unexpected. A relatively thin-walled long bone may even be stronger than a relatively thick-walled bone, if the bone tissue is placed far enough from the bending axis or centroid (thus greatly increasing SMAs and section moduli; see above) (Bouxsein, 2001; Ruff, 1992).

Age changes in some geometric section properties of the femoral midshaft over the entire age range of the Pecos Pueblo sample (4–55 years) are shown in Fig. 6.9. From childhood through early adulthood, the outer surface of the bone (TA) expands faster than the inner surface (MA), which leads to an increase in relative cortical area (%CA). From the fourth decade onward, however, this pattern is reversed, with MA increasing faster than TA and %CA declining. These patterns are broadly similar to those observed among modern living populations (Frisancho et al., 1970; Garn, 1970, 1972), although some subtle differences also exist, as discussed further below. The importance of considering normal age-related changes when interpreting juvenile skeletal morphology is illustrated in Case Example 6.3.

Case Example 6.3: The Nariokotome Juvenile Skeleton

An extraordinary, largely complete skeleton of an 11–12-year-old boy from about 1.5 million years ago was discovered and excavated in Kenya in the 1980s (Brown et al., 1985; Walker and Leakey, 1993). The museum catalog number given to him is KNM-WT 15,000, but he is also known informally as the Nariokotome Boy (after the site in which he was found). He is an early member of the genus Homo and shows, among other things, that our genus was already as large (if not larger) in body-size than modern humans, although relative brain size was smaller (Ruff and Walker, 1993; Ruff et al., 1997). Radiographs (see Fig. CE3) and cross sections derived from radiographs and external molds show that he had relatively thin long bone cortices.
compared with adult early Homo specimens. However, as shown in the main text, this is part of the normal pattern of cortical bone growth and does not necessarily imply a difference in behavior or mechanical loading. In fact, his outer (periosteal) bone area is not small (see Fig. CE3), and his relative bone strength is increased over that of modern juveniles, as are other archaic Homo juvenile femora (Ruff et al., 1994; Trinkaus et al., 2002), which is part of a pattern of generally increased long bone robusticity in premodern humans (see Fig. 6.2).

With aging, bone mass declines and cortices become thinner. However, in the Pecos Pueblo sample, the parallel increase in outer dimensions caused by continued subperiosteal expansion compensates mechanically for this loss of bone, which results in no decrease in bone strength in old age (Fig. 6.9, right graph). Although for simplicity the sexes are combined in Fig. 6.9, this pattern is characteristic of both males and females in the Pecos Pueblo sample (Ruff and Hayes, 1982, 1983b). This is contrary to age changes found in a sample of modern U.S. white lower limb bones, in which only males showed subperiosteal expansion, leading to no change in bone strength in males but to a decline in bone strength in females with aging (Ruff and Hayes, 1988). These differences are illustrated schematically in Fig. 6.10.

It is possible that the continued subperiosteal expansion of bone throughout adulthood in both Pecos males and females was caused by their relatively rigorous lifestyle, which stimulated...
continued bone apposition on that surface, whereas the activity stimulus was reduced in the modern sample, particularly among females. It is interesting in this regard that many more physically active populations in the world today do not suffer from the same increase in osteoporotic fractures as Western industrialized countries (Barss, 1984; Chalmers and Ho, 1970; Stott et al., 1980). Also, osteoporotic fractures generally are rare in archaeological samples from areas such as Europe that have high fracture rates today, even after adjusting for differences in average lifespan (Agarwal and Grynpas, 1996; Agarwal et al., 2004; Mays, 1999). A general decline in physical activity and, thus, stimulus for bone apposition and maintenance throughout life among modern industrialized societies may be a major contributing factor to the increasing incidence of osteoporosis observed in those societies (Obrant et al., 1989). In this sense, archaeological skeletal material can serve as a valuable baseline against which to compare modern populations (Pfeiffer and Lazenby, 1994; see Agarwal, this volume).

**Bilateral Asymmetry**

Studies of bilateral (right-left) asymmetry in long bone skeletal structure have provided important evidence on general mechanisms of bone functional adaptation, as well as specific behavioral characteristics of past populations (Auerbach and Ruff, 2006; Churchill and Formicola, 1997; Lazenby, 2002; Mays, 2002; Rhodes and Knüsel, 2005; Roy et al., 1994; Ruff and Jones, 1981; Ruff et al., 1994; Sakaue, 1998; Trinkaus et al., 1994). Because many systemic factors, such as body size, diet, and hormonal influences, are held constant, such comparisons may give a clearer picture of localized mechanical influences on bone structure. This general model is supported by many studies of living humans, particularly comparisons of athletes with increased asymmetric use of the upper limbs (summarized in Ruff et al., 2006a).

Figure 6.11 shows the average percent difference between right and left humeri in polar section modulus of the mid-distal shaft (35% from the distal end) and a measure of distal humeral articular size [articular M-L breadth squared, which has a strong correlation to articular surface area determined using more complex methods (Ruff, 2002)] in a series of recent human as well as Paleolithic samples (Churchill and Formicola, 1997; Trinkaus et al., 1994). The recent humans include a sample of professional tennis players (Jones et al., 1977; Trinkaus et al., 1994). All of the recent samples, except the tennis players, have bilateral asymmetry in humeral diaphyseal strength of between about 6% and 12%, with Aleuts showing the greatest asymmetry. The tennis players have an average asymmetry in this characteristic of about 40%, which is consistent with their very asymmetric upper limb loadings. Upper Paleolithic and Neandertal humeri have diaphyseal strength asymmetries approaching those of the tennis players, which...
also suggests strongly asymmetric upper limb use, possibly as a result of stereotypical use of tools such as spears (Schmitt and Churchill, 2003).

Distal humeral articular size shows less bilateral asymmetry and much less variation among these samples. Appropriate data for the tennis players are not available, but results for another peri-articular (i.e., near the articulation) dimension of the humerus indicate that their value would be \( \frac{10}{100} \). Values for the Upper Paleolithic and Neandertal samples are about 4–6%. Such findings are consistent with other evidence that articulations are much less responsive to mechanical loadings during life than are diaphyses (Churchill and Formicola, 1997; Lieberman et al., 2001; Trinkaus et al., 1994). The same appears to be true for bone lengths, which show only limited bilateral asymmetry among all human populations (Auerbach and Ruff, 2006; Trinkaus et al., 1994). Lower limb bone dimensions are less asymmetric than upper limb bone dimensions, because of locomotor constraints, but again cross-sectional diaphyseal dimensions are more asymmetric than either lengths or articular dimensions (Auerbach and Ruff, 2006).

These findings have more general implications regarding the functional interpretation of cross-sectional diaphyseal versus articular and length dimensions in the archaeological and fossil record. Cross-sectional diaphyseal dimensions, because of their greater plasticity during life, should more accurately reflect a person’s activities (mechanical loadings) than other long bone dimensions. In the current context, this includes degrees of behavioral laterality, or handedness within a population. Asymmetry was defined in Fig. 6.11 as maximum/minimum, rather than as right/left, to allow for possible left-handers. If right versus left differences are examined instead, it seems that Paleolithic humans shared about the same percentage of “right dominant” individuals as modern populations (Auerbach and Ruff, 2006). Thus, despite a greater magnitude of asymmetry in loading among earlier humans (Fig. 6.11), the frequency of right-handedness apparently has remained unchanged throughout at least more recent human prehistory.

Figure 6.11 Bilateral asymmetry [median of (maximum-minimum)/minimum] \( \times 100 \) of humeral shaft strength (polar section modulus) and distal articular size (articular breadth\(^2\)) in modern, archaeological, and paleontological samples. Question mark indicates no data available. (Data from Trinkaus et al., 1994 and Churchill and Formicola, 1997.)
CONCLUSIONS AND FUTURE DIRECTIONS

These examples give some flavor of the range of studies that are possible through biomechanical analysis of long bone diaphyses. This approach is ideally suited to behavioral reconstructions of past populations, because of the environmental plasticity of diaphyseal cross-sectional dimensions and the availability of relatively simple but informative biomechanical techniques that can be applied to many types of skeletal samples. Taking a functional rather than a purely statistical or typological approach (i.e., where dimensions are analyzed without regard to their functional significance) is critical to this process. Bones are living organs with important physiological roles during life, many of which leave their marks after death. Biomechanical analysis can help to decipher some of these marks in a more directly interpretable way.

Although the emphasis here has been on long bone diaphyseal structure, as indicated at the beginning of this chapter, there are many other areas of research within anthropology where biomechanical approaches are becoming common. Even limiting these strictly to studies of skeletal material, there are other potential ways in which bioarchaeologists can take advantage of biomechanical theory to extract more information from preserved morphological variation. For example, beam theory can be used to analyze structures other than long bones, including the mandible (Biknevicius and Ruff, 1992a; Daegling and Grine, 1991; Therrien, 2005), although this has not seen wide application in archaeological settings.

One promising new area of research is quantitative analysis of trabecular bone architecture. In this approach, sections of trabecular bone are obtained either invasively or noninvasively, and they are analyzed to determine such parameters as average trabecular density, thickness, orientation, and connectivity. Because evidence exists that trabeculae, like long bone cortical bone, respond to changes in mechanical loadings (Goldstein et al., 1991; Pauwels, 1976; Rafferty and Ruff, 1994), this has the potential to help reconstruct the mechanical loadings on these skeletal elements during life. Archaeological applications to date have been through physical sectioning of bones, mainly vertebrae, with an emphasis on explicating age changes, that is, osteoporosis, in past populations (Agarwal et al., 2004; Brickley and Howell, 1998; Kneissel et al., 1994, 1997; and Chapter 12 of this edition).

Noninvasive determination of trabecular architecture is possible using micro-CT, which is a technique that employs very high-resolution computed tomography, which together with image analysis programs, has the capability of accurately reconstructing a three-dimensional trabecular structure (Hildebrand et al., 1999). To date this technique has been applied in an anthropological context mainly to comparative studies of different species of primates (Fajardo and Muller, 2001; MacLatchy and Muller, 2002; Richmond, 2004; Ryan and Ketcham, 2002a, 2002b), but it could be applied to human archaeological material. Current drawbacks to this method are its relatively high-expense, time-consuming data collection protocol, especially for larger specimens, and the need for specialized equipment and software, limiting its current use for comparative analyses. However, like ordinary CT, which was originally limited mainly to clinical applications but is now fairly commonly used in bioarchaeological studies, it is likely that this technique will become more widely available in the future.

Both compact cortical and trabecular bone structure are also amenable to mechanical treatment using finite element analysis, or the finite element method (FEM), which is a more advanced technique for estimating stresses and strains that involves dividing a structure into a large series of smaller units or elements (Huiskes and Chao, 1983; Richmond et al., 2005). This approach is more flexible than traditional beam modeling and is particularly appropriate for mechanical analysis of three-dimensionally complex structures such as trabecular bone networks and the cranium...
(Ross et al., 2005; Ryan and van Rietbergen, 2005). Like micro-CT, however, it is computationally expensive and the level of structural detail needed to set up the models can be formidable (Richmond et al., 2005). To date it has not been used on archaeological material.

Finally, it is obvious that any mechanistic explanations of bone functional adaptation must involve consideration of the cellular processes underlying morphological variation (Martin et al., 1998). Histomorphometric studies of archaeological bone can give insights into such processes (Abbott et al., 1996; Stout and Simmons, 1979), although they require invasive (destructive) analysis. Additional details on bone histomorphometry are given elsewhere in this volume (Robling and Stout).

REFERENCES


Churchill SE, Formicola V. 1997. A case of marked bilateral asymmetry in the upper limbs of an
Upper Palaeolithic male from Barma Grande (Liguria), Italy. Intl J Osteoarch 7:18–38.


204 BIOMECHANICAL ANALYSES OF ARCHAEOLOGICAL HUMAN SKELETONS


Weaver TD. 2003. The shape of the Neandertal femur is primarily the consequence of a hyperpolar body form. Proc Natl Acad Sci USA 100:6926–6929.

Weidenreich F. 1941. The extremity bones of *Sinanthropus pekinensis*. Paleont Sinica (NS D) 5D:1–150.


INTRODUCTION

Like many new fields, morphometrics, or the quantitative assessment of form, has its roots in the desire to answer a specific scientific question. For morphometrics, this question was the nature of heredity and was a central question for many nineteenth-century biologists for many reasons, not the least of which was that a theory of heredity was the missing piece of Darwin’s theory of evolution by natural selection. Ironically, morphometrics and much of the basis for statistics has its roots in the failed attempt to discover the nature of inheritance through sophisticated means of measuring and analyzing quantitative phenotypic variation. Darwin’s cousin, Sir Francis Galton (1822–1911) devoted much of his life to this cause, on the way creating statistical tools such as correlation and regression (Bulmer, 2003). Similarly, Karl Pearson (1857–1936) who formalized these and other statistical concepts was motivated by the desire to understand heredity through the analysis of quantitative variation.

Although Mendel’s fundamental insights into heredity were made without the aid of sophisticated techniques for measuring and analyzing variation, the statistical and methodological foundation laid by the biometricians of the late nineteenth and early twentieth century was incorporated into the population genetics of the modern synthesis. The paper by Ronald Fisher (1918), for example, that began the building of the modern synthesis was entitled “The correlation between relatives on the supposition of Mendelian inheritance.” This paper, which first introduced the term “variance,” showed how continuous variation could be based on discontinuous inheritance and thus how the conceptual tools developed by the biometricians to analyze quantitatively phenotypic variation could be used in genetics. For much of the ensuing four decades, the measurement of phenotypic variation and the tools to analyze such measurements was an important part of the toolkit of mainstream biology.

This situation changed with the rise of molecular biology. As techniques developed to first isolate and study proteins and later map and clone genes, the phenotype as a level of analysis was downgraded in relative importance because aspects of the molecular machinery that underlie both development and disease could
come under more direct study. Immunohistochemical methods and later in situ hybridization allowed the anatomical localization of proteins and RNA. The development of transgenic methods allowed the creation of gene knockout (or “knock-in”) models. More recently, the development of the Cre-lox system to knockout genes in specific tissues and RNA interference to knock down selectively gene function has allowed the development of ever more complex developmental-genetic models. Not surprisingly, these techniques focused questions on the molecular players involved in development or genes and genetic pathways. Whereas the developmental biologists and geneticists of the 1940s, 1950s, and 1960s used complex analyses of phenotypic variation to probe the black box of development, molecular biology blew open this black box allowing biologists to pull out individual elements and expose them to scientific scrutiny. Refined analysis of phenotypic variation became unnecessary because the questions focus on the functions of specific genes or pathways, and therefore, one is usually only interested in knowing whether some gross and obvious phenotypic effect is or is not produced.

For these reasons, the quantitative study of phenotypic variation founded by Galton and developed by the biometric school is now marginalized within the biological sciences. Few biology departments and fewer medical schools have laboratories in which people routinely measure large numbers of individuals to study patterns of morphological variation. The continued emphasis on this type of research within biological anthropology is an anomaly within the biological sciences, and in the minds of some, this could be viewed as a symptom of marginalization within the broader biological community.

Ironically, however, the very success of molecular techniques for the mechanistic study of development and disease is creating a renewed interest in refined analysis of phenotypic variation of various kinds. Among these techniques is the quantitative study of morphology. The reasons for this change lie in the postgenomic revolution that has been sweeping through the biological sciences for the past few years. The sequencing of entire genomes and the development of tools for storing and mining vast quantities of molecular data and the development of theoretical frameworks for modeling systems based on these datasets is beginning to bring the phenotype back into focus for mainstream biology. When the questions of interest are not on the functions of individual genes or pathways but rather on how the many elements of a complex system interact to produce phenotypic variation in development or disease, the focus shifts away from individual genes and back to the phenotype. Within this “systems biology” framework, the quantitative analysis of subtle phenotypic variation, whether it is aspects of physiological function or of morphological variation, has a crucial role to play. This is because to understand how many factors in a system interact, one must compare many phenotypic variants that will often differ in subtle ways. This is the motivation behind the Mouse Phenome Project (Bogue, 2003; Grubb et al., 2004; Paigen and Eppig, 2000), and it is another reason that morphometrics is increasing in importance now after several decades of relative marginalization.

Biological anthropology is positioned to play an important role in this renaissance of morphometrics for several reasons. The first reason is that there is a strong morphometric tradition in biological anthropology and our field has been the source of much innovation in this area (Slice, 2005). A second reason is that biological anthropology contains expertise that is unique to the morphometric study of humans. A third is that the group studied by biological anthropologists—primates—includes humans and species that are closely related to humans. Several primates have or will soon have sequenced genomes (Mikkelsen et al., 2005; Rhesus Macaque Genome Sequencing and Analysis Consortium, 2007) and are represented by large, well-curated skeletal samples that, in some cases, have information on genealogy, life history, and genetics. Although obviously of interest in their own right, nonhuman
primates can also serve as bridging models between humans and the model species that are most commonly used to study developmental and disease processes such as mice. In fact, this has been the major impetus and justification for the primate genome sequencing projects undertaken thus far, such as the chimpanzee, gibbon, and macaque projects (http://www.genome.gov).

Much more importantly for biological anthropology, postgenomic biology presents important opportunities for new approaches to the core questions of the field. The development of the mouse as a model in developmental biology creates the opportunity to test hypotheses about the developmental basis for evolutionarily significant phenotypic variation. The use of the mouse and other models to study the evolutionary developmental biology of primates is necessary because experimental work with primate embryos is neither practical nor ethically acceptable. Importantly, such research should focus on developmental, process-level determinants of phenotypic variation since the developmental–genetic particulars will vary greatly among species (Hallgrimsson et al., 2007; Hallgrimsson et al., in press; Hallgrimsson et al., 2004). Knowing the developmental basis for evolutionary change is important because it tells us whether structures have changed as the result of direct selection or whether they are epiphenomenal to some other change (Lieberman, 2004). This information can also tell us how difficult evolutionary transitions are or have been in developmental–genetic terms, and this helps us to assess how strong selection would have needed to be to produce particular changes. For example, the study by Abzhanov et al. (2004) demonstrates that changes in beak length and width in Darwin’s finches can be produced by alterations in the expression of a single gene (Bmp4). In turn, this result helps explain how evolutionary changes in this character have occurred so rapidly (Grant and Grant, 2002). In biological anthropology, there is the potential to use mouse models to test long-standing evolutionary developmental questions in primate and human evolution. For instance, our study of epigenetic interactions among craniofacial modules in mouse mutants is directly relevant to developmental hypotheses explored through the work of Biegert, Ross, Lieberman, and others on the relationship between brain size and the cranial base (Biegert, 1963; Lieberman et al., 2002; Lieberman and McCarthy, 1999; Lieberman, 2000a; Lieberman et al., 2000b; Ross and Ravosa, 1993). Morphometrics has an important role to play in evolutionary developmental studies of this kind. To develop refined hypotheses about the developmental basis for evolutionary change, we will have to be able to quantify phenotypic variation that is produced by known developmental perturbations and on that basis make quantitative predictions about patterns of variation in humans and other primates in relation to known genetic variation.

To realize fully the opportunities of postgenomic biology, morphometrics needs to change in three significant ways. The first of these ways is that we need to develop standards for data collection, an infrastructure for data sharing, and a research culture in which the sharing of data is encouraged. The construction of communal databases is very important because it will allow the testing of ever larger and more complex hypotheses. Data standardization and sharing is, in fact, one foundation on which postgenomic biology is built (Peltonen and McKusick, 2001). This foundation is observed in the emergence of community databases such as Online Mendelian Inheritance in Man (OMIM) (www.ncbi.nlm.nih.gov/omim). The emerging field of bioinformatics is largely devoted to the management and mining of databases of this kind, and the vision for systems biology relies heavily on the availability of large amounts of accumulated data about complex systems. For morphometric data to relate to genomic data, as a field, biological anthropologists must develop standards for data collection and the infrastructure to support data sharing.

A related challenge is data standardization. For shared databases to be useful, data obtained
and deposited by different investigators will need to be comparable. As a community, therefore, we must agree on standard formats for data repository. One approach to this would be to agree on standard sets of 2D or 3D landmarks for different kinds of morphological data. Obtaining this agreement would be very difficult because devising a landmark set that suits all possible research aims is impossible and would require people to digitize landmarks not needed for their particular studies. Furthermore, among-observer error is a significant problem in landmark digitization. It would be very difficult to standardize landmarking protocols sufficiently to minimize among-observer error and virtually impossible to devise ways to assess the magnitude of error in a dataset contributed to by a large number of observers over long periods of time. A much more practical alternative would be to agree on formats for making available image and volumetric datasets. Combined with new techniques to automatically extract data from such repositories, this solution would eventually allow us to make use of shared data to address larger scale questions.

The third major challenge is to extend the morphometric toolkit to develop techniques for high-throughput analysis of morphological variation. The rate of gene discovery increased exponentially from 1981 to 2000 (Peltonen and McKusick, 2001) and that trend has continued since. To relate this growing wealth of genetic data to morphological variation as well as to environmental interactions, studies will focus on larger numbers of genetic or environmental factors, which will require larger sample sizes. Currently, the initial data-gathering step for most morphometric methods is very time and labor-intensive. Obtaining 3D landmarks from 40 adult mouse crania, for example, can take a single observer two weeks. At the very least, methods of automated data extraction are needed as initial screens of morphometric datasets if not as replacements for some existing methods.

In the following sections, we present a general overview of the major methods of morphometric data collection and analysis. This overview is intended as a practical guide and does not deal with the underlying theory or math. Those wishing to apply these methods should go beyond the broad outlines presented here. Slice’s introductory chapter to the volume Modern Morphometrics in Physical Anthropology (Slice, 2005) is a great general overview of morphometric theory, and the chapters contained in that volume such as Bookstein and Gunz et al. (Bookstein, 2005; Gunz et al., 2005) reveal current thought on the direction of morphometric theory. For more in-depth introductions, Zelditch et al. (2004) provide an easy-to-follow treatment of geometric morphometric theory for biologists, whereas the volume by Dryden and Mardia (1998) provides an in-depth background for those more mathematically inclined. Lele and Richtsmeier (2001) provide an overview of Euclidean distance matrix analysis (EDMA) and its theoretical basis.

**DATA TYPES AND LANDMARKS**

Traditional morphometrics, as defined by Marcus (1990), is the univariate or multivariate analysis of linear distances, areas, volumes, and angles. This type of analysis has dominated the quantitative assessment of form since Galton. Within biological anthropology, multivariate analyses of datasets of this kind were pioneered by W.W. Howells (Howells, 1973; Howells and Crichton, 1966), developed more by Charles Oxnard (1973, 1984), and developed even more within a quantitative genetics framework by Jim Cheverud (Cheverud, 1982, 1995; Cheverud and Routman, 1993).

Traditional morphometric methods are useful in many research contexts, but they have three fundamental limitations. The first is that size and shape are usually difficult to tease apart in such analyses, and second is that spatial relationship among measurements is typically lost in the course of measurement (Zelditch, 2004; Bookstein, 1991). The third limitation follows from the second: Results do not lend themselves to biologically meaningful
visualizations. Two types of solutions have been developed to address these limitations of traditional morphometrics. One is superimposition-based morphometrics, and the other is EDMA. Both are based on the analysis of sets of either 2D or 3D landmarks. Figure 7.1 shows examples of traditional measurements as well as of a 3D landmark set.

To be biologically meaningful, landmarks have to correspond to something biological and they have to mean the same thing across individuals or across groups or species, depending on the research design. In other words, landmarks have to be, in some sense, homologous across individuals and samples (Bookstein et al., 1985; Gunz et al., 2005). This can be tricky, especially if a large number of landmarks is desired, and species differ significantly in form or comparisons span very different ontogenetic stages. Bookstein’s (1991) classification of landmarks is useful here. In his scheme, Type I landmarks are those that are defined by anatomical discontinuities in all measured dimensions. Such landmarks are rare, but they would include those defined by the intersections of three sutures in the skull (e.g., bregma or asterion in human osteology). Landmarks based on foramina should count as Type I, although the center of the foramen or some location along its edge is an inferred point. Although Type I landmarks capture the locations of anatomical structures such as nerves, vessels, or bones, these structure may also vary in location even if the overall shape of interest is constrained. Type II landmarks are local curvature maxima. Thus, points of greatest curvature along some edge or the ends of bony projections would be Type II landmarks. The intersections of sutures with edges such as the intersection of a suture along the orbital rim or along a foraminal

**Figure 7.1** An example of traditional measurements and 3D (yellow and red diamonds) landmarks (used for a study of cranial integration in mice (Cooper et al., 2007). The traditional measurements include the solid lines (calculated as interlandmark distances) and endocranial volume (shown in red).
margin are Type II landmarks. Type III landmarks are external points that are defined with respect to distant structures such as the points at which maximum lengths occur. Such points are common in traditional morphometrics but are problematic as landmarks because their biological meaning and homology across individuals is often unclear. For both Type I and Type II landmarks, by contrast, a biological meaning is usually fairly obvious. Commonly, Type III landmarks have the property that they represent a biologically meaningful location in only one or two planes and that variation in one or two planes is essentially arbitrary or deficient (Bookstein, 1991; Gunz et al., 2005).

Confining datasets to Type I and II landmarks, such as those shown in Fig. 7.1, is statistically preferable because then one does not need to worry about the substantial arbitrary variation that Type III landmarks have in some dimensions of the landmark coordinate system. However, Type I and Type II landmarks often do not capture adequately the morphology of interest. For that reason, a method for dealing with large numbers of Type III landmarks in both two and three dimensions is valuable (Bookstein, 1997; Gunz et al., 2005). This method is based on defining a few meaningful landmarks and then placing other (often a large number) equally spaced landmarks along curves or surfaces between the biologically meaningful landmarks. These inferred landmarks are known as semi-landmarks. They are Type III landmarks in that their location at the surface of the structure that they represent is meaningful, but they vary arbitrarily along or across this surface. This variation is minimized by the specialized methods for superimposing semi-landmarks. Semi-landmark-based methods have the substantial advantage that very realistic graphical representations of the form of interest are possible. Gunz et al. (2005) provide an excellent discussion of the methods and challenges involved in semi-landmark-based analysis in three dimensions.

Landmarks are usually based on four different kinds of data. These data are raw coordinates, 2D images, 3D surface, or volumetric datasets. Raw coordinate data can be obtained in two dimensions from digitizing tablets and in three dimensions from 3D digitizers. Digitizing tablets record $x,y$-coordinates by clicking a pen or cursor over a physical image or an image that is projected onto the digitizing surface (Fig. 7.2a). These devices are now rarely used for morphometrics but were used commonly in the past. Much more commonly used now for direct recording of coordinate data are 3D digitizers such as the Microscribe series (Immersion Corporation, San Jose, CA) or Polhemus digitizers (Colchester, VT) (Fig. 7.2b). Such devices have a stylus attached to a base by an arm and record $x,y,z$-coordinates when contacting the surface of the object. Obviously, it is absolutely critical when digitizing objects using 3D digitizers that the object remains completely stationary during the digitizing trial. Achieving this goal can be a challenge for objects such as skulls. Morphometricians should put some thought into how their specimens will be fixed during digitizing. One should also collect a sample dataset with repeated trials to verify precision and the specimen-fixing protocol before embarking on a large data collection project. Devices such as these are relatively fast ways to obtain accurate 3D landmark data. For many research budgets and for accessing some museum collections, devices like a 3D digitizer are the only practical alternative because they are relatively inexpensive, portable, and can digitize fairly large skeletal elements. Unfortunately, these devices are difficult to use when specimens are fragile, subject to deformations when prepared, or small.

Another method for obtaining 3D landmarks directly from objects is the Reflex Microscope (Reflex Measurement, Cambridge, U.K.) (Fig. 7.2c). This device uses a stereoscope-based setup to align a light spot in the $x,y,z$-planes with accuracies up to a few microns. This technique is laborious, but it does have the advantage of working on very small and irregularly shaped objects and on objects in fluid such as cleared and stained specimens.
The main disadvantage of techniques that record landmark data directly is that the initial data-gathering step only generates the specific landmarks chosen and those landmarks are obtained in a potentially idiosyncratic way by a particular user. The datasets generated thus do not lend themselves to data-sharing or to data-mining applications.

Two-dimensional landmark data can be obtained from digital images using digital cameras or flatbed scanners. Flatbed scanners are cheap and have the advantage that there is little optical distortion in the images that they produce. As long as the structures to be measured are fairly close to the surface of the scanner, image data can be collected by this method very quickly. Young and Hallgrímsson (2005), for instance, used a flatbed scanner to obtain morphometric data from limb skeletal elements in several mammalian species. Figure 7.3a shows an example of a raw image from that dataset. Parallax error can be produced when objects vary in how they sit on the surface of the scanner. This error can be corrected for when there are planes of symmetry (and one is not interested in asymmetry) by averaging landmarks across the plane of symmetry affected by parallax. This is a problem, however, in directions (such as the anteroposterior direction for skulls) in which structures are not symmetrical.

Figure 7.2 Examples of equipment used to produce morphometric data. (a) digitizing tablet; (b) Microscribe 3D digitizer (Immersion Corporation, San Jose, CA); (c) Reflex microscope (Reflex Measurement, Cambridge, U.K.); (d) White light Scanner (Breuckmann GmbH, D, Meersburg, Germany); (e) Cyberware laser digitizer (Cyberware, Monterey, CA); (f) Scanco VivaCT 40; (g) Scanco MicroCT 40 (Scanco Medical AG, Basserdorf, Switzerland); (h) Skyscan 1172 MicroCT; and (i) Skyscan 1178 (in vivo) MicroCT (Skyscan, Kontich, Belgium).
Digital images of small structures such as small crania or cleared and stained skeletal elements of small mammals can be obtained using digital cameras attached to microscopes or stereoscopes, and larger samples can be imaged using digital cameras mounted on a tripod. Figure 7.3b and 7.3c show examples of images obtained in this way. In these setups, parallax error can be a serious problem and care must be taken to devise a protocol in
which the structures to be imaged are perpendicular to the camera axis of view (or at least in a standard position that does not vary across individuals).

For 3D data, an approach that is becoming increasingly popular is to obtain volumetric or surface data in the initial data collection step. Surface data can be obtained from white light or laser scanners (Fig. 7.2d and 7.2e). Figure 7.3d shows an example of such a reconstructed surface from Tocheri et al. (2005). There are many manufacturers of such scanners as this method is becoming very common in industry. Such scanners produce a dataset consisting of a dense sampling of $x,y,z$-coordinates from the surface of the scanned object. These point clouds are then converted to a triangular mesh that allows the construction of a topological model of the object surface. Morphometric analyses of datasets of this kind are increasingly common in anthropology (Tocheri et al., 2005). This method has some important advantages. The equipment needed to obtain datasets like this is not terribly expensive and so is within reach of many research programs. Large amounts of data can be obtained very quickly using such scanners so they are suitable for high-throughput applications. Unlike any other method, white light or optical scanning can generate color data as well as information about form, and this will be useful in some contexts. Such scanners are already being used to create archives of museum artifacts for both educational and research uses. Finally, 3D information can be extracted from volumetric data. True volumetric data can be obtained using X-ray or magnetic resonance-based tomography (Fig. 7.2f–7.2h). For small translucent samples, optical projection tomography can also be used to generate volumetric data (Fig. 7.2i). Such datasets have the advantage that they contain cross-sectional information and thus allow the extraction of information about internal as well as external morphology. Figure 7.3e–7.3g shows examples of volumetric data in reconstructed form. The three imaging modalities listed above are also increasingly being combined with molecular techniques to image structures that express particular proteins or mRNA. As argued later in this chapter, we also believe that volumetric datasets have significant advantages for enabling high-throughput morphometric analyses. The principal disadvantages are the high cost of the equipment involved and thus of the scanning time to produce datasets, the large amounts of data generated, and the lack of portability of the scanning equipment.

By far the most important advantage of digital imaging modalities, whether 2D images, 3D surface data, or volumetric data, is that the initial data collection step generates a digital representation of the object that can be saved and archived for data-sharing and data-mining applications. The same original dataset can be used to obtain different landmark sets, for example, to address multiple research questions. Given the critical importance of data archiving, sharing, and mining to the future of morphometrics, digital imaging is the preferred initial step in morphometric research.

SIZE AND SHAPE

Morphometric analyses deal with the quantification and visualization of form. Form refers to the combination of size and shape information that describe an object. A formal definition of form is those geometric properties of an object that are invariant to translation, rotation, and reflection. Form variables thus contain information about scale as well as about shape. Once scale is removed, the remaining information is shape, which is formally defined as the geometrical information left after variation in scale, position, and rotation have been removed (Kendall, 1977). Those parameters are most commonly removed by a Procrustes superimposition (Rohlf and Slice, 1990). A Procrustes superimposition centers the shapes at the origin, scales them to a common size, and rotates them to minimize the differences between corresponding landmarks across individuals. In general, analyses
are done in the space tangent to shape space because the space of Procrustes superimposed landmarks, like shape space itself, is curved (Dryden and Mardia, 1998).

Superimposition-based methods have been criticized extensively by Lele and Richtsmeier (Lele and Richtsmeier, 1991, 2001; Richtsmeier et al., 2002b) on the grounds that the variance–covariance matrix is rendered inestimable by the superimposition step. As they say, the true orientation of an object cannot be estimated from data, but superimposition requires orienting all specimens within a common coordinate system. As an alternative, Lele and Richtsmeier recommend the family of methods known as EDMA. This method is based on the analyses of matrices of all pairwise interlandmark distances. EDMA allows comparisons of form (size + shape) and shape (once the measurements are scaled for some measure of size). EDMA-based approaches have been criticized for producing a very large number of variables that substantially inflate degrees of freedom (Rohlf, 2000). For example, given a configuration of only 15 landmarks, which is in a 26-dimensional shape space, there are 105 interlandmark distances to analyze. Other criticisms of EDMA are more technical, relating to the methods to estimate means and variances (Rohlf, 2000).

The shortcomings and advantages of superimposition methods relative to EDMA have been the subject of acrimonious debates; readers do need to grasp the issues involved in this debate, but we caution readers against the view that there is only one correct way to quantify and analyze morphological variation. Most importantly, even if there are correctly identified shortcomings in a method, that does not automatically invalidate all analyses based on them. When the shortcoming of a method is germane to the biological question, such as when the subject of the analysis is the pattern of variation, one will have valid reasons to prefer results obtained using one method over another. In many cases, though, the use of both approaches can ensure that the results are not a methodological artifact.

**ANALYZING SHAPE**

Many analyses begin by exploring the structure of the data, seeking patterns that might elicit explanatory hypotheses. A common method for exploring the data is principal components analysis (PCA), which is used primarily to reduce the dimensions of variation to a few variables (see Pietrusewsky, this volume). These new variables can then be examined for evidence of biological causation. To visualize the method, imagine a 2D or 3D scatterplot of data. The first axis or principal component (PC) is the line aligned with the greatest dimension of variation (hence, it is the one that minimizes the squared distances between itself and all the specimens in the sample; Fig. 7.4). The next axis is the next longest, subject to the constraint that it be perpendicular to the first. Principal components are thus orthogonal to one another and explain progressively smaller amounts of variation in the data. These components are eigenvectors of the variance–covariance matrix, and the eigenvalues associated with the axes are the variances along that dimension. Principal component scores are the values for individuals along these axes, and it is these scores that are scattered to inspect the data for patterns. Because the mean shape is at the origin of the plot, principal component scores represent the distance between each individual and the phenotypic mean along that particular axis of variation.

Variation in superimposed datasets can be visualized in terms of displacement of landmarks in shape space so variation along PCs can be visualized as displacement of landmarks relative to each other. In other words, one can visualize what variation along PCs looks like in terms of morphological variation. The same can be done with other ordination methods such as canonical variates analysis (CVA). This property of superimposed landmark sets lends itself to intuitive visualizations. Figure 7.5 illustrates an example of this in a set of 3D landmarks taken to study craniofacial development in mouse embryos. In 2D analyses, the thin-plate spline can be used to
visualize the changes between landmarks implied by the changes at landmarks.

PCAs need to be interpreted cautiously for at least two reasons. First, there may be cases in which variation is distributed evenly across two or more dimensions, so the ordering of components is effectively arbitrary. To imagine such a case, visualize a sphere of variation—PC1 is the major axis of the variation, but spheres do not have major axes. This situation is particularly common when sample sizes are small (<30); Anderson (1963) presents a test of the null hypothesis that axes are the same length; this is implemented in PCAGen (for 2D data) and 3DPCA (for 3D data), both in the IMP statistical package (Sheets, 2004b). Another, perhaps more serious, problem is that principal components, with the possible exception of the first, do not necessarily have any biological meaning at all. There is no good reason to expect biological causes of variation to be statistically orthogonal to one another. Biologically relevant effects may be spread across multiple PCs, and individual PCs may confound multiple aspects of the biological variation structure. Thus, PCA yields a convenient low-dimensional space for exploring the structure of the data, but the PCs themselves should be interpreted cautiously in biological terms.

Another exploratory technique closely related to PCA is canonical variates analysis (CVA), which examines between-group differences relative to within-group variation. A CVA is similar to PCA except that the axes or canonical variates (CVs) maximize between group variation (relative to within-group variance) rather than variation of the entire sample. The first CV thus represents the morphological axis along which groups are best discriminated but not necessarily the axes along which their means are the most different because their means could differ in a direction that is highly variable within the samples. CVA can be regarded as a two-step PCA, in which the first step is a PCA on the within-group variance, followed by a rescaling step to circularize the within-group variation, after which a PCA can be done on the centroids of the groups (Campbell and Atchley, 1981). As in the case of PCA, there will always be a first axis, but that does not mean that groups differ significantly. That must be determined by a statistical test.

**COMPARING MEAN SHAPES**

The most common question addressed by morphometric studies is whether two or more samples of individuals differ in mean form or
shape. There is a long history of analyses in biological anthropology that address questions like this, ranging from the classic papers of Charles Oxnard (1973, 1984) to more recent analyses of morphological variation in humans and other primates (Bastir et al., 2006; Gunz and Harvati, 2007; Mitteroecker et al., 2004; Rosas and Bastir, 2004; Viðarsdóttir and Cobb, 2004; Young, 2006). In the univariate case, this question is answered by analysis of variance (ANOVA) and in the multivariate case by multivariate analysis of variance (MANOVA). Because shape data are inherently multivariate, MANOVA must be used to test the null hypothesis that the mean shapes do not differ by more than expected by chance. The most powerful approach available uses Goodall’s F-test, which has enormous degrees of freedom because they are a function not only of the number of groups (and individuals) but also of the dimensionality of the data. For example, in a comparison of two groups represented by 15 landmarks measured on 30 specimens per group, the degrees of freedom for the numerator is 26 (because the configurations are in a 26-dimensional space) and the degrees of freedom for the denominator is 728. The test, however, unrealistically assumes that variance is equal and uncorrelated across all landmarks; for that reason, resampling-based tests are preferred over analytic tests. MANOVA, using both analytic and resampling-based versions of Goodall’s F-test, are performed in tpsRegr (Rohlf, 2004) and by several programs in the IMP series of software written by David Sheets.

Comparisons of mean shapes can also be done using EDMA-based methods after scaling for some measure of size (Lele and Richtsmeier,
These methods are based on comparisons of the total set of Euclidean distances among all landmarks. Both resampling and bootstrap tests are available for comparing form or scaled (shape) matrices. These methods provide both an overall comparison of form or shape as well as confidence intervals for individual interlandmark distances. Figure 7.6 shows an EDMA-based comparison of mean form along with Procrustes-based analysis of the same sample. These, and many others kinds of analyses, can be performed using the WinEDMA software package written by Tim Cole (2003).

LOCALIZING DIFFERENCES

After obtaining the statistical result that groups differ in form, it is useful to localize the anatomical differences to particular regions or structures. Surprisingly, methods for localizing shape differences are a difficult and much contested issue in morphometrics. Superimposition methods have the unfortunate characteristic that differences at one highly variable landmark can be transferred to others even if no differences exist at those others. For instance, if one were to superimpose two mandibles, one of which is missing the coronoid process but is otherwise the same as the first, the superimposition process would make it appear that there are small differences at the other landmarks. A large difference concentrated at one landmark is known as the Pinocchio effect, and it can be a significant problem if shapes are defined by a small number of landmarks (Walker, 2000). Although Procrustes-based superimposition provides an excellent means of quantifying and visualizing overall shape variation, one
can occasionally encounter problems when attempting to localize differences in shape to particular structures or regions. EDMA is not affected by this problem because the analysis is based on interlandmark distances rather than on superimpositions. For this reason, some prefer this method to localize anatomical differences between shapes. The problem here is to interpret what can be very long lists of numerical outputs describing varying degrees of differences among forms. Determining whether variables differ significantly is difficult in light of the potentially high error rate for so many nonindependent variables. Much disagreement exists in the literature regarding the methods for localizing differences.

**ALLOMETRY**

Quantifying, analyzing, and removing shape variation that is correlated with size is a common task in morphometric studies. The easiest approach is to superimpose the data, which eliminates variation in geometric scale but not the consequences of scale for shape, and then to regress the Procrustes coordinates on centroid size (or log centroid size). The residuals from that regression are free of variation linearly related to both geometric scale and allometry. One can add these residuals to the mean landmark configuration to visualize variation in shape unrelated to size. Figure 7.7 shows an analysis of a sample of mouse embryos after the removal of the ontogenetic variation by regressing the Procrustes landmark data on stage and centroid size. Despite the fact that the ontogenetic variation within samples is much larger than the difference between samples, the differences in shape can be clearly seen once the ontogenetic variation is removed.

Two common pitfalls in this type of analysis exist, however. One is that when samples are small, or when the amount of variation in size is small, it is difficult to obtain a reliable estimate of the slope of the regression. In the face of a relatively large amount of error, one risks underestimating the true size-related variation or inappropriately correcting for it. Another problem arises when samples have little within-group variation in size but differ both in mean size as well as in aspects of shape unrelated to size. This kind of problem arises when comparing means of homogenous samples that differ in size, such as in comparisons of adults of

![Figure 7.7](image-url)
different species, and also when comparing conspecific samples that differ by a genetic factor that affects both size and unrelated aspects of shape. Here, correcting for within-group variation in size is of little help because the size-related variation within the groups is small. Correcting for size by regression can actually overcorrect the data because it can remove all the differences between groups because the groups differ in size. Figure 7.8 shows an example of this for a mutation in mice that affects both size and craniofacial shape. The only solution to this problem is to know how size is influencing shape for the morphology in question before analyzing the data. For the mouse example in Fig. 7.8, this could be an analysis of allometry in mutations that affect only size.

An alternative approach, which does not eliminate the problems caused by differences in shape confounded with (rather than dependent on) size, is to analyze form in size-shape space (Mitteroecker et al., 2004). This approach involves a PCA of a matrix of Procrustes coordinates plus log centroid size; size is then included in the analysis along with shape. Log centroid size will have the largest variance so PC1 will be closely associated

Figure 7.8  PCA plots of a dataset involving several mouse mutants and their wild-type controls before and after standardization for centroid size. Since several mutations result in decreased size as well as altered shape, much of the apparent difference in shape between the groups disappears in the size-standardized analysis.
with size. This size-related variation can be visualized in the same way as regular PCA analysis of Procrustes data as shown by Mitteroecker et al. (2004).

EDMA-based methods can also be used to study growth and allometric variation, as demonstrated by numerous studies (Richtsmeier and Lele, 1990; Richtsmeier and Lele, 1993; Zumpano and Richtsmeier, 2003). The analyses can be done using log-transformed interlandmark distances regressed on a measure of overall size, or they can be done using scale-free interlandmark distances that capture the variation in shape; these too can be regressed on a measure of size. Although these methods do not provide for the visualizations of shape variation that are possible for Procrustes-based methods, the same arguments alluded to above apply here to the localization of anatomical differences in terms of relative or scaled Euclidean distances.

**COVARIANCE STRUCTURE AND MAGNITUDE**

Morphological integration is an important target of many morphometric analyses, which typically infer integration from statistical correlations among measurements. Studies of covariation structure can be used to dissect patterns of developmental modularity (Hallgrimsson et al., 2002; Wagner, 1990) as long as these patterns are interpreted appropriately (Hallgrimsson et al., in press). Indeed, studies of covariation structure are fundamental to understanding how development structures phenotypic variation. Correlation matrices can be inspected for structure, such as high correlations among measurements within a putative module, and independence among measurements belonging to different modules. Correlation (or covariance) matrices can also be compared across different groups to determine whether they differ in structure.

One widely used method for comparing correlation or covariance matrices is the matrix correlation ($R_M$) between variance covariance matrices of two samples; the observed correlation is then compared with the correlations obtained from randomly permuted matrices or with the correlations between a matrix and resampled selves (Marroig and Cheverud, 2001). The first test, Mantel’s test (Mantel, 1967), determines whether two matrices are significantly correlated, i.e., that they are more similar than expected by chance. The second test determines whether two matrices are the same, i.e., that they differ no more than expected by chance. For Procrustes residuals, Klingenberg et al. (2004) have developed a variant of the Mantel’s test, which shuffles the $x,y,z$-coordinates for each landmark as single blocks and thus takes into account the nonindependence of the three coordinates for each landmark. This method is implemented by the programs Mace and Mace3D by written by Eladio Marquez.

As well as analyzing the structure of integration, and comparing it among samples, we often want to determine the degree to which variables are integrated. These analyses can be performed in several ways. When correlations are used, they should be subjected to the Fisher-z transformation because they are far from normally distributed. This approach is designed to compare individual correlations. A metric for integration among many variables is the variance of the eigenvalues (Wagner 1984, 1990). This metric is useful because, when variance is distributed evenly across principal components, as it would be when correlations are generally weak, the variance of the eigenvalues will be low. Conversely, when correlations are very high, variance will be concentrated along one or a few components; hence, the variance of the eigenvalues will be high, and the index will thus reflect a high degree of integration among traits. To remove variation in the index caused by the magnitude of the overall variance, one should scale the index to the overall variance, which is the sum of the eigenvalues. Code written by Nathan Young for performing this analysis using the R statistical programming environment is provided as an appendix to this chapter. This
metric, like other multidimensional indices, must be applied with caution when comparing datasets that differ in dimensionality.

Analyses of overall integration can provide different results depending on whether they are performed on interlandmark distances or Procrustes superimposed landmark data, especially if the allometric component of shape variation is not removed from the data. For raw, z-score-transformed, or log-transformed measurements, it is not surprising that size will be a major determinant of integration. When distances are scaled to some measure of individual size, allometry will often be a major determinant of integration if shape is highly correlated with size. For Procrustes superimposed data, geometric scale has been removed from the analysis but allometric shape variation remains in the data, and this is often the major determinant of integration. Removing this component sometimes substantially reduces integration. Figure 7.9 shows the same datasets compared using different methods to measure overall integration. In this case, the results are fairly concordant because the difference in integration between groups is not from size or allometry.

To compare the pattern of covariation among anatomical regions, one can use the trace correlation $r_t$ (discussed in more detail in Mardia et al., 1979). This metric estimates the strength of the relationship between two blocks of multivariate data measured on the same individuals. One block of variables is assumed to depend on the other. See Hallgrimsson et al. (2006) for an example of an analysis using this method in the context of a study of the developmental basis for canalization and morphological integration. This method is implemented by PLS3PAC in the IMP family of programs written by David Sheets.

One interesting question is whether some variables are highly correlated if their relationships to other variables are held constant. In this case, the objective is to test hypotheses about interrelationships among variables against the background of a multivariate set of correlations. One useful approach uses partial correlations, which examines correlations among traits while holding their correlations with other traits constant. See Magwene (2001) for a discussion of the use of partial correlations in the analysis of integration and Young and Hallgrimsson (2006) for an application of this method to patterns of limb covariance in primates and other mammals. Another method is to ask whether particular sets of correlations are significantly stronger than others within the matrix, which can be answered using resampling-based methods. This latter method was used by Hallgrimsson et al. (2002) to test hypotheses about patterns of integration in mouse fetal limb data.

**Comparing Shape Variance**

Comparing levels of variation is needed in studies of canalization or other aspects of variability. Van Valen’s (1978, 2005) papers on the statistics of variation provide an excellent point of departure for those wishing to analyze magnitudes and patterns of variation, but these do not cover variability of shape. Of the several methods that he discusses for comparing variances of traditional size measurement, Levene’s test is the most widely preferred (Hines and Hines, 2000; Levene, 1960; Van Valen, 1978). Both individual variances and the total variance of a collection of variables can be compared statistically, testing either the null hypothesis that specific variances do not differ between groups or the overall level of variance does not differ between groups.

For Procrustes superimposed data, variance can be measured by summing the variances across all landmarks, which gives the same value as found by calculating the squared Procrustes distance of individual from the mean (Hallgrimsson et al., 2006; Zelditch et al., 2004). The deviations from the mean can then be compared across groups by ANOVA or t-tests, which is equivalent to Levene’s test for differences in variance for
univariate traits. However, to avoid making unrealistic assumptions about the distributions of those distances, resampling-based methods can be used to test the null hypothesis that the groups are drawn randomly from the same population. Resampling-based tests of this hypothesis are used in DisparityBox (for 2D data) and in Simple 3D.

Figure 7.9 Comparisons of three aspects of morphological integration for mice with the brachymorph mutation compared with wild-type controls (Hallgrimsson et al., 2006). (a) shows integration of shape as measured by Procrustes coordinate data. (b) shows shape integration after removal of the allometric component of shape, and (c) shows form (size + shape) integration based on interlandmark distance data. The bar chart shows the integration values for the two groups, whereas the histogram shows the resampling distributions for both samples. All comparisons are significant at $p < 0.01$ based on 1000 resampling iterations.
Both programs are part of the IMP morphometrics package (Sheets, 2004a).

When samples are heterogeneous, from pooling sexes, ages, or across environments, deviations from the mean for each subgroup should be calculated separately; otherwise, differences in means across subgroups will add to the variance within each group. When subsets of landmarks are of particular interest, separate Procrustes superimpositions should be performed for each subset because the variance at landmarks is distributed by the superimposition process (Lele and McCulloch, 2002).

**FLUCTUATING ASYMMETRY**

Random deviations from bilateral asymmetry of bilaterally symmetric structures, i.e., fluctuating asymmetry (FA), offers considerable insight into causes of developmental noise. Such analyses, however, are difficult because the signal is often fairly small relative to the noise produced by measurement error as well as other sources of variance. For this reason, an established body of statistical methodology exists for the analysis of fluctuating asymmetry that tests for the significance of FA relative to measurement error (Palmer, 1994; Palmer and Strobeck, 1986, 2003). To use most of these methods, repeated measurement trials are highly desirable because these make it possible to estimate measurement error empirically. If those repeated measurements are impossible for some reason, repeated trials for subsets of the data can be used to estimate measurement error (Hallgrimsson, 1998).

For analysis of Procrustes superimposed data, FA can be quantified using the Procrustes distance between sides (Bookstein, 1991; Klingenberg et al., 1993, 2002). Two approaches are needed because the two sides might be far apart, as in the case of two limbs, or they could be two halves of a single structure with a midline. The first is a case of matching asymmetry, and the other is the object asymmetry method. In the case of object asymmetry, the analysis must take into account any asymmetry of the midline, by reflecting entire objects and measuring the Procrustes distance between the object and its reflected self (instead of between two sides). The object asymmetry method does not assume that the midline landmarks actually fall on the geometric midline and thus allows the assessment of asymmetry for landmarks that are not bilateral. To test the statistical significance of differences in FA, the object FA method adapts the Palmer and Strobeck mixed model ANOVA method (Palmer, 1994; Palmer and Strobeck, 1986, 2003) with the appropriate adjustment for the degrees of freedom (Klingenberg et al., 2002). Significance is tested using a permutation method as described by Klingenberg et al. (2002). This analysis can be run using the SAGE 3D program developed by E. Marquez (2004). (There is also Sage for 2D data.)

There is an EDMA based alternative for analysis of FA using landmark data (Richtsmeier et al., 2002a). For localizing differences in asymmetry to individual distances between landmarks, this method is very useful. A shortcoming of this method, as currently implemented, is that it does not incorporate explicitly a test against measurement error based on repeated trials, which means that the analysis relative to measurement error must be done for each measurement first. A Beta release of the WinEDMA software (Cole, 2003) performs this analysis.

**HIGH-THROUGHPUT ANALYSIS**

The most labor-intensive step in most morphometrics projects is obtaining landmark data whether it is from 2D image data or from 3D surface or volumetric data. Because of the high cost in time, high-throughput analyses will not be truly feasible until this step is automated or until alternative non-landmark-based morphometric methods are developed that will allow us to address questions that can currently only be addressed through landmark data. Various investigators are trying both approaches. Klingenberg et al. (2004), for
instance, developed and implemented a method for automated digitization of 2D landmarks in outlines of mouse mandibles. This task is much more difficult in 3D, but considerable progress has been made in this area, as is evident in the techniques underlying the semi-landmark methods.

The other major approach is the development of non-landmark-based morphometric methods. Bookstein (2005) has argued that the toolkit of landmark-based morphometrics is basically complete and that new methods that look beyond the use of landmarks should now be developed. We are working on one such extension—a non-landmark-based method for high-throughput analyses of volumetric data. Our method is based on well-established image-processing techniques. We retain the full 3D volumetric dataset and avoid the laborious manual digitization of landmarks. In our method, volumetric datasets are superimposed into a common orientation by rigid image registration with an isotropic scale factor (Boyd et al., 2006; Kristensen et al., submitted). An average sample shape or generalized shape image (GSI) is determined by averaging the intensities of spatially coincident voxels of the registered images (Fig. 7.10a). GSIs from different samples can then be superimposed on one another. Shape differences between two GSIs can be visualized using the surface-to-surface distance measures between the superimposed images (Fig. 7.11a). Within samples, shape variation can be quantified using image gradient of the average shape and can be visualized by superposition of the gradient magnitude on the isosurface representation of the average shape (Fig. 7.10b).

These tools are in the early stages of development, and the techniques for visualization have yet to be matched with methods for quantifying overall shape variation, for localizing that variation, and for statistical testing based on those quantifications. These methods are not intended to replace established morphometrics tools or landmark-based

---

**Figure 7.10** (a) shows images of mean shapes for 20 C57BL/6J and 20 A/WySnJ mouse skulls. (b) shows shape variation represented as a color map of the magnitude of the image gradient for both strains. Areas of high shape variation within each group are confined to the incisors of both strains as well as to the lateral mandibular body of the A/WySnJ strain. (See color insert.)
morphometrics. Rather, they are intended as a high-throughput first pass through large datasets that can be used to select subsamples for more refined analyses using landmark-based methods.

CONCLUSION

Despite its deep historical roots in biology, the quantitative study of morphological variation has largely been left out by the ongoing postgenomic revolution. This revolution was brought on by technological change and community-wide acceptance of common data formats and protocols for data storage. The advent of geometric morphometrics has been a great methodological advance for the analysis of phenotypic variation; yet similar agreements have been slow to evolve in this field. We argue in this chapter that these changes need to occur for biological anthropology and morphometrics to realize their potential place in postgenomic biology. In addition, there is an urgent need for development of high-throughput methods that do not rely on laborious manual selection of landmarks. Automated landmarking for volumetric data is a holy grail here, but intermediate screening methods are also available that bypass the need for landmarks entirely. As systems biology-based approaches to development and disease gain momentum, there will be an increased demand for more refined quantitative assessment of morphological variation and for intuitive visualization of morphometric analyses. Only through the development of a high-throughput toolkit and community-wide agreement on data-sharing and standards will the morphometrics community be able to meet these demands and thus regain a place at the high table of biological inquiry.

REFERENCES


Howells WW. 1973. Cranial variation in man; a study by multivariate analysis of patterns of


APPENDIX A: R CODE FOR DETERMINING AND RESAMPLING THE VARIANCE OF EIGENVALUES (WRITTEN BY NATHAN YOUNG)

1. Read table into R
   
   ```
   x <- read.table("filename.txt")
   ```

2. Determine Variance of Eigenvalues for sample
   
   ```
   evar <- function(x){
     x <- data.frame(x)
     covmat <- cov(x, use="pairwise.complete.obs")
     evals <- eigen(covmat, only.values=TRUE)$values
     evalstand <- evals/(sum(evals))
     evar <- var(evalstand)
   }
   ```

3. Run Bootstrap
   
   ```
   mxboot <- function(x){
     x <- t(x)
     x <- data.frame(x)
     rep <- sample(x, replace=TRUE)
     rep <- t(rep)
     covmat <- cov(rep, use="pairwise.complete.obs")
     evals <- eigen(covmat, only.values=TRUE)$values
     evalstand <- evals/(sum(evals))
     evar <- var(evalstand)
   out <- replicate(1000, mxboot(x), simplify=TRUE)
   ```

4. Export table to text file
   
   ```
   write.table(out, "results.txt", quote = FALSE, sep="", )
   ```
## APPENDIX B: GLOSSARY AND KEY TO ACRONYMS IN THIS CHAPTER

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>3DPCA</td>
<td>A program within IMP that conducts principal components analysis and determines the statistical distinctiveness of principal components for 3D data.</td>
</tr>
<tr>
<td>Allometry</td>
<td>Shape variation that correlates with size. Ontogenetic allometry refers to shape variation that is related to growth. Static allometry is shape variation that correlates with adult size or size at some particular developmental stage.</td>
</tr>
<tr>
<td>Bioinformatics</td>
<td>The intersection of computer science and the biological sciences. Although much of bioinformatics has focused on the development and mining of biological databases, this label is increasingly used to refer to work that seeks to construct predictive computer-based models of biological systems. Such models often use data generated within the more traditional areas of bioinformatics.</td>
</tr>
<tr>
<td>Canalization</td>
<td>The buffering of developmental processes against influences such as environmental perturbations or mutations. The population genetic definition of Wagner et al. (1997) is the reduction of the phenotypic effect of a mutation or environmental change (Hallgrímsson, et al., 2002).</td>
</tr>
<tr>
<td>CVA</td>
<td>Canonical variates analysis (see text)</td>
</tr>
<tr>
<td>DisparityBox</td>
<td>The program within IMP (see below) that compares mean shapes and variance in shape for 2D landmark datasets.</td>
</tr>
<tr>
<td>*Eigenvalue</td>
<td>Eigenvalues, ( \lambda_i ), are the diagonal elements of the diagonal matrix in the equation: ( SE = EA ). In the common data analysis case, ( S ) is a symmetrical variance–covariance matrix, ( E ) is a matrix of eigenvectors, ( \lambda_i \geq 0 ), and ( \sum \lambda_i = \sum s_i^2 ). The order of the columns of ( E ) is arbitrary, but by convention, they are usually sorted from largest to smallest eigenvalue. See eigenvectors and singular value decomposition (<a href="http://life.bio.sunysb.edu/morph/glossary/gloss1.html">http://life.bio.sunysb.edu/morph/glossary/gloss1.html</a>).</td>
</tr>
<tr>
<td>*Eigenvector</td>
<td>In the equation given to define eigenvalues, ( E ) contains the eigenvectors. In the common data analysis case, ( E ) is an orthonormal matrix (i.e., ( E'E = I ) and ( E'I = I )). When sorted by descending eigenvalues, the first eigenvector is that linear combination of variables that has the greatest variance. The second eigenvector is the linear combination of variables that has the greatest variance of such combinations orthogonal to the first, and so on (<a href="http://life.bio.sunysb.edu/morph/glossary/gloss1.html">http://life.bio.sunysb.edu/morph/glossary/gloss1.html</a>).</td>
</tr>
<tr>
<td>Fluctuating</td>
<td>The normally distributed differences around a mean of 0 between the sides of symmetrical biological structures. Fluctuating asymmetry is often used as a measure of developmental instability.</td>
</tr>
<tr>
<td>Asymmetry</td>
<td></td>
</tr>
<tr>
<td>Genomics</td>
<td>The study and analysis of genomes. This field came into existence with the large-scale mapping and sequencing of entire genomes.</td>
</tr>
<tr>
<td>GSI</td>
<td>Generalized shape image. This is a depiction of the mean shape for a sample of volumetric scans obtained through registration and thresholding. The GSI averages the voxel intensities of a group of scans and then is thresholded at the same level as individuals within the sample.</td>
</tr>
<tr>
<td>IMP</td>
<td>A package of morphometric software developed by David Sheets.</td>
</tr>
<tr>
<td>Isotropic</td>
<td>Equal in all dimensions (referring here to dimensions of voxels).</td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Landmark</td>
<td>A point used for measurement that corresponds to some biologically meaningful aspect of morphology. See the text for discussion of the types of landmarks.</td>
</tr>
<tr>
<td>Levene’s Test</td>
<td>A test of equality for the moments of a distribution. The test compares the set of mean deviations for two or more samples to obtain the significance of the difference in magnitude of variance.</td>
</tr>
<tr>
<td>Mantel’s Test</td>
<td>A permutation test for assessing the significance of matrix correlation between two matrices.</td>
</tr>
<tr>
<td>MACE</td>
<td>A program developed by Eladio Marquez for implementing Klingenberg et al.’s modified Mantel’s test for landmark data. The program shuffles landmarks in x,y or x,y,z blocks so that landmarks and not coordinates are shuffled through the matrix. This adapts the Mantel’s test to the structure of morphometric landmark data (<a href="http://www-personal.umich.edu/~emarquez/">http://www-personal.umich.edu/~emarquez/</a>).</td>
</tr>
<tr>
<td>Matrix</td>
<td>The correlation between two matrices.</td>
</tr>
<tr>
<td>Correlation</td>
<td></td>
</tr>
<tr>
<td>Morphological</td>
<td>The tendency for characters to covary as the result of common underlying developmental factors.</td>
</tr>
<tr>
<td>Integration</td>
<td></td>
</tr>
<tr>
<td>Partial</td>
<td>The correlation between two variables once the influence of one or more additional variables has been controlled for. The partial correlation is obtained as the correlation of the residuals of the variables of interest once they have been regressed on the controlling variables.</td>
</tr>
<tr>
<td>R</td>
<td>A freely available statistical programming language. R is a collaborative project that is coordinated through the R project (<a href="http://www.r-project.org/">http://www.r-project.org/</a>).</td>
</tr>
<tr>
<td>Procrustes</td>
<td>Methods based on the Procrustes superimposition. Here, variation related to size, orientation, and location is removed from a dataset by scaling, translating, and rotating a set of landmark configurations such that the differences between corresponding points in different configurations are minimized.</td>
</tr>
<tr>
<td>Method</td>
<td></td>
</tr>
<tr>
<td>PLS3PAC</td>
<td>A program within IMP for determining the trace correlation between two sets of variables.</td>
</tr>
<tr>
<td>SAGE3D</td>
<td>A program developed by Eladio Marquez for implementing Klingenberg et al.’s (Klingenberg and McIntyre, 1998; Klingenberg et al., 2002) adaptation of Palmer and Strobeck’s ANOVA-based method for analysis of FA to landmark data (Palmer and Strobeck, 1986; Palmer and Strobeck, 2003).</td>
</tr>
<tr>
<td>Simple3D</td>
<td>A program within IMP that compares mean shapes and variance for shape for 3D datasets.</td>
</tr>
<tr>
<td>Term</td>
<td>Definition</td>
</tr>
<tr>
<td>------------------------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>*Thin-plate spline</td>
<td>In continuum mechanics, a thin-plate spline models the form taken by a metal plate that is constrained at some combination of points and lines and is otherwise free to adopt the form that minimizes bending energy. (The extent of bending is taken as so small that elastic energy—stretches and shrinks in the plane of the original plate—can be neglected.) One particular version of this problem—an infinite, uniform plate constrained only by displacements at a set of discrete points—can be solved algebraically by a simple matrix inversion. In that form, the technique is a convenient general approach to the problem of surface interpolation for computer graphics and computer-aided design. In morphometrics, the same interpolation (applied once for each Cartesian coordinate) provides a unique solution to the construction of D’Arcy Thompson-type deformation grids for data in the form of two landmark configurations.</td>
</tr>
<tr>
<td>tpsRegr</td>
<td>This program is part of the tps series of programs developed by Jim Rholf for the analysis of landmark data (<a href="http://life.bio.sunysb.edu/morph/">http://life.bio.sunysb.edu/morph/</a>). This program performs multivariate regression of shape variables against an independent variable such as centroid size. The program also performs comparisons of mean shapes.</td>
</tr>
<tr>
<td>Trace Correlation</td>
<td>The trace correlation (discussed in more detail in Mardia et al., 1979) estimates the strength of the relationship between two blocks of multivariate data measured on the same individuals.</td>
</tr>
<tr>
<td>Variability</td>
<td>The tendency to exhibit variation as opposed to the observation of variation. Variability is a dispositional property (like solubility), whereas variation is a state (like solution) (Wagner et al., 1996).</td>
</tr>
<tr>
<td>WinEDMA</td>
<td>A software package developed by Tim Cole III to perform EDMA-based analyses.</td>
</tr>
<tr>
<td>Wireframe</td>
<td>A series of lines placed between landmarks to create a more representative depiction of the morphology of interest.</td>
</tr>
</tbody>
</table>

Definitions denoted by * are taken verbatim from the online geometric morphometrics glossary (http://129.49.19.42/morph/) maintained by Dennis Slice, Fred Bookstein, and Jim Rholf.
CHAPTER 8

READING BETWEEN THE LINES:
DENTAL DEVELOPMENT AND
SUBADULT AGE ASSESSMENT USING
THE MICROSTRUCTURAL GROWTH
MARKERS OF TEETH

CHARLES M. FITZGERALD and JEROME C. ROSE

INTRODUCTION

Accurate estimation of the age at death is a common problem and certainly one of the most difficult faced by anthropologists confronted with unknown skeletal and dental material. This difficulty is increased significantly when the bones or teeth belonged to someone who lived in the distant past—and who may not even have been an anatomically modern Homo sapiens. Particular methodological difficulties are associated with each of the many age estimation approaches available, but almost all share one essential problem. Most conventional methods must use growth and aging standards developed from living peoples to determine the ages of specimens from the past. A very troubling question develops from this approach: How applicable or appropriate can such modern standards be to an archaeological or early hominid population? Indeed, even the use of standards derived from one modern population for use on another from a different geographic region may be questionable, particularly since many standards are based on well-nourished populations of European descent.

This chapter presents an approach to age estimation that overcomes this essential difficulty. It is based on the interpretation of certain microstructures in the enamel and dentine of teeth that act as markers of growth, providing an endogenous record of development. This approach obviates the need to apply standards of any sort, allowing accurate assessments based on calibrations internal to the tooth itself. Several other histological methods of age determination from teeth exist, such as counting layers in cementum (Stott et al., 1982; Wittwer-Backofen et al., 2004), Gustafson’s method of multiple determination (Burns and Maples, 1976; Gustafson, 1950; Lucy et al., 1994; Lucy and Pollard, 1995), and dentine sclerosis (Bang and Ramm, 1970; Johanson, 1971; Lamendin, 1988; Lamendin et al., 1992; Lucy et al., 1994; Megyesi et al., 2006; Prince and Ubelaker, 2002), but this chapter is only concerned with the approach, actually more correctly a series of different techniques, that interpret development from dental microstructural
growth markers. A limitation of this method is that it can only be used to determine age in subadults, that is, in individuals who have not reached dental maturity and who still have at least one tooth that has not completed its growth. Given this limitation, a frequently asked question is “why would we choose to use such a laborious technique for age determination in circumstances other than the study of rare fossil specimens?”

Even a cursory review of the literature provides abundant answers to this question. Many fundamental problems are involved in using modern dental growth data to determine age at death of subadults recovered from archaeological excavations (Smith, 1991), but accurate ages with limited ranges of error are crucial to bioarchaeology and to the interpretation of demographic and palaeopathological data. Forensic anthropologists require exact ages at death and methods for determining the timing of past traumatic events observed in the teeth to match them with medical records or to establish the presence of past child abuse (e.g., Skinner and Anderson, 1991; Walker, 1997; Walker et al., 1997). Enamel hypoplasia and histological indicators of physiological perturbations, known as Wilson bands, have for some time played a major role in bioarchaeological analyses, and yet, their full potential as analytical tools has only been reached in a few recent studies that have used the techniques described in this chapter to determine accurately the ages at which they developed (FitzGerald and Saunders, 2005; FitzGerald et al., 2006; King et al., 2002, 2005). Furthermore, there are debates concerning the interpretation of hypoplasias of various widths and depths and how these may relate to the length and severity of the physiological stress that produced them that can only be resolved with methods that can precisely determine the timing of dental growth (see Hillson, this volume, as well as Hillson and Bond, 1997).

Before embarking on a description of these analytical techniques and of the methods of histological preparation of teeth required for their application, some background to their use in anthropology will be discussed. The literature providing the theoretical underpinning to dental histological age assessment is also surveyed briefly.

**BACKGROUND**

Dental histological aging first made a major appearance on the anthropological stage amidst a flurry of controversy. Bromage and Dean’s (1985) paper, which reevaluated the age estimates for several Plio-Pleistocene hominids using one of the enamel microstructures, introduced the histological aging approach to the field and immediately created contention. The particular technique that these authors used in their study, and the notion that dental microstructures were produced with regular periodicity throughout tooth growth, was challenged (e.g., Mann et al., 1987). The reasons for the controversy may have had less to do with technical issues than with the results of the study, which contradicted the then prevailing view that the extended period of childhood development characteristic of modern humans came early in hominid evolution (see Mann (1975) for the evidence supporting this point of view). Bromage and Dean’s results suggested that the pattern of early hominin dental growth was more like that of modern apes, who lack the particularly lengthy and delayed maturation and prolonged infant dependency of modern humans.

Since the beginning of the debate, there has been an expanding corpus of published material in support and few challenges in the literature presenting hard evidence against the use of histological aging techniques. Although there are still a few who disagree, the time dependency of dental microstructures is no longer seriously questioned (Antoine et al., 2000; FitzGerald, 1998; Smith, 2006). Most dental experts concur that microstructures record normal growth in a way that permits the developmental and chronological history of a tooth to be reconstructed accurately.
DENTAL ANATOMY AND THE HISTOLOGY OF TOOTH GROWTH

Although they did not come widely to the attention of anthropology until the publication of Bromage and Dean’s provocative paper, dental microstructures and their regular periodicity through tooth growth have an extensive history in the dental literature that extends back to the nineteenth century. Their pedigree as estimators of subadult aging is not as long, dating to the early 1960s. For those interested in pursuing this literature in more detail, the following references are recommended: 1) For the histology of tooth growth and dental microstructures, see (Aiello and Dean, 1990; Avery, 1987; Boyde, 1976; Hillson, 1996; Nanci, 2003; Ten Cate, 1989); and 2) for a review of the literature surrounding issues of time dependency, see (Dean, 1987, 1989, 1995; FitzGerald, 1995, 1998, 1999; Smith, 2006).

Hominid teeth comprise a crown and a root (see Fig. 8.1). The crown denotes that part of the tooth covered by the hard, whitish tissue called enamel. It contains no living cells and is the hardest biological substance known, consisting of about 97% (by weight) of inorganic material made up mainly of hydroxyapatite crystallites, a calcium phosphate. Enamel is secreted as an organic matrix by specialized columnar secretory cells called ameloblasts. Within 24 hours after secretion, this matrix undergoes initial mineralization, after which the proportion of hydroxyapatite in the tissue increases steadily through a process of maturation until the enamel reaches its final state of hardness.

The root and the bulk of the crown of the tooth are composed mainly of dentine. Dentine is a hard, elastic, avascular, vital, yellowish-white tissue that is less brittle than enamel. It is only about 70% (w/w)\textsuperscript{1} mineralized with hydroxyapatite crystals and contains a feltwork of elastic collagen, providing a resilient support. Dentine, like enamel, is a secretory product, manufactured by columnar cells called odontoblasts that produce an organic matrix that later mineralizes. Dentine encloses a central chamber, which is filled with a connective tissue called pulp that becomes more fibrous in nature throughout life. This pulp is lost in dried teeth.

Very early in tooth development, histodifferentiation results in the formation of the hard-tissue-producing cells. Odontoblasts first begin to form dentine along the future junction between the enamel and the dentine of the tooth crown (called the enamel–dentine junction or EDJ; see Fig. 8.1) at the site of future cusp tips. Inductive influences from the odontoblasts result in differentiation and formation of ameloblasts, which then begin to secrete enamel matrix. Both types of cells produce hard tissue in their wake as they move away from the EDJ in opposite directions toward their ultimate destinations—the surface in the case of ameloblasts and the pulp chamber in the case of odontoblasts. Crowns and roots thus increase appositionally (i.e., they grow thicker) at the same time that they increase in length by differentiation of new ameloblasts and odontoblasts. The rate at which length increases, i.e., the pace of recruitment of new ameloblasts and odontoblasts, is called the “extension rate.”

Enamel is composed of many hundreds of thousands of unbroken, interlaced prisms or rods that extend from the EDJ to the tooth surface (see Fig. 8.1). Prisms are formed by ameloblasts that secrete enamel matrix from their distal ends as they slowly make their way to the surface of the tooth. Very shortly after its secretion, this matrix begins to mineralize and mature into enamel. After the ameloblasts reach the tooth surface, they transform into a maturative phase and are eventually shed during tooth eruption. As a result, enamel cannot subsequently undergo repair during the lifetime of the tooth.

Dentine formation differs in several respects and is a far more complex process than enamel formation. The odontoblasts that form dentine

\textsuperscript{1}w/w is an abbreviation for “by weight.” Properly speaking, this means that the mass of hydroxyapatite crystals represents 70% of the total mass of dentine.
Figure 8.1 A tooth and its supporting tissues, illustrating some of the internal microstructures of enamel and dentine. A small section of enamel has been magnified to show the relationship between enamel prisms that follow a course from the EDJ to the enamel surface and striae of Retzius that cross cut prisms at an angle. The inset micrograph relating to the schematic (b) shows an actual longitudinal enamel section examined under polarized light with prisms running from right to left and three brown striae of Retzius cross cutting them diagonally from bottom right to top left. Cross striations, marking daily appositional increments, can be seen clearly as bands on the prisms in this photograph. The same number of cross striations can be counted between all adjacent regular striae in one tooth, and this circaseptan interval is also identical in all teeth in one dentition. The schematic (a) in the figure shows one prism highly magnified, with cross striations illustrated as dark bands on prism varicosities (one cross striation is defined as running from the end of one dark band to the beginning of another). The inset photomicrograph relating to box (a) is a taken with an SEM (scanning electron microscope), and it shows a small portion of enamel fractured in an oblique transverse plane. Prisms and their 3D relationship to each other can be seen as well as the varicosities on each prism, which are cross striations as they appear in the SEM (scanning electron microscope). The schematic (c) at the bottom of the figure depicts a section of dentine, with long period Andresen lines running diagonally from lower left to upper right and dentine tubules cross cutting them obliquely. The photomicrograph relating to (c) is a longitudinal section taken in polarized light.
first lay down a predentine matrix, which then goes through a process that may take several days and that involves the degradation and removal of some of the matrix’s components, the modification of others, and finally mineralization. Unlike ameloblasts, odontoblasts do not die after the primary formation of dentine is complete, but they remain alive and are able to add small amounts of tissue to the mature tooth throughout the lifetime of the individual.

The living odontoblasts remain lining the pulp chamber, and they have long cell processes that extend behind the cell along the route it has taken as it passed from the EDJ (or CDJ, the cement dentine junction, in the case of the root). Each odontoblast process lies in a dentine tubule that therefore reveals the former direction of travel of the odontoblast during tooth formation.

Teeth by their very nature are extremely durable. Because they are protected from any functional influences during their morphological development while contained within the jaws, and after emergence into occlusion there is only minimal influence on their development, great significance can be attached to studies of tooth morphology and microstructure (Aiello and Dean, 1990). Since enamel’s mineralization occurs so rapidly after secretion by ameloblasts, it is particularly useful since it retains an accurate histological record of its development unmodified by remodeling or turnover throughout life.

DENTAL MICROSTRUCTURAL GROWTH MARKERS

The premise that underlies the use of the histological aging techniques is that matrix secretion is governed by various metabolic rhythms, which change the rate and/or mineral density of secretion over a cycle with a regularly recurring periodicity. This leaves microstructural time markers that are preserved in the secretion products—enamel and dentine. Since remodeling of dental tissues is nil, or minimal in the case of dentine, these growth markers are permanently preserved and can be examined histologically and interpreted.

In the discussion so far, no distinction has been made between enamel and dentine, since both retain microstructural growth markers in their tissues that are presumed to have formed from the same underlying metabolic rhythms. However, dentine is a more complex and difficult tissue to analyze and only a few studies in anthropology have used it, mainly those by M.C. Dean (Dean, 1995, 1998b, 1999; Dean and Scandrett, 1995, 1996). A problem that has hindered the study of dentine by anthropologists is that the dentine of archaeologically recovered teeth is sometimes rendered opaque by taphonomic processes taking place during burial in the ground (Molnar and Ward, 1975). On the other hand, diagenetic changes, particularly if of a very long duration (as in the case of early hominoid fossils), sometimes have the opposite effect, making dentine microstructures more discernible. In any event, since there is a much larger body of work on enamel, and it is intrinsically an easier tissue to study, it is the prime focus of discussion through the rest of this chapter.

SHORT PERIOD MARKERS

The microstructural growth markers of teeth can be grouped into two basic categories: short period markers and long period markers. Short period markers result from the metabolic changes arising through one of the body’s central regulators, the circadian rhythm. Daily clocks are ubiquitous in all multicelled organisms, and such a rhythmic regulatory variation seems to be fundamental in the physiological activity of all cells (Hastings et al., 1991, 1997b; Scheving and Pauly, 1974; Scrutton, 1978). Since dentine and enamel are formed through secretory activity, it would therefore be surprising not to find evidence of such an elemental cycle in their structure.

The short period growth markers produced by this 24-hour rhythm in enamel are called cross striations and, in dentine, von Ebner’s
lines. Under polarized light microscopy (PLM), cross striations appear as bands consisting of fine, transverse, dark, striations along the length of enamel prisms, and under the SEM (scanning electron microscope), and sometimes also in PLM, these bands are observed to be associated with enlargements or varicosities of the prism (see Fig. 8.1 for examples). Cross striations are found in the enamel of all primates and many other mammalian groups, and their form varies only slightly in the relative width of bright and dark bands and in the regularity of their spacing (Hillson, 2005).

The evidence for circadian rhythmicity of cross striations and von Ebner’s lines is considerable. Perhaps most telling is direct experimental evidence, and this will be reviewed briefly. Schour and colleagues (Massler and Schour, 1946; Schour and Poncher, 1937; Schour and Hoffman, 1939a, 1939b) presented results that correlated daily rates of enamel production with cross striations. These data had been derived from studies that administered injections of substances that leave permanent tracer labels in growing enamel and dentine to infants with inoperable hydrocephalus. Another group of investigators working about the same, but in Japan, headed by Okada and Mimura (see Okada (1943), as well as Shinoda (1984) and Rosenberg and Simmons (1980a, 1980b)), provided quantified, experimentally established evidence of the daily nature of cross striations. Okada and colleagues used injections of sodium fluoride and lead acetate administered at irregular but known intervals to demonstrate that cross striations were daily increments in growing dogs, rabbits, pigs, and monkeys. These substances produced visible lines in the enamel of these animals, and the number of cross striations between labels were counted and were found to correspond to the number of days between injections. They also labeled dentine, and using a similar technique, this group established that von Ebner’s lines were also daily markers in dentine. Bromage (1989, 1991) administered three different fluorescent labeling compounds at known intervals to two postnatal pig-tailed macaques. The compounds chosen labeled both enamel and dentine. First permanent molars were sectioned and examined under ultraviolet light to reveal the fluorescent labels, but in only one specimen was there sufficient optical contrast in the lines to allow study of the enamel. Bromage counted the number of cross striations between the labels in this animal and the number from the last line to the occlusal edge. The tallies agreed precisely with the known intervals between dosages and the interval between the last dosage and the sacrifice of the animal. Finally, Smith (2006), using material originally prepared for a study on bone growth by Newell-Morris and Sirianni (1982), in which 17 immature Macaca nemestrina monkeys were injected three to five times each, with one to three different dyes during either prenatal or postnatal growth, provided the largest single body of conclusive evidence of the circadian rhythmicity of cross striations. The dyes left fluorescent labels in dentine that were then matched to lines in enamel that showed the same number of cross-striations between them as days between respective injections. These observations were confirmed in several sections from multiple individuals, using 98 histological sections. Not only was Smith able to verify the 24-hour periodicity of cross striations, but she also provided support for the regular periodicity of several other microstructures in enamel and dentine, in particular for the long period lines of enamel, which will be discussed shortly.

Several hypotheses have been put forward to account for the appearance and particular mineral properties of cross striations, but perhaps the most plausible is from Boyde (1964, 1979, 1989), which was also elaborated by Risnes (1998). Boyde suggested that the carbon dioxide (carbonate/bicarbonate) available for incorporation in the enamel mineral component would be greater during those phases of the 24-hour cycle when metabolic activity was the greatest. This variation in the rate of metabolic activity will result in variation in the secretion rate of the enamel matrix, with enamel being secreted faster at
times of most intense metabolic activity. During the time when enamel is forming most quickly (and ameloblast movement is fastest), because of its effect on the distribution and orientation of growing crystallites that make up the prism, the prism body will in fact become wider. This mechanical explanation adds to the fundamental biochemical one, which also explains the variation in mineral density that is found to be associated with cross-striation areas along the prism.

LONG PERIOD MARKERS

The long period markers of enamel are called brown striæ of Retzius (1836, 1837), and those of dentine are called Andresen lines. Some striæ of Retzius project onto and crop out at the surface of the crown, and these can be thought of as constituting separate long period markers of enamel, called perikymata.

In longitudinal sections (i.e., in two dimensions), brown striæ of Retzius appear to form successive layers (or “caps”) around the dentine horn (see Fig. 8.2c), which after the first cap reaches the occlusal surface, become discontinuous cuspal y on either side of it (see Fig. 8.2c). Striae layers continue down on either side of the crown all the way to the cervix of the tooth; most striæ appear to run obliquely from the EDJ to the occlusal surface so that they form an acute angle with the prisms they cross (see Fig. 8.1b). In transverse sections, striæ of Retzius appear as concentric rings that encircle the tooth. However, it is necessary to mentally reconstruct the views in Fig. 8.2 in three dimensions to understand the real architecture of the striæ. In three dimensions, the continuous striæ around the dentine horn are imagined more aptly as “domes,” a term used by Hillson (1986, 2005) rather than caps, and the striæ down the crown that are discontinuous cuspal y, as circumferential “sleeves” (Hillson, 1986, 2005).

Figure 8.2 Schematic illustrating striæ of Retzius and perikymata in three views. Top left (a) is a transverse section through a tooth in which striæ appear as concentric rings. Top right (b) is a view looking at the outside surface of the crown in which perikymata appear as ridges encircling it. The inset is an actual photograph of perikymata. On the bottom (c) is a longitudinal section showing striæ as continuous domes around the dentine horn. The first discontinuous layer to emerge at the surface terminates as a perikyma, and striæ and perikymata continue down either side of the crown to the cervix. The inset photomicrograph is a longitudinal section of enamel.
Striae differ in their visibility, and they are variably expressed, even within one tooth. They are most clearly discernable in the outer enamel and particularly in the cervical third of crowns. In modern humans, striae are spaced more widely in the occlusal part of the crown, and they become closer together toward the cervical portion. This difference in spacing—30–45 μm at the widest and 15–20 μm at the narrowest—reflects differences in the rate of enamel formation, which are fastest cuspally and slow toward the cervix where striae become more closely packed together. The angle that striae form with the surface (and with the EDJ) also changes, being most acute cuspally and becoming more obtuse through the crown cervically. Again, this reflects changes in the rate of enamel production and the enamel extension rate.

Striae of Retzius are only visible in sectioned or broken teeth; however, those striae that terminate at the enamel surface are projected onto it as fine, transversely oriented, circular “wrinkles” consisting of a series of ridges separated by corresponding grooves (Risnes, 1985a, 1985b). The totality of the excrescence, i.e., both the ridge and the groove (see Fig. 8.2), is known as a perikyma (Risnes, 1984), which is a compound word derived from the Greek peri, meaning around, and kyma wave, perikymata in the plural. From Fig. 8.2 it is clear that only the sleeved striae crop out at the surface as perikymata; all but one of the domed striae remain buried within the cusp. These hidden striae are referred to as cuspal (or appositional) striae and those terminating at the surface as lateral (or imbricational) striae.

It is now well established that the number of daily increments between adjacent long period lines in enamel is uniform within one tooth and is consistent among all of the teeth in one dentition (FitzGerald, 1998; Smith, 2006). This is likely also to be the case with von Ebner’s and Andresen lines in dentine (Dean, 1999). However, the number of cross striations between adjacent striae, although uniform in all teeth in one individual, commonly varies from 8 to 10 days among different individuals. The range of this period, called a circaseptan interval (because it is “around seven [days]”), is even greater than this, and intervals have been recorded as low as 5 and as high as 14 in modern humans (FitzGerald, 1995; Hillson, 1996). Fukuhara (1959) in a very careful study using ground sections of teeth, calculated circaseptan intervals that ranged between mean values of 2 and 8 for ten other primate species (with a modal value of 7 or 8).

Such a range of rhythms does not seem to be related to any known astronomical or natural cadences, and although a number of possible explanations have been proposed, none has yet been accepted unreservedly as correct. Among the likeliest are several put forward by (Newman and Poole, 1974, 1993) who suggest that a near-weekly rhythm might arise from interference beats between several interacting rhythms. For instance, an eight-day periodicity will result from interference between two rhythms, one of 24 hours and the other of 27—two rhythms running independently of each other interacting to produce a third. It has also been hypothesised that melatonin, known to be one of the major regulators of the circadian rhythm (Hastings, 1997a, 1997b; Hastings and Loudon, 2006; Hastings et al., 1989), affects the production of some hormones as well as the maintenance of several physiologic cycles (Haus and Touitou, 1997; Vollrath et al., 1975) and may therefore also show a circaseptan rhythm. It has also been suggested that the circaseptan interval may be chaotic in origin (FitzGerald, 1995).

Although the precise nature of the rhythm is unclear at this time, the etiology of striae of Retzius formation is better understood. Boyd (1964, 1979, 1989) has extended his hypothesis for the formation of cross striations (already discussed) to embrace the formation of striae of Retzius and perikymata. Risnes (1990, 1998) has made meticulous SEM observations of striae and has incorporated some of Boyd’s ideas to come up with a convincing explanation for their production. In the same
way that prism varicosities (see Fig. 8.1a) are regarded as a rhythmic, reciprocal expansion and reduction of prism growth, regular striae of Retzius may be regarded as an accentuation of the same phenomenon (Risnes, 1998:345). Striae result from growth discontinuities (i.e., disturbances in mineralization) affecting all ameloblasts secreting at the time of the disturbance or perturbation (and it is likely that the same underlying rhythm that produces striae in enamel also produces Andresen lines in dentine). These long period lines, therefore, mark successive layers or sheets of enamel (or dentine) formed at regular circaseptan intervals throughout tooth development (see Fig. 8.3). Each layer (or, more correctly, shell in three dimensions) represents the external profile of the tooth as it existed at the end of each circaseptan interval.

Because of the layering that is observed, tooth growth is usually described as being appositional or incremental, but in its strictest sense, this is incorrect, and in fact, it may contribute to the confusion that exists in the minds of some people on this topic. Layers of enamel are not “added”; rather individual prisms continuously increase in length, growing away from the EDJ, and from time to time, their growth is perturbed slightly, producing a discontinuity. The collective result of this discontinuity, which occurred coevally in all developing ameloblasts, assumes the appearance of a layer. Thinking of striae of Retzius as epiphenomena (as represented in Fig. 8.4) makes it easier to understand the rhythmic superposition on growth that in fact is occurring. It is the circaseptan rhythm that is the phenomenon, not the stria. Cross striations may be seen in precisely the same way, as epiphenomena of circadian regulatory rhythms.

Perikymata are formed by the same slowing down or discontinuity of enamel production that occurs at the end of a circaseptan interval. The systemic trigger that produces striae causes ameloblasts at the surface to cease their matrix secretion phase slightly prematurely. Those cells that were about to cease production do so immediately and slightly earlier than their cervical neighbors. These cells are associated with the troughs of the perikymata, and in fact,
the plane of projection of the floor of a trough corresponds with the plane of the associated stria. The ridges represent ameloblasts undergoing a normal terminal phase process in the interval before the next perturbation in growth (Boyde, 1979, 1989).

**IRREGULAR STRIAE OF RETZIUS**

Not all striae of Retzius are formed through the action of the regular circaseptan rhythm. One very prominent stria, called the neonatal line, can be observed in the enamel of deciduous teeth and permanent first molars (and sometimes permanent incisor crowns, which can occasionally begin mineralizing before birth) (Eli et al., 1989; Schour, 1936; Weber and Eisenmann, 1971; Whittaker and Richards, 1978). This line develops at parturition and is thought to relate to the physiological stresses of birth, and in fact, a relationship is said to exist between the width of the line and babies who suffered severe birth trauma (Eli et al., 1989). This neonatal line is essentially the first accentuated stria of Retzius, since prenatal enamel does not normally contain striae, except in cases where extensive prenatal stress is associated with birth defects (Hillson, 1996). Because they record the birth event and thus “zero” developmental chronology, neonatal lines are important in age estimation techniques, as will be discussed later.

A second group of irregular striae have been associated with various types of physiologic stresses, such as infection, disease, and malnutrition, and these are called Wilson bands or sometimes pathological striae (FitzGerald and Saunders, 2005; FitzGerald et al., 2006; Goodman and Rose, 1990; Rose, 1977, 1979; Wilson and Schroff, 1970). Wilson bands are thought to relate to the same growth-disrupting factors that cause enamel hypoplastic defects on crowns. Because it is believed that these disruptions in enamel growth arise from metabolic disturbances, enamel hypoplasias and Wilson bands have been studied intensively in many living and past populations by anthropologists who use them as indicators of environmental stress (see Hillson, volume). This means that accurate timing of the duration and frequency of the stress episode are important in helping to isolate the cause of the stressor.

Wilson bands can be distinguished from regular striae of Retzius because they are visible along more of their length than regular striae. Wilson bands may in fact be visible along their full length from the EDJ to the surface, which is not often the case for striae of Retzius. Goodman and Rose (1990) considered, in their minimum definition of a Wilson band, that it need be continuous for only three quarters of its length, a definition that has been used in several recent studies (FitzGerald and Saunders, 2005; FitzGerald et al., 2006). The same uncertainty surrounding causation that pertains with regular striae exists with Wilson bands. However, it is likely that irregular striae are formed like regular striae as a response, albeit a heightened or exaggerated one, of the active ameloblast sheet to some
disruption to growth. In the case of Wilson bands, the trigger clearly does not arise from a regular circaseptan rhythm, but instead it can be attributed to an external stressor (FitzGerald and Saunders, 2005; Goodman and Rose, 1990; Hillson et al., 1999; Simpson, 1999).

Together with regular striae, irregular striae serve an important function in histological aging. Because irregular striae are distinctive, and because they are triggered by a stimulus that affects all growing teeth at a point in time, they can be used as ‘registration lines,’ which trace positions of contemporaneity from tooth to tooth in the dentition (Boyde, 1963; Condon and Rose, 1992; Reid et al., 1998).

There is usually a well defined sequence in variation in width and appearance of striae (of both regular & irregular types) within one crown, and a number of authors have noted that this sequence is the same in any crown that was forming in an individual at the same time (Fujita, 1939; Gustafson, 1955; Hillson, 1992; Hillson and Bond, 1997; Takiguchi, 1966). The sequence of striae in different teeth from one individual can be matched, and an extended sequence of striae can be built up covering the whole length of crown formation for all of the dentition.

Irregular striae appear to have counterparts in dentine. Neonatal lines can be observed in dentine, and accentuated Andresen lines have been matched with counterpart Wilson bands (Dean, 1995; Smith, 2006). This effect on both hard tissues implies that the upset causing irregular striae and Andresen lines has a systemic origin, adding weight to the belief that it is a stressor that influences the dental development environment.

CLASSIFICATION OF ENAMEL MICROSTRUCTURES

Although for heuristic reasons we have chosen to present individual microstructures as separate entities so far in this chapter, we believe that they are intimately interconnected; they may manifest themselves diachronically and at different spatial scales, but they all share the same fundamental etiology. We mentioned the interrelationship of cross striations and striae of Retzius as epiphenomena of an underlying metabolic trigger and tried to convey this in Fig. 8.4. We can now broaden this concept to embrace all microstructures, whether their trigger arises from an internal metabolic rhythm with a regular periodicity, or whether the trigger comes from some event, like the stress of birth, that is occurring external to the organism. Figure 8.5 shows a scheme that groups together the major microstructures. The black diamond on the left of Fig. 8.5 labeled “trigger” is the factor that all share: their formation as a result of a deviation, or even stoppage, of enamel matrix secretion from the ameloblasts that are ‘operational’ at the time of the deviation. This disruption to the ameloblast in its normal course of secretion may be initiated by something that arises directly from the genetic program of growth or indirectly from some constraint on the availability of required nutrient raw material. This shortage may affect the composition of the enamel matrix, the proper functioning of the ameloblast cells themselves, or the microenvironment around the ameloblasts.

Whether the microstructure emerges on the surface of the tooth or remains within the enamel² depends on the stage of crown growth in which the disruption occurs. For instance, if the deviation takes place before cuspal development is complete, the microstructure will be buried within the jacket of enamel. On the other hand, if the ameloblast sheet is disrupted toward the end of crown growth, it may result in ‘two’ microstructures, one on the surface and the other within the mantel. (In fact, it will of course not really be two distinct microstructures but merely an internal and external manifestation of one disruptive event.) The trigger may have a regular periodicity, resulting in the formation of perikyma, or it may be episodic,

²The layer of enamel below the surface that constitutes the crown is sometimes called “the mantel” (cf. the portion of a rocky planet surrounding the core) or “the jacket.”
producing an external enamel defect (EED, called in the literature enamel hypoplasia) or its inside equivalent, an internal enamel defect (IED), a Wilson band.

**PREPARING TEETH FOR HISTOLOGICAL EXAMINATION**

To interpret internal microstructural growth markers, teeth must be examined microscopically, and this means cutting thin, longitudinal sections from them and mounting these on glass slides. This is a technical challenge and there are many ways to accomplish it, which are well described in the literature (e.g., Schenk et al., 1984; Schmidt and Keil, 1971; Beasley et al., 1992; Marks et al., 1996)—also particularly recommended is Hillson (1996, 2005), who discusses sectioning and allied techniques that are particularly useful for archaeological specimens, and who in the same volume also gives a primer on microscopy for novices.

Here we will discuss one method for section preparation that is intended to give some idea of what is involved for those who have never sectioned teeth. This methodology was developed because it is comparatively easy to carry out, and because it cuts down the time involved not only in preparation, but also in sectioning, grinding, and polishing, since teeth are not embedded. However, it must be stressed that it is intended to be used in large-scale investigations, and although the quality of the finished slide will likely be acceptable, it will be more difficult to achieve outstanding results using this “quick and dirty” approach. Despite these limitations, these are the ideal circumstances for most bioarchaeological investigations. It is not appropriate for use with very precious specimens, like fossil teeth, nor for teeth that are very brittle and prone to damage, nor for teeth that will be subjected to additional investigative approaches (such as, for instance, transmission electron microscopy). A more conventional procedure (see the references in the previous paragraph) that embeds the tooth in a block before sectioning is recommended for specimens of this sort, or for any tooth whose state of preservation is uncertain.
Before sectioning, a line marking the cutting plane should be drawn around the whole tooth with a permanent fine-tipped marker since cutting accuracy is vitally important: Sections must pass through the maximum thickness of enamel at the tip of the dentine horn and run nonobliquely through the long axis of the longest growing cusp (normally the mesiobuccal cusp of molars). It is usual to embed specimens in a hard supporting material before sectioning, but tooth preparation in this approach consists only of soaking teeth in cyanoacrylate cement (Super Glue, Super Glue Corporation, Rancho Cucamonga, CA, or Instant Krazy Glue, Krazy Glue, Columbus, OH), which fills in surface microcracks and provides support to fragile enamel. In other words, the cement keeps the enamel from falling off during the sectioning process. After immersing and rotating teeth in Super Glue, it is easiest to allow the film to harden on a layer of “dental sticky wax” (Dental Model Cement Sticky Wax, which is readily available from most dental supply companies) that has been melted and poured into a petri dish. They can then be removed easily simply by softening the wax slightly on a hot plate.

Teeth must be sectioned using a special low-speed saw, which is available from several manufacturers and in a variety of models. One of the simplest, and probably the type in widest use (but not the most technically superior), are peripheral-bladed saw models made by Buehler Ltd. (Lake Bluff, IL). These saws have rotating diamond-edged blades against which the specimen is held by a counterweighted arm. The tooth should be oriented precisely and attached to the cutting chuck by coating with dental sticky wax that has been applied with a small bladed hot knife (the wax also provides additional support for brittle enamel). Distilled water serves as an adequate cutting coolant, and after sectioning, the chuck is dismounted and the cut face of the tooth is dried and coated with a film of Super Glue, which is allowed to set. A microscope slide is then attached to this coated face (parallel to the tooth surface) with a layer of dental sticky wax. The chuck is then remounted on the saw and a slice about 250–300 μm thick is cut.

The tooth section, still attached to its slide, is then lapped (or polished) with a medium-grade (e.g., 3 micron) aluminium oxide slurry (an abrasive compound mixed with water) to remove any saw marks. Lapping may be done by hand on a glass plate or with a rotary machine, which is much faster. If a grinding/polishing wheel is employed, then a diamond paste of the same size spread on a nylon cloth covered wheel may also be used. After lapping, the slide is cleaned in an ultrasonic bath and then dried in a vacuum chamber over silica gel. A small amount of adhesive is spread over the surface of the section, and another glass slide is placed onto it (several adhesives are available for cementing specimens to glass, including Super Glue and different types of epoxy resins, but highly recommended is a UV-cured resin, acrylic acid and hydroxypropyl methacrylate, which is obtainable from Logitech Ltd, Glasgow, U.K.). The two slides should be clamped together until the new slide is firmly cemented to the specimen. The slide ‘sandwich’ is then placed on a hot plate, and the sticky wax attaching the specimen to the first glass slide is melted off. The first slide can then be removed and discarded, leaving the polished face of the tooth now attached permanently to a glass slide.

The other face of the section is then lapped with a coarse aluminium oxide slurry (e.g., 9 micron) to remove any remaining wax, the Super Glue layer, and any saw marks, to within about 40 μm of the desired 100 μm thickness. Lapping is then continued with a fine grade of aluminium oxide (e.g., 0.5 micron) until the remaining excess material has been removed. The slide and section are then cleaned thoroughly in an ultrasonic bath. We do not routinely etch our specimens, but if it is desired to do so, then it should be done at this stage. Etching is done to enhance cross
striations or striae of Retzius and is also particularly valuable for identifying Wilson bands. After polishing and cleaning, the slide and section are suspended in any one of a variety of acid solutions, for instance, hydrochloric, phosphoric, or EDTA (ethylenediaminetetraacetic), depending on the purpose of the treatment (for details on alternative treatments to accentuate various structures, see Boyde et al., 1978; Grine, 1986; Hillson, 1986; Schmidt and Keil, 1971; Wilson and Schroff, 1970). The specimen on its slide is dehydrated in absolute ethyl alcohol and thoroughly dried under vacuum over silica gel. It should then be immersed in xylene and kept under vacuum until air bubbles stop coming from the specimen. After removal from the xylene, the coverslip is mounted by floating it over the section on a drop of mounting medium (FitzGerald, 1995). Finished slides should be left for several weeks to set.

AGE ESTIMATION AND TIMING OF DEVELOPMENTAL EVENTS

Several approaches are available, depending on which incremental growth markers are used, for estimating crown development time, chronological age, or establishing the timing of certain developmental events, like the formation of Wilson bands. Fundamentally, they may be grouped into two categories: those that interpret internal markers, the most accurate, and those that employ (surface) perikymata, the least accurate. The former, which have the advantage of yielding often very exact chronologies, suffer the downside of requiring that teeth be sectioned. This is neither always practical nor appropriate, as in the case of rare early hominin teeth, and for these, no other recourse is available but to use the less-accurate perikymata approach.

Determinations of crown formation time and developmental event timing are made on single teeth, which can be either fully mature or still immature. Subadult age estimation requires a more specific set of criteria, the primary one being that at least one tooth in the dentition not be fully grown. Also, the most accurate age estimates can be made when at least one of the teeth being analyzed has a neonatal line (these are all deciduous teeth, first permanent molars, and sometimes permanent lower central and lateral incisors and upper central incisors). Lastly, if the dentition is nearly complete, for instance, if the only crown not fully developed is the M2, then examples of all teeth that provide a full chronological history of the dentition must be histologically analyzed, and nonoverlapping results must be summed (e.g., in the example cited, this might involve analyses of an M1, a C, and a PM3—the method for determining the nonoverlapping portions of their development time will be described later). The focus of the following section will be a discussion of techniques involving single teeth; age estimation will simply require extrapolation of the method being discussed to the relevant number of teeth in the dentition.

A special case of age estimation arises where crowns in the dentition are fully mature, but apical closure of all teeth has not yet occurred (i.e., root development of at least one tooth is incomplete), which means that a summation of enamel development will fall short of total age by the amount of root development that extends beyond it. Since we have confined ourselves here to techniques that analyze enamel microstructures, interested readers are referred to Dean and Beynon (1991), Dean et al. (1993), Dirks (1998) and Dean and Leakey (2004) for information on how to estimate incomplete root development.

NONINVASIVE ESTIMATES USING PERIKYMATA

In their 1985 paper, Bromage and Dean counted all perikymata on the surface of the Plio-Pleistocene hominin teeth that they were studying. They assumed that the circaseptan interval in all of these teeth was seven days, and by multiplying the number of perikymata they counted by this number, they calculated the time in days
taken to form the imbrical enamel portion of the crown (the white area in Fig. 8.2c). To arrive at total crown formation time, they had to add an estimate of cuspal enamel time to this figure (the shaded area of enamel in Fig. 8.2c). Since the crowns of all of their specimens were not yet fully mature, this permitted the age of the specimens rather than simply crown formation times to be determined. However, it meant that for incisors they had to estimate a third factor—the time elapsed from birth to initial calcification (since these teeth do not usually begin to mineralize prenatally). These authors estimated the total time for these two factors for the incisors in their study to be six months. This estimation was based on some of their own investigations of modern human teeth and on the limited data that were then available in the literature.

The obvious weaknesses to this method are the errors originating from the estimates that are required. Modal values and means for circaseptan intervals have been established for many studies, and it is known that they occur commonly in a range from 8 to 10 days. However, these data are almost derived wholly from modern humans. Although the number of sectioned early hominid teeth has grown since 1985, the numbers of teeth remain small, and they are not likely to expand significantly because of the reluctance to subject rare fossils to a potentially destructive procedure. Therefore, the data for circaseptan intervals and cuspal enamel formation for other hominid species are too poor to allow any potentially helpful statistical conclusions to be drawn. This finding is ironic since nondestructive estimates are most appropriate for early hominid teeth, the very teeth where the most apposite data are simply not known with certainty.

The confidence that can be attached to estimates for these three factors is therefore not great, and this problem is irremediable without more research data, particularly on early hominid species. This is not likely to be substantial without an enhancement in technology that will permit these variables to be established from fossil teeth nondestructively. Nonetheless, in the absence of any other noninvasive endogenous approach for estimating the development of extinct species, this method still has an important role to play and it continues to be used in many palaeoanthropology studies (Dean and Reid, 2001; Dean et al., 2001, 2003; Dean and Leakey, 2004; Guatelli-Steinberg et al., 2005, 2006; Reid and Ferrell, 2006; Schwartz et al., 2003; Smith et al., 2004).

However, it is strongly recommended that when perikymata-based approaches are used, a range of estimated ages should be provided. This range will reflect the recognition that there is variation within the three parameters of circaseptan interval, cuspal enamel, and initial calcification that have been input into the calculation. It is also imperative that perikymata counts be taken from photomontages of the whole buccal/labial surface of the tooth being analyzed, and these should preferably be taken in an SEM. Attempting to count perikymata “by eye” under a low-power stereomicroscope is a very difficult task since the tooth surface must be illuminated by direct light, the angle of which is critical and changes down the curved surface of the tooth. Perikymata should be marked on the photomontage, and counts should be repeated until a clear consensus figure is reached.

**HISTOLOGICALLY ESTABLISHED ESTIMATES**

To view dental microstructures, a good light microscope with polarizing attachments is required. These attachments consist of two polarizing filters, one fitted to the condenser below the ground section being viewed, called the polarizer, and other above the section in the objective, called the analyzer. These filters polarize the transmitted light; that is, they “comb out” incoming light and only permit light rays vibrating in one plane to emerge. The polarizer and analyzer are adjusted so that the vibration planes of light transmitted
through them are perpendicular to each other. In this “crossed” position, with no tooth section on the microscope stage, no light is transmitted through the analyzer and the field appears dark when viewed through the objective. However, crystalline substances, like the hard tissues of teeth, exhibit a property called birefringence or the ability to be “doubly refracting.” Light waves traversing them are reorganized into two sets of waves vibrating in perpendicular planes, with the refractive indices of the two planes usually being different. In other words, the light vibrations in one plane travel faster through the tooth section than those in the plane perpendicular to them. The effect of this is to “twist” the emerging rays, which will now no longer be polarized in the same plane. When resolved by the analyzer and viewed through the objective, these twisted rays from the ground tooth section will appear in a range of interference colors that is related to the amount of birefringence and the thickness of the section.

Enamel microstructures are therefore observed more easily under polarized light, with polarizer and analyzer “crossed” to give maximum contrast (Hillson, 1986, 2005; Schmidt and Keil, 1971). Ideally, the microscope used should be able to magnify in powers as low as 100x, with at least one or two more increments through to 400x. A camera attachment, or a device capable of capturing digital images, is also a necessity.

**Counting All Cross Striations**

Boyde (1963) was the first to use dental histological analyses to estimate subadult age. He matched distinctive striae in different growing teeth in an archaeological specimen and counted cross striations from one tooth to the next. The objective was to count the total number of all cross striations formed from birth (registered by the neonatal line) until death (the end of enamel production). In his paper, Boyde sectioned the first permanent molar and the maxillary central incisor of a child from an Anglo-Saxon grave. He identified the neonatal line in M1 and counted the cross striations along a prism from it to the surface (see Fig. 8.6, which also elaborates on the rationale). He then traced the stria of Retzius that intersected with this prism back to the EDJ. Picking up the prism from the point of intersection, he again counted cross striations along it to the surface and repeated the process until he reached a stria that was distinctive enough to be matched with a homologous stria in the incisor. Carrying the total count of days forward, he continued the same procedure on the incisor until the last prism in the cervix was counted. He was then able to arrive at the total number of days from birth to death of this child.

Although highly accurate, this methodology, which involves counting all cross striations,

**Figure 8.6** Schematic illustrating Boyde’s histological age estimation technique. The two outline drawings are longitudinal sections of teeth, the upper a mandibular first molar and the lower a mandibular central incisor. The numbers are cross-striation counts. The dotted lines represent significant striae of Retzius that could be identified in both teeth. The vertical lines labeled, e.g., AA’, represent prisms. Boyde in fact used more prisms than shown, but those illustrated demonstrate the technique using the significant Retzius lines. Cross striations are first counted from the neonatal line in M1 out to the surface along prism AA’. Note that when the ameloblast forming prism AA’ reached the surface at A’, the incremental layer was situated along the Retzius line approximated by A’B, and the next ameloblast to start secreting enamel was at point B. Therefore, the next cross-striation count is made along prism BB’, the next along CC’, and so on. Looking at the lower drawing the pattern of significant striae shows the same cross-striation counts up to the incremental line at 1122. Counting is picked up from point D in the upper diagram and then transferred to point J in the lower diagram and is continued until the final total of 1692 days is reached at prism KK’. This is the point where enamel formation ceased at death. (Adapted from Boyde, 1963).
is tedious and it requires excellent visibility of both cross striations and striae of Retzius throughout the whole enamel, which is not often the case. Several other approaches are easier, some of which exploit the uniformity of the circaseptan interval, and these will be examined next.

Using Short-Period and Long-Period Markers

Techniques for calculating crown formation times that rely on uniformity of circaseptan intervals must derive the two components of enamel growth, imbricational enamel formation and cuspal enamel formation, and these are usually obtained separately. Imbricational growth can be calculated by multiplying the total number of striae of Retzius to emerge at the surface, by the circaseptan interval, the number of cross striations between adjacent striae of Retzius (see Fig. 8.7). Since, as has been said, the circaseptan interval may differ from individual to individual, but is identical within all the teeth of one individual, it is therefore only necessary to determine the number of days between striae once within one tooth from any dentition being studied. This is done in whichever tooth cross striations and striae are discernible most clearly in one field of view. Cross striations should be counted along a prism between two adjacent regular striae.

Once the circaseptan interval has been established, the first stria of Retzius to terminate at the surface is identified (at P₁ on Fig. 8.7). In a permanent tooth, this stria will emerge as a distinct perikymata, but this is not the case with deciduous teeth, in which perikymata are not well defined. In either instance, this point, which marks the junction between cuspal and imbricational enamel, is often difficult to determine and the section should be scrutinized carefully to ensure that the correct location is identified accurately. Once established, all of the striae terminating at the surface are

\[
\text{Crown Formation Time} = Pr A_{(\text{days})} + (CI \cdot P)
\]

Figure 8.7  Schematic diagram of a longitudinally sectioned tooth showing derivation of formula for calculating total crown formation time and the timing of developmental events, like Wilson bands.
counted away from it cervically down the face of the crown (labeled P₂ to Pₙ on Fig. 8.7). Particular care should be exercised to ensure that the enamel at the CEJ, which is fragile and may have broken away postmortem, is present.

Imbricational growth is then determined according to the following formula:

\[ CI \cdot Pₙ \]

(where \( CI \) is the circaseptan interval in days and \( Pₙ \) is the total number of striae to terminate at the crown surface).

There are several ways to estimate the time taken to form cuspal enamel, but in all cases, the highest accuracy is obtained by working from photomontages of the cuspal area, rather than by eye directly through the microscope. Montages have been assembled traditionally by pasting together individual photomicrographs, taken across the area of interest. However, if digital images are available, captured either directly through the microscope using a digital video camera, by photographing with an analog video camera and converting the signal with a digitizing frame grabber, or by digitally scanning photomicrographs, then photomontages can be assembled on the computer. This process offers the very real advantage of being able to use image analysis software to take measurements and counts from the digital montage. Montages may be put together using the computer program Adobe Photoshop. A variety of sophisticated image analysis packages are available for purchase, but there are also a number of freeware programs that can be downloaded from the Internet, with more of these being developed all of the time. The one with the longest pedigree is called NIH Image, which is a public domain image processing and analysis program for the Macintosh developed at the Research Services Branch of the National Institute of Mental Health (NIMH). There is an equivalent open-source, Java-based program called ImageJ that runs on Linux, Mac OS 9, Mac OS X, and Windows, which the makers say was “inspired” by the NIH Image analysis program.

Dean (1998a) evaluated four ways of determining cuspal enamel formation times, but he found few differences among them. Each may be “expected to provide equally valid results, with each having different practical advantages in different situations” (Dean, 1998a:460). A variant of one of these approaches will be described here.

A prism running from the tip of the dentine horn (or near to it) to the point at the junction of cuspal and imbricational enamel (Prism A at P₁ in Fig. 8.7) represents the time elapsed since the initiation of crown growth. A montage should be pieced together that visualizes the full run of this prism from the EDJ to the occlusal surface (digital images or photomicrographs taken at 200x to 400x magnification will probably be most appropriate). If cross striations are easily discernible along the full length of this prism, then counting all of them will yield cuspal enamel formation in days. If cross-striation clarity is good but not complete on Prism A, then cross striations need not be counted along its full length; counts can be transferred from one adjacent prism to another at different levels along prominent striae of Retzius (see Fig. 8.8 and Risnes (1986) and Beynon et al. (1991).

However, counting all cross striations can be a time-consuming process, and sometimes they cannot be visualized through the full length of the enamel mantle, particularly in deciduous teeth. In these circumstances, the time taken to form Prism A (in Fig. 8.7) can be estimated, and this is most easily done using image analysis software and a computer spreadsheet. First, the apparent path taken by the prism on the digital montage is traced out, and this path is then measured using image analysis software. The prism is scrutinized closely, and wherever cross striations are discernible, either on the prism or, more commonly, on adjacent prisms, their positions are recorded on the montage. Next, a cross-striation repeat interval, representing the average daily appositional growth rate of as large a group of cross striations as can be clearly discerned, is calculated at each of these points on the prism path. For example, a group of prisms, say 6, is measured and then
divided by 6 to arrive at an average cross-striation repeat interval (representing one day’s mean growth of enamel at this point on the prism). The distance between each of these recording points is measured. Using a computer spreadsheet, the average rates can then interpolated in 5% lengths along the prism. Finally, knowing the total distance of the prism and the average rates of production at 19 points along it, the total time taken for its formation can then be calculated. If image analysis software is not available, then estimates of length of both the prism and the groups of cross striations can be made by using needlepoint callipers on the photomontage, and the number of percentiles can be reduced if necessary (e.g., to 10% or 20% increments). The resulting estimate will not be as accurate, but it will still be very good. After cuspal enamel time is calculated, it is added to imbricational enamel to arrive at total crown formation time.

As illustrated in the schematic Fig. 8.7, it is possible using this technique to time particular developmental events like the neonatal line (in absolute terms from prenatal initiation of calcification) and Wilson bands. The timing of the neonatal line, which will always cross the cuspal/imbricational prism, will simply involve subdividing the estimate of Prism A made above into prenatal and postnatal portions.

This split will also permit the chronological age of the specimen (rather than the total crown formation time) to be established.

Depending on where they are found in the crown, Wilson bands can be timed by modifying the techniques for arriving at imbricational and cuspal enamel timing just discussed. In the example shown in Fig. 8.7, for instance, where the Wilson band is found in the imbricational enamel, the number of striae from P1 to the striae just cuspal to the Wilson band are counted and multiplied by the circaseptan interval (in Fig. 8.7 there are none). The prism that intersects the junction of the next regular striae after the Wilson band is identified (Prism B at P2 in Fig. 8.7) and the number of cross striations along this prism from the cuspal striae is counted. The age of occurrence of the Wilson band in this case would therefore be calculated as

\[
\frac{Pr A_{(days)}}{C0} + (CI \cdot P_x) + (NCS)
\]

(where \(P_x = \) number of striae from \(P_1\) cervically up to the Wilson band and \(NCS = \) number of cross striations from the striae at \(P_x\) to the Wilson band).

It is known that prisms do not pursue straight courses through the enamel mantel; they weave

Figure 8.8 A longitudinal photomicrograph (at 200x magnification) of enamel showing how cross striations (which have been marked by thin white lines) can be tracked across a series of striae and transferred from one striae to another so that clearer fields of view are used (Dean and Beynon, 1991).
or decussate from the EDJ to the surface (Boyde, 1989; FitzGerald, 1995; Osborn, 1990; Risnes, 1986). Under the cusp, directly over the tip of the dentine horn and on the true axis of the cusp, prisms spiral toward the surface in an area known as “gnarled enamel.” More laterally away from the centre of the cusp, decussation is reduced, but prisms still weave from side to side in both the longitudinal and the semi-transverse dimensions, which means that the prism path that seems to be continuous in a two-dimensional longitudinal section is actually a composite of several prisms that weave into and out of the plane of section. Decussation is greatest in the inner third of the cuspal enamel lateral to the cusp center, and prisms pursue a much straighter course in the outer two thirds. Prisms also become straighter and straighter moving laterally and cervically in a tooth. This movement has represented a point of potential concern in that it was thought that the effect might be significant enough to underestimate cross-striation counts in areas at the junction of cuspal and imbricational enamel. Suggestions on how to correct this count upward have been made (FitzGerald, 1995). However, the paper by Dean (1998a) that compared cuspal enamel formation times determined from cross-striation counts with times independently determined in three other ways found no significant differences. Dean established cuspal enamel formation times four ways in Pan and Pongo as well as in Homo and discovered that the cross-striation count was understated against the other methods only in Homo, which implies that decussation was greatest in our species. However, the difference in Homo was not significant (less than 3%) and may well have reflected measurement error. On this basis, therefore, no adjustments to cross-striation counts need be made to account for possible decussation.

Cuspal enamel formation time may be calculated in other ways, for instance, by counting all striae under the cusp tip (Bullion, 1987; Dean, 1985) and then by multiplying this number by the circaseptan interval. Several other approaches rely on a formula derived by Shellis (1984) that determines the enamel extension rate (also see Dean, 1998a). This is the daily rate at which, during crown development, new enamel formation extends in the cervical direction. The extension rate can be estimated at each point where a stria of Retzius intersects the EDJ, and it can therefore be used to calculate not only cuspal enamel formation time, but also total crown formation time, if the calculations are continued cervically along the EDJ all the way to the CEJ.

**CONCLUSIONS AND FUTURE RESEARCH DIRECTIONS**

More research needs to be done by dental biologists, anatomists, histologists, and chronobiologists into the nature of dental microstructures and into the fundamental causes of their observed regular periodicity. An abundance of both direct experimental data (see the earlier discussion of Smith (2006) for the most telling example) as well as deductive evidence (Antoine’s work (2000) is an exemplar of a quality study) exist that verify the regular periodicity of some microstructures. These data provide ample justification for the use of microstructures by anthropologists and others in assigning chronology to dental development events, but many unknowns still surround their essential formation and growth.

There are still problems to be resolved in the anthropological application of dental histological techniques. One of the most significant is the requirement to partially destroy the tooth being analysed, which involves the technical challenges of producing a mid-longitudinal section that can be viewed through a microscope. However, the seemingly relentless march of technology appears to have cracked this “Holy Grail” of odontochronology, the ability to visualize internal microstructures nondestructively and well enough to incorporate them into odontochronological analyses. Astonishing
advances in synchrotron technology achieved this remarkable result just as this volume was going to press. Tafforeau et al. (2006) developed an approach that uses X-rays produced from a synchrotron source, which are monochromatic, nearly parallel, and have a high flux (unlike X-rays produced from conventional sources), to produce high-quality, microtomographic 3D images, even from densely fossilized specimens. Smith et al. (2007) have used these techniques to image the cross striations between adjacent striae of Retzius within the enamel of an early Homo sapiens juvenile from Morocco dated at 160,000 b.p. They were thus able to establish a circaseptan interval (of 10 days for this child) and to determine the cuspal enamel thickness from microtomographic slices noninvasively. Using this information, they could then calculate from the number of perikymata on the crown surface accurate ages of tooth development through the dentition, from which they were able to conclude that this juvenile is currently the oldest known member of Homo with a developmental pattern (degree of eruption, developmental stage, and crown formation time) that is more similar to modern H. sapiens than to earlier members of Homo.

Unfortunately, it will be sometime before most researchers can have routine access to a synchrotron (which is a particular type of cyclic nuclear particle accelerator). Until this becomes a reality, more large-scale studies need to be done to increase knowledge about the variability of dental microstructures within and between modern human and ape populations, and to increase the database of information that will lead to better estimates of some microstructural variables. This is happening (Dean and Reid, 2001; Reid and Dean, 2000, 2006; Schwartz et al., 2001), but many more studies are needed. These studies will increase the accuracy of perikymata-based approaches and increase the simplicity of other methods by allowing population-specific standards to be substituted for components that must now be calculated for each tooth being analyzed.

This chapter has focused on histological analyses of enamel, although the information provided on the microstructures of dentine and their incremental growth will have suggested that this tissue also has trapped in it growth and development data, if only we can improve techniques for its interpretation. So far dentine has only had limited exploitation, but clearly, since it comprises the bulk of a tooth, it offers perhaps a greater potential than enamel. A better understanding of the complexities of dentine growth and advances in techniques for its histological analysis are areas of research that hold out enormous future promise.

However, perhaps the most exciting areas of research are not associated with validating histological techniques or with making them easier to apply. They lie instead in the potential that these techniques have for resolving some fundamental anthropological problems. Quite simply, dental histological assessment using microstructures can provide information that is not obtainable in any other way. No better illustration of this potential can be found than the predicament facing biological anthropologists trying to assess growth and development from skeletal material. A fundamental problem centers on the use of appropriate exogenous reference standards, which was a question raised at the beginning of this chapter. However, the problem extends beyond the correctness of the use of one population standard on a different population—there are also concerns about the accuracy of the standards themselves. Tooth formation as traditionally measured by the appearance of growing teeth on radiographs is fraught with methodological problems (e.g., Aiello and Dean, 1990; Beynon et al., 1991, 1998; Risnes, 1986), and in addition, studies have used idiosyncratic approaches in the way that they collect data and assess tooth maturity; some have been cross-sectional, others semi-longitudinal, and some have used as little as three fractional stages to gauge tooth development, others as many as 20. Histological techniques that provide an accurate endogenous record of
tooth development can overcome these difficulties. Not only do they obviate the need to use radiographs, but they also furnish a continuous history of both rate and total time taken for dental development. Moreover, importantly, they also avert the necessity of having to select an appropriate reference population, or the best modern analog in the case of an extinct species.

While finding their greatest application so far in evolutionary ontogeny and studies of pre-modern hominid growth, dental microstructural analyses are beginning to be exploited in bioarchaeological studies (e.g., Guatelli-Steinberg, 2001; Hoppa and FitzGerald, 1999; Huda and Bowman, 1995; King et al., 2005; FitzGerald and Saunders, 2005; FitzGerald et al., 2006). The time and labor required to make assessments on large samples of subadult individuals have so far precluded widespread use of these techniques for bioarchaeological studies. But the potential benefits are clear, and improvement and standardization of methods and incorporation of computer-based digital imaging procedures will make dental microstructural analysis more accessible to nonspecialists and will make assessment of large samples feasible. Ultimately, of course, the ability to assess accurately dental formation patterns and timing is crucially necessary to reconstructing with confidence the age at death and other palaeodemographic variables of recently dead, long dead, or fossil specimens.

ACKNOWLEDGMENTS

We would like to gratefully acknowledge funding support from a Postdoctoral Fellowship from The Social Sciences and Humanities Research Council of Canada (CF). Particular thanks are also extended to Professor David Beynon and to Dr. Don Reid, both of the University of Newcastle upon Tyne Dental School, UK, for generously sharing their expert technical knowledge of tooth sectioning during development and modification of the approach briefly outlined here.

REFERENCES


Boyd A. 1963. Estimation of age at death of young human skeletal material from incremental lines in the dental enamel. Third International Meeting in Forensic Immunology, Medicine, Pathology, and Toxicology, Plenary Session 11A, London.


Guatelli-Steinberg D, Reid DJ, Bishop TA, Larsen CS. 2005. Anterior tooth growth periods in Neandertals were comparable to those of modern humans. Proc Natl Acad Sci USA 102:14197–14202.


Reid DJ, Ferrell RJ. 2006. The relationship between number of striae of Retzius and their periodicity

REFERENCES 261


INTRODUCTION

Dental morphology means different things to different people. Some authors consider tooth size one aspect of morphology, whereas others include shape under this rubric. I distinguish size from morphology because the methods of study and general underlying principles for each are distinct. If learning the intricacies of measuring teeth is your goal, there is a vast literature on the subject (cf. Hillson et al., 2005; Kieser, 1991). Although morphology and shape have more in common than morphology and size, shape also shows noteworthy differences (Morris, 1981, 1986; Taylor, 1969a, 1969b). Methods developed for ascertaining "tooth shape" for anthropological and forensic purposes have not been adopted widely, partly because they are difficult to replicate, diminishing their utility in comparative studies. Odontoglyphics, which is another approach to the study of tooth morphology, focuses on the complex pattern of furrows in multicusped teeth. Developed by A.A. Zubov (1968, 1977), this method involves the examination of negative relief on waxbite impressions. Although widely used in the study of Russian populations (cf. Zubov and Khaldeeva, 1979), it rarely has been used in other countries where the standard procedure is to make negative alginate impressions and positive plaster casts.

In this chapter, I focus on what most anthropologists refer to as dental morphology; that is, distinct features or traits of the crowns and roots that are present or absent and, when present, exhibit variable degrees of expression. Common examples include shovel-shaped incisors, upper and lower molar cusp number, Carabelli’s cusp, three-rooted lower first molars, and more. My goal is to provide background and a protocol for advanced undergraduates, graduate students, and professionals who are interested in the study of dental morphological variation.

To repeat the mantra I learned in graduate school, teeth are remarkably useful in anthropological research. Their advantages include preservability (in the fossil, archaeological, and forensic records; Fig. 9.1), observability (in the living, skeletons, and fossils), variability (dozens of measurements and discrete attributes of the crowns and roots that vary within and between populations), and heritability (a strong genetic basis underlying tooth development and trait expression). Teeth also provide geographic flexibility—no matter where you go, people have teeth. Whether you are studying crown wear, oral pathology, tooth size, or crown and root morphology, if there is
an interesting problem to pursue using teeth, all you need is an airplane ticket and a protocol.

A BIT OF HISTORY

The foundation for dental morphological research was laid by dentists, paleontologists, and physical anthropologists. Pioneers who made significant contributions to the field include Aleš Hrdlička (1920, 1921), William King Gregory (1922), Albert A. Dahlberg (1945, 1951, 1963), P.O. Pedersen (1949), C.F.A. Moorrees (1957), Bertram S. Kraus (1963), Kraus and Jordan (1965), Gabriel Lasker (1950, 1957), Kazuro Hanihara (1960, 1963), and Stanley M. Garn (1977), to name but a few. These workers focused on a number of key issues, including 1) long-term evolution of primate and hominid tooth form, 2) geographic variation in trait expression, 3) morphogenetic fields, 4) genetics and development of crown morphology, and 5) interaction between morphology and crown size and tooth number.

Despite the contributions of dentists who focused on dental morphology in the broader context of oral biology, the full potential of morphology was not realized until physical anthropologists started using crown and root trait frequencies to assess population origins and relationships. Dentists and anthropologists look at teeth through different lenses. Dentists bring their expertise in oral biology to the table when assessing crown morphology, but they are not as tuned in to the range of historical questions that can be addressed using this morphology. Anthropologists have less background in oral biology, but the mysteries of human history are their purview. Although either side may claim primacy in some areas of dental morphology, the contributions from both have been significant and syncretic and the principles that follow were developed by dentists, physical anthropologists, and scholars with doctorates in both fields.

FUNDAMENTAL ISSUES

Dental Anatomy

One item every student of dental morphology should have is a good textbook on dental
anatomy. Most texts are written by dentists who focus primarily on modal crown and root form, but some authors illustrate variations around the modal theme (cf. Kraus et al., 1969). There is, however, no better starting point than the standardized tooth (cf. Ash and Nelson, 2003).

A tooth is divided into two primary components (crown and root) and four primary tissue types (enamel, dentine, cementum, and pulp). The visible portion of the crown is covered by enamel, which is an extremely hard and durable substance that is ca. 97% inorganic. Underlying the enamel is dentine, which is a substance more in line with bone in terms of organic/inorganic composition. The pulp chamber, at the core of the tooth, is made up mostly of soft tissue (blood vessels and nerves). Cementum is the tissue that covers the surface of the root that, in concert with the periodontal ligament, anchors the tooth in its socket (Carlsen, 1987; Hillson, 1986, 1996).

The primary units that constitute the crown are called cusps, whereas the primary units of the root are cones. Following the terminology of Carlson (1987), the crown of a single-cusped tooth and each cusp of a multicusped tooth are made up of essential and accessory lobes and lobe segments that, in turn, can exhibit essential and accessory ridges. Root number is defined by the presence or absence of inter-radicular projections that separate or fail to separate root cones. Often, root cones are fused together but show a developmental groove (root groove) that separates the two cones. Such unseparated cones are called radicals.

Humans and other primates have four types of teeth, and each type is represented in both the upper and the lower jaws, albeit with their own distinct morphology. The anterior teeth include spatulate incisors and conical canines. All are characteristically one-cusped and one-rooted teeth. The posterior teeth, or cheek teeth, consist of premolars and molars. Premolars usually have two cusps, although in the lower jaw, the lingual cusp of the first premolar often lacks a free apex and the second premolar frequently exhibits three cusps (one buccal, two lingual). There is no modal root number for premolars in either jaw—the first premolars have either one or two roots (occasionally three), whereas the second is mostly a single-rooted tooth. For the upper molars, modal cusp number is four, whereas modal root number is three. For lower molars, modal cusp number is five and modal root number is two. The exceptions to these normative characterizations constitute a significant part of dental morphology. As an adjunct to this basic description, refer to the Sidebar for brief definitions of terms used to describe location and direction in teeth in particular and jaws in general.

**TERMS USED TO DESCRIBE ELEMENTS OF THE HUMAN DENTITION**

- **Midline**: line that follows the sagittal plane between the central incisors that divides the upper jaw and lower jaws into left and right sides
- **Quadrant**: one side (right or left) of either jaw
- **Antimeres**: corresponding teeth in the two quadrants of a jaw (e.g., left UM1, right UM1)
- **Isomerese**: corresponding teeth in opposing quadrants of a jaw (e.g., left UM1, left LM1)
- **Anterior teeth**: incisors and canines
- **Posterior teeth (or check teeth)**: premolars and molars
- **Upper**: maxillary
- **Lower**: mandibular
- **Incisal**: cutting surface of the anterior teeth
- **Occlusal**: chewing surface of the posterior teeth
- **Mesial**: toward the midline
- **Distal**: away from the midline
- **Lingual**: toward the tongue (used for all teeth)
**Labial**: toward the lip (used for anterior teeth)

**Buccal**: toward the cheek (used for posterior teeth)

### Types of Traits

Morphological traits take several different forms. In Carlsen’s (1987) terminology, the general forms that subsume one to many variants are 1) lobes, 2) marginal ridges, 3) tuberculum projections, 4) occlusal tubercles, 5) cervical enamel lines, 6) root cones, and 7) supernumerary radicular structures. Not included in this classification are variants in supernumerary occlusal cusps, accessory ridges, and socket orientation.

- Lobes (or major cusps; e.g., hypocone or hypoconulid)
- Marginal ridges (e.g., shoveling, double-shoveling, or marginal tubercles)
- Tuberculum projections (e.g., *tuberculum dentale*, Carabelli’s trait, or protostylid)
- Occlusal tubercles (e.g., tuberculated premolars or odontomes)
- Cervical enamel line (e.g., enamel extensions and pearls)
- Supernumerary occlusal cusps (e.g., cusp 6 or cusp 7)
- Accessory ridges (e.g., canine distal accessory ridge or premolar mesial and distal accessory ridges)
- Socket orientation (e.g., UI1 bilateral winging)
- Root cones (e.g., two-rooted UP1 or two-rooted LC)
- Supernumerary radicular structures (e.g., three-rooted LM1)

### Morphogenetic Fields

In 1939, Butler introduced the concept of morphogenetic fields to the study of the mammalian dentition. He recognized three primary fields—incisors, canines, and molars. In this scheme, premolars were considered anterior extensions of the molar field. The significance of the field concept is that neither genes nor evolution act on individual teeth but on overall fields. Selective pressures to enhance the grinding surfaces of the molars, for example, impact not only the molars but also the premolars. The distal premolars of grazers grade imperceptibly into the molars. By contrast, carnivores need slicing and dicing teeth, so premolars are reduced frequently or even lost altogether in this order as are some molars (cf. Hillson, 1986; Peyer, 1968).

A.A. Dahlberg (1945) adapted Butler’s field concept to the human dentition, and the organizational principles he laid out are invaluable to students of dental morphology. He defines four rather than three fields in the human dentition by adding premolars. In considering size, morphology, and tooth number (missing or extra teeth), each tooth district (I, C, P, M) has stable and variable members. Usually, the key tooth (most stable) in each district is the most mesial member (i.e., UI1, UC, UP1, UM1, LC, LP1, and LM1, where U = upper and L = lower). The only exception to that generalization is LI2, which is more stable than LI1. The most distal member in each district is typically the most variable (UI2, UP2, UM3, LP2, LM3). In the molar tooth district, where three rather than two members exist, the second molar is intermediate in variability between the mesial and the distal elements. This point is not obscure, but it is one that is relevant to the operational definitions of key morphologic variables.

### Interactions Among Size, Morphology, and Tooth Number

Tooth morphology has never been linked to anthroposcopic, anthropometric, dermatoglyphic, or serologic variables. However, some relationships should be considered when studying morphological traits. The strongest correlations are between the same trait expressed on different members of a single tooth district and intra-trait–intra-field interactions. An example
would be hypocone expression on UM1, UM2, and UM3 (Fig. 9.2). To go one step more, examples of intra-trait–inter-field correlations exist, including shoveling on UI1, UI2, LI1, and LI2, which show six significant pairwise correlations (Scott, 1977a). Additional examples of significant intra-trait–inter-field correlations include *tuberculum dentale* of UI1, UI2, and UC (Fig. 9.3) and the distal accessory ridge of the upper and lower canines (Scott, 1977b, 1977c). An example of an inter-trait–intra-field relationship has been found for Carabelli’s trait and the hypocone of the upper molars (Keene, 1968; Scott, 1979). One of the few examples of a significant inter-trait–inter-field relationship is between Carabelli’s trait (UM1) and the protostylid (LM1) (Scott, 1978).

Because of consistently strong and positive intra-trait–intra-field correlations, population comparisons should focus on only one member of a field for a specific trait. Usually, the most stable member of a district should be used because trait expression on this tooth shows the highest degree of genetic control (and the least amount of environmental plasticity). Shoveling expression, for example, is best represented by UI1 as this tooth shows greater symmetry than UI2 and a broader range of variation than LI1 and LI2. In population comparisons, shoveling can be scored on all four teeth, but only UI1 shoveling frequencies should be used in a distance analysis to avoid the inclusion of redundant information. Although the key tooth is usually the best field member to represent a specific trait, exceptions
to this rule exist. For some traits, such as the hypocone of UM1 where prominent expression is the norm, it is preferable to compare UM2 hypocone frequencies as this tooth exhibits more variation in terms of total presence and degree of expression.

**Dental Genetics**

In the 1950s, when dental morphology was gaining momentum in anthropological studies, some authors believed that discrete crown traits might have simple modes of inheritance (Lasker, 1950). Kraus (1951) and Tsuji (1958) studied Carabelli’s trait in families and found parent–offspring segregation patterns consistent with simple autosomal codominant and dominant inheritance, respectively. The ramifications were exciting. If crown trait frequencies could be reduced to gene frequencies, the population genetics models of Wright–Fisher–Haldane could be applied to the teeth of extant and extinct populations. In the 1960s, Turner (1967, 1969) did exactly that when he converted phenotype frequencies to presumed gene frequencies for prehistoric and living Hopi Indians and Koniag Eskimos and for living Europeans. His goal was to compare living and prehistoric populations to estimate the level of European admixture in the modern Native American groups.

In the 1970s, additional attempts to corrobate or refute simple genetic models of inheritance for crown traits led workers down a different path. Given the nature of dental morphological expression, where traits can be absent, slight, moderate, large, and so on, one could test simple models of dominant-recessive inheritance only under a certain set of assumptions. For example, one had to assume that trait absence represented one homozygous genotype. For trait presence, the assumption was that intermediate expressions were associated with heterozygous genotypes, whereas pronounced expressions were associated with the dominant homozygous genotype. These assumptions required a substantial contribution from the environment to smooth out the range of variation of the two genotypes for trait presence.

In pondering the question of how environmental factors could influence trait expression, Grünberg’s (1952) model of quasicontinuous variation became the more parsimonious explanation for the nature and inheritance of morphological traits. The quasicontinuous model holds that some discontinuous traits have continuous genotypic distributions with underlying (absent) and visible (present) scales separated by a physiological threshold. Inheritance under this model is polygenic with genes at many loci interacting to produce the final phenotype. Edwards (1960) noted that many medically significant traits such as cleft lip/palate, pyloric stenosis, and spina bifida, which were traditionally thought of as simple Mendelian traits, were quasicontinuous variants with complex modes of inheritance. Even though they were not Mendelian traits, they simulated the behavior of simple inheritance models.

In analyzing segregation patterns of six traits in 53 American white families, Scott (1973, 1974) found two that conformed to simple models of inheritance. Carabelli’s trait segregated like a simple Mendelian dominant, whereas cusp 7 followed a pattern consistent with autosomal recessive inheritance. On closer examination, it was found that low-frequency traits in general (e.g., cusp 7) followed segregation ratios consistent with or close to the expectations of recessive inheritance, whereas high-frequency traits (e.g., Carabelli’s trait) simulated simple Mendelian dominant patterns. As total trait frequency in a population should not have anything to do with inheritance patterns, a closer perusal of the data suggested that morphological traits, which also show a high correlation between total trait frequency and degree of expression, were threshold dichotomies (Wright, 1934, 1968) with complex modes of inheritance, not point dichotomies with simple modes of inheritance.

One of the more telling lines of support for the quasicontinuous model comes from the distribution of shovel-shaped incisors in
Indians of the American Southwest. This trait attains or approximates a frequency of 100% in living Southwest Indian groups. When the distribution of graded trait expressions is plotted, it approximates closely a normal curve, which is the precise expectation one would have when a quasicontinuous variant becomes a continuous variant (Scott and Turner, 1997). In 1977, Harris tested genetic models on a wide array of crown traits in Melanesian families and concluded these traits were quasicontinuous variants with complex modes of inheritance.

Kolakowski et al. (1980) and Nichol (1989, 1990) applied methods of complex segregation analysis to the problem of morphological trait inheritance and found some indications that major genes are involved in the inheritance of some traits. However, even if traits have major genes underlying their development, it is still not practical to tease these genes out and to reduce trait frequencies to gene frequencies. From a practical standpoint, the ramifications of dental genetic studies are as follows:

• Twin and family studies indicate dental morphology has a strong heritable component.
• Inheritance is not simple, so traits cannot be reduced to gene frequencies.
• For quasicontinuous traits, the best way to characterize a population is through the use of total trait frequencies. As total trait frequency represents the threshold separating trait absence from trait presence, this one number specifies the entire continuous and normal distribution of the genotypic variation underlying trait expression (Falconer, 1960).
• In population studies, either total trait frequencies or frequencies defined by breakpoints (e.g., individuals showing grade 3 or higher/total number of individuals) are equally useful parameters for characterizing the genetic variation underlying trait expression in a particular group.

METHODS

Trait Classifications

The collection of data on crown and root morphology begins with operational definitions of trait expression. These definitions focus primarily on the precise location of the trait and its range of variation in size and form. Early workers often scored only obvious and pronounced trait expressions. The problem with this approach is that observers often disagreed over what constituted obvious and pronounced expressions.

In his classic paper on shovel-shaped incisors, Hrdlička (1920) appreciated the problems inherent in ill-defined trait definitions. Shoveling was clearly not just present or absent. Marginal lingual ridges ranged from absent to pronounced, but between these extremes, many gradations existed. Hrdlička ultimately arrived at a four-grade classification: 1) no shovel, 2) trace shovel, 3) semi-shovel, and 4) shovel or full shovel.

In 1956, A.A. Dahlberg released a series of standardized plaques from the Zoller Clinic in Chicago that were intended to bring order and consistency to the study of morphological variation in human populations. One of these plaques codified Hrdlicka’s four-grade shoveling scale. For other variables, such as Carabelli’s trait and the protostylid, Dahlberg developed scales with more grades of trait expression. The Carabelli’s plaque, for example, had absence (grade 0) and seven degrees of trait presence (from very slight grooves and pits to large free-standing cusps). The protostylid plaque included absence, buccal pits (grade 1), and five degrees of positive expression. For the hypocone, four grades of expression were codified as 4, 4−, 3+, and 3.

Building on the pioneering efforts of Dahlberg, Christy G. Turner II released two standard plaques for cusp 6 and cusp 7 in 1970. These traits represented the first step in what would become the Arizona State University Dental Anthropology System or ASUDAS (Turner et al., 1991). Turner, in
concert with his students, developed an extensive series of trait classifications that included not only crown traits but also root traits. In collaboration with ASU, Dahlberg sent his original plaques to be duplicated and distributed along with the new ASU trait classifications. Turner et al. (1991) provide a summary of the entire set of classifications developed at Chicago and ASU. These plaques have been distributed throughout the world and are often employed in the study of living, prehistoric, and fossil hominid populations.

I developed several ASU standard plaques (shoveling UI1, UI2; tuberculum dentale UI1, UC; distal accessory ridge UC, LC; multiple lingual cusps LP1, LP2) as part of my dissertation (Scott, 1973). The principles in setting up a trait classification are as follows:

- Find a distinct trait. This can take the form of a supernumerary occlusal cusp, an accessory ridge, a marginal tubercle, etc.
- This trait should be expressed in a consistent manner on all members of a tooth district (in some cases, between tooth districts as well, e.g., accessory ridges on occlusal surface of premolars).
- Once a trait is isolated, its variation should be examined in numerous casts and skeletons representing geographically diverse populations.
- When the full range of expression has been sampled, find teeth that express absence and the most pronounced degree of trait expression.
- After the lower and upper boundaries have been established, find additional examples that gradually increase in expression from grade 0 (absence) to maximal manifestation (grade X), making some effort to keep the grade differences approximately equal (Fig. 9.4).
- The number of grades for a specific trait is dictated by its overall range of expression, e.g., premolar mesial and distal accessory ridges vary to a modest degree; in this case, absence and three degrees of presence are adequate. For a trait that shows a large range of expression, such as shoveling, more grades are called for (anywhere from five to seven as a rule).
- Use examples from the specific tooth you are interested in, not just examples from the tooth district where the trait is expressed. The ASUDAS hypocone plaque includes expressions of this trait on UM1 and UM2, and this has required workers to adjust their observations. To avoid such adjustments, each upper molar should have its own plaque for hypocone variability.

The Hrdlička classification for shovel-shaped incisors provides an object lesson for a pitfall to avoid in setting up a standard plaque (see Fig. 9.5). Data on American Indian shoveling collected on the basis of the Hrdlička–Dahlberg scale shows a distribution that is strongly skewed to the right. Typically, the full shovel category ranges from 75% to 95% with correspondingly small values for trace and semi-shovel and basically no absence expressions. After I developed an eight-grade scale for shoveling and used it as the basis for

![Figure 9.4](image.png)

Figure 9.4  Ranked scale for canine tuberculum dentale; note the absence of trait for grade 0 and the pronounced expression of tubercle on grade 7. Intermediate ranks approximate even spacing between any two grades. Only left upper canines were used in this classification.
collecting data on Southwest Indian tribes, a different picture emerged. With absence and pronounced shoveling (grade 7) framing the overall pattern of variation, the distribution of shoveling in American Indians went from strongly skewed to approximately normal (Scott, 1973; Scott and Turner, 1997).

Another practice to avoid in trait classifications is the use of different manifestations of a single trait as separate variants. A.C. Berry (1976) adopts this approach for Carabelli’s trait where she lists “Carabelli’s pit present” and “Carabelli’s cusp present” as two separate variables. She also considers “groove pattern Y on molar 1” and “groove pattern X on molar 1” as separate traits. Each trait should be represented only once in a population analysis and on a single member of the field it occurs in. That is, any given individual should have only one grade of Carabelli’s trait, and for distance analysis, only the trait frequency on UMI should be used.

Crown and Root Traits

The following set of traits can be used as a starting point in any population study. The addition or subtraction of variables from this list should be dictated by the problem at hand. The classifications of most traits, with descriptions of each grade, can be found in Turner et al. (1991). Here, I provide a brief characterization of each variable along with some cautionary notes.

Maxillary Incisors

UI1 winging. This trait is one of the few morphological characteristics that is not reflected directly on a crown or root. Instead, winging is dictated by the orientation of the upper central incisor sockets (Dahlberg, 1959). Rather than presenting a flat or slightly parabolic surface, the distal border of the labial faces of the central incisors are rotated outward (Fig. 9.6).

Winging, a frequent trait in Asian and Asian-derived populations (especially American Indians), is not a function of anterior tooth crowding. In all but a few instances, genetically based bilateral winging can be distinguished readily from cases that result from crowding (typically, these are much more asymmetrical). A substantial range of variation occurs from slight winging to pronounced bilateral winging, but this has not been codified on a plaque. Minimally, an observer should classify the trait as absent, slight, moderate, or pronounced. Counter-winging, where the distal borders of the incisors turn in rather than out (Dahlberg, 1963), is not as useful as bilateral winging because this condition is often the result of crowding.
Shoveling. Technically, the descriptive term for this trait should be lingual marginal ridging (Fig. 9.7). More precisely, this single trait could be considered two separate, albeit highly correlated traits. That is, it is the combination of lingual ridges along the distal margin and mesial margin that form shoveling. Ordinarily, if you have a large mesial

Figure 9.6 Pronounced bilateral winging in an American Indian dentition; incisors also exhibit grade 6 shoveling expression.

Figure 9.7 The maxillary dentitions of a European (a) and American Indian (b) showing the two extremes of upper incisor shoveling variation. In both dentitions, UI1 and UI2 exhibit comparable levels of lingual marginal ridge development.
marginal ridge, you will also see a large distal marginal ridge on the same tooth. However, in populations with low frequencies of shoveling, some individuals may have incisors with one ridge (usually distal) but not the other. Mizoguchi (1985) observed these ridges separately and found interclass correlation coefficients of 0.4788 for males and 0.9167 for females. Given the extensive data already available for shoveling, it would be impractical to go back and reanalyze populations by separate mesial and distal marginal ridges.

For making observations on shoveling, the classic four-grade scale of Hrdlička (1920) and Dahlberg (1956) has largely been replaced by the two eight-grade scales of the ASUDAS (separate plaques for UI1 and UI2). Some workers have had success in treating shoveling like an interval scale trait by measuring the depth of the lingual fossa. This approach, pioneered by Dahlberg and Mikkelsen (1947), has been adopted with some success by several later workers (Aas, 1979; Aas and Risnes, 1979; Mizoguchi, 1985). This method is most applicable to the study of young individuals with unworn dentitions.

Double-shoveling. Lingual marginal ridge development is the primary characteristic of shoveling, whereas labial marginal ridging is the hallmark of double-shoveling. These ridges are never as pronounced as the lingual ridges, but the ASUDAS plaque has absence and six degrees of trait presence. Mesial and distal marginal ridges on the labial face of the incisors are often asymmetrical, with more and greater expression on the mesial marginal ridge. Snyder (1960) used the phrase “3/4 double-shoveling” to denote cases of teeth that exhibited both lingual ridges and mesial marginal labial ridges but not distal marginal labial ridges.

Tuberculum dentale. Tuberculum projections on the basal eminence of the two upper incisors traditionally have been referred to as tuberculum dentale, a Latin phrase for dental tubercles. These cingular derivatives take the form of either ridges or tubercles, both of which can vary from slight to pronounced. Another aspect of this trait’s variation is the number of ridges, especially on UI1 but also UI2, where one, two, or three ridges might be expressed. This trait is exhibited in a myriad of forms involving degree of expression and number, so additional work is required to standardize observations.

Despite difficulties in standardization, this trait is relatively common in European and European-derived populations, making it one of the few positively expressed traits found in this geographic group. Fossil hominoids and hominids, especially Neanderthals, often exhibit basal tubercles and lingual ridges making an evaluation of their evolutionary significance a fruitful research topic.

Interruption grooves. Along the lingual marginal borders or basal eminence of UI1 and UI2, one can find developmental furrows that some call corono-radicular grooves because they involve both the crown and the root (Fig. 9.8). However, in many cases, the grooves occur along the marginal ridges so the term “interruption groove” has more general application. The classification of this trait focuses on location (mesiolingual, distolingual, or both) rather than on degree of development (Turner et al., 1991).

Maxillary Canines

Tuberculum dentale. This trait, which was discussed for the incisors, is also expressed on the upper canines (Scott, 1977b). Although the canine can exhibit ridges, expression more often takes the form of slight-to-pronounced tubercles on the basal eminence. In some cases, these tubercles have free apexes.

Distal accessory ridge. This extra ridge is expressed on the lingual surface of the upper canine between the essential ridge and the distal marginal ridge. It is one of the few traits that shows a significant male–female difference in frequency and degree of expression (Scott, 1977c). It also wears off fairly quickly,
making observations of this trait difficult in middle-aged adults.

**Bushman canine.** To avoid geographic designators, Turner et al. (1991) refer to this variable as the canine mesial ridge. Nevertheless, the Bushman canine is used widely. This trait is related closely to the canine tubercle, which occurs commonly among the Bushmen (Morris, 1975). It is considered present when the mesial marginal ridge intersects and joins a dental tubercle without a developmental groove separating the two features.

**Maxillary Premolars**

**Uto-Aztecan premolar.** Upper premolars typically have two major cusps—one buccal and one lingual. For the second premolars, the main axes of the two cusps usually run in parallel. For first premolars, however, the distal margin of the buccal cusp often shows a buccalward rotation. Although this rotation is common, a more pronounced buccalward rotation of the distal margin with an associated fossa or pit is rare. Morris et al. (1978), who first described this trait in Papago Indians, believed it was limited to Uto-Aztecan speakers in the American Southwest. Although a few instances have since been found outside the Southwest, most cases have been observed in living or prehistoric Uto-Aztecan speakers (e.g., Pima, Hopi, or Papago). In Turner et al. (1991), this rare but interesting variant of UP1 is called the distosagittal ridge.

**Accessory ridges.** On the occlusal surface of the buccal cusp of the upper premolars, accessory ridges are often expressed on either side of the essential or median ridge. Distal accessory ridges are more common than mesial accessory ridges, and the second premolar has a higher frequency of ridges than the first premolar. These ridges vary in size, but the range of expression is too limited to divide this trait into any more than absent and two or three degrees of trait presence. Accessory ridges are sometimes evident on the lingual cusp, but these are far less common than ridges on the buccal cusp.

**Odontomes.** This accessory coronal tubercle can be expressed on the upper and lower premolars. Located in the central occlusal groove, between the buccal and the lingual cusps, this cone-shaped tubercle starts off as a pointed structure, but it does wear through time. Even though it seems to be primarily enamel, a worn odontome shows a dentine component.

St. Lawrence Island Eskimos have perhaps the highest frequencies of odontomes yet recorded. For the upper premolars, the incidence is 20.5% (18/88), whereas the lower premolar frequency is 14.1% (10/71) (Scott

---

**Figure 9.8** Interruption groove on UI2; the groove crosses both the crown and the root, hence, the alternative name of coronoradicular groove.
and Gillispie, 2002). Prehistoric Kodiak Islanders have frequencies of 2.2% (5/229) for upper premolars and 3.7% (9/242) for lower premolars (Scott, 1991). Based on these two series, it seems that upper or lower premolars are about equally likely to exhibit odontomes. In most world populations, this trait is rare.

**Upper premolar root number.** Upper premolars can express one, two, or three independent roots. This variation is, however, limited for the most part to the upper first premolar. UP1 should only be considered to have two roots when the inter-radicular projection extends from 1/4 to 1/3 of total root length (Turner, 1981). In the unusual case of three-rooted UP1, two roots are buccal and one is lingual, paralleling the pattern of UM1 roots.

**Maxillary Molars**

*Hypocone.* Early in primate evolutionary history, upper molars had three major cusps (protocone, paracone, and metacone) that made up the trigon. The hypocone was an additional cusp that originates from the distolingual cingulum (Gregory, 1922). In recent times, the hypocone has been treated as one of four major cusps of the upper molars, but it is, in fact, a late evolutionary addition that is in the process of being subtracted from the upper molars of modern humans (Fig. 9.9).

Dahlberg’s classification of the hypocone included grades 3 (hypocone absent), 3+ (conule form), 4− (hypocone reduced), and 4 (hypocone normal in size). This scale was the foundation for a new classification that had absence and five degrees of trait presence (Larson, 1978). For this scale, grade 1 involved a hypocone outline but no free-standing cusp on the distolingual surface of the protocone. Grade 2, like Dahlberg’s grade 3+, involved the presence of a small conule with a free apex. Grades 3 through 6 represented hypocones of increasing size although the use of UM1 and UM2 examples made the plaque difficult to use in some instances.

**Carabelli’s trait.** This tuberculum projection is a cingular derivative on the mesiolingual cusp of the upper molars. It ranges in size from a slight deflected developmental groove to a large free-standing cusp that is almost the size of the hypocone (Fig. 9.10). The threshold expressions of this trait are subtle, so different observers may disagree on what constitutes absence or presence. In population studies, grade 3 is often used as a breakpoint to avoid this problem.

In upper molars, the typical wear gradient is buccal to lingual. When the protocone is worn down to 1/3 or more of total crown height, observations on Carabelli’s trait are difficult to make. If this factor is not taken into account, the overall frequency of Carabelli’s trait would be seriously underestimated.

**Cusp 5.** This accessory cusp is located on the distal marginal ridge of the upper molars between the metacone and the hypocone. Harris and Bailit (1980) refer to this variable as the metaconule. Once again, a large cusp 5 is very distinctive, but slight expressions of this trait are subtle, manifested as small vertical grooves extending down the distal surface of the molar. As with many supernumerary occlusal cusps, cusp 5 can wear down relatively quickly, making observations difficult in older adults.
Mesial marginal accessory tubercles. Three tubercles can develop along the mesial marginal ridge of the upper molars. These tubercles are associated, in turn, with the paracone (mesial paracone tubercle), the protocone (protoconule), and the marginal ridge between the paracone and the protocone (mesial accessory tubercle) (Kanazawa et al., 1990, 1992). Although these tubercles can be distinct, they are often obscured by wear in early adulthood. Observations on children and adolescents yield the most reliable results for these marginal tubercles.

Enamel extensions. The cervical enamel line of the molars typically runs a straight or
horizontal course from the mesial to the distal margins of the upper molars. In some instances, however, the enamel line deflects toward the root in the area separating the mesial and distal roots. These extensions range from slight to pronounced deflections, with the latter sometimes terminating in an enamel pearl. This variable is manifest on both the buccal side of the upper and lower molars, although focus in population studies is usually on the upper molars (Turner et al., 1991).

**Upper molar root number.** The upper molars are anchored by three major roots (two buccal and one lingual). This ancestral condition has characterized primates for millions of years. This three-rooted condition is basically invariant on UM1, so this tooth fails to exhibit the type of variation that is useful for making population comparisons. However, UM2 is polymorphic, with one-, two-, or three-rooted teeth possible. Reduced root number is not accompanied by actual loss of root cones but rather by the fusion of two or more cones. To record root number, each distinct root should have an inter-radicular projection (bifurcation) at least 1/4 to 1/3 along the total length of the root complex.

**Mandibular Incisors**

**Shoveling.** Mesial and distal lingual marginal ridges are not limited to the upper incisors. They can also be expressed in the lower incisors but to a much lesser extent. Part of this is attributable to tooth size as the upper incisors are about 50% larger than the lower incisors. As shoveling expression shows a high correlation between the upper and the lower incisors (Scott, 1977a), UI1 is the key tooth for population studies.

Two other traits of the upper incisors—winging and tuberculum dentale—can be found in low frequencies, but because of rarity, researchers rarely score these traits on the lower incisors.

**Mandibular Canines**

**Distal accessory ridge.** Lower canines are incisiform and morphologically simplified. However, one relatively common discrete trait is expressed on the lingual surface between the median ridge and the distal marginal ridge. As it is found on the distal lobe segment, it is called a distal accessory ridge. Like the corresponding ridge on UC, this trait shows a significant sex difference (Kieser and Preston, 1981; Noss et al., 1983; Scott, 1977c). As crown wear often impacts the distal slope of the incisal edge first, this feature can wear off quickly in life.

**Root number.** Upper and lower incisors, along with upper canines, are one-rooted teeth. Of the anterior teeth, only the lower canine exhibits multiple roots in polymorphic frequencies. To be scored as a two-rooted tooth, the bifurcation of the buccal and lingual root cones should occur in no less than 1/4 to 1/3 of total root length (Fig. 9.11). This trait is one of the few that is found most commonly in Europeans (Scott and Turner, 1997).

**Mandibular Premolars**

**Lingual cusp number.** Lower premolars invariably exhibit a single buccal cusp with a well-developed free apex. However, extensive variation occurs in the expression of the lingual cusp (or cusps). In contrast to the upper premolars, where the lingual cusp is centered effectively on the median ridge of the buccal cusp, lingual cusp expression on the lower premolars is far more asymmetrical and variable. Some lower first premolars, for example, have a lingual cusp fused to the buccal cusp that lacks a free apex. More commonly, the lingual cusp is well-developed and assumes a mesial position on the crown. If an extra lingual cusp occurs, it is usually situated on the ridge extending distally from the primary lingual cusp. This extra lingual cusp is typically smaller in overall size and lower on the occlusal plane than the primary lingual cusp. Although these principles apply to both LP1 and LP2, the extra lingual cusp is far more common on LP2 (Fig. 9.12).

Kraus and Furr (1953) describe 17 discrete variables for the lower first premolars, including
extra lingual cusps. Following their liberal
definition of what constitutes a lingual cusp
(any independent apex, no matter how slight),
many lower premolars would be scored as “mul-
tiple lingual cusps” even though patterned vari-
ation in the size or position of lingual cusps is
not taken into account. Ludwig (1957) defines
seven traits, including multiple lingual cusps,
for the lower second premolars. These workers
have shown that useful untapped variation
may occur in the lower premolar field that has
not been exploited fully to date. For developing
new trait classifications, the lower premolars
are a good place to begin.

Accessory ridges. Like the upper premolars,
the occlusal surfaces of the buccal cusp (proto-
conid) can exhibit either mesial or distal acces-
sory ridges on the accessory lobe segments of

the lower premolars. Mesial and distal ridges
should be scored separately, even though they
do exhibit some level of interaction. With their
limited range of variation in size, absence and
two or three grades of presence are considered
adequate to score these ridges.

Odontomes. These coronal tubercles are
described under maxillary premolars.

Tomes’ root. Although lower first premolars
can exhibit two roots, the expression of the
secondary root is different from upper first

Figure 9.11 Two rooted lower canines can often be scored
even when the tooth is missing. The arrow points to the
vacated lower canine socket that shows the presence of
two distinct roots. In this case, the two-rooted tooth was
also available.

Figure 9.12 Lower premolars always have a prominent
buccal cusp, but lingual cusps vary in both size and
number. In (a) both LP1 and LP2 effectively lack a lingual
cusp; by contrast, (b) shows cases where both lower premo-
lars have two lingual cusps. The case shown is the most
common, where the mesiolingual cusp is typically larger
than the distolingual cusp.
premolars where distinct roots are associated with the two major cusps. The lower first premolar, by contrast, rarely exhibits a distinct bifurcation along a buccolingual plane. When two roots with slight-to-moderate separation are evident, this usually involves a division of the distobuccal and lingual root components. C.S. Tomes (1889) was the first to describe this root variant that bears his name.

**Mandibular Molars**

**Hypoconulid (cusp number).** Lower molars have five major cusps: 1) protoconid, 2) metaconid, 3) hypoconid, 4) entoconid, and 5) hypoconulid (Fig. 9.13). When there is variation in cusp number, it involves the hypoconulid or cusp 5. As with the hypocone, the hypoconulid was the last addition to the primate lower molar crown and is the first to be reduced in size or eliminated altogether. Although the hypoconulid varies in size, most workers report the frequency of four-cusped or five-cusped LM1 and LM2.

Early workers like William King Gregory (1922) reported cusp number sequence for the lower molars. Higher primates and early hominids retained the hypoconulid for the most part so they were characterized as 5-5-5. Over the course of hominid evolution, the fifth cusp was often lost, especially on the second molar. A cusp number sequence of 5-5-5, 5-4-5, or 5-4-4 would characterize most living populations.

**Groove pattern.** In the lower molars, the major cusps are separated by deep fissures or developmental grooves. The fissure pattern varies in a manner that impacts cusp contact at the central occlusal pit. In higher primates, the most common pattern involves contact between cusps 2 and 3. During hominid evolution, this pattern started changing, especially for LM2. In recent humans, it is common for cusps 1 and 4 to come in contact. Between these two patterns, there is a third and less common alternative where all four cusps come in contact at a single point. Groove pattern, defined by major cusp contact, is classified as Y (2–3 contact), X (1–4), and +(1–2–3–4) (Jørgensen, 1955).

Although no evidence supports a strong correlation between cusp number and groove...
pattern, Gregory (1916) linked the two traits together in his famous *Dryopithecus* Y5 pattern. For Miocene hominoids, the description was accurate as most early apes had five major cusps and 2–3 cusp contact. In fact, most modern humans exhibit five cusps and a Y pattern on LM1. The noteworthy changes impact LM2 and to some extent LM3.

**Cusp 6.** The modal number of cusps in the lower molars is five. However, a relatively common supernumerary cusp is positioned between the entoconid (cusp 4) and the hypoconulid (cusp 5) referred to as cusp 6, or the *tuberculum sextum*. On the ASU plaque, which has grade 0 and five degrees of trait presence, C6 is judged relative to the size of C5. If C6 is small relative to C5, grade 1 is assigned. If it is about half the size of C5, then it is assigned to grade 2, which is by far the most common expression on the lower first molars. Grade 3 is used when C5 and C6 are equal, irrespective of absolute size. Grades 4 and 5 are reserved for those uncommon instances where C6 exceeds C5 in size.

One complaint leveled against the ASU classification focuses on the issue of whether the presence of C5 is required to have a C6. The hypoconulid (C5) is immediately distal to and closely associated with the hypoconid, hence the name hypoconulid. Cusp 6 is immediately distal to the entoconid and, correspondingly, is referred to as the entoconulid (Turner, 1970). Cusp 5, especially when large, can be centered perfectly between the hypoconid and the entoconid. More commonly, however, the hypoconulid lies more directly behind the hypoconid, with a position more buccal than lingual. What if the single distal cusp is situated behind the entoconid, so that position is more lingual than medial or buccal? When this occurs, some observers feel this signals the presence of cusp 6, even though there is only a single distal cusp. If this observation is valid, the effect on population studies would be relatively minor because it rarely applies to LM1, the key tooth for cusp 6 frequencies. However, there could be a secondary ramification for the LM2 cusp number. That is, workers could classify LM2 as five-cusped even though the fifth cusp may be C6 and not the hypoconulid, the defining distal cusp for the variable of cusp number.

**Cusp 7.** A second supernumerary cusp of the lower molars is cusp 7, or the *tuberculum intermedium*, situated between cusps 2 and 4. In the ASU classification, the plaque has absence (grade 0), grade 1A, and four grades of trait presence (1–4). Grade 1A has always been a problem as it is associated with the distal lobe segment of the metaconid—hence the name post-metaconulid. It is also the most common expression taken by cusp 7 on the deciduous lower second molar. Grine (1981) pointed out that some lower molars exhibit both a post-metaconulid and a wedge-shaped cusp 7. If they co-occur on the same tooth, they cannot be considered a single trait. For that reason, 1A expressions should be scored, but for population comparisons, workers should focus on the traditional wedge-shaped cusp 7 forms evidenced by grades 1 through 4 (Fig. 9.14).

**Protostylid.** This cingular derivative, often associated with the buccal groove between the protoconid and the hypoconid, is located on the mesiobuccal cusp of the lower molars (Dahlberg, 1950). In addition to positive cingular expressions, Dahlberg (1956) included buccal pits as grade 1 in his classification. Although the ASU standard follows Dahlberg’s classification faithfully, including buccal pits as grade 1, no evidence supports the inclusion of negative pits with positive cingular expressions. Although buccal pits should be scored, their use in estimating population frequencies of the protostylid is questionable.

**Deflecting wrinkle.** The essential ridge of the metaconid usually runs a straight course from the cusp tip to the central occlusal pit. However, in some cases, this ridge begins with a more mesial orientation before changing course toward the center of the tooth, hence,
the name deflecting wrinkle. Although the ASU standard includes grade 0 and three degrees of grade presence (Turner et al., 1991), most workers focus on grade 3 expressions. The key tooth for deflecting wrinkle observations is LM1. It is unusual to find this trait on LM2 or LM3.

**Root number (LM1).** The modal root number for lower molars is two, with one mesial and one distal root. Root grooves are often present on either root producing three or four radicals. World populations almost always adhere to this modal root number, but there is a supernumerary root that can develop on the lingual border of the distal root, hence, the name distolingual root. As this extra root makes the tooth three-rooted, the trait is often referred to in shorthand form as 3RM1 (three-rooted lower first molar). This extra root, which is most common in North Asians and Eskimo-Aleuts, is rare on LM2 and LM3.

**Root number (LM2).** Rather than exhibiting extra roots, LM2 often shows a reduction in root number. There are three ways that a two-rooted lower molar’s roots can fuse to produce a single-rooted tooth. The root cones may fuse along the buccal axis, along the lingual axis, or both. When root fusion precludes the expression of independent roots for at least $1/4$ to $1/3$ of total root length, the tooth is classified as single rooted.

**Observational Methods**

Once a trait set has been chosen for a population study, the worker has to decide how to best observe each trait. With the exception of shoveling, where some workers have had success
measuring the depth of the lingual fossa (Aas, 1979; Aas and Risnes 1979a, 1979b; Blanco and Chakraborty, 1977; Dahlberg and Mikkelsen, 1947), it has been difficult to measure teeth on an interval scale (but see Taverner et al., 1979). For most variables, no landmarks are consistently present because of a combination of size and shape considerations, not to mention complications introduced by crown wear, casting error, or tooth breakage.

Hanihara et al. (1970) quantified the size of the hypocone through the use of photographs, but this method is time consuming, expensive, and does not address shape concerns, especially for low grades of expression. The method works well for UM1, which consistently expresses a hypocone but is less useful for the more variable UM2 and UM3, where the hypocone is often reduced or lost altogether. Of course, this method requires teeth that show very little wear. Because of these limitations, no workers have adopted this method for scoring variation in hypocone size.

Because most dental traits are either present or absent and exhibit variable degrees of expression when present, there are different ways to characterize a population for a particular trait. At the most basic level, traits can be dichotomized as present (+) or absent (−). This approach is popular among researchers who focus on nonmetric cranial traits. However, the dividing line between present and absent is always “fuzzy.” Berry (1976:259) notes “the dubious presence of a variant was scored as 0.50 provided that the scoring difficulty was due to a genuine minimal expression of a variant rather than to attrition.” During my dissertation research, I noted questionable expressions that fell on or close to the threshold by the letter L (Scott, 1973). Anyone who has scored crown and root traits has faced this conundrum—“to be or not to be,” that is the question.

Because of the difficulties introduced by threshold expressions, Turner (1985a, 1986a; Turner et al., 1991) adopted the strategy of using breakpoints on ranked scales to dichotomize groups. This was not applied to all traits but primarily to those where the line between presence and absence was blurred (e.g., shoveling or Carabelli’s trait). If reference is made to a specific expression on a standard plaque and you use this point as a basis for sample frequencies, this increases the comparability of datasets collected by different workers. Of course, whenever possible, numbers and frequencies for all grades of expression should be reported.

Although the use of standardized ranked scales is recommended strongly, the use of all ranks in population comparisons is usually unnecessary. Turner (1985a) developed an expression count method for dental traits that was aimed at calculating frequencies based on weighted ranks. Although this method uses data on expressivity more thoroughly, it was never adopted by other workers in the field nor used by Turner in later population analyses. In the context of genetic analysis, I used a value called MSA or the “mean score of affected” individuals (Scott, 1973). However, distance statistics typically are designed to analyze frequencies or sample means, so I did not use the MSA in population comparisons. Although some quantitative method to take all ranks into account might provide information above and beyond total trait frequency, it should be emphasized that quasicontinuous traits are best characterized by a single frequency as this specifies the nature of the entire normal distribution (Falconer, 1960).

Asymmetry and Counting Methods

Although the dentition is characterized by mirror imagery between the two sides of the jaw, teeth on opposite sides of the jaw—antimeres—can exhibit asymmetry in size, morphology, and hypodontia (congenitally missing teeth, especially M3). Many workers have shown this asymmetry shows no side preference so it is referred to as fluctuating asymmetry rather than as directional asymmetry (Mizoguchi, 1988; Saunders and Mayhall, 1982; Scott, 1980; Van Valen, 1962).
Given the lack of directional bias in trait expression, there are three methods for scoring and tabulating crown and root trait frequencies:

- **Side count:** With this method, followed often by those who measure teeth but occasionally by those who observe morphology, a worker scores only left or right antimeres for trait expression. The caveat is often stated that only one side will be scored, but if the antimere on that side is missing, the tooth from the other side can be used. This method is the most efficient in terms of scoring and tabulation because the issue of asymmetry is removed from the process.

- **Tooth count:** With a large collection of loose teeth, this is the only method that can be followed when antimeres cannot be matched. Basically, all teeth, both left and right, are scored for the expression of a trait. When the data are tabulated, the population frequency is based on all teeth scored, whether they are antimeres or not. Some workers use this method for deriving trait frequencies even when studying whole dentitions (casts or skulls).

- **Individual count:** This method combines elements of the side and tooth count methods and provides a rationale for dealing with the issue of asymmetry. For deriving frequencies based on individual counts, all teeth are scored for crown and root traits. However, when individuals are scored for a trait on both antimeres, only the antimere with the highest grade of trait expression is used to characterize that individual. The rationale for this procedure is that an individual has only a single genotype directing dental development. If two antimeres exhibit grades 1 and 4 phenotypes, which of the two best reflects the underlying genotype of that individual? Adherents of this method contend it is the more pronounced phenotype that best reflects the underlying genotype. The hypothetical case above would be scored at 4 in the final tabulation (not 1, not 2.5, and not 1 and 4, the three other alternatives) (Turner and Scott, 1977; Turner et al., 1991).

Although the debate over appropriate counting methods has been contentious at times, so little asymmetry exists in dental morphology that any method can be used to arrive at similar results (cf. Scott, 1980). The only method that has a major weakness is the tooth count method. Given the high degree of symmetry in nonmetric dental traits, a tooth count artificially increases sample size with very little new information. As sample size is a critical element of estimating sample variance and testing hypotheses, this method should not be followed except in the case of loose teeth where there is no other choice. In the ASU system, the individual count method is preferred although results differ little from those obtained by a side count.

**Intraobserver and Interobserver Error**

It is a straightforward matter to teach students how to observe a distinct example of shoveling, Carabelli’s trait, cusp 7, or any other crown or root trait. However, to make consistent observations on ranked scales and to distinguish the subtle expressions that separate absence from presence, research on dental morphology requires experience and caution. In the literature several egregious examples of observations are so far off base that they are impossible. The reasons for such observational error may be from inexperience, the lack of adequate operational definitions, bad lighting, poor magnification, hurried observations, or some combination thereof.

Although researchers rarely address such things, there is an element of “art” in the study of dental morphology. I do not mean the study of such traits is not scientific—it is simply an acknowledgment that some experience is required to make hundred or thousands of individual decisions on “present or absent”
and “grade of expression when present.” Before setting off on any dental morphological research, it would behoove any worker to observe a sample of casts or skeletons for a particular set of traits. Then, after a month or more, return to the same sample and repeat your observations. This assessment of intra-observer error provides a benchmark for your research. You will find that some observations can be replicated at a very high level (>90%), whereas other variables pose problems in consistent scoring (Nichol and Turner, 1986; Scott, 1973). At that point, you either want to hone your skills on traits that cause the most problems or eliminate them from your trait list.

After gaining experience and testing your ability to replicate observations, you are ready to begin your study. However, the issue of error is not entirely behind you. When you tabulate your data and evaluate your trait frequencies, you should also assess your results in the context of comparisons with other observers. If you find a total shoveling frequency of 50% in American Indians, you should evaluate your breakpoint as most workers have found this frequency to be 90–100%. If you find a cusp 7 frequency of 80% in a Native American sample, you should reflect on your scoring methods as this trait is less than 10% in most populations and only 40–50% in Africans who have by far the highest frequency of this trait in the world (Scott and Turner, 1997).

POPULATION STUDIES

Defining a Problem


When workers make observations on dental morphology, this is usually in the context of a problem or set of problems. However, some morphological studies are descriptive in nature. To observe, tabulate, and report the frequency of a trait in a population is not a problem-oriented approach, even though such descriptive studies are useful when others assess the data in a broader population analysis.

Although most dental morphological analyses focus on between-group comparisons, some within-group analysis is possible. For example, if you had a sample from one population, several issues could be addressed: 1) sex differences—do males and females show significant differences in the frequency and expression of a morphological trait? 2) fluctuating asymmetry—do different morphological variables show the same levels of asymmetry or are some traits more asymmetric than others? and 3) trait interaction—do different morphological traits in the same and different fields show significant correlations or are they expressed independently of one another? Should crown dimensions be measured along with morphological traits, one can address whether these traits are expressed independently of tooth size. If your sample is composed of families or twins, you can address genetic issues through intrafamilial correlations and
concordance–discordance rates in MZ and DZ twins, either of which can be used to estimate heritability. The point is that you can address some problems if you only have a single population sample at your disposal. However, if you observe multiple samples, you can employ methods of comparative analysis and address a wider range of problems.

Over the past three decades, most studies of dental morphological variation have focused on problems of population origins and relationships. In a pioneering effort, Turner (1971, 1983, 1984, 1985a, 1985b, 1986a, 1986b) traversed museums across the Western Hemisphere to observe skeletal series for crown and root morphology. His goal was to determine whether teeth could add insights to the problem of Native American origins. It was in this context that Turner ultimately collaborated with the linguist Joseph Greenberg and the geneticist Steven Zegura who collectively formulated the three-wave model for the peopling of the Americas (Greenberg et al., 1985, 1986). In their massive tome on genetic markers, Cavalli-Sforza et al. (1994) found strong support for the three-wave model using over 100 nuclear alleles. Although recent mtDNA and Y chromosome studies may challenge the three-wave model, it remains a common starting point for discussions of New World settlement.

When he turned his attention to Asia, Turner (1976) found dental morphology did not support the long-held notion that the Jomon, Japan’s Neolithic population, were ancestral to modern Japanese populations. He argued that the Jomon were ancestral to the Ainu who had been peripheralized slowly to the northern islands of Hokkaido and Sakhalin. The ancestors of the modern Japanese population were derived more likely from the Asian mainland during the Yayoi invasion of 2200 B.P. Despite initial resistance to this idea, Japanese dental anthropologists subsequently supported this finding (K. Hanihara, 1992; T. Hanihara, 1990b, 1991a, 1991b).

During the course of more intensive work in Asia, Australia, and the Pacific, Turner (1987, 1989, 1990) found a dental divide between North Asians and Southeast Asians. The morphological patterns for these two groups were referred to as Sinodonty and Sundadonty, respectively. Comparative analysis indicates New World populations were derived from Sinodont ancestors in North Asia, whereas Polynesians and Micronesians were derived from Sundadont groups of insular and mainland Southeast Asia.


Although most recent studies focus on population relationships, two areas where dental morphology has great but untapped promise are 1) admixture analysis and 2) ethnic identification in forensic contexts. Edgar (2002) has demonstrated this promise in her research on European, African, and African-American dentitions. Admixed groups show intermediate crown and root trait frequencies, but the precision with which we can use these traits requires more research. In his textbook on forensic anthropology, Byers (2001) has one small table illustrating how a few dental traits differ among Europeans, Africans, and Asians. It represents only the tip of the iceberg on how morphology could be used to identify individual ethnic affiliation in a forensic context.

Living or Dead?

The first 10 years of my career involved the analysis of thousands of dental casts from Southwest Indian tribes and American white
family and twin samples (cf. Scott, 1973; Scott and Dahlberg, 1982; Scott and Potter, 1984; Scott et al., 1983, 1988). Over the past 20 years, my focus has shifted almost entirely to skeletal collections from the American Arctic and North Atlantic (cf. Scott, 1991, 1992; Scott and Alexandersen, 1992; Scott and Gillispie, 2002; Scott et al., 1992) and, more recently, from Spain. I have also taken advantage of the special quality of teeth and compared the dentitions of extinct and extant Kodiak Islanders (Scott, 1994). With this experience, I have gained several insights into the advantages and disadvantages of studying casts and skulls.

Casts

Advantages. As casts are obtained from living populations, you can establish your own agenda in terms of sample size and age profile. Because of fillings, wear, and tooth loss, many workers who collect impressions focus on children and teenagers whose crown morphology is least impacted by the ravages of time and disease. If you have the resources and a population’s approval, you can obtain thousands of impressions like Al and Thelma Dahlberg did among the Pima Indians. With less time and more limited resources, it is still feasible to collect one or two hundred impressions, which is a large enough sample to provide a good population characterization.

Disadvantages. Until I studied Basque and Spanish casts in the summer of 2005, I had forgotten one major disadvantage of casts—you can only observe crown traits. Moreover, although conditions vary between populations, casts from living populations pose significant problems for scoring some traits. This problem is especially true for groove pattern and other variables located primarily on the occlusal surfaces of the premolars and molars. In recent populations, caries has been rampant, so a significant fraction of a sample is lost because fillings obscure a trait's locus. Finally, if you have only casts, that means you have no temporal variation to speak of, so problems of microevolution cannot be addressed.

Skulls

Advantages. A big advantage in studying skeletal remains is the inclusion of root traits in your morphological protocol. Root variants can even be observed when a tooth is absent. Equally important is the ability to study skeletal series distributed through time, which makes it possible to evaluate microevolutionary changes in the dentition.

Disadvantages. Downsides exist to the study of skeletons. For example, a paleodemographic fact of life is that once individuals survive infancy, they usually live until middle-to-late adulthood. Any given skeletal collection has a relatively small percentage of children and teenagers, exactly those individuals whose teeth are least impacted by disease and wear. Dental attrition is a major limiting factor in the study of skeletons because earlier human populations wore down their teeth at a much faster rate than modern populations. Even in relatively young individuals (late teens, early 20s), when the protocone is worn, precise observations on Carabelli’s trait are difficult to make. In addition to wear, postmortem tooth loss also acts to reduce sample size. As the single-rooted anterior teeth are especially subject to postmortem loss, the ultimate sample size for traits such as shoveling and tuberculum dentale are reduced greatly. The multirooted molars, especially root traits, typically have the largest sample sizes. Thus, it should be borne in mind that 100 skeletons will not provide a sample of 100 for the traits you are observing.

To illustrate sampling limitation of skeletal remains, observations on 155 St. Lawrence Island Eskimo skeletons yielded the following sample sizes for a standard set of traits: UI1 shoveling (29), UI1 tuberculum dentale (34), UC distal accessory ridge (34), UP1 odontomes (51), UM2 hypocone (70), UM1 Carabelli’s trait (54), UM1 cusp 5 (34), LP1 multiple
lingual cusps (54), LM1 deflecting wrinkle (46), LM2 hypoconulid (70), LM2 groove pattern (93), LM1 cusp 6 (43), LM1 cusp 7 (81), and three-rooted LM1 (59) (Scott and Gillespie, 2002). These numbers indicate that traits on the anterior teeth can be scored in about 20% of the total skeletal sample. Premolar traits are somewhat better, yielding samples sizes of about 30%. The multirooted molars yield the biggest samples, but even here only 35–60% of the teeth from the total sample are scorable. The variation in frequencies between traits indicates they are impacted variably by crown wear, tooth loss, and other environmental effects. By way of contrast, in a sample of 1528 Pima Indian casts, about 80% of the total sample could be observed for UI1 shoveling, UM1 Carabelli’s trait, UM1 hypocone, LM1 cusp number, and LM1 cusp 6. The only trait on LM1 that was relatively limited was groove pattern at 50%, supporting the earlier point on the effect of fillings on the occlusal surfaces of living populations.

A final point in the cast–skull debate is whether it is easier to score subtle shades of morphology on casts or on real teeth. For either, the use of a good light and a 10x magnifying lens are highly recommended. Given the same observational conditions, it is hard to say either has a major advantage. Casting error can be an issue as air bubbles can obscure the site where a trait is scored. Untreated casts present a very dull appearance, so many dentists rub them with baby powder to develop sheen for enhanced contrast. Personally, I find it easier to make observations on casts treated with baby powder, but this may be a matter of taste. Although real teeth have no issues of casting error, the enamel does reflect light making subtle expressions difficult to observe. One has to be careful to examine the trait’s locus under good light from all possible angles or subtle trait expressions can be overlooked easily.

One method for studying crown morphology that I do not recommend is intraoral observations in living individuals. Not only do real teeth reflect light, but the reflectance is exacerbated by the presence of saliva. Although large expressions of traits are observed easily in vivo, any slight manifestations are hard to score because of light issues and the limited number of views available during such examinations. Fortunately, most anthropologists appreciate the limitations of intraoral observations, so this is not a major problem in dental morphological studies.

Permanent or Deciduous?

Over 90% of the papers published on human dental morphology focus on the permanent dentition. Some researchers, in particular Kazuro Haninara (1956, 1960, 1963, 1966) and Paul Sciulli (1977, 1990, 1998), plus a few others (Grine, 1986, 1990; Jørgensen, 1956; Lukacs and Walimbe, 1984), have devoted time to the analysis of deciduous morphology, but they are exceptions to the rule.

In an early influential review of the field, Lasker (1950) actively discouraged the study of deciduous dental morphology. He argued that the traits observable on primary teeth are the same as those evident on permanent teeth, so few new insights could be gained from the study of deciduous morphology. Although Lasker was correct to note the overlap between primary and permanent crown and root morphology, deciduous teeth do offer another perspective on morphological variation.

Deciduous teeth, as the developmental precursors of the permanent teeth, presage morphological characteristics that are exhibited in the permanent dentition. For example, children from Asian and Asian-derived populations often show shovel-shaped primary incisors, even though the teeth are smaller and trait development never reaches the same level. Traits exhibited on dm2 also foreshadow the expression of variables in the permanent molar field (Fig. 9.15). It is common to see Carabelli’s cusp, the hypocone, cusp 5, and other features of the permanent upper molars on dm2. The same is true for groove pattern, cusp 6, cusp 7, the protostylid, deflecting wrinkle, and distal trigonid crest on the lower molars. Dahlberg (1945) and others often
considered dm1 the first member of the molar field.

The most unusual teeth in the primary dentition are the upper and lower first molars. The upper first molar is typically a two-cusped tooth (one buccal and one lingual) that shows a general similarity to upper premolars, albeit somewhat wider in the BL dimension. The lower first primary molar is the most unusual baby tooth. It is elongated in an anteroposterior direction and shows some similarities to the sectorial third premolar of monkeys and apes. Although these teeth are also succeeded by lower premolars, they do not look like the premolars of modern humans.

Although Lasker’s comments may have discouraged a few workers from considering deciduous tooth morphology, a more important limiting factor is sample size. Most researchers, who collect casts, prefer individuals with all their permanent teeth (except for M3). Thus, they typically focus on school-aged children around the age of 10 years or older. Some deciduous teeth are present in such a sample, but numbers are limited. As noted, there are relatively few children in archaeological assemblages, so this limits the availability of primary teeth in this context.

If a researcher wants to address issues with deciduous dental morphology, and there is a niche out there for new students, they have to collect impressions from preschoolers or kindergarten-aged children. This avoids a mixed dentition and enhances sample size for teeth that are lost early (mostly the incisors). The second alternative is to examine primary teeth in skeletons, but this does require a large skeletal collection with very good preservation to provide enough cases to make up a statistically valid sample.

**How Many Traits?**

In 1950, W.C. Boyd used three blood group systems (ABO, Rh, and MN) to assess worldwide genetic variation. Four decades later, Cavalli-Sforza et al. (1994) analyzed 120 alleles to delineate regional and global genetic differentiation. Livingstone (1991) would argue that the expanded analysis is more reliable and urges researchers to use as many variables as possible in the evaluation of population relationships.

Early dental researchers usually focused on a limited number of dental variables—primarily shovel-shaped incisors, upper molar cusp number, Carabelli’s trait, and lower molar cusp number and groove pattern. Some augmented this battery with a few additional traits such as incisor winging and lower molar cusp 6,
cusp 7, the protostylid, distal trigonid crest, and deflecting wrinkle. If only a single tooth is used to represent each trait, the number of “classic” variables falls between five and ten.

K. and T. Hanihara have traditionally used six to nine morphological variables in their population studies, with their entire focus on crown traits (cf. K. Hanihara, 1968; T. Hanihara, 1992b). A recent study by Matsumura and Hudson (2005) on Asian dental variation employed 16 traits following some parts of the ASUDAS, but they included only one root trait (UP1 root number). In a study of East Asian dentitions, Turner (1987) observed 28 dental variables, including six root traits. In an analysis of ancient Egyptians, Irish (2006) used 31 dental variables, not counting palatine and mandibular tori and rocker jaw. When I had to address whether the early period inhabitants of Kodiak Island (Kachemak) differed in any significant way from later period peoples (Koniag), I used 22 dental variables, of which 16 were crown traits and 6 were root traits (Scott, 1991, 1992).

If your observations focus on dental casts, I would consider the following list of traits a minimal set (key tooth also noted):

1. Bilateral winging (UI1)
2. Shovel-shaped incisors (UI1)
3. Double-shoveling (UI1)
4. Interruption grooves (UI2)
5. Tuberculum dentale (UI1 or UC)
6. Carabelli’s trait (UM1)
7. Hypocone (UM2)
8. Cusp 5 or metaconule (UM1)
9. Distal accessory ridge (LC)
10. Multiple lingual cusps (LP2)
11. Cusp number/hypoconulid (LM1, LM2)
12. Groove pattern (LM2)
13. Cusp 6 (LM1)
14. Cusp 7 (LM1)
15. Protostylid (LM1)
16. Deflecting wrinkle (LM1)

If skeletal remains are analyzed, the following crown and root traits should be added to the list above:

1. Enamel extensions (UM1)
2. UP1 root number
3. UM2 root number
4. LC root number
5. Tomes’ root (LP1)
6. LM1 root number (3RM1)
7. LM2 root number

Depending on geographic region, a researcher might add one or more of the following traits to their protocol:

1. Labial convexity (UI1)
2. Premolar odontomes (North Asia, American Arctic) (UP1, UP2, LP1, LP2)
3. Uto-Aztecan premolar (North and South America) (UP1)
4. Bushmen canine (Africa) (UC)
5. Distal trigonid crest (LM1)
6. Marginal tubercles (UP1, UP2)
7. Mesial and distal accessory ridges (UP1, UP2, LP1, LP2)
8. Protoconule, mesial accessory tubercle, mesial paracone tubercle (UM1)
9. Anterior fovea (LM1)

Concluding Thoughts

Hopefully, I have provided some direction and guidelines for future students of human dental morphology. There are a plethora of interesting historical and processual issues to address using crown and root variation. I strongly encourage students to evaluate the current status of morphological studies and to think of ways to push them forward. The development of new trait classifications would be one place to start. The ASUDAS is a foundation, like the Dahlberg system was an earlier foundation, but it should not be considered a finished project. Like a large medieval cathedral, it is a
work in progress. To complement methodological advancements, new types of problems should be addressed. As noted, the use of tooth morphology in admixture studies and the ethnic identification of forensic remains are fertile grounds for expansion. Evaluating the possible action of natural selection on morphology is another research avenue to pursue. Already, many scholars are beginning to look at the same morphological features in fossil hominids that we examine in living and recent skeletal populations. Delineating long-term evolutionary trends and assessing how diet, behavior, and climate impacts crown and root morphology are on the horizon. The future study of dental morphology is bright, so for all of you contemplating research in this area, to misquote a famous Vulcan, “go forth and prosper.”

REFERENCES

Dahlberg AA. 1956. Materials for the establishment of standards for classification of tooth characters, attributes, and techniques in morphological studies of the dentition. Chicago, Ill.: Zollar Laboratory of Dental Anthropology, University of Chicago (mimeo).


Hanihara K. 1960. Standard models for classification of crown characters of the human deciduous dentition. Chicago, Ill.: Dept of Legal Medicine, Sapporo Medical College and Zoller Memorial Dental Clinic (mimeo).


REFERENCES


Mizoguchi Y. 1988. Degree of bilateral asymmetry of nonmetric tooth crown characters quantified


Wright S. 1934. The results of crosses between inbred strains of guinea pigs, differing in number of digits. Genetics 19:537–551.


CHAPTER 10

DENTAL PATHOLOGY

SIMON HILLSON

INTRODUCTION

The diseases and injuries of teeth and jaws are among the most common conditions observed in human remains. Many of them occur so frequently that they can be regarded practically as normal. Among the most frequent are dental defects, preserved lifelong unless worn away, resulting from disturbances to the formation of teeth during childhood. Other conditions reflect the uses to which the teeth are put during life, such as anomalous or exceptionally heavy tooth wear, fracturing or chipping, or the way in which the bone of the jaws remodels around their roots. The long-term presence of micro-organisms in dental plaque deposits in the mouth gives rise to a range of conditions such as dental caries and periodontal disease, which cause damage to the tissues of the teeth and loss of supporting bone in the jaws. Plaque-related disease is in turn strongly related to the progress of dental wear and to the remodeling of the jaws, which takes place independently, in response to changing mechanical forces. All these processes are, therefore, best observed as an integrated whole (Fig. 10.1), and their pattern of progression within a population is strongly indicative of the nature of the diet, the mode of subsistence, and the tasks carried out each day. The linkages are complex, however, and it is important to recognize that the different dental conditions include a variety of lesions and defects that interact in a variety ways. There are several different categories of dental caries, and loss of supporting bone, for example, just as there are several different types of tooth wear with contrasting implications for the life history of the teeth. Recording schemes for use in the study of human remains need to reflect these differences, even at the expense of complexity.

DEFECTS OF DENTAL DEVELOPMENT IN THE ENAMEL OF THE TOOTH CROWN

The development (Smith, 1991) of permanent tooth crowns (roots continue to form after crowns are completed) takes place in three phases:

- Incisors, canines, and first molars, initiated during the first year after birth (or just before birth) and completed between 3 and 7 years
- Premolars and second molars, which start formation during the second and third years after birth and are completed between 4 and 8 years
- Third molars, initiated any time between 7 and 12 years, and completed sometime between 10 and 18 years of age.
A whole range of factors (Goodman and Rose, 1990; Hillson, 2005) may disrupt crown development during this period, including dietary deficiency, childhood fevers, and major infections such as congenital syphilis (Hillson et al., 1998; Jacobi et al., 1992). Each disruption is expressed as enamel defects (Fig. 10.2) in all the crowns that were being formed at the time of the episode so, in theory, the pattern of defects (which is preserved into adulthood) provides a detailed record of growth disturbance. This, together with the generalized nature of growth disturbing factors that can cause such defects, makes them ideal indicators when fitting Selyean stress models to interpretation of human remains in archaeology (Goodman et al., 1988). Care needs to be taken, however, because the mechanisms by which the defects are created are complex, poorly understood, and vary depending on which particular part of the tooth crown is being formed at the time (Hillson and Bond, 1997; Hillson, 2005). The defects are best studied in relation to the microscopic pattern of growth layering, which is preserved within the enamel.

The most prominent expression of growth layering, which is observed in microscope sections under low magnifications (around ten times), is a concentric pattern of lines that radiates out from the enamel underlying the cusp tips (see FitzGerald and Rose, this volume). They are called the brown striae of Retzius, and each represents the momentary position of the enamel-forming front at a particular point in development (Boyle, 1989). In the deeper part of the enamel jacket that coats the tooth crown, the brown striae are often poorly defined and irregular, but where they angle up to meet the crown surface, they become sharply defined and regularly spaced at 20–40 μm (or one quarter the diameter of a human hair) apart. At the point where each brown stria reaches the surface, a groove runs around the circumference of the crown side. This gives a rippled or wave-like appearance to the crown side when observed under modest magnification, and the grooves are collectively known as perikymata (Greek peri- “around,” kymata “waves”). They vary in spacing from about 150 μm (a little wider than a hair) at the cusps, down to 30 μm or less at the base of the crown, largely (but not entirely) as a result of the changing angle at which the brown striae meet the surface. The brown striae are superimposed over an even finer set of incremental structures, the prism cross striations, which represent a 24 hourly growth rhythm. Where both they and the brown striae are sharply defined and regular, there seems to be a constant number of cross striations between neighboring striae throughout the enamel of each tooth for one individual. This number varies between 7 and 11 among individuals and represents some fundamental, but little understood,
The defects observed at the crown surface most commonly take the form of furrows, steps, or pits arranged circumferentially in bands around the crown side (Berten, 1895; Hillson, 2005; Hillson and Bond, 1997). They parallel the perikymata and are thus clearly both episodic and developmental in origin. It is important to distinguish them from the much rarer inherited enamel defects, usually affecting most teeth in both permanent and deciduous dentitions in a similar way (and so not attributable to particular periods of disturbance), and collectively known as amelogenesis imperfecta (Bäckman, 1989, 1997; Winter and Brook, 1975). For this reason, the phrase “Developmental Defects of Enamel (DDE)” is often used (Commission on Oral Health, 1982), although they are just as frequently called “enamel hypoplasia.” Various types have long been recognized. By far the most common are furrow-form defects, formed partly by a disruption to the even spacing of perikymata (Hillson et al., 1999; King et al., 2002) and ranging from microscopic furrows that involve just one pair of perikymata up to defects that can be seen with the naked eye and involve 20 or more perikymata. These large furrow defects are often called LEH or “Linear Enamel Hypoplasia” in anthropological literature (Goodman and Rose, 1990). A considerable body of clinical literature links them with medical histories involving dietary deficiency (Infante and Gillespie, 1974;
Mellanby, 1929, 1930, 1934, 1941; Sweeney et al., 1971), fevers, infectious disease (Kreshover, 1944; Kreshover and Clough, 1953; Kreshover et al., 1954; Kreshover, 1960; Pindborg, 1982; Sarnat and Schour, 1941) and a whole host of other childhood conditions. Much less commonly, whole layers of enamel matrix are missing, to expose the plane of a brown stria curving down into the crown, or even the underlying dentine. Such plane-form defects are usually bounded by a prominent step running around the crown, marking the resumption of normal enamel matrix formation. It is usually assumed that they represent a more marked growth disturbance than the furrow-form defects, but there is no direct evidence for this. In between, there are the pitted defects where the exposure of a brown stria plane is more intermittent, with normal enamel matrix in between. Such pits may decorate the edge of a full plane-form defect, or they may accentuate a furrow-form defect. Occasionally a band of pits may occur on its own.

Enamel hypoplasia contributes to the study of human remains in several ways. First, the pattern of defects may allow isolated teeth from a commingled burial to be matched. Second, the prevalence of defects may suggest the importance of childhood infectious diseases and/or dietary deficiency in the health of an ancient community. The sequence of defects, if this can be determined from their position in the sequence of tooth crown development, would give a unique record of the seasonality of such conditions, and their distribution with age. The difficulty, however, lies in devising a simple system for recording the presence of the defects, their size, and their timing. They vary in size from a microscopic disruption to a single perikyma groove, to a pronounced deformity of the whole crown (Hillson and Bond, 1997). For the most common furrow-form defects, the size and prominence has more to do with the period of time over which growth was disturbed, and where the defect is situated, on which tooth crown, than it has to do with the severity of the disruption. The only way to resolve these questions satisfactorily is to study the furrow defects under a microscope, as part of the whole sequence of layered growth (Hillson et al., 1999). For plane-form or pit-form defects, there is still a problem because they are very difficult to relate to the sequence of crown development shown at the surface by the perikymata. It has been common to measure the size and position of furrow-, plane-, and pit-form defects, using calipers, and then to estimate timing and duration from a table that assumes linear growth in crown height, calculated on the basis of an average final crown height and standard ages for the start and completion of crown formation, all drawn from the literature (Goodman and Rose, 1990; Swärdstedt, 1966). Usually a single tooth class, the canine, is selected. This has been recommended as a standard approach (Buikstra and Ubelaker, 1994), but it has formidable difficulties (Hillson, 2005; Hillson and Bond, 1997). First, although the size of a growing tooth germ (measured by the height from the developing base to the cusp tip) has a relationship with age up to 5 years or so, this relationship does not follow a simple straight line (Liversidge et al., 1993; Liversidge and Molleson, 1999; Reid and Dean, 2006). Furthermore, the first 20% to 50% of crown growth is hidden under the cusp tips (see FitzGerald and Rose, this volume), whereas the crown surface is strongly curved and grows more rapidly at the base of the cusps than near the cervix. Use of the canine may minimize the latter difficulty, because the middle part of its tall crown sides has a fairly constant extension rate, but still too many assumptions are made. Instead, for a rapid estimate of the place of a defect in the development sequence, and its duration, it is best to use a histologically derived chart in which counts of cross striations and brown striae of Retzius in enamel sections have been used to determine the sequence of formation of the crown surface (Fig. 10.3). With such a chart, matches of defects between teeth in the dentition confirm the systemic nature of the growth disruptions that caused them, and allow the age of disruption to be confirmed.
Wear and abrasion, in any case, make hypoplasia difficult to study—modern teeth are so heavily abraded by tooth brushing that they are often almost useless for comparative purposes. In fact, the heavy occlusal wear of archaeological teeth is preferable, because at least the cervix and protected approximal parts of the crown can still be studied, and it is possible to find equivalent parts of other teeth that can continue the sequence. In addition to all the above difficulties, the size and prominence of furrow-form defects vary with their position on the tooth crown (this may be ascertained by comparing the effects of the same period of growth disturbance on different teeth), and with the angle of view and orientation of lighting. Particular care needs to be taken with plane- or pit-form defects. In many cases, they seem to represent a single episode of growth disturbance that did not necessarily continue for a long time (Hillson and Bond, 1997). For this reason, their size and prominence does not necessarily bear any relationship to the duration of the growth disturbance.

As a basic record, the Fédération Dentaire Internationale DDE Index (Commission on Oral Health, 1982) is well defined in many ways and has been recognized as a standard (Buikstra and Ubelaker, 1994). The main difficulty with the DDE Index, and any other

Figure 10.3 Formation sequence for crown surfaces of all permanent teeth, except premolars. Redrawn from figures 10.3 and 10.4 in Reid and Dean (2006). The crown surfaces are each divided into 10 equal zones of crown height, and the mean ages at which they were formed were determined by counting cross striations and brown striae of Retzius in sections under an optical microscope. The ages are quoted in years after birth. For most, two ages are quoted, separated by a slash. The first age was determined from a sample of teeth from southern Africa. The second age was determined from a tooth sample from northern Europe. Where only one figure is given, the two samples provided identical results. It is possible to use the sequence to estimate the ages at which furrow-form defects were initiated by drawing on the positions of the defects for different teeth. Ages can be confirmed by matching the position of defects on one tooth with another.
scoring system, is that the minimum size of defect to be recorded as hypoplasia has not been defined. It is difficult to see how consistency can be maintained when defects vary in a continuous series, from those that are microscopic up to those that can be observed clearly at a distance of some yards (Hillson and Bond, 1997). If detailed defect sequences are required, then there is little alternative to a microscope and the histological approach (Hillson et al., 1999; King et al., 2002). Microscopic studies of crown surfaces supported, in an ideal world, by sections and cross-striation counts are time consuming and are most successful when applied to little worn teeth from children in their late infancy or early teenagers. Although sections give the most reliable chronology, it is possible to estimate timing and duration of defects from counts of perikymata at the crown surface. This can be recorded by taking dental impressions of the labial/buccal surfaces of the teeth. At the time of writing, the highest resolution impression material for this work is Coltene President (Beynon, 1987; Hillson, 1992). The impressions are then used to cast epoxy resin replicas that can be examined either by light microscopy or in the scanning electron microscope (see FitGerald and Rose, this volume). It is possible to use an engineer’s measuring microscope to measure the spacing of perikymata along a transect down the crown side, from the cusp tip to the base of the crown at the cement–enamel junction. There is a trend of perikyma spacing down the crown, and furrow-form defects (which are by far the most common) are defined as groups of perikymata with a wider spacing than expected for the part of the crown in which they are situated.

To build a defect/perikyma groove sequence for teeth in which some occlusal layers are worn away, counts are taken back from the last increment formed at the cement–enamel junction of each tooth (Hillson, 2005; King et al., 2002). The repeat interval of cross striations between pairs of brown striae (and thus perikyma grooves) is unknown without sectioning, but the relative timings of defects can be reckoned in terms of percentage total counts of perikyma grooves between, say, the end of first molar crown formation and the end of second molar crown formation. It is then possible to attach approximate age estimates by using the figures shown in Fig. 10.3, which are calculated for tenths of crown height.

One particular form of hypoplasia is that associated with congenital syphilis. Three characteristic defects have been defined for over 100 years—Hutchinson’s incisors, Moon’s molar, and mulberry molars (Hillson et al., 1998; Jacobi et al., 1992). A mulberry molar is a plane-form defect (initiated a few weeks after birth) surrounding the cusps of permanent first molars, leaving the cusps themselves as small nodules set into the normal-sized occlusal area. These nodules usually break away, and the occlusal surface is rapidly worn or develops a carious lesion, so they are observed clearly only in children. The defect overlaps the possible range of general hypoplasia, but its position and sharpness are at least distinctive. By contrast, a Moon’s molar (also permanent first molars only) shows no clear hypoplastic defect and is normal in appearance in every way except that the cusp tips are abnormally close together, compared with the main bulge of the crown side. The mechanism by which this morphological anomaly is formed has not yet been investigated. In Hutchinson’s incisors, the central one of the usual three mamelons (small cusplets along the incisal edge) is little developed. This leads to a notch and a drawing-in of the incisal edge so that the labial outline of the incisor crown is more oval than normal. The notch wears away rapidly, so that the evidence is again lost in early adulthood. Again, the disruption is probably confined to a few weeks or so after birth, and the defect is outside the wider range of hypoplasia because of its position.

**TOOTH WEAR: CHIPPING AND FRACTURES**

Many collections of archaeological human remains show evidence of rapid tooth wear,
both on the occlusal surfaces and at the contact points between teeth in the same jaw (Fig. 10.4). The rate, pattern, and orientation of wear seem to be strongly related to the nature of subsistence, and for example, hunter-gatherer groups and early agricultural groups have markedly contrasting wear patterns. Loss of crown height from this wear, together with the decreased length of the crowns along the tooth row, causes remodeling of the bone and other supporting tissues in the jaws and results in continuous eruption of the teeth throughout life. As more and more of the root is exposed, then the nature of mechanical forces acting on the jaws changes, further initiating remodeling of the bone. In many past populations, wear was rapid enough to progress down the roots during older adulthood, and eventually teeth were lost through lack of bony support. Heavy tooth wear populations were also typically

![Figure 10.4](image-url)

**Figure 10.4** Two views of a left upper jaw from an Anglo-Saxon site in Winchester, England. The central part of the occlusal surface in the molars has been worn away to expose the softer dentine, leaving a higher rim of the harder enamel. Approximal attrition has broken through the mesial and distal sides of this rim in the first molar, and has thinned it on these sides in the second molar. The second premolar was fractured during life, and subsequently become worn. There is little evidence of bone loss from periodontal disease, but remodeling of the alveolar process has led to thinning of the buccal plate. This is particularly evident over the sockets of the incisors, canines, and premolars, which bulge outward. The first molar is loose and has dropped down slightly in the lower picture, to reveal an opening into the smooth-walled cavity of a periapical granuloma around the apex of its distobuccal root. The wafer-thin opening suggests that the granuloma has been exposed by a fenestration due to the buccal plate thinning, and not by the sinus of an abscess. There is no sign of an open pulp chamber or root canal on the worn occlusal surface of this tooth, so it may be that a fine crack allowed the pulp to be infected. In the fractured second premolar, there is also no evidence root canal exposure or periapical inflammation, and the fenestration that exposes the root surface must again be from general remodeling of the buccal plate. The exposed root surfaces are roughened and bulbous, suggesting hypercementosis. Eroded remnants of supragingival calculus are visible on the roots of the molars. This specimen is in the collection of the Natural History Museum, London, and has been photographed with their permission.
subject to dental fractures, from minor chipping to major cracks and breaks that exposed the dental pulp to infection. In addition, many such groups show evidence of abnormal abrasion that is apparently from the use of teeth for preparing food or in the manufacture of artifacts. All these factors would have had a profound effect on the development of dental disease.

Tooth-on-tooth wear that creates wear facets on both teeth involved is described as attrition. The important features to record from a dental pathology point of view are the exposure of new sites for initiation of dental caries, or the modification of existing sites, the progressive loss of crown height in relation to continuous eruption (below), and the distribution and orientation of wear facets in relation to remodeling of the alveolar process as it responds to forces acting on the dentition. Occlusal attrition is usually recorded by scoring the pattern of dentine exposed as the wear progresses. The most frequently used scheme is that of Smith (1984), and this is a recognized standard (Buikstra and Ubelaker, 1994). The similar method of Molnar (1971) has some advantages for the study of dental pathology, in that it also records the exposure of secondary dentine and the shape and orientation of wear facets. Both of these methods are satisfactory in relation to the study of dental caries. A detailed method defined by Scott (1979) is confined in use to the molars and is therefore too limiting in this context. It is difficult to define measurements for occlusal wear, although attempts have been made to measure the reduction in crown height (Molnar et al., 1983; Tomenchuk and Mayhall, 1979; Walker et al., 1991). The difficulty here is that the height of unworn crowns is variable within any population, and it is hard to interpret results because of this. In addition, many archaeological teeth are so worn that the crown is missing altogether down one side. It might be more appropriate, in relation to studies of continuous eruption versus root exposure caused by periodontal disease, to measure the height of the occlusal facet in relation to the crest of the alveolar process. This can be done simply using a graduated periodontal probe, at the same time as measuring the root exposure (below). In the mandible, another suggestion is to measure the height of both the occlusal facet and the alveolar crest in relation to the mandibular (inferior alveolar) canal, in an attempt to find an independent landmark on which to base the dimension (Whittaker, 1986; Whittaker et al., 1987, 1990). This makes it possible to see whether the occlusal plane has changed its position with increasing wear, and the few studies that have been done in this way suggest that the rate of continuous eruption approximately matches the rate of wear (in heavy wear populations) so that the occlusal plane remains in about the same position with respect to the body of the mandible (Fig. 10.5). Such measurements require X-rays, including the whole mandibular body, and there is the difficulty that the mandibular canal varies in position between individuals, relative to the boundaries of the body as a whole, so that it is not entirely a homologous reference point. Still other measurements have been devised to record the angle of occlusal wear facets (Smith, 1984, 1986; Walker et al., 1991), and these have shown that there may be differences between hunter-gatherers and agriculturalists in the progression of angular changes in relation to stage of molar wear.

Wear between neighboring teeth in the same jaw is known as approximal, interproximal, or interstitial attrition (Wolpoff, 1970). The rate of approximal attrition, relative to occlusal attrition, is particularly high in hunter-gatherer groups when compared with early agriculturalists (Hinton, 1982). It is more difficult to record, but the edge of approximal attrition facets is a common site for the initiation of dental caries. One method is to measure the buccal–lingual length of the facet, as exposed at the edge of the occlusal facet, with a pair of needle-point calipers. This has proved useful in studies of the pattern of wear, but from the dental pathology viewpoint, it is more appropriate to record the exposure of dentine in the wear facet, the extension of wear down to the cervical
margin of the tooth, and the eventual loss of approximal contact between teeth because of the combined results of heavy occlusal and approximal wear. For a detailed study, therefore, it makes sense to record both the length of the facet and a simple score for its appearance (below).

Wear that results from contact with objects other than teeth is known as abrasion. General abrasion occurs in all teeth with increasing age, especially on the buccal/labial and lingual surfaces (Hillson, 1996; Scott et al., 1949). Nowadays, toothbrush abrasion is observed in almost all teeth as a very glossy polish to the crown surface. As the roots become exposed with increasing age, the softer cement and dentine may be abraded more rapidly, to undercut the enamel (Soames and Southam, 1993). The very heavy abrasion observed today results from the use of toothpastes containing abrasives, and nothing like it is observed in most archaeological or
museum collections. However, one particular feature of European jaws from the eighteenth and nineteenth century is clay pipe facets—semicircular notches worn usually in the canines and premolars by holding the pipe stem (Kvaal and Derry, 1996). Some assemblages of Native American remains from the North also have a characteristic pattern of very heavy abrasions on the labial/buccal surfaces of teeth. These abrasions relate to the wearing of labrets: plugs of stone, bone or ivory fitted into openings cut into the lips and cheeks (Cybulski, 1994; Milner and Larsen, 1991; Moss, 1999). Labret facets are readily identifiable as flat areas on the labial/buccal surfaces of incisors, canines, and premolars, usually in the lower jaw, but there may be more general polishing or more limited grooving of a single tooth. Wear of this type presumably went slowly enough for secondary dentine deposition in the pulp chamber to repair the damage, because cases of pulp exposure and periapical inflammation are not recorded. The same is true of deliberate tooth mutilation by cutting, drilling, or filing, as is observed particularly in human remains from Central and South America (Arcini, 2005; Milner and Larsen, 1991; Romero, 1958, 1970; Williams and White, 2006). Many different patterns are known, often involving deep penetration into the crown, and some ethnographic accounts suggest that the cuts were made bit-by-bit, with long intervals between operations that must have allowed reparative secondary dentine to build up slowly in the pulp chamber. Restorative procedures in dental surgery, starting mainly in the nineteenth century, are in effect also a form of abrasion, and an important branch of dental studies is the matching of jaws with dental records in identification for forensic purposes (Clark, 1992; Cottone and Standish, 1981; Stimson and Metz, 1997; Whittaker and MacDonald, 1989).

Another range of common abrasions results from causes that are more difficult to identify. One of these is heavy scratching, particularly of the labial surface of incisors, and especially in hunter-gatherer groups (Bermúdez de Castro et al., 1988; Fox, 1992; Fox and Pérez-Pérez, 1994; Fox and Frayer, 1997; Frayer and Russell, 1987). It is usually assumed that these represent the cutting of objects held between the teeth, but various forms and orientations of marks imply a variety of causes. Another such feature is so-called approximal grooving—broad grooves running from buccal to lingual along the cervical margin of cheek teeth (Bermúdez de Castro et al., 1997; Brown and Molnar, 1990; Formicola, 1991; Frayer, 1991; Ubelaker et al., 1969; Milner and Larsen, 1991). These grooves are now commonly recognized in a range of archaeological and fossil material. The cause is unknown and the subject of considerable debate. Still other grooves may be observed on the occlusal wear facets of particularly incisor, canine, or premolar teeth (Larsen, 1985; Molleson, 1994). It is assumed that these result from the holding of fibers or yarn when processing materials or weaving artifacts.

Several standard terms are recognized in the study of abrasions, but the only recording scheme is that of Romero (above) for classifying tooth mutilation. For a general study, good photographs and descriptions are the main requirement along with, perhaps, dental impressions (see above).

Fractures of the teeth and jaws are also common in archaeological assemblages. Teeth often show chipping around the edge of occlusal attrition facets, especially in molars (Milner and Larsen, 1991; Pedersen, 1947; Turner II and Cadien, 1969), and there are sometimes more major breaks that can be distinguished from postmortem damage by the presence of wear on the exposed surface and rounding of exposed edges. These breaks may open the pulp chamber to infection (below). It is also possible for teeth to be cracked, and there may be little surface sign that this has taken place (Fig. 10.4). Sometimes an apparently intact tooth has associated periapical inflammation (see below), and one possible explanation is that a fine crack has exposed the pulp chamber (Alexandersen, 1967; Pedersen, 1938). Heavy wear on crowns may predispose them to
cracking and more serious breaks. Still other fractures may occur in the roots (Pindborg, 1970), with little sign at the surface even in living people. Fractures of the bone of the mandibular body, or alveolar process in the maxilla, are sometimes observed in archaeological collections. Such fractures tend to track down the socket of a tooth, as a line of weakness. Not only does this involve the teeth in the processes of inflammation and repair, but it also commonly makes the fracture a compound one in which the wound penetrates to the surface and is thus exposed to infection. Cracking and chipping of the teeth themselves can be recorded by simple coding, supported by photographs and description.

**PLAQUE-RELATED DISEASES**

Dental plaque deposits form on the surfaces of all teeth. They consist of large colonies of micro-organisms and associated extracellular material, and collections of human remains show abundant evidence for the presence of plaque, in the form of mineralized plaque, known as dental calculus or tartar (Figs 10.4 and 10.8). Long-standing accumulations of micro-organisms and their extracellular products next to the gums provoke an immune response. This follows a well-established pattern through life, and episodes of inflammation alternate with periods of recovery, leading to a cumulative loss of support for the affected tooth, and remodeling of the bone of the jaw, resulting in a progressive reduction in the height of the alveolar process and eventual loss of teeth. The extracellular products of plaque micro-organisms also affect the surfaces of the teeth themselves. Organic acids form as by-products of plaque physiology, and result in localized demineralization of the tooth surface, to produce the characteristic lesions of dental caries. A cavity may develop, and ultimately this may penetrate the pulp and expose it to an infection that may result in inflammation of the tissues around the apex of the root. The nature and age-related development of dental caries lesions are indicative of the diet and, particularly, its carbohydrate component.

![Figure 10.6](image_url)

*Figure 10.6* Lower first molar, premolars, and canine of a post-Medieval jaw from London. On either side, the empty sockets denote teeth that have been lost postmortem. Alveolar bone (lining the socket) has been lost around the roots of the molar and, most noticeably in this view, the second premolar to create a deep crater between it and the molar. This is the classic appearance of bone loss caused by periodontal disease. The crests of the interdental plates between other teeth are much less affected, but they are still somewhat porous. Along the CEJ of all four teeth is a lesion that looks very much like root surface caries, but in archaeological material like this, it is important to consider possible diagenetic effects.
Dental Calculus

The presence of dental calculus indicates long-standing plaque accumulations, but it is difficult to deduce anything more because the factors that initiate mineralization are little understood. Two types of calculus are recognized: supragingival and subgingival. The irregular clay-like deposits often observed on the surface of crowns (and sometimes roots) are all supragingival calculus. They can be so large that they overhang the gums, but they do not seem to have any direct role in irritating the periodontal tissues (below), except in providing an extended surface on which living plaque can accumulate. At a population level, there is a slight inverse relationship between calculus and caries prevalence (Manji et al., 1989b), which is understandable in that the first is a mineralization phenomenon, whereas the second is mostly a demineralization phenomenon. The relationship is not at all strong, however, and in an individual jaw, it is common to see both carious cavities and supragingival calculus deposits (Thylstrup et al., 1989). Sometimes they are found on the same teeth (Jones and Boyd, 1987). Both are long-term conditions, and the lesions could represent different phases of plaque biochemistry, but the development of caries in any case relates to very local biochemical changes, which may continue even underneath a calculus deposit.

Subgingival calculus occurs only in association with periodontal disease. It is a much thinner, less obvious, layer that coats the surface of the roots in a periodontal pocket (see below) that extends down beside the root inside the gingivae. There is usually an area of exposed root, bare of subgingival calculus, around the cervix of the tooth marking the opening into the pocket where the gingival cuff (border of the gums) rested. A subgingival calculus deposit frequently has a somewhat more strongly pigmented apical border, but this can only be used to infer the minimum size of the pocket as the deeper parts of the pocket may not be underlain by calculus (Richardson et al., 1990). It is important to remember that supragingival calculus may also be present on a root surface. It forms where the root surface is exposed above the border of the gums. The term subgingival applies only to the special case of a periodontal pocket.

Should calculus be recorded? It is certainly difficult to deduce much from its presence. In any case, great care needs to be taken when recording archaeological material and museum specimens, because the supragingival deposits can easily be dislodged to leave only a vestige on the crown side. On the other hand, calculus is relatively quick and simple to record. Most anthropologists use the Brothwell (1981) three-stage score for supragingival deposits, and a graduated probe can be used to measure the maximum length of root covered by subgingival deposits (Powell and Garnick, 1978) although this gives only a minimum depth for the periodontal pocket.

The histology of calculus demonstrates the plaque architecture and the forms of microorganisms (Arensburg, 1996; Linossier et al., 1996), which are preserved as voids within the mineralized matrix, and it is possible to distinguish between supragingival and subgingival deposits (Bercy and Frank, 1980; Friskopp and Hammarstrom, 1980; Friskopp, 1983). Ideally, tooth and calculus are resin impregnated and sectioned together, producing a flat surface that can be examined by back-scattered electron microscopy or confocal light microscopy (Boyd and Jones, 1983; Jones, 1987). The biochemistry of calculus is complex and not well understood. One potentially useful aspect is that calculus is an effective preserver of DNA (see Stone, this volume). Human cheek cells are incorporated into plaque and subsequently mineralized, and it is possible to determine the sex of the owner of a tooth by DNA extraction from calculus and PCR amplification of appropriate sequences on the X and Y chromosomes (Kawano et al., 1995). Removal of small calculus samples would be relatively simple and nondestructive, although contamination is likely to be a problem with archaeological material.
Dental Caries

Dental caries is a progressive demineralization of the enamel, cementum, and dentine of the tooth by organic acids, which are produced through the fermentation of dietary carbohydrates by some plaque bacteria. During the twentieth century, sugars were established as the main factor in caries rates (Navia, 1994; Rugg Gunn, 1993a, 1993b; Sheiham, 1983; Thylstrup and Fejerskov, 1994). This was shown in a particularly striking way by the decrease in caries rate among children as a result of sugar rationing in Japan, Norway, and the Island of Jersey during the war of 1939–1945. Sucrose is often regarded as the sugar mainly responsible for caries, but this is probably because it is the one that is eaten in greatest quantity and there is little difference in cariogenicity among fructose, glucose, lactose, and sucrose. It should be noted, however, that it is possible for an individual to develop caries even if they do not eat sugar (Lingström et al., 2000; Rugg Gunn, 1993b; Schamschula et al., 1978). Starches seem generally to have a low cariogenicity relative to sugars, but they do cause caries, and a mixture of starches and sugars is as cariogenic, weight for weight, as pure sugar. The role of dietary proteins and fats is currently not well understood, but milk products seem to have a protective effect (Bowen and Pearson, 1993; Mundorff-Shrestha et al., 1994), and the caries rate was very low among recent Inuit who ate almost entirely animal-based foods, with little or no carbohydrate (Mayhall, 1970, 1977, 1978; Pedersen, 1966). On an archaeological timescale, caries was rare until the adoption of agriculture (Larsen, 1995; Larsen et al., 1991; 1997), and fermentable carbohydrates from cultivated crops were added to the diet. From that point onward, there was a steady rise in caries with, in Britain at least, a sharp rise during the nineteenth century accompanied by a change in the pattern of lesions that may be related by an increase in sugar consumption (Corbett and Moore, 1976; Moore and Corbett, 1971, 1973, 1975a, 1975b). Caries, along with dental attrition and stable isotope studies (see Katzenberg, this volume), thus has considerable potential for identifying dietary changes in the archaeological record.

Recording and interpretation of dental caries statistics need to keep in mind the nature of the disease. In most cases, caries is very slowly progressive, with alternating phases of stability and activity over many years (Pine and ten Bosch, 1996). Not all lesions recorded will currently be active, and even when most of the crown has been destroyed, the tooth is not necessarily just about to be lost. In fact, the main cause of tooth loss in caries is deliberate extraction to treat tooth pain. When modern clinical studies record teeth as “missing due to caries,” they mean that a tooth that was painful or had a clear cavity in it was extracted previously (Manji et al., 1989a, 1989b). Sensitivity and pain result from an acute inflammation of the pulp and periapical tissues, but such inflammation does not produce the type of bone resorption that would cause a tooth to be lost without human intervention. Nevertheless, tooth extraction is one of the oldest surgical procedures and would possibly have been available to many people in the ancient past.

The slowly progressive nature of caries leads to a pattern of development that is related strongly to age (Manji et al., 1991, Thylstrup and Fejerskov, 1994). It is therefore essential to divide a collection of jaws up into different age groups when preparing caries statistics. It is also necessary to recognize that carious lesions fall into several different categories, in relation to their site of initiation on the tooth surface. These categories have contrasting etiologies and develop in different ways with increasing age, so that they also need to be kept separate. Typically, the average rate of caries (the word “prevalence” has deliberately not been used—see below) in a population is from a few individuals with a high caries experience, balanced by the bulk of the population who have relatively little experience of caries. This is a difficult situation for archaeology, where collections are often small, and affected by taphonomy, variable preservation, and
uneven recovery (see Stodder, this volume). Not only does the potential for differential survival of particular age groups need to be considered but also males versus females and different parts of the dentition. In many studies of caries, females show higher rates than males, and there frequently seems to be a bias against the burial and survival of female remains in archaeology. Similarly, whereas caries is usually symmetrical between left and right sides, there are considerable differences between upper and lower teeth, and between incisors, canines, premolars, and molars. This means that such factors as relative survival of upper and lower jaws, and between incisors, canines, premolars, and molars, will greatly affect caries statistics in archaeological and museum assemblages. Separate statistics are needed for each category, and this increases the complexity of recording, but there is no alternative.

Carious lesions can be divided into two main forms (Thylstrup and Fejerskov, 1994): coronal caries and root surface caries. The root surface lesions (Fig. 10.7) are initiated along the cement–enamel junction (CEJ) at the base of the crown, or on the cement of the root, as they are exposed in adults by periodontal disease (below). They appear to be initiated only on root surfaces exposed above the margin of the gingivae (gums) and not inside periodontal pockets (above and below). Coronal lesions may be initiated at any age, in the enamel surface of the crown, or in dentine exposed by wear. Two important sites of initiation in modern populations are in the occlusal fissures and fossae of molars (occlusal caries) and in the protected sides of the crown just below the contact points between neighboring teeth (approximal or interproximal caries).

In the more worn teeth of many archaeological collections, other coronal sites appear to have been created (or modified) in wear facets or from damage by chipping (Hillson, 2001). In the twentieth century, a clear hierarchy of coronal carious lesion sites was demonstrated for permanent teeth (Batchelor and Sheiham, 2004; Sheiham, 1997). The sites most at risk were the occlusal fissures of first molars (Table 10.1). These were followed by the fissures of second molars and second premolars. Next came the fissures of first premolars and

![Figure 10.7](image_url)

**Figure 10.7** Gross approximal caries that has destroyed the mesial side of an upper incisor in an early Medieval skull from London. The lesion may well have been initiated at the contact point but has advanced so far that it is not possible to be sure.
contact points of first molars. After that came the contact points of second molars, premolars, and incisors. In a high caries rate population, all these sites might commonly be involved, but in a low caries rate population, just those sites most at risk. From this, it is clear that the teeth involved and the site of lesions are important factors in caries epidemiology. Archaeological assemblages also show a hierarchy of sites, although their location is somewhat different (below).

From its original site of initiation, a coronal lesion may progress from a white or brown spot into a cavity, which may in turn grow into the dentine and ultimately penetrate the secondary dentine defenses of the pulp (above), to expose the soft tissues of the pulp to infection (Fig. 10.6). Untreated, this sequence seems in most cases to take many years. There are relatively few clinical studies of the development of caries in a living population where dental treatment is very limited (the so-called natural history of caries). What evidence there is (Luan et al., 1989; Manji et al., 1988, 1988a, 1988b, 1990, 1991, Matthesen et al., 1990) suggests a lifelong development, with occlusal and approximal caries dominating in children and young adults, and these forms being replaced gradually by dentine caries and pulp exposure, loss of teeth (presumably through both extraction and periodontal disease), and root surface caries. This sequence is shown in caries statistics for the population as a whole, but it reflects mostly the contribution of a relatively small, caries-prone, group of individuals. Caries is more common in women than in men and shows a stronger age-related development. Cheek teeth are more strongly affected by caries than canines and incisors and upper teeth more than lower. The populations studied in this way were predominantly rural agriculturalists, but although studies in the

<table>
<thead>
<tr>
<th>Risk of Developing Caries</th>
<th>Second Molars</th>
<th>First Molars</th>
<th>Second Premolars</th>
<th>First Premolars</th>
<th>Canines</th>
<th>Second Incisors</th>
<th>First Incisors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Highest risk</td>
<td>Fissures</td>
<td>Upper pits</td>
<td>Fissures</td>
<td>Upper pits</td>
<td>Fissures</td>
<td>Upper pits</td>
<td></td>
</tr>
<tr>
<td>Fissures and lower pits</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Upper pits</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other surfaces</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lowest risk</td>
<td>Lower contacts</td>
<td>Any surface</td>
<td>Lower contacts</td>
<td>Any surface</td>
<td>Lower</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contacts and other surfaces</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Data from Batchelor and Sheiham (2004) for twentieth-century children. Some potential caries sites have different susceptibilities in upper and lower teeth, and these are indicated in the table. “Fissures” means the fissure systems of the occlusal surface. “Pits” means the extensions of the fissure system over on to the buccal side of lower molars and lingual side of upper molars that results in a marked depression in some teeth (but this potential caries site is not present in all teeth in all individuals). There is sometimes a similar pit on the lingual surface of upper incisors. “Contacts” refers to the points at which neighboring teeth in the same jaw meet. “Other surfaces” are anything else—usually buccal/labial or lingual surfaces along the line of the gingivae. See text for explanation.
1940s and 1950s of recent hunter-gatherer groups suggest an equally strongly age-related development, the nature and pattern of caries was greatly different. Among the Australian aborigine people (Barrett, 1953; Campbell, 1925; Campbell and Gray, 1936; Campbell and Barrett, 1953; Cran, 1959; Moody, 1960), and !Kung Bushmen (du Plessis, 1986; Van Reenen, 1964, 1966), carious lesions of any type were rare in children and young adults, and caries was predominantly a disease of older adults associated with advanced wear, abrasions, and chipping of teeth. Until recently, East Greenland Inuit (Pedersen, 1938, 1947, 1949, 1966) ate no carbohydrate at all and experienced almost no caries. It is thus not surprising that archaeological collections show a range of lesion types and distributions, reflecting the nature of the diet. Some, such as Archaic Native Americans (Larsen, 1995, 1997; Leigh, 1925, 1928) or Upper Palaeolithic/Mesolithic Europeans (Brothwell, 1963; Lubell et al., 1994), have heavy wear but very little evidence of caries and conform well with the hunter-gatherer model. Others, such as eighteenth-century Londoners (Molleson and Cox, 1993), have slight wear and a pattern of caries little different to recent times. Still others, such as British Anglo-Saxon or Native American agriculturalists (Larsen, 1995, 1997), may have heavy wear in association with coronal caries in many adults.

Most modern studies of living people use the DMF score (a simple count of decayed, missing, or filled teeth) to summarize an individual’s experience of caries, and the arithmetic mean DMF score to represent a caries index for the population (Thylstrup and Fejerskov, 1994). There are many difficulties with this approach, even in living populations, particularly when the study includes all age groups. Mean DMF score is not a prevalence (Waldron, 1994) in epidemiological terms (the number of individuals affected divided by the overall number in the population), but it is instead the average number of teeth per individual affected. It implicitly assumes that all populations have the same average starting number of teeth, that all missing teeth were lost because of a similar balance of caries and other factors, and that equivalent proportions of males and females, and different age groups, are present in the study groups. The latter assumption could be built into the selection of groups for modern clinical studies, although it makes for difficulties in comparing different studies, but it is an insuperable problem with archaeological collections where there is much less scope for control of selection. The other two assumptions are equally problematic, particularly when groups with widely differing diets and patterns of subsistence are compared. For archaeological and museum material, there is the added difficulty that many teeth are missing postmortem, simply because they have fallen out of the jaws, with no evidence whether they were carious. However widespread they may be in the clinical literature, DMF scores are therefore not appropriate for archaeology. The most commonly used alternative is the percentage of teeth with caries, out of the total in the collection (Hillson, 2001). This takes some account of postmortem tooth loss and has the advantage of allowing isolated teeth to be used. It ignores, however, the varied survival of different tooth classes, which must have a profound effect on the apparent caries rate because of the greatly differing susceptibilities of anterior and cheek teeth. The inclusion of isolated teeth makes it difficult to distinguish between males and females, and in its most common usage, the percentage of teeth does not distinguish between different age groups or types of carious lesion. For these reasons, the basic overall percentage is also inappropriate for archaeological material. Various modifications have been proposed (Costa, 1980; Lukacs, 1995), but the fundamental development has been the work of Moore and Corbett (1971, 1973, 1975, 1976), who recorded separate percentages for different lesion categories, different tooth classes, and different age groups. For each lesion category, they calculated the percentage out of the number of “teeth at risk” of developing that type of lesion, because these teeth retained the appropriate initiation sites (for
example, only counting premolars and molars for occlusal fissure caries). They also introduced a separate category of “gross caries,” where the lesion had removed so much of the crown that its original site of initiation could not be determined with any certainty. Their procedure has been proposed as the standard (Buikstra and Ubelaker, 1994) for presentation of caries statistics in anthropology, and it can be developed in several of ways.

For this approach, the carious lesion initiation site on the tooth is the basic unit of recording rather than the tooth as a whole. In every tooth, of each individual, separate records are kept for the presence of occlusal fissures, enamel just below the approximal contact points, and the cement–enamel junction or exposed root surface. These are recorded as normal, bearing a white or stained spot (possibly a carious lesion), or bearing a clear cavity, and whether this cavity exposes dentine or the pulp chamber. Where the expected initiation site is missing or obscured for any reason (breakage, wear, calculus, or even a gross carious lesion), it is counted as absent because its status as a site of caries initiation is unknown. Carious lesions are described as gross when so much dental tissue has been lost that it is not possible to determine exactly where on the tooth the lesion started. Gross caries is recorded as the combination of sites in which it might have been initiated—gross occlusal/approxiomal, gross coronal/cej, gross gross (when none of the crown is left to tell whether it was initiated there or in the roots). Both occlusal and approximal attrition have to be recorded, with the additional initiation sites that they open up through exposure of dentine. The scoring methods of Smith or Molnar (above) are suitable, although a new caries score needs to be added to describe the nature of these lesion sites. For interpreting root surface caries, it is useful to have some record of the exposure of root surfaces. In a modern clinical study, the exposure of roots by gingival recession would be noted (Fejerskov et al., 1993; Luan et al., 1989), but this is not possible in archaeological material because the position of the gingival margin cannot be reconstructed. Subgingival calculus deposits merely indicate the presence of a periodontal pocket (below), and it is not thought that carious lesions are initiated within such pockets (above). Where there is a band free of subgingival calculus at the cervix of the tooth, however, this may indicate the position of the gingival margin (above). The simplest way to keep a general record of root exposure in dry bone specimens is to take measurements from the cement–enamel junction to the crest of the alveolar process with a graduated probe (Watson, 1986), without taking into account intra-bony defects (below). It needs to be made clear, however, that this does not correspond to the level of gingival attachment, and such measurements can only be used on a comparative basis. In clinical radiography, 2–3 mm between cement–enamel junction and alveolar crest would be regarded as a normal value (Whaites, 2006), so this could be used as a cut-off point when tabulating root exposure against root surface caries. Finally, some consideration needs to be given to antemortem tooth loss. This is important because, in a population where caries is common, the most heavily affected teeth are those that are lost because of extraction. The difficulty is that the cause of antemortem tooth loss cannot reliably be determined and caries cannot be assumed. Postmortem and antemortem tooth loss, and potential factors that might initiate tooth loss, therefore need to be recorded as separate categories, in such a way that alternative causes can be assessed. For this reason, it is important that congenital absence of teeth, impaction, trauma, alveolar bone loss, and periapical bone loss are all recorded. These will be present, in any case, in any detailed dental study.

Diagnosis of caries is also a problem, although archaeology presents both advantages and added difficulties. Archaeological teeth are exposed fully and dry, may be rocked to one side in their sockets to show the approximal area, can be brightly lit from any direction, examined under a microscope, and X-rayed without...
restriction. They may, however, be subject to
diagenetic changes that mimic caries (Poole
and Tratman, 1978), and the patient cannot be
asked for a history such as “did the tooth hurt
or was it wobbly before you lost it?”,
Examination is by eye, with the aid of a lens
or a low-power stereomicroscope and routine
X-rays, as used in trials of clinical methods on
isolated teeth in the laboratory (Ekstrand et al.,
1995). A sharp dental probe or explorer is unli-

tely to improve the diagnosis and may indeed
scrath delicate specimens (Ismail, 2004;
Penning et al., 1992). If a critical diagnosis
must be confirmed, even at the expense of the
specimen, relatively nondestructive sections
may be made by impregnating a tooth in methyl-
methacrylate resin (Hillson, 1996), cutting it
open with a slow speed saw, and then examining
the polished surface for signs of dentine invol-
ue and secondary dentine reaction in the
pulp chamber. This can be by a simple examin-
ation under low-power magnification, but it is
also possible to examine such specimens using
backscattered electron imaging in the scanning
electron microscope or confocal light
microscopy (Jones, 1987; Jones and Boyd,
1987). The methylmethacrylate can be
removed in acetone, and only a thin layer at
the center of the tooth is lost. In diagnosis, the
main difficulty is in recognizing the early
stages of coronal caries, in which there is only
a small opaque white (or brown stained) spot
in the translucent enamel, with no evidence of
a cavity or even surface roughness. Many clini-
cal studies do not record such “incipient caries”
because of the difficulty in attaining consistency
between observers (Burt, 1997; Lussi, 1996).
More recently, it has been suggested that there
are fewer interobserver problems than had
been thought, and the exclusion of the incipient
stage must underestimate caries considerably
(Ismail, 1997, 2004). Where white or brown
spots are recorded, it is probably best to note
them as a separate category rather than including
them in the overall caries statistics. Additional
difficulty arises in occlusal caries, where the
initiation site may be very deep in the fissures
(Dowker et al., 2006; Ekstrand et al., 1995,
1998; Espelid et al., 1994; Penning et al.,
1992; Tveit et al., 1994; van Amerongen et al.,
1992). The lesion may progress to a cavity
with little sign at the surface, and standard
X-rays are of limited use because of the
convoluted nature of the (much more radio-
opaque) enamel in this part of the crown, so
that they only record the lesion when the under-
lying dentine is demineralized significantly.
Nonetheless, a combination of microscope exami-
nation and X-ray catches most developing
lesions in the fissures (Hopcraft and Morgan,
2005; Lussi, 1996). Most “stained fissure”
lesions or minor cavities in the occlusal surface
are accompanied by underlying dentine caries
(Ekstrand et al., 1995). A variety of additional
methods for measurement and observation has
been tested for clinical use (Hall and Girkin,
2004), and some of these may prove appropriate
in archaeological material.

Several factors predispose to caries. One of
the best known relationships is with enamel
hypoplasia, where the defects act as lines of
weakness through the enamel. The deminerali-
ization of dental caries occurs selectively along
such lines and may potentially reach the
dentine faster than in sound enamel.

Indivduals with hypoplasia are therefore more
likely to have dental caries (Mellanby, 1927,
1934). If hypoplasia is recorded in occlusal,
mid-crown, and cervical parts of the crown as
described above, then this can be linked to the
position of lesions on the crown. Fluorine also
needs to be considered, but it poses much
more difficult questions. The presence of fluor-
ide ions in the hydroxyapatite crystal lattice of
the enamel has an important inhibiting effect
on caries, and fluoridation of the water supply
together with fluoride additives in toothpaste
are probably the main reasons for the reduction
in dental caries in Europe and North America
over recent years (Thylstrup and Fejerskov,
1994). High levels of fluoride during growth
cause hypomineralized and hypoplastic
defects in the enamel, but it is difficult to use
these as an index of fluoridation (Fejerskov
et al., 1988; Lukacs et al., 1985). Nor is it
possible to use measurements of fluorine in the
enamel, because the element is highly mobile in the soil and, indeed, gradually accumulates in buried bones and teeth. The best line may be to consider current water analyses from the region and assess the potential water supply for the people whose remains are being studied.

Bone Loss from the Alveolar Process

The alveolar process is a site of very active bone turnover and remodeling at all ages. Within the arch of the process in both jaws, the sockets (or alveolae) change position slowly with remodeling to allow changes in the position of the teeth in response to the developing occlusion and progress of wear. Even after the initial eruption of permanent teeth during childhood and young adulthood, they continue to erupt gradually throughout life. Part of this is from deposition of cement at the apex of the roots, but this is generally only a thin layer (although see hypercementosis below) and a more important factor seems to be the remodeling of the alveolar process, with migration of the sockets upward through it (Fig. 10.5). Even where there is no bone loss from periodontal disease (below), therefore, the cervical part of the root is exposed progressively with increasing age, while the sockets become shallower. At the same time, the alveolar process stays at approximately the same height so that, in populations where there is little tooth wear, the face becomes slightly taller with age (Whittaker et al., 1990). Some evidence suggests that the rate of remodeling may adapt to the rate of wear, and that rapidly wearing teeth are accompanied by a more rapid rate of continuous eruption that maintains the occlusal surfaces (and face as a whole) at a similar height through life (Whittaker et al., 1982, 1985). In addition to vertical migration of the sockets, there is mesial migration in response to approximal wear between teeth. Such “mesial drift” is a particular feature of the rapidly wearing dentitions of hunter-gatherers (Begg, 1954; Corruccini, 1990, 1991; Kaul and Corruccini, 1992). Another common feature in such dentitions is that teeth, together with their sockets, may become inclined progressively as the alveolar process remodels around them. As this takes place, the buccal/labial plate of cortical bone is often thinned so that there is a depression between the positions of each pair of sockets (Fig. 10.4). The combined socket and buccal plate may become so thin that an aperture exposes the root in dry bone specimens. If the aperture appears around the apex of the root, it is known as a fenestration, but if it simply extends as a notch down from the cervix, it is called a dehiscence (Clarke and Hirsch, 1991; Muller and Perizonius, 1980). These are usually covered by soft tissue during life, although there may be a large area of exposed root surface in some cases. It is assumed that all these features relate to the forces (MacPhee and Cowley, 1975) acting on the teeth (see tooth wear above), although there is mixed evidence that dehiscences and fenestrations are in any way related to heavy tooth wear (Rupprecht et al., 2001; Stahl et al., 1963).

Bone may also be lost from the alveolar process (Clarke and Hirsch, 1991; Hildebolt and Molnar, 1991) because of inflammation of the supporting tissues of the teeth (known as the periodontal tissues). Two distinct types of bone loss are observed. One is the preferential removal of the alveolar bone proper (Figs 10.5 and 10.8)—the lining of the sockets—in a narrow trench extending down from the tooth...
cervix (Kerr, 1991; Soames and Southam, 2005). This relates to the condition known as periodontal disease, which involves inflammation of the gums. The other form of bone loss (Fig. 10.4 and 10.6) is concentrated around the apex of the roots (Dias and Tayles, 1997) and relates to so-called periapical inflammation, which originates in infections of the pulp. Both forms are very common in archaeological specimens, and they may be so extensive that they are hard to distinguish from one another. They may also be difficult to distinguish from the general remodeling of the alveolar process described above.

The micro-organisms present in dental plaque produce a variety of antigens that trigger an immune response in the supporting tissues of the teeth (Lehner, 1992; MacPhee

![Figure 10.8](image_url)  
**Figure 10.8** Two views of part of the right upper jaw from an Anglo-Saxon skull from Winchester, England. Lower view shows a gross gross lesion (so large that it cannot reliably be distinguished as coronal in origin) that has exposed the open root canal of the upper first molar. The second premolar next to it has been lost postmortem. Upper view shows the buccal side of the same specimen. The thinned buccal plate of the alveolar process has broken away (the sharp broken edge is clearly visible) to reveal the smooth cavities of two periapical granulomata around the apices of the buccal roots of the molar. There are eroded remnants of supragingival calculus on the crown and roots of the first molar and first premolar. This specimen is in the collection of the Natural History Museum, London, and has been photographed with their kind permission.
and Cowley, 1975; Marsh and Martin, 1999; Schluger et al., 1990). In most people’s mouths there are constantly at least some areas of low-level inflammation in the gums at all times. With long-standing plaque deposits, there are intermittent local escalations of this inflammatory reaction in particular areas of the mouth. This higher level of inflammatory response involves the underlying tissues and disrupts the periodontal ligament that binds the roots into their sockets. The continuity of fibers linking the cement of the root surface with the alveolar bone of the socket is broken, starting at the cervical part of the root and extending gradually down toward the apex. A so-called periodontal pocket forms down the side of the root, and subgingival plaque accumulates within it. Once the connection with the root is broken, the alveolar bone is resorbed, and this is what causes the characteristic pattern of bone loss. The condition as a whole is called periodontal disease, and in its most common form, it is episodic in nature, with intervals of inflammation alternating with periods of relative normality during adult life. The loss of the alveolar bone, therefore, progresses gradually and eventually reaches the point at which the tooth becomes unstable. This mobility of teeth is characteristic of periodontal disease and is the prelude to the eventual loss of the teeth. Once the teeth have gone, the dental plaque, the source of irritation, is also lost and the lesion heals over. The underlying bone develops a smooth surface, with no sign that tooth sockets were ever there, and the alveolar process is greatly reduced in height. Periodontal disease particularly affects the molars, and it is common to find mandibles in which the anterior teeth are still present in a high “prow” of bone, whereas the cheek tooth area is much lower and smoothed over. This pattern strongly suggests that periodontal disease had a major part in the loss of the teeth, but it must be remembered that dental caries is also common in populations with periodontal disease, and molars are particularly affected by it, so that teeth may also have been extracted for this reason.

The initial phase of bone loss in periodontal disease is therefore a trench-like removal of alveolar bone around the cervix of a tooth. This may be monitored by examining the interdental walls (Fig. 10.5 and 10.6) of bone that separate neighboring sockets (Kerr, 1991). In effect, these are the combined alveolar bone of the sides of the sockets, with the trabecular bone in between. The tops of these plates normally bulge up between incisors, or run level between cheek teeth, and have a relatively smooth cortical bone surface. When affected by inflammation in the overlying soft tissues, this surface becomes porotic. With a deeper involvement of supporting tissues, the normal contour of the plate is disrupted to create a concave top, with a sharp, ragged texture. Additional development of the disease is marked by increased loss of bone at the top of the interdental walls, making a deep, steep-sided “intra-bony” defect. Phases of active inflammation are marked by the sharp, ragged texture, whereas quiescent phases are marked by a smoother, but still porotic texture. As more alveolar bone is lost, the surrounding trabecular bone and buccal and lingual cortical plates are also remodeled. This creates a more general reduction in the height of the alveolar process and, depending on how many teeth are involved, may create a relatively regular outline (“horizontal bone loss”) or a strongly irregular one. Karn et al. (1984) devised a classification to describe irregular defects. Alveolar bone is not necessarily lost equally around all sides of the roots, so they distinguished a “crater” that involved just one side, a “trench” that involved two or three sides (Fig. 10.6), and a “moat” that involved all four sides. When both alveolar bone and outer cortical bone plates were lost, they distinguished a “ramp” with margins at different levels and a “plane” with margins all at a similar level. Thus, irregular defects may be used to suggest the presence of periodontal disease, particularly the crater, trench, and moat forms. Diagnosis of periodontal disease, however, is difficult even in
the living, because no one feature is pathognomonic (uniquely diagnostic). In dry bone specimens, there is also the difficulty of distinguishing the remodeling of the alveolar process, which relates to tooth wear, occlusal forces, and continuous eruption and has nothing to do with periodontal disease. Where a large proportion of the alveolar process has been lost, it is additionally difficult to distinguish the defects of periapical inflammation.

When dental caries, attrition, and fracturing expose the pulp, micro-organisms from the mouth enter and cause an inflammation of the pulp or pulpitis (Dias and Tayles, 1997; Soames and Southam, 2005). Sometimes this causes short bouts of sharp pain and is described as acute pulpitis, or it may cause longer bouts of dull pain and be described as chronic pulpitis. Some patients, however, have no history of tooth/jaw pain. Indeed, the symptoms experienced apparently bear little relationship to the size or type of lesion in the pulp. Several defense mechanisms come into play, including deposition of secondary dentine, and the rate of progression of the lesion varies between individuals or even between different roots of a single molar. It is possible for the inflammation to resolve, but with its narrow connection to the vascular system through the root apex, the ability of the pulp to heal is limited and the usual end result is that it dies, right down the root canal to the apical foramen. It allows the products of inflammation, bacterial toxins, or perhaps bacteria themselves, to emerge from the foramen and initiate an inflammatory response in the tissues around the apex of the root (periapical inflammation). The response may cause severe pain from even the lightest pressure on the tooth and be described as acute periapical periodontitis. In such cases, inflammatory exudate has accumulated in the periodontal ligament and, being noncompressible, may cause the tooth to protrude slightly. In some cases, pus is formed, and the lesion becomes an acute abscess—even more painful. The pus drains through spaces in the bone, usually without causing changes detectable by radiography, and the inflammation resolves. Sometimes there is an area of diffuse bone loss (Goaz et al., 1999) which is reversed once the acute phase subsides.

Acute periapical lesions may pass into a more chronic phase of inflammation (or alternate with it), in which granulation tissue is produced around the apex of the root (a periapical granuloma). Chronic periapical periodontitis of this kind is usually symptomless, although there are often intervals when the tooth is sensitive to pressure. It is also possible for pulpitis to progress into a chronic periapical lesion without going through an acute phase. As the small mass of granulation tissue grows, bone is resorbed around the apex of the root, to create a smooth-walled chamber up to 3 mm or so in diameter (Dias and Tayles, 1997). This is detected in dental X-rays as a radiolucency, which may be the first sign of the lesion. In dry bone specimens, such chambers may not be visible on the surface, but if remodeling has made the buccal plate of the alveolar process thin enough, they are often observed through a small opening—in effect a fenestration (above; Figs 10.4 and 10.6). This can be enlarged in archaeological or museum specimens because the edges of such openings are thin and fragile. A periapical granuloma may remain in this state for some considerable time without the patient being aware of it. It may, however, develop in two ways that make it more noticeable. One of these ways is the development of an apical periodontal cyst (radicular cyst). It is a replacement of the granulation tissue by a fluid and results in progressive growth of the bone cavity. Differential diagnosis in dry bone specimens is only possible from the size of the cavity because, in both granuloma and cyst, the lining of the cavity is smooth. Most granulomata are small, less than 2 or 3 mm in diameter, and most cysts are larger than that although many remain less than 5 mm (Goaz et al., 1999; Wood and Goaz, 1997). Looking at dry bone specimens, Dias and Tayles (1997) identified such cavities as a granuloma when the diameter (measured from the side of the root) was 3 mm or less and as a cyst when larger. Radiology textbooks (Goaz et al., 1999;
Wood and Goaz, 1997) use a 15 mm diameter as the dividing line for interpreting X-rays. The outer “shell” of a cyst may be very thin and can bulge outside the surface of the alveolar process, but it is frequently broken through in dry bone specimens. Periodontal (radicular) cysts are by far the most frequent type of cyst and are almost always associated with permanent teeth of adults, especially the upper anterior teeth. They may remain in the jaw after the tooth around which they were initiated has been extracted. It should be borne in mind, however, that a variety of other cysts with very different origins may be found (Soames and Southam, 2005).

The other way in which a periapical granuloma may develop is through the accumulation of pus (a periapical abscess). In acute cases, this produces symptoms of pain, swelling of the gums, and sensitivity of the tooth to pressure, but here the pus usually migrates through the spaces in the alveolar process without leaving a clear track, and the lesion usually heals once the pus has discharged. Chronic cases result in an intermittent discharge of pus through a well-defined channel (commonly called a sinus or fistula) in the buccal plate of the alveolar process or, in some cases, through the lingual plate, or into the nose or maxillary sinus. Apart from this noxious discharge, a chronic periapical abscess is usually pain free, and accounts of recent aboriginal Australians suggest that they might well have been common among older people in the absence of dental treatment. In dry bone specimens, an abscess of this type can be difficult to distinguish from a granuloma, because the sinus takes the path of least resistance, which is often through the much-thinned buccal plate of the alveolar process. Where there is an opening of this kind, there may be no particular grounds for suggesting that it is an abscess sinus rather than the “window” of a granuloma exposed by remodeling, unless it there is a clear tube-like hole with substantial walls instead of an edge of irregular, wafer-like remnants of thinned cortical bone. At the apical end of the sinus in a chronic abscess, there may be a small cavity with a rough wall (Dias and Tayles, 1997). Slightly roughened walls in a better defined cavity may be an acute abscess that developed in a granuloma or cyst. Larger areas of bone loss within a jaw (usually the cheek tooth area of the mandible), possibly associated with a sequestrum of dead bone, multiple sinuses or cloacae, and subperiosteal new bone formation (involutrum) at the surface, are defined as osteomyelitis. They may originate in a periapical infection, or a compound fracture, and they are distinguished from a localized periapical abscess by their involvement of the marrow spaces of the bone.

A periapical granuloma may be associated with hypercementosis (above). Even when not associated with inflammation of this kind, hypercementosis is accompanied by loss of bone around the apex of the root to accommodate its increased bulk. This may create a fenestration, but a granuloma is distinguished by the cavity, which leaves a gap between the root and the alveolar bone of the socket.

Chronic periapical inflammation causes bone loss mostly at the apex but, unless this area of loss becomes very large or is located at the side of the root (see below), leaves the crest of the alveolar process intact. Mechanically, therefore, it usually has less effect on the stability of a tooth than does periodontal disease. Loosening of teeth is not commonly associated with gross caries, pulpitis, or periapical inflammation, and bone loss around the apex is unlikely to lead on its own to tooth loss. There is a widespread idea that dental caries might be a major cause of antemortem tooth loss in ancient populations, but it is hard to see how this might work unless the teeth were extracted deliberately to treat acute pulpitis or periapical bone loss became combined with loss from the crest of the alveolar process from periodontal disease (and see below). It is also possible that migration of the tooth sockets through the jaw, as part of the process of continuous eruption (above), might bring a small zone of periapical bone loss closer to the alveolar crest, and more general remodeling around it
might seriously weaken support for the tooth. Another point of potential confusion is that the apical foramen may not be the only opening of a root canal. Some teeth have so-called lateral canals that emerge as a tiny foramen part way down the side of a root, and it is perfectly possible for a granuloma to be initiated around one of these without involving the root apex. In such cases, the cavity of the granuloma is much closer to the crest of the alveolar process than one at the apex and may become exposed by remodeling to contribute to the loss of tooth support. Clarke (1990; Clarke and Hirsch, 1991) has suggested that such “alveolar defects of pulpal origin” have been underestimated as a potential cause of tooth loss.

The main requirement for a comprehensive study of periapical inflammation is appropriate X-rays (Goaz et al., 1999; Whaites, 2006). Although many archaeological specimens display ample surface evidence, some cases with apparent pulp exposure also show no such signs of periapical inflammation and may well have a granuloma deep inside. It can also be difficult to examine the walls of a chamber inside the alveolar process, even when there is an aperture open to the outside. Bright lights and a low-power stereomicroscope help, and it may be possible to withdraw the tooth slightly from the socket. No appropriate recording schemes have been defined, and it would seem best to make notes, with detailed photographs and X-rays, including the evidence for pulp exposure (caries cavity, wear facet, or tooth fracture), size of radiolucency or visible chamber (radius measured from the side of the root), the texture of its wall, nature of any opening (bearing in mind that it may have been widened by postmortem damage), and the presence of any local resorption or new bone formation around it. Other important details might include general remodeling of the alveolar process that might tend to expose granulomata or evidence for alveolar bone loss relating to periodontal disease. Where the height of the process has been reduced considerably, it may be very difficult to determine the most important contributing cause.

The Place of Dental Paleopathology in Archaeology

Few archaeologists would even dream about justifying a cemetery excavation on the basis of dental paleopathology. This author is probably biased, but maybe they should. At a site where the artifactual finds are unexceptional and the bones fragmentary, the dental pathology may be just about the most exciting finding. Teeth have a variable pattern of recovery in excavations and survival in storage, however, and this may strongly affect the results. When a jaw is lifted from the ground, it needs to be checked by a knowledgeable eye to make sure that the full surviving set of teeth is lifted with it. Sieving the soil from around the jaws (to a relatively fine mesh) is another good idea, and any washing should automatically take place over a 1-mm sieve. Teeth do not always look particularly like teeth, especially when they are not fully formed, heavily worn, or with a gross carious lesion removing most of the crown. In such states they can also be difficult to identify and side, even for a specialist. Errors at this point may have an important effect on the results of a paleopathological study. Cleaning really needs to be fairly gentle, with the minimum of soaking and no rough scrubbing; and teeth need to be allowed to dry slowly or they may split along natural lines of weakness. In a similar vein, the air in most storage facilities is a little too dry for teeth and there is progressive damage over the years (the recommended relative humidity is 50–65%), particularly affecting dentine (the problem is even more intense for elephant and hippopotamus tusks). The anterior teeth fall out easily and are lost, so it is as well to bag any loose teeth up separately within the box (pierce small holes in plastic bags to keep the air moving through) and label them before they disappear. The traditional skull box holds the lower jaw in occlusion, under the skull, and has probably contributed more to the damage of important dental pathology specimens than any other factor because the tooth crowns are crushed and chipped. It is better
to bag or wrap the lower jaw separately within the box or, at the very least, provide a pad between the jaws. Crumpled-up acid-free tissue paper is a good way to stop loose bones rattling about.

Providing that care is taken, so that the surviving specimens of teeth and jaws really do represent what was preserved in the ground, then the distribution of dental disease is a valuable source of evidence for the activities and subsistence of a past population. In a society with limited artifactual technology, the dentition is an important piece of equipment, in everyday use for making objects and handling materials, as well as for preparing food and eating it. Even emotional state can have an effect—today abnormally heavy wear is common among stressed-out young businessmen and women who habitually hold their jaws clamped together. The advent of arable agriculture and the type of staple crop grown, its consistency, and carbohydrate and protein content all leave their mark as a clear horizon in the frequency and type of dental caries and periodontal disease as well as the pattern of tooth wear. Dental pathology should have a central place in many discussions of change in subsistence and technology. It fits well with evidence derived from the analysis of stable isotopes; artifactual evidence for the gathering, production, and processing of food; and the remains of food animals and plants.

**SUGGESTED SCORING SCHEMES FOR CARIES AND PERIODONTAL DISEASE**

The following detailed scheme was initially proposed in Hillson (2001), as an extension of the standard methods recommended elsewhere (Buikstra and Ubelaker, 1994). Experience in the field and museum has shown that it is better to record pathology on paper, rather than directly into a database, because this is more secure (allowing for dust, mistakes, theft, power supply fluctuations, etc.) and enables nonstandard features to be drawn and described in words. Some kind of recording sheet helps. It is, however, sensible to transfer the scores and measurements to a computer database regularly; otherwise, the task becomes overwhelming. It is straightforward to design a Microsoft Excel worksheet, or Access form (Microsoft Corporation, Redmond, WA), or SPSS data table (SPSS Inc., Chicago, IL) that has the appropriate fields. The database can have a field for each of the categories below, so that each tooth has one record that is identified by a code for the individual and an FDI code for the tooth (Fédération, 1971; Hillson, 1996). This makes it possible to extract information either per tooth type or per individual.

The scores in the list below are numbered to match across different categories, so some of the numbers are missing in particular categories where the scores do not match in this way (e.g., presence/absence of tooth below). To the list below, it will be necessary to add general information such as an identification code, aging and sexing data. It is far better to record all the original data, such as measurements, pubic symphysis scores, and so on, rather than a simple diagnosis of male/female, and age. The age/sex diagnosis is vital for interpreting the pathological data, but it is a matter of opinion and should be capable of being questioned.

**Score for Gross Gross Caries**

BLANK = missing postmortem and jaw with socket missing too

0 = tooth present, without gross gross caries

7 = gross gross carious cavity, involving the loss of so much of the tooth that it is not possible to determine whether the lesion was initiated in the crown or root

8 = gross gross carious cavity, defined as in 7 above, in which there is a clear opening into an exposed pulp chamber or root canal

10 = tooth missing, leaving an empty socket in the jaw without any sign of remodeling (postmortem tooth loss)
11 = tooth missing, leaving an empty cavity in which there are signs of remodelling, but the bone is not fully remodeled to a level contour
12 = tooth missing, with full remodeling of the jaw to leave a level contour
13 = no evidence that the tooth has ever erupted (because of young age, impaction, or agenesis)
14 = tooth partly erupted (crypt communicating with crest of alveolar process, or tooth not yet in wear)
15 = anomalous eruption, so that the tooth has not reached its normal position in the tooth row

**SUPRAGINGIVAL CALCULUS**

Brothwell (1981) score. Where the level of the gingival margin has been lowered, there may be supragingival calculus on the root as well as the crown, and Brothwell’s maximum score of 3 can be applied to this.

**OCCLUSAL SURFACE CARIES IN PREMOLARS AND MOLARS**

Fissure system, groove, and fossa sites in the occlusal surface. Count the whole occlusal fissure system of each premolar or molar as one site, when any part of it remains and can be seen unobscured. Score the most developed lesion, if there is more than one.

BLANK = sites missing for any reason, or fully obscured
0 = sites present but enamel is translucent and with a smooth surface
1 = white or stained opaque area in enamel of fissure/groove/fossa with smooth glossy or matte surface
2 = white or stained opaque area with associated roughening or slight surface destruction
3 = small cavity where there is no clear evidence that it penetrates to the dentine
5 = larger cavity that clearly penetrates the dentine
6 = large cavity that was clearly initiated in a fissure/groove/fossa site within the occlusal surface (it does not involve the contact areas), within the floor of which is the open pulp chamber, or open root canals
7 = gross coronal caries involving the occlusal crown surface and a contact area or pit
8 = gross coronal caries, defined as in score 7 above, within the floor of which is the open pulp chamber, or open root canals

**CARIES IN PIT SITES OF MOLARS AND UPPER INCISORS**

Score each discrete pit present. Not all dentitions have them, but there is often one buccal pit on molars and sometimes a lingual pit tucked in above the lingual tubercle of upper incisors (rarely canines). It would be uncommon for there to be more than one pit site per tooth, but it may happen.

BLANK = pit site not present or not visible (for any reason)
0 = site or sites present but enamel is translucent and with a smooth surface
1 = white or stained opaque area in enamel of pit with smooth glossy or matte surface
2 = white or stained opaque area with associated roughening or slight surface destruction
3 = small cavity where there is no clear evidence that it penetrates to the dentine
5 = larger cavity that clearly penetrates the dentine
6 = large cavity that was clearly initiated in a pit site, within the floor of which is the open pulp chamber, or open root canals

7 = gross coronal caries involving a pit and the occlusal crown surface (occlusal surface caries scores 6 and 7 above)

8 = gross coronal caries, defined as in score 7 above, within the floor of which is the open pulp chamber, or open root canals

**OCCLUSAL ATRITION SCORE**

The Smith (1984) system is simplest to use.

BLANK = occlusal surface not present, or obscured, for any reason
1 to 8 = Smith attrition stage
10 = tooth fractured, leaving a surface that shows some wear

**OCCLUSAL ATRITION FACET DENTINE CARIES AND PULP EXPOSURE**

Count whole facet as one site, and record the most severe lesion if there is more than one.

BLANK = worn dentine surface either not yet exposed, missing, or obscured (for whatever reason)
0 = dentine exposed in occlusal attrition facet but without any stained areas, or cavitation
4 = stained area of dentine and/or enamel that may or may not be a carious lesion
5 = clear cavity in dentine
6 = pulp chamber, exposed in the attrition facet, that is stained or appears to have been modified by the development of a cavity

8 = exposed pulp chamber in which there is no sign of either staining or irregular formation of a cavity

**OCCLUSAL ATRITION FACET ENAMEL EDGE CHIPPING AND CARIES**

This category may only be important in some collections—particularly hunter-gatherer groups. Count the enamel rim of the attrition facet as one site, because there may be many small areas of chipping around it.

BLANK = worn enamel rim not yet exposed at any point on the perimeter of the occlusal surface, missing, or obscured (for whatever reason)
0 = enamel rim of occlusal attrition facet exposed at any point, but intact with no chipping
1 = chipping that appears to be postmortem in origin
2 = chipping that appears to be antemortem but is not affected by caries
3 = chipping associated with carious lesion
7 = gross carious lesion (scores 7 or 8 in occlusal surface caries, occlusal attrition facet caries, pit caries, contact area caries) involving the enamel rim of the occlusal facet but not clearly associated with any chipping
8 = gross carious lesion as defined in score 7 above, involving the enamel rim, within the floor of which is the open pulp chamber, or open root canals

**APPROXIMAL ATRITION FACET MEASUREMENT**

In the occlusal plane, the maximum buccolingual breadth of the facet (measured in millimetres with needle-point calipers). See
Hinton (1982). Enter separate measurement for mesial and distal sides of each tooth.

**MESIAL AND DISTAL APPROXIMAL ATTRITION SCORE**

Enter separate score for mesial and distal sides of each tooth.

BLANK = contact point missing (for whatever reason)
0 = no attrition facet around contact point
1 = approximal attrition facet confined to the enamel
2 = approximal attrition facet exposing dentine at its center
3 = approximal attrition facet exposes dentine all the way down to the CEJ
4 = occlusal attrition has proceeded down into the roots of the teeth so that there is no longer any contact between neighboring teeth

4 = discoloration in exposed dentine of an approximal attrition facet
5 = larger enamel cavity that clearly penetrates the dentine (or clear cavity in dentine of approximal attrition facet)
6 = large cavity, clearly initiated in the contact area or approximal attrition facet, within the floor of which is the open pulp chamber, or open root canals
7 = gross cavity in the the contact area or approximal attrition facet that involves neighboring occlusal sites (occlusal surface caries, occlusal attrition facet caries above) and/or root surface sites (below)
8 = gross cavity, defined as in score 7 above, within the floor of which is the open pulp chamber, or open root canals

**MESIAL AND DISTAL CONTACT AREA CARIES**

Enter a separate score for mesial and distal sides of each tooth.

BLANK = contact area missing, or not visible, for any reason
0 = contact area present but enamel is translucent and with a smooth surface (and any exposed dentine is unstained and not cavitated)
1 = white or stained opaque area in enamel with smooth glossy or matte surface (or stained patch in dentine)
2 = white or stained opaque area of enamel with associated roughening or slight surface destruction
3 = small enamel cavity where there is no clear evidence that it penetrates to the dentine

4 = discoloration in exposed dentine of an approximal attrition facet
5 = larger enamel cavity that clearly penetrates the dentine (or clear cavity in dentine of approximal attrition facet)
6 = large cavity, clearly initiated in the contact area or approximal attrition facet, within the floor of which is the open pulp chamber, or open root canals
7 = gross cavity in the the contact area or approximal attrition facet that involves neighboring occlusal sites (occlusal surface caries, occlusal attrition facet caries above) and/or root surface sites (below)
8 = gross cavity, defined as in score 7 above, within the floor of which is the open pulp chamber, or open root canals

**BUCCAL/LABIAL AND LINGUAL ENAMEL SMOOTH-SURFACE CARIES SITE**

One site, just above the margin of the gingivae in life. Count as present only when it is clearly separate from cement–enamel junction (CEJ). Only score if the lesion clearly does not involve the CEJ, the fissure system, a pit, or any worn occlusal attrition facet. Rare in archaeological material.

BLANK = site not present or not visible (for any reason)
0 = site present but enamel is translucent and with a smooth surface
1 = white or stained opaque area in enamel with smooth glossy or matte surface
2 = white or stained opaque area with associated roughening or slight destruction of the enamel surface
3 = small enamel cavity where there is no clear evidence that it penetrates to the dentine
4 = larger cavity that clearly penetrates the dentine
6 = large cavity that has exposed the open pulp chamber, still without involving the cement–enamel junction
7 = gross cavity that involves neighboring occlusal sites and/or root surface sites
8 = gross cavity, defined as in score 7 above, within the floor of which is the open pulp chamber, or open root canals

**ROOT SURFACE CARIES**

Count one site per buccal/labial, lingual, mesial or distal surface of each tooth, and score separately for each. The site may run into other root surface sites.

BLANK = no part of root surface or CEJ preserved, or at least not visible if present
0 = root surface/CEJ present and visible, with no evidence of staining or cavitation
1 = area of darker staining along CEJ or on root surface
5 = shallow cavity (stained or unstained) following the line of the CEJ, or confined to the surface of the root
6 = cavity involving the CEJ or root surface alone, within the floor of which is the open pulp chamber, or open root canals
7 = gross cavity, including the CEJ or root surface, which involves the neighboring contact area site, occlusal sites, or occlusal attrition facet sites (above)
8 = gross cavity, defined as in score 7 above, within the floor of which is the open pulp chamber, or open root canals

**SUBGINGIVAL CALCULUS ON BUCCAL/LABIAL, LINGUAL, MESIAL, AND DISTAL SIDES**

Maximum extent (measured to the nearest millimeter) of subgingival calculus deposit, down the long axis of the root from the CEJ, separately for each side of the root. There is usually a calculus-free space along the CEJ, but this is irregular and it is simplest to ignore it.

**GENERAL REMODELLING OF THE BUCCAL AND LINGUAL PLATES**

There are no published standards for describing the thinning of the buccal and lingual cortical plates of the alveolar process, which forms part of the process of remodeling for continuous eruption and occlusal adaptation. The changes are usually more common in the buccal plate than in the lingual plate.

**Thinning of the Buccal Plate**

The main surface manifestation of buccal plate thinning is that pronounced grooves develop on the surface of the alveolar process, between
the roots. In advanced cases, the whole outline of the roots can be clearly made out through the thin “skin” of bone. Some roots are in any case particularly prominent, most notably the upper canines, followed by the upper incisors, but the level of prominence becomes much greater. At the same time, the bone around the cervix of the tooth shows wafer thin against the root, rather than making a stout edge.

Dehiscence
A dehiscence is a “V”-shaped opening in the alveolar process, extending down a root from the cervix of the tooth. It may be deep or shallow or narrow or wide (Clarke and Hirsch, 1991; Muller and Perizonius, 1980). The margins are usually wafer thin. Care must be taken to distinguish it from postmortem damage (look for a sharp edge with a different color) and from irregular defects of periodontal disease (look for involvement of the alveolar bone in the socket either side).

Fenestration
A fenestration is similar to a dehiscence, but it is a circumscribed opening, further down the root (Clarke and Hirsch, 1991; Muller and Perizonius, 1980). It may expose a granuloma, and it is important to distinguish it from the sinus of an abscess (below).

Hypercementosis
Hypercementosis is common in heavy wear populations, where the alveolar process is usually strongly remodeled in later life. The irregular swelling of the apical half of the roots is accommodated by resorption, and they may be exposed in a fenestration or dehiscence. Hypercementosis can also be detected in X-rays as a swollen and irregular root.

Marginal Lipping
Many archaeological collections of jaws show a ridge along the buccal margin of the crest of the alveolar process, in the cheek tooth region of the mandible. This may become sharp and prominent. It is accentuated by the alveolar bone loss of periodontal disease, but it is thought to be more an adaptation to changing occlusion and the functional loads placed on the jaw (Clarke and Hirsch, 1991; Newman and Levers, 1979).

PERIAPICAL BONE LOSS
Various descriptive terms can be used, and it is important to distinguish between the different ways in which an area of bone loss is exposed to view. Both direct observation and X-ray interpretation are ideally used.

Visible Periapical Cavity
Defined as a cavity around the apex of a root or roots in multirooted teeth. Occasionally seen at the side of a root, where there is a lateral canal and foramen.

Visibility
- Visible through an opening in the thinned buccal or lingual wall of the alveolar process (that is, not a clear sinus)
- Seen through a sinus
- Seen on extracting the tooth from its socket
- Seen through a postmortem break in the bone (look for a line of different colored bone along the break)

Location
- Most granulomata are at the apex of one or more roots of a tooth
- Some occur in relation to a lateral canal part way up the root
- Others may be up toward the cervical part of the root (Clarke and Hirsch, 1991)

Size
- Less than 3 mm diameter (make allowance for root apex protruding into cavity)
• More than 3 mm diameter
• More than 15 mm diameter

Wall of Cavity
• Smooth and well demarcated
• Slightly roughened
• Clearly roughened, with a ragged margin

Sinus. Defined as a well-demarcated tube or opening, with substantial walls, which is clearly cutting through the bone cortex rather than representing a thinning of the cortical plate—most apertures around the apex are probably not a sinus but a fenestration. Care needs to be taken to recognize an opening made or enlarged by postmortem damage. A sinus may have new bone deposition around its point of emergence, which can be

• Through the buccal/labial or lingual plate of the alveolar process
• Through the angle of the palatine process of the maxilla into the floor of the nose
• Into the maxillary sinus
• Onto the buccal or lingual aspect of the mandibular body

Interpretation. Following the diagnostic criteria of Dias and Tayles (1997), a periapical granuloma is a small cavity (<3 mm) with smooth walls, a cyst is a larger cavity of the same type, an acute abscess developing in a granuloma or cyst shows as a slight roughening of the wall, and a chronic abscess is a small cavity with rougher more ill-defined walls and a sinus, whereas a larger cavity of this type is identified as chronic osteomyelitis (it may also have multiple cloacae, sequestra, and involucrum).

Periapical Radiolucency
The following terminology follows Whaites (2006) as well as Goaz et al. (1999). Care needs to be taken, because the size and prominence of the radiolucency depends to some extent on the projection of the image and on the presence of other spaces such as the maxillary sinus.

Normal. The root and apex are outlined by a narrow radiolucent band that represents the space occupied in life by the periodontal ligament. This may be narrower in archaeological specimens because the ligament is no longer in position. Outside the ligament space is a thin, clearly demarcated line of radio-opacity (called the lamina dura) marking the position of the alveolar bone lining the socket. Between adjacent sockets, and within the body of the alveolar process, an irregular, dense granular texture of radio-opacity is created by the trabeculae (particularly large trabeculae may sometimes be visible individually, especially in the mandible).

Ligament Space. In living patients, the edema of acute inflammation within the ligament causes the space to increase as the tooth protrudes out of the socket. Clearly, in archaeological specimens, this is not a useful consideration.

Lamina Dura. As bone is resorbed, there is local loss of the radio-opaque line of the lamina, and this may be apparent without a wider area of radiolucency.

Radiolucency. Within the trabecular texture surrounding the lamina dura, there may be a radiolucency of varying size and shape. It may be circular to pear-shaped, and centered on the apex of a root (occasionally the side), or more widespread. It may be diffuse, without clear margins, or sharply defined with a clear radio-opaque line around the margin. In dental radiography, the diagnostic size cut-off point is 15 mm in diameter (below).

Sclerosis. Thickening of trabeculae and deposition of isolated areas of denser bone are seen as patches of radio-opacity. Sometimes,
individual trabeculae may be made out, and the area of opacity as a whole may be diffuse, or clearly defined. It may occur right next to the lamina dura of the apex or root side or around the edge of a radiolucency. Occasionally, sclerotic patches occur adjacent to teeth that do not have any evidence of pulp exposure—the cause is unknown.

**Secondary Dentine.** There is no difference in opacity between primary and secondary dentine, but the pulp chamber/root canal size is reduced, starting at its roof (infilling of the horns or diverticles), followed by more general reduction in width. The chamber/canal may appear to be infilled completely, but there may still be a thread of living pulp running down the root, so it cannot be assumed necessarily that the tooth is dead.

**Interpretation.** Circular or pear-shaped radiolucencies centered on a root apex are interpreted as a granuloma or a cyst. In radiographic terms, a diagnosis of a radicular cyst is made only when the radiolucency is larger than 15 mm in diameter, even though it is recognized that most are probably smaller than 5 mm. Cysts are also more likely than granulomata to have a well-defined, thin radio-opaque border. Sclerosis is taken to be evidence of chronicity of the infection. A periapical abscess may not give rise to any radiographic features, or there may be a diffuse radiolucency around the root apex. Osteomyelitis (involving the marrow spaces of the bone) is almost always confined to the cheek tooth area of the mandible and is seen as a much larger, diffuse, ragged radiolucency of “moth-eaten” appearance.

**DEFORMITIES OF THE ALVEOLAR PROCESS**


**APPROXIMAL WALL DEFECT SCORE**


**PRESENTATION OF CARIES RATES**

Caries percentages should be calculated separately for each lesion site (or combination of sites in the case of gross caries), and for each tooth type (left and right sides can be combined). Males and females, and different age groups, need to be tabulated separately. The tables are big, because separate percentages also need to be presented for opacities/stains, cavities, dentine penetration, and pulp exposure. A minimum list of categories could then be as follows:

1. Occlusal caries as a percentage of fissure and fossa occlusal sites surviving.
2. Occlusal attrition facet caries as a percentage of facets present.
3. Pit caries as a percentage of pit sites preserved.
4. Smooth-surface caries as a percentage of sites preserved (combine buccal and lingual figures).
5. Contact area caries as a percentage of contact area sites surviving (combine mesial and distal figures).
6. Root surface caries as a percentage of root surface sites preserved (combine mesial, distal, buccal, and lingual figures). Tabulate root surface exposure separately.
7. Gross occlusal caries as a percentage of occlusal crown surfaces surviving (worn or unworn—one per tooth).
8. Gross occlusal/contact area caries as a percentage of the number of mesial or distal crown elements surviving (two per tooth).
9. Gross root surface/contact area caries as a percentage of the number of mesial or
distal tooth elements surviving (two per tooth).

10. Gross root surface/contact area/occlusal caries as a percentage of the number of mesial or distal tooth elements surviving (two per tooth).

11. Gross gross caries as a percentage of teeth surviving.

Caries is usually symmetrical, so left and right teeth can be included together, but that still means that separate percentages for all these categories of lesion sites are needed for both permanent and deciduous, upper and lower first and second incisors; canines; third (first) and fourth (second) premolars; and first, second, and third molars—a total of 26 separate percentages in all. An archaeological assemblage also needs to be divided into age and sex groups, because of the age-progressive nature of dental disease. Ideally this would be done with nondental aging criteria so, for example, the dental data for adults could be tabulated against Suchey–Brooks pubic symphysis stages (Brooks and Suchey, 1990). Where this is not possible, it might be necessary to tabulate disease rates against dental development and occlusal attrition stages. So long as this is made explicit (i.e., they are not converted into estimates of ages in years), it is at least clear what is the basis of comparison. At a minimum, the development stages might for example be as follows:

1. Children with a deciduous dentition only
2. Children with a mixed dentition
3. Children or subadults with permanent dentition only, but third molar not in occlusion yet
4. Young adults (males counted separately from females) with third molar in occlusion but first molar occlusal attrition less than Smith stage 5
5. Older adults (males separated from females) exceeding this stage of wear

REFERENCES


Goodman AH, Thomas RB, Swedlund AC, Armelagos GJ. 1988. Biocultural perspectives of stress in prehistoric, historical and contemporar-
REFERENCES


CHAPTER 11

ANALYSIS AND INTERPRETATION OF SKELETAL TRAUMA

NANCY C. LOVELL

INTRODUCTION

One of the most commonly observed pathological conditions in skeletal remains, trauma, has been recognized by paleopathologists since the origins of the discipline for its value in reconstructing the past lifeways of humans and other organisms. Indeed, traumatic lesions have been identified in the remains of a variety of life forms, including dinosaurs (Rothschild and Martin, 1993), fossilized plants (Moodie, 1923), and non-human primates (Bramblett, 1967; Jurmain, 1989; Lovell, 1990). Aside from its anatomical importance, evidence of trauma provides information about an organism’s or population’s interaction with its physical and sociocultural environment. The objectives of trauma analysis, therefore, include the identification and description of lesions in both individual specimens (i.e., case studies) and populations; the interpretation of social, cultural, or environmental causes of trauma; the relationship of trauma to age and sex (which may, of course, also have social or cultural relevance); and temporal and geographic patterning. As such, paleopathological research on trauma includes studies of inter- and intra-group conflict (e.g., Jurmain, 1991; Liston and Baker, 1996; Martin and Frayer, 1997; Owsley and Jantz, 1994; Walker, 1989), environmentally or occupationally related accident or misadventure (e.g., Cybulski, 1992; Grauer and Roberts, 1996; Judd and Roberts, 1999; Kilgore et al., 1997), and the experience of trauma among juveniles (e.g., Blondiaux et al., 2002; Glenncross and Stuart-Macadam, 2000; Jimenez-Brobeil et al., 2007; Walker et al., 1997).

Trauma refers to injury to living tissue that is caused by a force or mechanism extrinsic to the body, whether incidental or intentional. The word “trauma” is also conventionally used as a synonym for force, however, and interpretation of the type of force relies on characteristics of the fracture. Blunt force trauma, for example, is caused by a blunt object such as the ground, a club, or a piece of furniture (comprehensive coverage of blunt force trauma as exhibited in skeletal remains can be found in several excellent forensic anthropology texts, especially Galloway (1999), which includes the description of the many and varied eponymous fractures, i.e., those fractures named after the person, usually a physician, who described or gave the name to a fracture). Sharp force trauma is a penetrating injury and refers to chopping, stabbing, slashing, or incising wounds inflicted by a sharp object.
Most fractures and dislocations are incidental injuries, resulting from an accident, whereas surgical practices and injuries caused by weapons are the result of intent. These are interpretive categories, however, that can only be reached after a process of trauma description and analysis. Proper description of a lesion is the first step in trauma analysis and is the basis for determining the mechanism of injury, which is either direct or indirect trauma. In turn, an understanding of the mechanism of injury is crucial for the identification of the cause of trauma. Proper description of observed lesions also provides other scholars with an opportunity to agree or disagree with the diagnosis and/or inferences that are made about the sociocultural or environmental context of the trauma. Trauma description should improve the accuracy and reliability of interpretation without exceeding the limits of inference that are set by the descriptive data themselves. The principal aim of most protocols for trauma analysis (e.g., Buikstra and Ubelaker, 1994; Lovell, 1997) has been to establish standardized descriptions for lesions observed in dry bone, although additional objectives, such as the evaluation of evidence for treatment of traumatic injuries are sometimes also stated (e.g., Grauer and Roberts, 1996; Neri and Lancelloti, 2004). It is likely that no one system of trauma description will suit all investigators since some will be more or less concerned with the affected body part, specific complications, or possible causative behaviors. Most protocols share similar basic categories of description, but it should be recognized that one element of description, that is, inventory of the observable skeletal elements and the parts thereof, varies by researcher and may influence the interpretation of patterns of trauma (Judd, 2002).

Regardless of the protocol selected for the examination of trauma in skeletal remains, the primary objective is a good description of the lesion and documentation by photographs and radiographs where possible. The proper analysis and interpretation of trauma in skeletal remains originates with the classification of injuries according to their predominate characteristic, that is, (1) the ossification of soft tissues, (2) extrinsically induced abnormal shape or contour of a bone, (3) dislocation (the displacement of one or more bones at a joint), and 4) fracture (any break in the continuity of a bone). Extensive photographic illustration of paleopathological examples of these conditions can be found in Ortner (2003).

**EXPLANATORY BOX 11.1. METHODS**

Visual observation is usually the first method of analysis that is used when examining archaeological remains for traumatic lesions. In many cases, it may be the only method required, whereas in some circumstances, it may be the only method available. Radiography has proven to be a valuable adjunct to visual observation because it can, in essence, see “inside” bodies and bodily structures such as bone. Radiography is a transmission imaging technique: Radiation in the form of X-ray energy photons is emitted from a source and transmitted through an object, such as a bone. The object attenuates the X-rays through processes of absorption and scatter and produces a pattern of transmitted X-rays that is imaged by a detector. Projection radiography produces an image that is life-sized, appears as a “negative” on conventional radiographic film, and superimposes the three-dimensional structure of the object onto a two-dimensional film, so that information about depth is lost. The technique is nondestructive and inflicts no damage on an inanimate object such as archaeological human remains.

The need for radiographic evaluation of fractures to determine the amount of healing and the particulars of deformity and/or displacement often is stressed, as is the importance of radiography for detecting and interpreting well-remodeled fractures (e.g., Glencross and Stuart-Macadam,
The value and application of standardized fracture descriptions is illustrated by description of radiographs from four clinical cases (Lovell, 1997); with known mechanisms of injury, treatment, and follow-up, these cases unambiguously illustrate the skeletal effects of trauma and their variability of expression.

Unfortunately, radiographic equipment is not always available, especially in field settings, and the interpretation of radiographs may be made difficult by postmortem alterations common in archaeological contexts, such as soil inclusions that affect density or the differential identification of osteoporosis versus diagenetic bone loss (Roberts, 1991).

Another noninvasive method of analysis is computed tomography (CT) scanning, which accomplishes three-dimensional imaging by measuring the simultaneous transmission of X-rays through an object in different directions. The resulting data are computed to construct a cross-sectional image in electronic form, which can be used as digital information or converted to a pictorial display. Rühli et al. (2002) have demonstrated the value of multislice CT (MSCT) in the differential diagnosis of cranial lesions and the potential of this method for providing relatively small scale estimates of the duration of healing in cases of perimortem trauma. Although there are various types of CT systems in use today, the method is time-consuming and expensive, and logistical issues make the method untenable for field application. The application of CT scanning in cases of questionable lesions may, however, make destructive analysis unnecessary (Rühli et al., 2002).

Microscopic analysis, including transmitted light, stereobinocular, and scanning electron microscopy, provides a clearer picture of cellular activities in response to trauma than does visual observation or radiography. The use of microscopic analysis to clarify, confirm, or refute the nature of lesions that have been identified macroscopically is of tremendous importance in paleopathology and is particularly helpful in identifying postmortem alterations to bone (Schultz, 2001, 2003). The most common microscopic technique used by paleopathologists is transmitted light microscopy. It is a destructive technique that uses thin sections of bone mounted on glass slides, onto which light is directed from below. The examination of bone thin sections usually ranges from 20× to 200× magnification. At this power the primary changes in bone that can be detected include the differentiation between woven and lamellar bone and an increase and/or decrease in mineralized bone mass, all of which may be important in evaluating evidence of trauma and trauma healing. Polarized light (obtained by adding a polarizer and special analyzer to the microscope) is generally considered more useful than ordinary transmitted light for obtaining a detailed picture of cell and tissue structure in transparent specimens and is considered to be superior to MSCT for diagnostic accuracy (Schultz, 2003). Stereobinocular or dissecting microscopes use reflected light and have a relatively great depth of field, which enables them to be used in the examination of uneven surfaces of bones. The technique is essentially nondestructive in that the bone does not need to be sectioned, although in some cases, the structural components of the bone can be more easily observed if it is first sectioned and the section is then polished and chemically etched. These microscopes usually have a range of from 40× to 400× magnification. Scanning electron microscopy (SEM) resembles reflected light microscopy in that it is possible to examine the surface features of bone microstructures with three-dimensional resolution, but with the use of electrons a higher magnification (up to about 40,000×), it is possible without loss of detail. Features such as microfractures and the evidence of cellular destruction and repair of fractures may be readily imaged by SEM (Schultz, 2001, 2003; Wakely, 1993). Unfortunately this is usually a destructive technique when applied to bone in that only small samples of about 1 cm² can be accommodated in the apparatus and a nonremovable coating must be applied to nonconductive materials such as bone.
Furthermore, the sample is placed under vacuum, which may damage porous bone. It is therefore recommended that a replica of the specimen be examined if there is risk of damage or if the specimen is too large to be scanned.

The most recent technological innovations that show promise for the imaging of lesions are three-dimensional scanning for modeling objects and three-dimensional printing from computer (CT) models. Researchers at the University of Manitoba (http://home.cc.umanitoba.ca/~hoppard/BDIAL/index.html) have used the technology successfully to reconstruct the faces of wrapped mummies and to reproduce archaeological artifacts, and there are numerous other applications to cultural heritage (e.g., Koller et al., 2006; Landon and Seales, 2006; Levoy et al., 2000) and medicine (e.g., Chen and Zq, 2005; Lauer et al., 2002; Majid et al., 2005).

OSSIFICATION OF SOFT TISSUES

Extra-skeletal ossification that occurs in muscles and other soft tissues is referred to as heterotopic ossification. The most common form is myositis ossificans traumatica (MOT), which results from a muscle tear or blunt trauma that results in a bruise in the muscle that ossifies. The ossification may appear as a smooth but irregularly shaped mass of mature bone that is attached to the surface of a skeletal element and that may be mistaken for a cancerous tumor (DiMaio and Francis, 2001; Ortner, 2003), although MOT is far more common, and hence a more likely diagnosis (Ortner, 2003). Alternatively an isolated, calcified mass can be produced (Hendifar et al., 2005) that in an archaeological case likely would be unidentifiable unless recognized in situ. The quadriceps femoris muscle is most often affected clinically, usually as a result of a contact sports injury (King, 1998), followed by the muscles of the buttocks and upper arm (Carlson and Klassen, 1984; DiMaio and Francis, 2001; Hendifar et al., 2005). MOT also can occur as a complication of joint injury, where the tendons and ligaments that attach to bones and joint capsules ossify and leave characteristic projections around the joint that are recognizable in dry bone, although similar lesions may be from chronic and hard muscle use (Ortner, 2003); the ankle is the most common site of ossified ligamentous attachments caused by acute joint injury. Myositis ossificans progressiva is a hereditary condition that occurs without injury and in a predictable pattern (it will not be discussed in this chapter). Heterotopic ossification is also documented in cases of paraplegia because of spinal cord trauma (Ortner, 2003) and thus may serve as a diagnostic tool as well as a tool for personal identification (e.g., DiMaio and Francis, 2001).

EXTRINSICALLY INDUCED ABNORMAL SHAPE OR CONTOUR

Abnormal shape or contour is usually classified as an incomplete fracture and may be incidental or intentional in origin. For example, intentional cultural modifications of the body, such as corseting, head binding, and foot binding, cause deformities without a break in the bone. These “pressure fractures” result when the bone responds to the application of direct force. Unintentional modification also may be from cultural practices, such as the deformity of an infant’s cranium that is referred to as positional plagiocephaly and that results from lying in a persistent sleeping position. In addition, exaggerated curvatures of long bones may be the result of incomplete fractures, particularly if the trauma occurred in the juvenile skeleton in which bones are still pliable and may bend rather than break (Ogden, 2000; Stuart-Macadam et al., 1998). Bowing deformity typically is from traumatic
longitudinal compression and can be identified histologically by a series of oblique microfractures (Schultz, 2003). As illustrated in Fig. 11.1, the two common classifications of incomplete fractures in children are the torus fracture, a buckling or localized bulging of bone cortex from longitudinal compression/impaction that is usually observed at the junction of the metaphysis and diaphysis, and the greenstick fracture, an incomplete transverse fracture in a bone that has been subjected to bending or angulation forces (Ogden, 2000). A greenstick fracture of the clavicle can occur during childbirth when the infant’s bi-acromial breadth is too large to pass easily through the mother’s pelvic outlet. Greenstick fractures also are referred to as plastic bowing deformities (“plastic” referring to the fact the bones are pliable and can be deformed permanently without breaking or tearing). They are clinically common forearm injuries in children from the force of a fall onto an outstretched hand; the younger the child the more likely the bowing can occur (Ogden, 2000), but the bowing usually will correct itself (Jones, 2003).

DISLOCATION

Traumatic injuries to joints may result in partial or complete dislocations. A dislocation, or luxation, occurs when the articular surfaces of a joint are totally displaced from one another. A partial dislocation, or subluxation, results when the articular surfaces are partially displaced but retain some contact. Dislocations are most commonly caused by trauma and often are associated with a fracture. The glenohumeral joint is very susceptible to displacement because of its shallowness. By contrast, traumatic dislocation of the femoral head from the acetabulum requires considerable force, such as experienced in motor vehicle accidents, and thus when observed in archaeological human remains, it is more commonly caused by congenital dislocation. Dislocations tend to be more frequent in young and middle-aged adults, since a similar force in juveniles instead causes epiphyseal separation and in older adults causes fracture of fragile bones.

Joint displacement cannot occur without damage to the joint capsule and ligaments and hence may be accompanied by complications such as ossification of membrane, ligament, and tendon attachments (i.e., myositis ossificans traumatica). Persistent instability of the joint may result from a dislocation, particularly in the shoulder and ankle, although this cannot be identified easily in dry bone. Osteoarthritis is one of the more common and recognizable complications and results from damage to the articular cartilage itself or from prolonged incongruence of the joint surfaces. For dislocations to be recognizable in dry bone, the injury must have occurred some time before the death of an individual and displacement must have persisted long enough for bone modifications to take place (e.g., Miles, 2000). Some dislocations, such as of the digits, often can be put back quickly and easily into place (but see Dreier, 1992), whereas others, such as of the vertebrae, may cause immediate death. In either of these cases, no evidence of the injury will be observable in archaeological skeletons.

**Figure 11.1** The two common types of traumatic injury in children, the greenstick or longitudinal fracture (left) and the torus or buckling fracture (right).
A fracture consists of an incomplete or complete break in the continuity of a bone. In addition to being complete or incomplete, all fractures can be classified as intra-articular (involving a joint) or extra-articular (not involving a joint) and as open (where there is an open wound between the skin surface and the fracture; also known as a compound fracture) or closed (also known as a simple fracture). Fractures can be caused by either direct or indirect force. Table 11.1 summarizes fracture types and their mechanisms of injury. A break at the point of impact is referred to as direct trauma injury and results in transverse, penetrating, comminuted, or crush fractures (Fig. 11.2). For example, crush fractures result from direct force to cancellous bone, which collapses, such as when an object is dropped on the hand or foot. A break in a place other than the point of impact is said to represent indirect trauma (Fig. 11.3). Several types of fractures typically affect tubular bones and result from indirect trauma: an oblique fracture line angles across the longitudinal axis and is confined to one plane, whereas a spiral fracture line winds down around a long bone shaft, like a corkscrew, because of a twisting stress on the longitudinal axis. An impacted fracture results when the bone ends at a fracture site are driven into each other by the force of injury. An avulsion fracture is caused when a joint capsule, ligament, or

**TABLE 11.1 Fracture Types and Mechanisms of Injury (adapted from Lovell, 1997)**

<table>
<thead>
<tr>
<th>Type of Fracture</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Penetrating</td>
<td>Partial or complete penetration of bone cortex</td>
</tr>
<tr>
<td>2. Comminuted</td>
<td>Bone is broken in more than two pieces; most common in long bone diaphyses</td>
</tr>
</tbody>
</table>
| 3. Crush         | Most common in cancellous bone  
| a. Depression    | Crushing force on one side of the bone  
| b. Compression   | Crushing force on both sides  
| c. Pressure      | Force applied to growing bone |
| 4. Transverse    | Force applied in a line perpendicular to long axis of the bone |
| 5. Spiral        | Rotational stress on long axis; often confused with oblique fracture |
| 6. Oblique       | Angular stress on long axis; often confused with spiral fracture |
| 7. Torus         | Buckling of the bone from longitudinal compression (impaction); common in children |
| 8. Greenstick    | Incomplete fracture from angular or rotational stress; common in children |
| 9. Impacted      | Bone ends are driven into each other |
| 10. Burst        | Found in the spine from vertical compression |
| 11. Comminuted   | Force splits in several directions and forms a T or Y shape |
| 12. Avulsion     | Fracture from tension at ligament or tendon attachment |
| 13. Stress/fatigue | Because of repetitive force, usually perpendicular to long axis |
| 14. Secondary/ pathological | Secondary to localized or systemic disease that has weakened the bone |

Types 1 through 4 are caused by direct trauma; types 5 through 14 are caused by indirect trauma.
tendon pulls away from its attachment to the bone, tearing a piece of bone with it (Fig. 11.4).

Additional classification of fractures relates to the degree of fragmentation that occurs. A comminuted fracture is where the bone has been broken, splintered, or crushed into several pieces. A segmental fracture refers to fracture of a bone in two places. Repetitive force can result in a stress or fatigue fracture, which commonly is visible as a nondisplaced line or crack in the bone, called a hairline fracture; the usual areas of occurrence are the necks of the second and third metatarsals, the calcaneus, and tibia (Early, 2006; Teague, 2006; Wilson and Katz, 1969).

Any of these types of fractures can occur secondarily to a disease already present in the body. Systemic diseases such as metabolic disturbances and nutritional deficiencies leave bone vulnerable to spontaneous fracture or to fracture from minor trauma. For example, fractures in past populations have occurred over underlying lesions of leprosy (Judd and Roberts, 1998) and are caused by preexisting osteoporosis (a condition characterized by a decrease in bone mass producing porosity and brittleness) and osteomalacia (a condition characterized by softening of the bones) (Brickley, 2006); neoplastic fractures are observed when the break is through or adjacent to a cancerous tumor that is in, or of, bone; and the collapse of vertebral bodies is not an uncommon consequence of bone destruction caused by tuberculosis in the spine (Pott’s disease).

**EXPLANATORY BOX 11.2. DESCRIPTION OF FRACTURES**

To adequately describe a specimen and its lesion/s, consider the following as examples of the types of information you would consider under ideal conditions (e.g., well-preserved bone):

1. What bone is it? If a paired element, what side? Adult or juvenile? Male or female?
2. Where on the bone is the lesion located? Is a joint involved? Is an identifiable anatomical feature involved? Where are the fracture lines located (e.g., proximal end or proximal third of the shaft)?

3. What does the lesion look like? Is there evidence for shortening? Lengthening? Angular deformity? (Comparison with the contralateral element for paired bones is a help here.) How big is the lesion? (Provide length and breadth or diameter measurements, taken with calipers, if appropriate.)

4. Can you venture an opinion as to the type of injury (e.g., transverse or oblique)?

5. Can you venture an interpretation as to the biomechanics of injury (e.g., compression or torsion)?

6. Is there any evidence for a predisposing (i.e., preexisting) condition?

7. Is there any evidence for complications resulting from the injury (e.g., arthritis or infection)?

8. Is the fracture callus represented by immature, woven bone or by mature, well-remodeled bone? Can you hazard a guess as to the minimum length of time that has elapsed since the injury?

9. Are there any other similar lesions in the skeleton that may indicate multiple injuries from one traumatic incident?

Much debate exists over how common are different fractures (Court-Brown and Caesar, 2006), and many reviews of fractures emphasize contemporary clinical practice in industrialized countries without consideration of fracture data for other parts of the world. In addition, data presented for clinical contexts usually are based on hospital admissions after serious injury and/or self-reporting of traumatic events to a medical practitioner (e.g., Court-Brown and Caesar, 2006; Koval and Cooley, 2006). Thus, in addition to problems inherent in the nature of an archaeological sample, a direct correlation of contemporary clinical data with paleopathological data is questionable. However, the descriptions below take clinical data that are most likely comparable with past circumstances (for example, ignoring trauma frequencies related to motor vehicle accidents and high-velocity gunshot wounds, and cautioning when frequencies appear related to advanced ages that are not well documented in past populations) and provide a framework within which the interpretation of past trauma can be situated (for comprehensive descriptions of fractures to different bones of the skeleton, the reader is referred to Adams and Hamblen, 1992; Beaty and Kasser, 2006a; Browner et al., 2003; Bucholz et al., 2006; Galloway, 1999; Green and Swiontkowski, 2003; and Ogden, 2000). Although vehicular trauma is the most prevalent mechanism of injury according to clinical data in the United States, falls constitute the next most common cause of injury (Koval and Cooley, 2006), which has clear relevance for the interpretation of trauma in the past.

EXPLANATORY BOX 11.3. PALEOEPIDEMIOLOGY

Determining the pattern of trauma within a skeletal sample, and comparing patterns of trauma between sexes or among age groups or different populations, relies on the statistical analysis and presentation of trauma data (for a detailed discussion, the reader is referred to Waldron, 1994).

Paleoepidemiology is the study of the factors affecting the distribution of health and illness of past populations, with the population, rather than the individual, being the unit of study. Thus, data
obtained for all individuals or bones in a skeletal sample (which is treated as a population) are combined for the total sample or for subgroups (e.g., age) within the sample. The confidence with which patterns can be ascertained and compared depends on several factors, including an assessment of the size of the original, living population (usually unknown); the proportion of that population that was likely buried; the proportion of the buried population that survived burial; the proportion of the surviving burial population that was excavated; and the proportion of the excavated population that was examined (i.e., the skeletal sample). Rarely can all of these factors be determined with any degree of certainty, but for paleopathologists, the principal concern would be whether it is reasonable to assume that the studied sample is representative of the once-living population. Fracture types and frequencies in a skeletal sample from a battleground, for example, would not be considered representative of any living population other than the individuals involved in the battle.

Although the skeletons in a cemetery may represent deaths occurring over decades or even hundreds or thousands of years, skeletal data are treated as if it were cross-sectional in that all individuals are treated as if they died at the same time. Measures of disease frequency in living populations are usually presented as incidence data, which indicates the number of new cases of a disease in a specified population over a certain period of time. Thus, an outbreak of influenza may be reported by health authorities as the number of new cases per 1000 persons in the month of January, in a certain location or age group. Typically outbreaks such as influenza are tracked by the number of new cases that are reported by school districts, and “flu season” may be identified by a sudden increase in new cases from one week to the next. Because paleopathologists do not have that kind of temporal control, frequencies of pathological conditions are reported as prevalence, not as incidence, data. Prevalence refers to the frequency of a condition in a skeletal sample and is presented as percentages. Thus, it may be reported that the prevalence of femoral shaft fractures in a skeletal sample is 3%. However, this crude prevalence rate needs additional elaboration. How was the 3% calculated? If three left femoral fractures were observed in a skeletal sample of 100 observable individuals, the prevalence would be 3%. That is considered the “individual count” rate. But if fractures of two left femora and one right femora were found in two individuals, then only 2 individuals out of 100 would have been affected and the individual count rate would be 2%. Similarly, if the skeletal sample included both femora from all 100 individuals, the total number of observable femora would be 200. Then the “bone count” (or “element count”) rate of femoral fractures would be 3 out of 200, or 1.5%. Prevalence (%) is thus calculated as

\[ n/N \times 100, \]

where \( n \) is the number of affected individuals or elements and \( N \) is the number of observable individuals or elements.\(^1\) Prevalence can also be calculated and reported as age-, sex-, or status-specific frequencies as required.

Most paleopathologists consider the element count to be most appropriate, but this also assumes that all elements were complete and equally observable for fractures, which would not be the case if there were extensive postmortem damage to the skeletal remains. Thus, it may be more informative when studying trauma to calculate frequencies based on the preservation and observation of parts of elements, such as diaphysis, proximal third, and distal third. (Judd, 2002).

Prevalence or frequency measures are considered descriptive statistics. By contrast, inferential statistical techniques are used to compare frequencies in samples or populations. Published reports may therefore provide the results of t-tests or chi-squared tests to indicate whether observed
differences in averages or in proportions are statistically significant. The use of statistics is not always warranted, however, especially when dealing with very small sample sizes, as seems typical for skeletal collections from past populations. Students with interests in paleopathology are well advised to become conversant with basic statistical procedures and assumptions so that they are capable of critiquing the published literature as well as of conducting their own analyses.  

When scoring arthritis, the “joint count” may be an appropriate calculation and the analysis of dental disease may rely on a “tooth count” measure of frequency.

**SKULL FRACTURES**

Skull fractures are fractures of the cranial vault and skull base (facial fractures are considered separately), and these typically result from significant force to the head. Brain injury should always be considered as a possible complication of skull fracture. The biomechanics of skull fractures depend on several variables, including the surface area of the force, the velocity of the force, the location of impact, and the age of the individual. With respect to age, bone is very elastic in young children and can be deformed without fracturing. Since the structurally weak areas of the cranium are most prone to develop fracture lines, unfused cranial sutures in children will readily separate to accommodate the forces of impact. Alternatively, in very young children, the cranial bones may bend inward without fracturing and the depressed deformity may persist. Diastasis (the separation of normally joined parts) in adults with fused sutures, however, is a sign of significant trauma.

The most common fractures of the bones of the cranium affect the vault and are caused by direct trauma (Fig. 11.5). These fractures can be described according to their basic type: linear, depressed, or penetrating, or a combination of these. Simple linear fractures are fractures that follow one linear pattern. A single linear fracture line indicates less force than does a pattern of multiple or complex fracture lines. For example, skull fractures from a fall to a flat surface generally show an impact site with one or several linear fracture lines radiating from the point of impact. With low-velocity blunt trauma, the curve of the skull at the point of impact flattens at impact, distributing the force over a relatively large area and leading to the outward bending of the bone surrounding the area of impact. If deformation is great enough, fracture lines begin in the bone that is subjected to outward bending. Linear fracture lines tend to sweep around the thick, bony buttresses of the cranium (i.e., the petrous bones or the mastoid process) unless they approach these areas perpendicularly. The degree and direction to which these lines extend depend also on the size of the force applied: In cases of blunt trauma, the force is on the outside of the cranium and concentric fractures are associated with beveling of endocranium, whereas in

![Figure 11.5](image_url) Typical cranial vault fractures, from left to right: linear, depressed, and penetrating.
high-velocity projectile trauma (e.g., from a bullet), the force is caused by pressure from within the cranium and so the beveling occurs on the ectocranium. (Adequate coverage of gunshot trauma is beyond the scope of this chapter, but the interested reader is referred to Galloway, 1999; Gellman et al., 2003; and Koval and Zuckerman, 2006.)

Blunt force to the cranium usually causes linear fractures of the vault, and the appearance of these fracture lines may help identify the point of impact. Two specific types of linear fractures are longitudinal and transverse linear fractures. The former are front-to-back skull fractures often caused by blunt force trauma to the head around the face, frontal, or occipital regions, but these also can be caused by compression of the head from front to back. Transverse fractures are side-to-side fractures of the skull caused by impact or compression to the sides of the head. In children, linear fractures of the calvarium are uncommon because of the flexibility of the bone (Duncan, 1993). Identification of the point of impact and the direction of the force becomes increasingly difficult with more severe trauma, but the sequence of multiple impacts usually can be determined since a subsequently produced fracture will not cross a preexisting one. Blunt trauma to the frontal bone produces fracture lines that radiate through the frontal sinus, the cribriform plate, and the orbital roofs, although transverse fracture lines affecting the temporal regions may also appear. Anterior temporal impact leads to fracture lines that radiate down, across either the orbital plate or the sphenoid–temporal region. In contrast, lateral or posterior temporal impact produces fracture lines that radiate downward either in front of or behind the petrous portion of the temporal bone and extend across the cranial base. Impact to the occipital bone usually produces fracture lines that radiate down to the foramen magnum or the jugular foramen, and that may extend anteriorly across the cranial base. Trauma to the cranial base must be severe to cause a fracture, since the bone here is buttressed heavily. After vault fractures, sphenoid fractures are the most common clinical result of blunt trauma to the cranium (Unger et al., 1990). Unfortunately, sphenoidal structures are very fragile and thus prone to postmortem damage as well as to fatal consequences of fracture, and therefore, it may be difficult to identify sphenoid fractures in archaeological skeletons.

Depressed fractures are those in which bone fragments are pushed inward. The depressed area indicates the point of impact, from which linear fractures radiate; these are usually irregular and may be comminuted, producing a cobweb or mosaic pattern. Depression fractures of the cranial vault are commonly observed in archaeological human remains and are caused by low-velocity direct trauma (Fig. 11.6). Generally these fractures result when a person falls on a sharp corner or on a flat surface with a raised object on it, but they also may result from assaults with weapons such as clubs or other objects that may be used in a fight (Galloway, 1999). Rarely, depressed cranial fractures can be caused by a punch to the head (Fenton et al., 2003). In cases of blunt force assault, the depressed fracture is most often located over the frontoparietal region because this location is readily accessible to the assailant (Walker, 1997). In children, this type of fracture

Figure 11.6 Depressed cranial fracture.
commonly results from a fall but may also represent a birthing injury. Lesser force is indicated by the lack of displacement of bone fragments, whereas greater force is characterized by inward displacement. A portion of bone might be detached completely if great force is applied, particularly if the object has a small striking surface, but more often observed in archaeological remains is the incomplete detachment of the bone.

Penetrating injuries of the cranium are characterized by a small area of impact with a localized area of distortion and usually are caused by sharp-edged objects such as knives or swords or by projectiles. This distortion includes the partial or complete penetration of bone cortex by cutting, piercing, drilling, or scraping for the excision of pieces of cranial vault bones in the practice of trepanation. Heavy cutting-edged weapons that are used in a chopping manner will produce crush injuries in addition to penetration, and more injury may be caused if the embedded weapon is removed with a twisting motion. This damage is often indicated by splintering of the bone with outward displacement near the initial impact site. The type and size of wound produced by a projectile depends on the size of the projectile, the speed at which it strikes the bone, and the distance it travels. Historical skeletons may exhibit evidence of gunshot trauma (Bailey and Mitchell, 2007; Willey and Scott, 1996), although these injuries would be less severe in terms of bone fragmentation and destruction than those typically found in a metropolitan trauma center today. Typically, with a higher velocity of impact, the stresses are more localized but the depth of penetration is greater.

Injuries to the cranium from indirect impact are relatively rare, but they may result from vertical loading forces transmitted from below when a person falls from a height. A basilar “ring” fracture around the foramen magnum is an example of such an injury; it reflects impact forces transmitted up through the cervical spine and occipital condyles. Basilar fractures through the petrous bones have been observed to result from impact to the chin (Harvey and Jones, 1980). Skull base fractures typically are described according to the fossa/e (i.e., anterior, middle, or posterior) that are affected.

Possible complications of cranial fractures include displacement of bone fragments, indirect trauma injuries elsewhere on the cranium from the transmission of impact force, and soft-tissue damage. The location of the impact determines the subsequent consequences of the injury from the different anatomical structures in the cranium. Linear fractures usually involve both the inner and the outer tables of the cranium but do not involve displacement or depression of the bone and thus are often not considered as serious as those injuries resulting from greater force. Complications can develop, however, from damage to a nerve or transmission of the force of impact, such as when direct trauma to the back of the cranium produces indirect effects on the orbital plates. The consequences of cranial fractures can be fatal if the blood vessels running along the inner tables of the cranium (e.g., anterior and middle meningeal arteries) are torn, although this complication is unlikely to be detected with certainty in archaeological remains. Because of the proximity of the paranasal sinuses and mastoid air cells to the dura mater (the outermost of the three layers of tissue covering the brain), basilar fractures are usually compound, and hence, post-traumatic infection is possible. Penetrating injuries are open fractures with an attendant risk of infection, but the most immediate complications relate to damage to the dura mater. Penetrating trauma caused by a sword, axe, or machete is usually associated with dural laceration.

**FACIAL FRACTURES**

For the purposes of clinical description, the maxillofacial region is divided into four parts: upper face (frontal bone and sinus), upper
midface (nasal, nasoethmoidal, zygomatico-maxillary complex; orbital floor), lower midface, and mandible. Facial fractures may be unilateral or bilateral and are commonly caused by assault, including domestic violence and abuse of children and the elderly. Recreational activities and industrial accidents are also known causes of facial fractures in clinical settings, as are motor vehicle accidents.

Because of their density and buttressing, the mandible and the frontal bone (particularly the supraorbital rim) require a high-impact force to be damaged; by contrast, the zygomas and nasal bones require only a low-impact force. Facial fractures, whether from direct or indirect trauma, are often very complex but typically heal adequately without medical treatment.

René Le Fort, a French army surgeon working at the turn of the twentieth century, described three classic patterns of facial fracture after studying what happened when he inflicted low-velocity blows to cadaver crania (Le Fort, 1901). Simply stated, the Le Fort I fracture is a mid-face fracture that separates the palate from the maxilla; the Le Fort II is a pyramidal fracture through the sinus wall laterally and the nasal bones medially, separating the maxilla from the face; and the Le Fort III fracture is through the frontozygomatic sutures and orbits, detaching the facial skeleton from the base of the skull (Fig. 11.7). Since the zygoma, maxilla, and orbital margin are mutually supportive, a fracture of one of these usually involves a fracture of at least one other. Although not reported widely in the paleopathological literature, archaeological examples of various Le Fort fractures are illustrated in Figure 2.4 of Gregg and Gregg (1987).

Fracture of the mandible frequently occurs in fistfights and can occur from a fall or similar accident. The mandible is a ring structure, and thus, any fracture on one side may be accompanied by a fracture on the other (Alexandersen, 1967). For example, a fracture of the angle of the mandible may be accompanied by a fracture to the opposite condyle. Indirect trauma can fracture the neck of the mandibular condyles if direct trauma impact

![Figure 11.7 Le Fort fractures of the face.](image-url)
occurs to the chin because the condyle is the weaker part of the bone (Harvey and Jones, 1980). Mandibular fractures in infants typically affect the symphysis, whereas in older children the most common site is the condyle. In adults, mandibular fractures resulting from a punch tend to appear at the mandibular angle. Fracture lines in the mandible may communicate with the roots of one or more teeth, which makes the fracture a compound fracture and introduces the risk of infection from oral bacteria. Asymmetrical tooth wear and osteoarthritis at the temporomandibular joint are possible complications of jaw fractures and ankylosis of the joint after blunt trauma has been reported clinically (Ferretti et al., 2005). Other complications of facial fractures include injuries to the sensory organs and vital structures in the head and neck. Personal and social consequences of disfigurement also may be important.

Uncommonly discussed as trauma in the paleopathological literature are injuries to teeth, but these should be rigorously assessed, especially in cases of suspected interpersonal violence. Lukacs (2007) has described tooth fractures in 11 specimens from the Canary Islands in which the fracture occurred at the alveolar margin, reparative root growth maintained root integrity and inhibited infection, and proliferative alveolar bone growth formed over the broken root. The sequestration of the root can be identified radiographically, and thus, it may be possible to identify dental fracture up until the time that the root is completely resorbed. Depending on the cultural and environmental context, traumatic injuries to teeth may be caused incidentally by accidental falls, interpersonal violence, or the use of teeth in food preparation or as tools (Alexandersen, 1967; Andraesen, 1982; Lukacs, 2007; Lukacs and Hemphill, 1990; Milner and Larsen, 1991; Patterson, 1984; Pindborg, 1970) or intentionally through cultural modification involving the removal of part or all of a tooth (Cook, 1981; Milner and Larsen, 1991; Pietrusewsky and Douglas, 1993). Fractures to teeth also may result indirectly through direct trauma to the mandible.

FRACTURES OF FLAT AND IRREGULAR BONES

Hyoid

The hyoid bone is not always recovered during archaeological excavations, but it can be a useful indicator of strangulation or hanging (Ubelaker, 1992) if displaying a perimortem fracture (that is, a fracture that was sustained at or about the time of death). Hyoid fractures are more likely observed in cases of strangulation of older individuals, in whom the synchondroses have fused (Pollanen and Chiasson, 1996; Ubelaker, 1992). Isolated hyoid fractures are reported rarely in the clinical literature, with a recent review noting only 31 cases; most resulted from traffic accidents, but three resulted from assault and were accompanied by other injuries (such as mandibular fractures), three resulted from accidental direct blows, and two were caused by falls onto an object (Dalati, 2005). A case resulting from indirect trauma has been attributed to self-induced vomiting (Gupta et al., 1995).

Vertebrae, Sacrum, and Coccyx

Direct trauma injuries to the vertebrae are rare, but fractures of the transverse processes can result from direct blows, either from interpersonal conflict or from falls, and hairline transverse fractures on, or posterior to, the superior articular processes of the second cervical vertebra are worth noting since they can result from strangulation (Maples, 1986). Cervical trauma is also common in cases of hanging (Spence et al., 1999; Ubelaker, 1992). Isolated sacral fractures also are rare but can be caused by direct blows to the sacroiliac region or to a fall onto the buttocks (Boachie-Adjei, 1992; Kane, 1984; Levine, 1992). Fractures of the coccyx occur when an
individual falls onto the buttocks or lands in a sitting position, and are more common in females, whose pelvic morphology leaves the coccyx more exposed to injury (Kane, 1984).

The most common fractures of the vertebrae and sacrum are from repeated physical stress, indirect trauma, or preexisting disease. A very distinctive vertebral fracture for which there may be a genetic predisposition (Merbs, 1996) is the traumatic separation of the neural arch from the vertebral body known as spondylolysis. Spondylolysis may be unilateral or bilateral in expression, but it predominates in the lumbar-sacral region, particularly L5 and, to a lesser degree, L4 (Merbs, 1996). Although complete separation is observed most often in clinical settings, the condition apparently begins as incomplete stress fractures in adolescents that may either heal or progress to complete lysis (Merbs, 1995). The condition may be initiated by an acute overload event that causes microfractures, but it is generally agreed that the determining factor is chronic trauma, with repeated stress promoting nonunion of the microfractures. Clinically these fractures are most often observed among athletes and laborers whose activities involve frequent and large stress reversals between lumbar hyperextension and lumbar flexion, and in past populations could be attributed to paddling, wrestling, hauling large stones in building projects, and so on (Arraza, 1997; Merbs, 1989, 1996). “Clay-shoveller’s fracture” is a historically documented traumatic separation of the tip of the spinous process of the seventh cervical or first thoracic vertebra and, like spondylolysis, may result from strenuous muscle action, which in this case is associated with shoveling clay, cement, or rocks (Jordana et al., 2006; Knüsel et al., 1996).

More common in archaeological vertebrae than stress fractures are indirect trauma injuries. Sudden vertical compression that ruptures the intervertebral disk though the end plate forces disk tissue into the vertebral body (Fig. 11.8). Schmorl’s nodes are a milder form of this injury and result from bulging of the disk’s nucleus pulposus, which puts chronic pressure on the vertebral end plate and leads to bone resorption in the affected area. Herniation of the disk tends to occur gradually in adults because the nucleus has lost resiliency, whereas it may occur suddenly in younger individuals in whom the nucleus is still gelatinous (Bullough and Boachie-Adjei, 1988). Indirect vertebral damage can be from force applied to the head, but also it can occur in a fall or jump onto the feet, since the force of impact is carried up from the lower limbs through the spine. Crush fractures that result in a wedge appearance to the anterior portion of the vertebral body (Fig. 11.9) occur most commonly in the thoracic and lumbar vertebrae and are from compression associated with flexion, such as when a person falls from a height into a crouched position or falls and lands on the shoulders. Fractures of the sacrum are most commonly from indirect trauma, whereby forces are
transmitted through the spine or the pelvis (Boachie-Adjei, 1992; Levine, 1992).

Fractures secondary to preexisting disease are also common in the vertebral column. The best known clinical example results when intervertebral disks expand into the superior or inferior surfaces of vertebral bodies that have been weakened by osteoporosis (Eastlack and Bono, 2006). Similarly, compression flattening of vertebral end plates from sparse, coarse trabeculation in the vertebral body is a classic feature of sickle cell disease. Because of the greater strength of the rims of the vertebral end plates, they may be spared even when the vertebral body is compressed. Pronounced collapse of vertebral bodies that occurs secondary to destruction by tuberculosis infection is known as Pott’s disease.

Ribs and Sternum

Ribs are known to incur stress fractures as a result of habitual labor (Pavlov and Freiberger, 1978), as a consequence of persistent coughing (De Maesener et al., 2000; George et al., 2000; Hanak et al., 2005; Roberge et al., 1984) especially if there is underlying disease (Lampe, 2007) or, allegedly, from repeated vomiting such as induced in the eating disorder bulimia. Most often, however, rib fractures result from direct trauma, such as a blow or a fall against a hard object (Lampe, 2007). Transverse rib fractures are caused by direct blows to the chest, whereas oblique fractures are caused by crushing or bending (Galloway, 1999). The direction of the impact usually can be determined from the location of the fracture: Ribs are usually fractured near the angle if the force is applied from the front; beside the spine if the force is applied from the back; and beside both the spine and the sternum if the force is applied from the sides. A single blow, however, can cause multiple fractures at various locations (Love and Symes, 2004), and the posterior angle is a common site regardless of the type of impact because it is the structurally weakest part of the rib (Lampe, 2007).

The fourth to eighth ribs are most often fractured (Galloway, 1999; Lampe, 2007). Because they are well protected by the pectoral girdle, fracture of the first to third ribs is usually from a high-energy force (Lampe, 2007) and thus would not often be observed in archaeological human remains. The ninth through twelfth ribs are relatively mobile, and fractures at this location would suggest associated intra-abdominal injury (Lampe, 2007).

Because of the flexibility of the rib cage, particularly in the antero-posterior dimension, the degree of inward displacement at impact may have been much greater than that discernable postinjury, and thus soft-tissue damage may have been more severe than might be inferred from the damage to the bones themselves. Children have a more flexible skeleton, and thus, bony injuries may not be observed despite injury to the underlying heart, lungs, and other organs; incomplete rib fractures also are not uncommon in younger adults who retain good quality cartilage (Love and Symes, 2004).

Possible soft-tissue damage includes laceration of the pleura (the membranous lining of the upper body cavity and covering for the lungs), the lungs, or blood vessels, which would have been largely untreatable in earlier times and thus may point to a possible cause of death. Pneumothorax (the presence of air in the pleural cavity) and hemothorax (the presence of blood in the pleural cavity) may be caused by rib fractures at any level in the thoracic cage and may be similarly life-threatening. Another serious complication of rib fracture occurs when there are at least two breaks in one rib. This fracture produces a free-floating fragment and thus a risk of internal damage. Severe blunt injury to the chest is one of the leading causes of morbidity and mortality among trauma victims clinically, and it may be caused by motor vehicle accidents, falls, and assaults. If the soft-tissue damage caused by rib fractures is serious enough to cause death, the bony injuries observed in the archaeological skeleton would be identifiable only as perimortem fractures. Healed rib fractures in the
archaeological record thus probably represent non-life-threatening trauma. Although rib fractures cause severe pain and varying degrees of disability, people of the past may not have been able to avoid work or domestic tasks and thus there may be evidence of disruption of callus formation and delayed healing in archaeological specimens (Brickley, 2006).

Most sternal fractures in clinical cases result from vehicular accidents, but fractures also can result from falls and assaults, particularly assaults that involve kicking or stomping the chest (Galloway, 1999). Direct blows and crushing injuries to the chest may cause transverse fractures of the sternum, although longitudinal fractures also may occur. Displacement of fracture segments is a possible complication of sternal fractures.

**Clavicle**

Clavicular fractures most often occur at the junction of the middle and lateral thirds and are usually caused by indirect trauma, typically a fall from a standing or moderate height (e.g., from a horse) onto the lateral aspect of the shoulder, when forces are transmitted through the acromion process of the scapula (Court-Brown and Caesar, 2006; Lazarus and Seon, 2006; Wurtz et al., 1992). Downward and medial displacement of the lateral fragment is a common complication. Since clinical treatment of clavicular fractures is usually limited to the use of a sling for one or two weeks for pain relief, healed fractures often exhibit some deformity and thus malaligned clavicular fractures in archaeological skeletons do not necessarily point to an absence of medical treatment. The clavicle is easily broken at birth because of the need for the shoulders to rotate in the birth canal in order for the body to clear the ischial spines, and the fracture is often a greenstick fracture because of the flexibility of the bones.

**Scapula**

Scapular fractures are uncommon but can result from indirect trauma through axial loading onto an outstretched arm or from direct trauma (Butters, 2006; Court-Brown and Caesar, 2006). Both the flat and the irregular portions of the scapula may be involved. The body of the scapula is relatively protected in assaults and falls because it is protected by overlying muscles, but a fall onto the shoulder could cause fractures of the glenoid fossa and/or acromion process. Four types of scapular fractures are commonly described in the clinical literature: fracture of the scapular body, which may be comminuted but rarely displaced because of the large muscles holding the bone in place; fracture of the neck, which may lead to downward displacement of the glenoid; and fractures of the acromion and coracoid processes, although acromioclavicular separation is more common than is fracture of the acromion itself (Butters, 2006). Fractures range from simple cracks to comminution and may be associated with downward displacement. Although none of these injuries is usually considered serious, a possible complication of scapular fracture, especially if the injury occurs on the left side, is pneumothorax (Butters, 2006; Carrero and Wayne, 1989), and the articulating humerus should be examined for evidence of impaired function.

**Pelvis**

Pelvic fractures are relatively uncommon without high-energy force but can result from low-energy falls, falls from heights, and crushing injuries (Court-Brown and Caesar, 2006; Koval and Zuckerman, 2006). Fractures of the pubic ramus are not uncommon in the elderly and typically are caused by a fall from standing height (Court-Brown and Caesar, 2006). More serious pelvic fractures are those that disrupt the pelvic ring through the rami or at the pubic symphysis, with associated dislocation of the sacro-iliac joint. The mechanism of injury in these cases tends to be anterior–posterior crushing, lateral compression, or vertical shearing.

---

2 Acromial fractures can be confused with a developmental defect that prevents fusion of the acromion process to the body, producing an *os acromiale* (Butters, 2006).
force (Koval and Zuckerman, 2006; Starr and Malekzadeh, 2006). Complications are usually serious and would likely be life-threatening in the absence of modern medical treatment. A fracture-dislocation of the hip occurs when the head of the femur is driven through the floor of the acetabulum. This result usually is from a heavy blow on the lateral femur, caused by a serious fall or a similar impact (e.g., vehicular trauma in clinical cases). The injury tends to comminution and serious complications. A poorly understood lesion that may result from the incomplete and temporary dislocation of the hip is the acetabular flange lesion, which appears as a flattening of the supero-posterior rim of the acetabulum (Knowles, 1983). Acetabular fracture can also be from a simple fall in the elderly and may be visible as a crack in the osteoporotic bone (Galloway, 1999). Osteoarthritis is a common sequel to any injury that involves the acetabulum.

**Patella**

Fractures of the patella may be caused by direct or indirect trauma. Indirect trauma is caused by the sudden and violent contraction of the quadriceps muscle of the thigh, which leads to a clean, transverse separation of the bone that is called an avulsion fracture. In contrast, direct trauma from a fall or a blow onto the patella tends to cause a crack fracture or a comminuted fracture. Undisplaced crack fractures usually heal without complication because the fragments are held in position by the aponeurosis (a sheet of connective tissue) of the quadriceps muscle, but fractures that involve separation of the fragments and those that are comminuted will produce an irregular articular surface unless surgically repaired, and osteoarthritis is then an obvious sequel. Patellar fractures are not often described in the paleopathological literature, but a transverse patellar fracture with growth modification is illustrated in Figure 2.12 of Gregg and Gregg (1987). A congenital abnormality called bipartite patella produces a transverse separation of bone and could be confused with a lesion of traumatic origin.

**Hand, Wrist, Foot, and Ankle**

Clinically, the irregular bones most commonly fractured are the scaphoid in the hand and the calcaneus in the foot (Court-Brown and Caesar, 2006; Gaebler, 2006; Koval and Zuckerman, 2006). Scaphoid fractures account for 50–80% of carpal injuries and are second only to the distal radius among fractures of the upper extremity (Gaebler, 2006). Fracture of the scaphoid usually is caused by indirect trauma from a fall onto an outstretched hand with the arm straight and the body weight concentrated across a hyperextended wrist (Gaebler, 2006). The typical injury is a transverse break through the "waist" of the bone and should not be confused with a congenitally bipartite scaphoid. Scaphoid fractures are accompanied by a high frequency of complications, including delayed union, non-union, avascular necrosis, and osteoarthritis (described later in this chapter). Healing of a non-displaced scaphoid fracture in a cast takes approximately 8 to 12 weeks (Gaebler, 2006), so in past populations there would have been associated pain and loss of function.

Almost all calcaneal fractures are caused by a fall from a height (Koval and Zuckerman, 2006). A fall from a tree, cliff, or roof of a dwelling are possible causes of such injury in past populations, although it has been suggested that these injuries would be uncommon in falls of less than 4 m (Wells, 1976). The fracture may be observed as a split or crack in the sub-talar tuberosity, but more often the articular surface of the calcaneus fails to withstand the stress and results in a crush injury. Fracture lines may radiate to the front and appear also on the calcaneo-cuboid joint. Associated crush fractures of lower thoracic or upper lumbar vertebrae or other lower extremity injuries may be noted (Sanders and Clare, 2006). Individuals with calcaneal fractures may be incapacitated for up to five years after injury (Sanders and Clare, 2006).

Metacarpals, metatarsals, and phalanges are common sites of traumatic injury. Metacarpals are often fractured because of longitudinal compression impact such as from boxing (Fig. 11.10). If the fracture line enters the joint,
then osteoarthritis is a likely complication. The neck and distal shaft of metacarpals also are prone to transverse or oblique fractures, often complicated by displacement. Manual phalanges tend to exhibit transverse or spiral fractures of the shaft (Fig. 11.11) or oblique fractures of the base. Comminution is most likely in the distal phalanges. Dislocations of phalanges are mainly from forced hyperextension as may occur in a fall. An archaeological example seems to result from forced adduction of the tip of the fifth finger, with associated tearing of the medial collateral ligament of the joint between the first and second phalanges (Dreier, 1992). Metatarsal fractures are common injuries that usually result from the direct trauma of a heavy object dropped onto the foot although spiral fractures of the middle third metatarsals are caused by indirect trauma where torque is applied when the toes are fixed and the body twists around the foot (Early, 2006). The base of the fifth metatarsal is a common site of avulsion fracture caused by a twisting injury (Court-Brown and Caesar, 2006; Early, 2006). Fractures of the pedal phalanges, particularly those of the great toe, are relatively common and more than half are comminuted crushing injuries from a kick or stubbing injury leading to a spiral or oblique fracture or a direct blow from a heavy object dropped on the foot leading to a transverse or comminuted fracture (Court-Brown and Caesar, 2006; Early, 2006). Fracture of the proximal phalanx of any toe is much more common than fracture of the middle or terminal phalanx (Early, 2006).

LONG BONE FRACTURES

Humerus

The most common humeral fractures in adults affect the proximal humerus (Koval and Zuckerman, 2006). Classification systems for humeral fractures are many and varied, and they are well reviewed by Galloway (1999). Neck fractures are most commonly from indirect trauma from a fall onto an outstretched hand, especially in older women in whom osteoporosis has weakened the bone. Usually
the fracture is self-stabilized through impaction. Direct trauma, in the form of a fall onto the shoulder or a blow, may cause fracture of the greater tuberosity or may drive the humeral head into the glenoid fossa of the scapula. Because of the anatomy of the joint, the humerus is susceptible to anterior dislocation at the shoulder, which may result in compression fracture of the humeral head and rim of the glenoid fossa (Fig. 11.12). Shaft fractures are usually transverse or spiral and are most common in the middle third of the bone; in antiquity they would be most commonly caused by indirect trauma, including simple falls or rotational injuries in older individuals, although direct trauma from a blow can also fracture the shaft (direct trauma from motor vehicle accidents is the most common cause in clinical cases; Koval and Zuckerman, 2006; McKee, 2006). The proximal half of the shaft is a common site of fracture secondary to a preexisting disease, such as when the bone has been invaded by cancer. Distal humerus fractures are uncommon but can be caused by falls from a standing height in elderly individuals with bones weakened by age-related bone loss (Robinson, 2006). Complications of humeral shaft fractures include displacement, nonunion, and injury to the radial nerve.

In contrast to the pattern observed in adults and described above, humeral fractures in children tend to occur at the distal end, affecting the supracondylar, epicondylar, and condylar regions (Beaty and Kasser, 2006b). Supracondylar fractures are indirect trauma injuries caused by a fall onto an outstretched arm, with displacement occurring posteriorly, although some of these fractures are incomplete fractures of the bending or buckling variety.
Epicondylar fractures are usually medial and may result from direct trauma or from avulsion by flexor muscles in a fall. This injury may also cause damage to the ulnar nerve. Condylar fractures are uncommon but also tend to result from a fall. In contrast to the lateral involvement in epicondylar fractures, the lateral portion, or capitulum, is usually involved in condylar fractures. Displacement of distal humerus fractures is not uncommon and may be complicated by deformity, nonunion, and osteoarthritis. Fractures of the distal humerus in children often are associated with forearm fractures or dislocations (Beaty and Kasser, 2006b).

EXPLANATORY BOX 11.4. SPECIAL FEATURES OF TRAUMA IN JUVENILE SKELETONS

The unique part of the juvenile skeleton that has relevance for the analysis and interpretation of trauma is the cartilaginous growth plate (the physis), the region in a long bone between the epiphysis and the diaphysis (i.e., between the primary and the secondary ossification centers) where growth occurs. The growth plate is the weakest area of the growing skeleton, weaker than the nearby ligaments and tendons that connect bones to other bones and muscles. In a growing child, a serious injury to a joint is more likely to damage the growth plate than the ligaments that stabilize the joint. An injury that would cause a sprain in an adult can be associated with a growth plate injury in a child (Koval and Zuckerman, 2006). Growth plate injuries can also result from overuse, so children in antiquity could have developed growth plate injuries from manual labor related to household or agricultural chores or industrial work.

Clinically, 15% to 30% of skeletal trauma in children are physeal injuries, with 71% of these occurring in the upper extremity, mainly in the fingers and distal radius (Koval and Zuckerman, 2006; Rathjen and Birch, 2006). These injuries also occur frequently in the tibia and fibula and less often in the femur or in the ankle, foot, or hip. Growth plate injuries at the knee are at greatest risk of complications because nerve and blood vessel damage occurs most frequently there and can lead to premature growth arrest and crooked growth (Jones, 2003; Xian and Foster, 2006).

Fractures of the growing ends of long bones are classified clinically according to the Salter–Harris system (Fig. 11.13), which distinguishes fractures according to the degree of involvement of the physis, metaphysis, and epiphysis (Salter and Harris, 1963):

Type I: Epiphyseal slip—no fracture, but the epiphyseal plate is completely separated from the end of the bone. The growth plate remains attached to the epiphysis. Unless there is damage to the blood supply to the growth plate, the likelihood that the bone will grow normally is excellent.

Type II: Clinically, this is the most common type of growth plate fracture. The epiphysis, together with the growth plate, is separated from the metaphysis.

Type III: Fracture through epiphysis extending into epiphyseal plate. This fracture occurs only rarely, usually at the lower end of the tibia. It happens when a fracture runs completely through the epiphysis and separates part of the epiphysis and growth plate from the metaphysis. The outlook or prognosis for growth is good if the blood supply to the separated portion of the epiphysis is still intact and if the fracture is not displaced.

Type IV: This fracture runs through the epiphysis, across the growth plate, and into the metaphysis. Surgery is needed to restore the joint surface to normal and to align perfectly the growth plate. Unless perfect alignment is achieved and maintained during healing, prognosis for growth is poor. This injury occurs most commonly at the end of the humerus.
Type V: Crush injury causing obliteration of the growth plate. This uncommon injury occurs when the end of the bone is crushed and the growth plate is compressed. It is most likely to occur at the knee or ankle. Prognosis is poor, since premature stunting of growth is almost inevitable.

The Peterson classification (Peterson, 1994) adds a type VI fracture, in which a portion of the epiphysis, growth plate, and metaphysis is missing. This is a very serious injury and would not likely have been survived in antiquity because of the amount of blood loss; clinically this injury often involves lawn mowers, farm machinery, snowmobiles, or gunshot wounds.

Although juvenile skeletal remains may provide clear evidence of trauma, it is more difficult to recognize the remnants of childhood fractures in adult bones because of the growth and remodeling that transforms immature bones into mature bones: One complete skeletal turnover occurs during a child’s first year of life, declines to about 10% per year in late childhood, and continues at about this rate for life (Jones, 2003). Thus, the younger the age at which the injury occurred, the greater the amount of remodeling that can be expected. However, some unique patterns of injury facilitate the recognition of childhood fractures (reviewed by Glencross and Stuart-Macadam, 2000), including the possibility of accelerated growth of the injured bone (and of surrounding bones) as well as of eccentric overgrowth (Jones, 2003; Xian and Foster, 2006). In particular, the special features of trauma in children described above can be most helpful in identifying childhood injury in adult skeletal remains, since remodeling has little or no effect on displaced fractures that involve the growth plate (Ogden, 2000), and growth arrest lines (transversely oriented trabeculae), if present, would indicate physeal damage by not paralleling the physeal contour (Jones, 2003).

Ulna

Ulnar fractures are not common but can involve the olecranon or the shaft. Olecranon fractures are more common in adults and result from the direct trauma of a fall onto the point of the elbow. Severity ranges from a simple crack to comminution, and the injury may be complicated by nonunion and osteoarthritis. Diaphyseal fractures can result from either direct or indirect trauma. They are prone to severe displacement, malunion, nonunion, and infection because of the bone’s proximity to the surface of the skin. Fracture of the proximal shaft of the ulna is often associated with the dislocation of the radial head and is usually caused by a fall onto an outstretched hand with forced pronation. This injury is referred to as a Monteggia fracture-dislocation, after Giovanni Monteggia, an Italian surgeon, who first described this type of fracture in 1814. An ulnar shaft fracture may also be caused by a blow to the back of the upper forearm (i.e., a “parry” fracture, caused by raising the forearm to fend off a blow). Deformity commonly characterizes the fracture/dislocation injury in archaeological skeletons since the fracture

Figure 11.13  The Salter–Harris classification system for fractures of juvenile long bones.
cannot be properly reduced, positioning the bone fragments in correct alignment (colloquially, “set”), without surgery.

**Radius**

The most common radius fracture, and perhaps the most common fracture overall, occurs at the distal shaft and is called the Colles fracture. (This fracture was first described by an Irish surgeon and anatomist, Abraham Colles, in 1814; hence the name). Clinically, it is the most common of all fractures in adults over the age of 40 years, especially females, and is nearly always caused by the indirect trauma of a fall onto the hand (Court-Brown and Caesar, 2006). The break usually occurs about 2 cm above the distal articular surface of the radius, and the distal fragment is posteriorly displaced and usually impacted (Fig. 11.14). This injury may be associated with fracture of the styloid process of the ulna.

Fractures of the radial shaft are not common, but typically they are caused by a fall onto the pronated hand. Fractures of the radial head and neck are common in adults and children, respectively, when the force of a fall onto the hand is transmitted proximally. The radial head is driven into the capitulum of the humerus, which also may be fractured. Comminution of the radial head is possible, and in such cases, the injury may be complicated by osteoarthritis.

Malunion is the most common complication of radius fractures. Deformity in an archaeological specimen does not necessarily mean that there was no medical treatment since there is a high rate of redisplacement in clinical cases if shaft fractures are not surgically fixed (Hertel and Rothenfluh, 2006). At least one study of archaeological specimens has reported that radius fractures rarely healed without deformity (Grauer and Roberts, 1996). Mays (2006) provides a detailed review of the Colles fracture and notes that long-term residual disability is not uncommon and likely had a significant impact on past populations in which grip strength and hand/wrist mobility were important aspects of physical labor and routine domestic tasks.

**Femur**

Clinical data indicate that 50% of femoral fractures involve the femoral neck (Koval and Zuckerman, 2006), but these are most often observed in the elderly as a consequence of osteoporosis and therefore appear most often among older females because of weight-bearing on thinned bone. In these individuals, the fracture may result from very mild trauma, such as a stumble, or may be spontaneous, with no apparent cause (Karlsson et al., 2006). Neck fractures usually are caused by rotational force and cause lateral rotation and upward displacement of the shaft. Avascular necrosis (death of the bone from a lack of blood supply) is a serious complication of femoral neck fracture that is caused by damage to vessels in the neck that supply the femoral head. The result is that the blood supply to the femoral head may rely too much on vessels in the ligamentum teres, which is only one of three routes of blood supply and is usually inadequate on its own. Bone death is usually sufficiently advanced within two years postinjury that the head collapses (Leighton, 2006). Unless rigidly immobilized, femoral neck fractures are prone to nonunion (Koval and Zuckerman, 2006). Osteoarthritis is another common sequel and may be from mechanical damage at the time of injury, impairment of blood supply to the basal layer of articular cartilage, and/or to malalignment of a united fracture. Impacted abduction fractures of the femoral neck are uncommon; these usually self-stabilize but may result in slight shortening of the limb and/or arthritis.
Fractures of the trochanteric region (i.e., roughly between the greater and lesser trochanters) are common clinically but are not likely to be observed in archaeological populations because they are almost always observed in adults over 75 years of age. Decreased physical activity among older individuals is likely a contributing factor (Nilsson et al., 1991), further limiting their occurrence in past populations who followed traditional, nonmechanized, lifeways.

Femoral head fractures may occur early in life, with traumatic separation of the proximal epiphysis sometimes resulting from a difficult birth (Salter and Harris, 1963). In the absence of surgical repair, such injuries would be clearly recognizable in archaeological human remains, especially if complicated by avascular necrosis.

Severe trauma, whether direct or indirect, usually is needed to fracture the femoral shaft because it is so well protected by the thigh muscles, although shaft fractures may result from a low-energy fall in cases of osteoporosis (Koval and Zuckerman, 2006). Shaft fractures are complicated by simultaneous hip dislocation; arterial damage that can compromise the viability of the limb; nerve damage, especially to the sciatic nerve; and delayed, mal- or nonunion. Three to six months is considered the average healing time in clinical cases of shaft fractures (Nork, 2006). Associated deformity of shaft fractures tends to be shortening or angulation, the latter having a tendency to facilitate development of osteoarthritis at the knee.

Supracondylar fractures of the femur are more or less transverse and located just above the epicondylar region. Malunion is a possible complication if the knee was not immobilized satisfactorily. Condylar fractures are uncommon, but when these do occur, they are almost always from direct trauma. The severity of the fracture can range from an undisplaced crack to complete separation of a condyle with marked upward displacement. In the absence of treatment, a displaced fracture is likely to heal out of alignment and osteoarthritis of the knee joint is a probable consequence. Occasionally a transverse supracondylar fracture combines with a condylar fracture and forms a T-shaped fracture line that splits apart the two condyles.

Tibia and Fibula

Injuries to the knee joint usually involve the menisci and ligaments, not bones, and therefore, no evidence of trauma may be found other than soft-tissue ossification after ligament strain or tear, or avulsion of the tibial spine from injury to the cruciate ligaments. However, the lateral tibial condyle can be fractured by a lateral force against the knee, such as observed clinically in football players and pedestrians struck by a vehicle bumper. If the articular surface is fragmented in the injury, then osteoarthritis is a predictable sequel. When fractures do occur around lower limb joints, they tend to affect the ankle rather than the knee (Court-Brown and Caesar, 2006). Isolated fractures of the malleolus of the tibia or fibula are especially common, and they occur with or without dislocation of the talus. The usual mechanism of injury for these is either abduction and/or lateral rotation for fractures of the lateral malleolus of the fibula and adduction for fractures of the medial malleolus of the tibia. Vertical compression forces can lead to fracture of the anterior margin of the distal tibia or, if severe, fragmentation of the distal tibial articular surface, with the latter injury being prone to osteoarthritis. Fractures and dislocations at the ankle are often complicated by ligament damage, which could be identifiable in archaeological skeletons from soft-tissue ossification.

Most diaphyseal fractures of the leg involve both the tibia and the fibula (Court-Brown, 2006). If the mechanism of injury is an angular force, it will lead to transverse or short oblique fractures of the shafts at roughly the same level. If the injury is from a rotational force, spiral fractures will result and will occur at different levels in the two bones. Distal tibial shaft fractures above the medial malleolus are commonly accompanied by proximal fibular shaft fractures. Full union of tibial shaft fractures may take as
long as four to six months (Court-Brown, 2006), and thus, malunion is often observed in archaeological specimens because of the greater likelihood of fractures remaining unreduced and the difficulty of immobilizing the leg. Because it is so close to the skin, the anterior tibial shaft is commonly the site of a bruise to the shin, which may lead to an ossified subperiosteal hemorrhage. For the same reason, it is the most common site of an open fracture and hence infection from contamination (Court-Brown and Caesar, 2006; Rodner et al., 2003). In contrast to tibial shaft fractures, those of the fibula are not considered serious by most clinical practitioners because they unite readily and are of little functional importance.

**EXPLANATORY BOX 11.5. ANATOMICAL SUMMARY OF FRACTURES**

**Skull**
- Direct trauma, from falls or interpersonal violence, mainly affects the cranial vault.
- Produces linear, penetrating, or depressed fractures (or a combination of these).
- Indirect trauma from falls from a height may lead to basilar ring fracture; basilar fractures also result from impact to the chin.

**Face**
- Direct trauma, often from interpersonal violence, commonly fractures nasal bones and zygomas.
- Indirect trauma from impact to the chin may fracture mandibular condyle or angle.

**Teeth**
- Direct trauma through interpersonal violence, falls, use of teeth as tools or in food preparation, intentional removal or modification.
- Indirect trauma from impact to the mandible.

**Ribs**
- Direct trauma, from falls or interpersonal violence; located near the angle.

**Sternum**
- Direct trauma, or vertical compression with simultaneous fracture of thoracic spine; usually occurs at the manubrium/corpus junction.

**Clavicle**
- Usually caused by a fall onto the shoulder and located at junction of the middle and lateral thirds.

**Scapula**
- Direct trauma; may affect the body, neck, acromion process, or coracoid process.

**Humerus**
- Usually caused by a fall onto an outstretched arm and involves the neck, greater tuberosity (from a fall on the shoulder), shaft, medial epicondyle.
In children often occurs at the lateral condyle or above the lateral condyle.

Neck fracture is most common in older women and often is impacted (i.e., telescoped).

A complication is that if the brachial artery is damaged, then circulation to the hand will be affected.

**Ulna**

- Olecranon process may be fractured by a fall on the point of the elbow.
- Coronoid process is seldom fractured except in association with dislocation of the elbow.
- Fracture of uppermost third on the ulnar shaft may occur with dislocation of the radial head and is usually caused by fall with forced pronation but may also be caused by a direct blow (e.g., a “parry” fracture).

**Radius**

- Head is commonly fractured by a fall onto an outstretched hand.
- Shaft may be fractured and is often fractured at the same time as is the ulnar shaft.
- Fracture of the distal radius almost always caused by fall onto an outstretched hand (called a “Colles” fracture), which is common in older women but rare in children and which is characterized by posterior displacement and tilting.

**Hand**

- Scaphoid fracture is common in young adults, usually from a fall. It is usually a transverse fracture, and a common complication is avascular necrosis and osteoarthritis.
- Base of first metacarpal from a longitudinal force, such as a boxing blow.
- Shaft of metacarpals may have transverse or oblique fractures that may cross the joint surface (watch for complications); telescoping is not uncommon.
- Neck of the fifth metacarpal commonly associated with displacement.
- Phalanges—shaft may show spiral or transverse fracture, whereas base is commonly obliquely fractured; distal phalanges commonly affected.

**Dislocations (luxations) of the upper limb**

- Sterno-clavicular joint: rare.
- Acromio-clavicular joint: common, caused by fall.
- Gleno-humeral joint: common in adults and usually an anterior displacement caused by fall onto an outstretched hand.
- A “dent” on the postero-lateral articular surface of the humeral head indicates recurrent dislocation.
- Elbow: usually caused by a heavy fall on an outstretched hand.

**Pelvis**

- Two main types: 1) isolated, and 2) displaced.
- In displaced fractures, the pelvis breaks at two points that are opposite each other; there is a likelihood of severe soft-tissue complications involving the internal organs.
• Ischio-pubic ramus—superior or inferior ramus may be fractured.
• Acetabulum—a fracture line may enter the joint.
• Ilium—wing or blade of the ilium most commonly involved.

Femur
• Neck—common in adults; avascular necrosis a common complication; nonunion occurs in one quarter to one third of all cases despite treatment.
• Trochanteric—usually caused by a fall, common in older adults, rarely has complications.
• Shaft—common; 4 months is usual time for union.
• Supracondylar—sometimes from carcinomatous metastasis.

Patella
• Two main causes: 1) fall or blow directly to the knee cap, resulting in a crack or comminuted fracture; or 2) from the violent contraction of the quadriceps tendon, leading to a clean break with separation of the superior and inferior parts of the bone.
• Do not confuse this latter fracture with a congenitally bipartite patella.

Tibia
• Condyle—usually only lateral condyle; common in clinical cases when car bumper strikes outside of knee of a pedestrian; most common is a comminuted compression fracture of lateral condyle.
• Shaft—most commonly the tibial and fibular shafts are fractured together, usually from angular or rotational force.
• Angular force leads to transverse or short oblique fracture with the two bones fractured at the same level.
• Rotational force leads to a spiral fracture with the two bones broken at different levels (e.g., distal tibia and proximal fibula).
• Stress or fatigue fractures are common in the tibia.
• Tibial shaft fracture is associated frequently with a communicating skin wound, and infection is a common complication.

Foot
• Calcaneus—fall from a height onto the feet.
• Metatarsal shafts—heavy object falling on foot.
• Base of fifth metatarsal—twisting injury; often avulsion fracture.
• Stress or “marching” fractures.
• Phalanges—usually crushing injuries from falling heavy objects.

Dislocations of the lower limb
• Hip—posterior displacement a common result of auto accidents, related to sitting position.
• Knee—rare, but ligament injuries common.
• Ankle—Injured more often than any other bone except the distal radius; typically the medial wall of the tibia and lateral malleolus of the fibula; ligament injuries common; vertical compressive force can lead to fracture of the anterior margin of the tibia.

Spine

• In clinical practice, fractures are more common in thoracic and lumbar vertebrae than in cervical.
• Wedge compression of vertebral body—occurs on a flexed spine and is usually caused by vertical force associated with hyperflexion or combined flexion and rotation; posterior ligaments remain intact so the fracture is stable, and usually there is no complication involving the spinal cord; almost all occur in thoracic or TL transitional regions.
• Burst fracture of vertebral body—not common; occurs by vertical compression on a straight spine; the vertebral end plate is ruptured, and the disk is forced into the vertebral body; usually produces a comminuted compression fracture.
• Avulsion of vertebral end plates.
• Fracture of atlas: caused by vertical compression.
• Fracture of dens process (C2)—forward displacement is most common; true frequency of this injury is unknown, since damage to the spinal cord at this level often causes death.
• Fracture of spinous process—typically C7 or T1; “shoveller’s fracture”.
• Fracture of transverse process—almost always in lumbar region; typically on one side; usually caused by a heavy blow or a fall.

Dislocations of the spine

• Extension or flexion subluxations and dislocation injuries are unstable.
• When associated with a fracture is referred to as a “slice” fracture of the vertebral body; most occur in mid-thoracic or at the thoraco-lumbar junction; nearly always involve cord injury leading to paraplegia.

FRACTURE HEALING

Duration of Healing

Fractures begin to heal immediately after the bone is broken, but the process varies according to the type of bone involved, the nature of the fracture, and the age of the individual (described in Adams and Hamblen, 1992; Buchwalter et al., 2006; Jones, 2003; Ogden, 2000; Paton, 1992). For example, a phalanx may develop a solidly united fracture in less than one month, whereas the same transformation may take up to six months in a tibia or femur. Bones of the upper limb tend to heal twice as fast as do those of the lower limbs and long bones heal faster than does the skull. Spiral and oblique fractures heal faster than do transverse fractures because of their greater amount of surface contact. Similarly, cancellous bone heals faster than does tubular bone because its mesh-like structure provides more contact between fracture ends and can be penetrated easily by bone-forming tissue. Fractures in children heal rapidly because bone is already being formed beneath the periosteum as part of normal growth, and this process is easily accelerated; children repair fractures about twice as quickly as do adults (Paton, 1992).
The healing process is a continuous one that is divided arbitrarily into stages to facilitate discussion (for a detailed review of the biology of fracture healing, the reader is referred to Schenk, 2003). Table 11.2 summarizes the activities of three overlapping stages in tubular bone healing and the approximate time after injury that each is observed. The first two stages constitute approximately 30% of the total healing process and are involved in healing the area, removing dead tissue, and restoring function to the bone. The final, remodeling, stage lacks the initial importance of the first stages but acts to return the bone to its initial integrity. Many published estimates of the time required for healing are difficult to

<table>
<thead>
<tr>
<th>TABLE 11.2</th>
<th>Summary of Fracture Healing in Adult Tubular Bone (after Adams and Hamblin, 1992; Buchwalter et al., 2006; Schenk, 2003)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Generally, three stages (five substages) in fracture healing are recognized:</td>
<td></td>
</tr>
<tr>
<td>Stage 1—Cellular</td>
<td></td>
</tr>
<tr>
<td>1) tissue destruction and hematoma formation</td>
<td>a) within 24 hours</td>
</tr>
<tr>
<td>b) blood clot forms from torn vessel leakage</td>
<td></td>
</tr>
<tr>
<td>c) fractured bone ends die from lack of blood supply</td>
<td></td>
</tr>
<tr>
<td>d) periosteum is stripped for variable length</td>
<td></td>
</tr>
<tr>
<td>e) not visible in dry bone or on radiographs</td>
<td></td>
</tr>
<tr>
<td>2) inflammation and cellular proliferation</td>
<td>a) begins within 48 hours of fracture and continues for approximately three to four weeks</td>
</tr>
<tr>
<td>b) osteoblasts form collar of osteoid around each fragment</td>
<td></td>
</tr>
<tr>
<td>c) osteoid forms on periosteal and endosteal surface</td>
<td></td>
</tr>
<tr>
<td>d) fracture is bridged</td>
<td></td>
</tr>
<tr>
<td>e) may be visible in dry bone</td>
<td></td>
</tr>
<tr>
<td>f) not visible on radiographs</td>
<td></td>
</tr>
<tr>
<td>3) callus formation</td>
<td>a) begins at two to three weeks and continues until the fracture gap is bridged, which may last eight to nine weeks</td>
</tr>
<tr>
<td>b) mineralization of osteoid</td>
<td></td>
</tr>
<tr>
<td>c) callus of woven bone forms</td>
<td></td>
</tr>
<tr>
<td>d) easily visible on dry bone</td>
<td></td>
</tr>
<tr>
<td>e) first appearance on radiographs</td>
<td></td>
</tr>
<tr>
<td>Stage 2—Metabolic</td>
<td></td>
</tr>
<tr>
<td>4) consolidation</td>
<td>a) highly variable time of weeks to months</td>
</tr>
<tr>
<td>- small bones (e.g., phalanx) fastest, as little as three to four weeks</td>
<td></td>
</tr>
<tr>
<td>- upper limbs may be twice as fast as lower limbs; humerus in three to four months, tibia and femur in approximately six months</td>
<td></td>
</tr>
<tr>
<td>- the greater the contact surface, the faster the healing; spiral or oblique fracture of the humerus in six weeks while transverse fracture 12 weeks.</td>
<td></td>
</tr>
<tr>
<td>b) mature lamellar bone forms from callus precursor</td>
<td></td>
</tr>
<tr>
<td>c) results in solidly united fracture</td>
<td></td>
</tr>
<tr>
<td>Stage 3—Mechanical</td>
<td></td>
</tr>
<tr>
<td>5) remodeling</td>
<td>a) gradual, over years, starting from the beginning of bony response</td>
</tr>
<tr>
<td>b) follows lines of biomechanical stress</td>
<td></td>
</tr>
<tr>
<td>c) fracture site usually marked permanently by surface contour or by abnormal radiographic density</td>
<td></td>
</tr>
</tbody>
</table>
translate to evidence in archaeological bone since many are based on radiographic findings and do not specify whether to callus formation or consolidation is the stage referred. A review of the dating of fractures in children, however, places the radiographic appearance of periosteal new bone usually at between one and two weeks (but as early as four days) after injury (Cramer and Green, 2003; Xian and Foster, 2006), with complete formation of a hard callus between three and six weeks and remodeling anywhere from several months to several years (Cramer and Green, 2003). Femoral shaft fractures in neonates heal in three to four weeks, whereas the same fracture in teenagers and adults heals in three to five months (Adams and Hamblin, 1992; Jones, 2003; Ogden, 2000). Injuries to the growth plate, while at greater risk of complication, heal more rapidly than do shaft fractures (Jones, 2003).

At the time of fracture, blood seeps out through damaged vessels and is contained by the periosteum (the membrane that covers the surface of bones) and other soft tissue, forming a bruise or hematoma. A few millimeters of bone adjacent to the fracture site loses blood and dies. After this initial process, the hematoma is pushed aside by growth of osteoblast progenitor cells that originate in the periosteum and produce osteoid; there is simultaneous activity via the endosteum, the thin layer of cells lining the internal surface in bones with a medullary cavity. This process ends with the resorption of the hematoma and the formation of woven bone in the subperiosteal region (Fig. 11.14). This callus of woven bone results from the mineralization of osteoid and imparts rigidity to the fracture. In children the periosteum is attached loosely to the diaphysis and is slipped easily from the bone by blood collecting under it; this may result in a relatively large bony callus even when there has been little displacement of fragments.

Healing normally is not visible on dry bone or radiographs until approximately one to three weeks after the injury, when the callus appears around the site of injury. Because there is a delay before healing is visible macroscopically or radiographically, unhealed premortem fractures may be hard to distinguish from postmortem breaks that occurred while the bone was still relatively fresh; these lesions are referred to as perimortem fractures because they occurred around the time of death. Delayed death after significant trauma might be identifiable: In clinical cases, death can be delayed by weeks (Templeman and Marinelli, 2006), which would be long enough for bony evidence of healing to appear, although the relevance of clinical statistics to past situations can be questioned on the grounds of medical treatment. Delayed death after blunt trauma occurs at the same frequency (45%) as does immediate death and results from the effects of head injury, respiratory problems (resulting from the injured individual being bedridden instead of upright and mobile for most of the day), and multiple system organ failure caused by the effects of

Figure 11.15 Fine woven bone indicating healing at a fracture site.
systemic inflammatory responses to the bodily insult (Templeman and Marinelli, 2006).

Subsequent consolidation represents the transformation of the woven bone into mature lamellar bone. The remodeling process is a gradual one that strengthens the bone along the lines of mechanical stress. A fracture in adults is usually permanently visible and identified by an area of altered volume or contour or increased radiographic density, although decreased radiographic density that mimics vascular grooves can persist in adult skull fractures several years after the injury.

**COMPLICATIONS OF HEALING**

Complications of fracture healing include infection, vascular damage, nerve injury, hematoma ossification, arthritis, nonunion, and malunion, which are discussed in detail below. These should always be assessed when examining fractures because they may provide information regarding an individual’s mobility or other functions, morbidity, mortality, and medical treatment or the lack thereof. Brickley (2006) has noted that rib fractures carry a high level of pain and disability and thus would have had a serious impact in terms of morbidity in past populations.

**Infection**

Open fractures are prone to infection, which hinders the union of the fracture and creates instability. A bacterium may be introduced to the body through an open fracture from surface contamination or from a penetrating instrument or contaminant, although there is a greater risk of infection with penetrating or crushing force from outside of the body as opposed to a broken bone end breaking the skin from the inside (Rodner et al., 2003). Several bacteria are commonly found on the skin, and of these, *Staphylococcus aureus* is both common and a potential pathogen; other bacteria (e.g., *Streptococcus sp.*) are found in the mouth and nose and are possible pathogens if those body parts are affected by open wounds (Moholkar and Ziran, 2006). An unusual case of probable bacterial meningitis after an arrow wound has been reported for the late Neolithic period (Schutkowski et al., 1996). Posttraumatic infection may be present either as a localized inflammatory condition or as a systemic infection, the latter caused by the spread of the pathogen via the bloodstream. Convention in paleopathology differentiates among periostitis, osteitis, and osteomyelitis (Ortner, 2003). The localized response is visible in the form of periostitis, which is bone formation that occurs under the periosteum and may continue to develop and form a plaque-like sheet over the bone cortex. Osteitis is an inflammation that affects the inner structures of compact bone, such as the vascular channels and the endosteum. Systemic infection appears as osteomyelitis, a severe bone infection that involves the medullary cavity and is identified by a thickened contour in the area of the fracture and often by a heavier feel to the bone. Pathognomonic evidence of osteomyelitis results from the development of subperiosteal abscesses that eventually drain pus through the formation of one or more sinuses (cloacae) in the bone. Posttraumatic osteomyelitis is most commonly observed in the long bones of archaeological skeletons.

**Vascular Damage**

Fractures inevitably result in the rupture of minor blood vessels. This rupture is not usually a serious complication although bone displacement can compress or twist blood vessels and lead to ischemia (a restriction in the blood supply), delaying the healing process and possibly leading to bone death. The femoral head, proximal scaphoid, and body of the talus are prone to ischemia and necrosis. Death of the tissue begins a week after

---

3Medical classification of posttraumatic infection varies and includes a category of superficial osteomyelitis that is confined to the bone surface (Moholkar and Ziran, 2006).

4A sign or symptom that is distinctively characteristic of a disease or condition.
the nutrient supply is reduced and may continue for several years. During this time, the bone loses its trabecular structure and eventually disintegrates because of muscle stress or body weight. The adjacent articular cartilage also dies as a result of deficient nourishment, usually resulting in osteoarthritis. In the condition called osteochondritis dissecans, a loose piece of bone and cartilage separates from the end of the bone because of a loss of blood supply. The loose piece may stay in place or may fall into the joint space, making the joint unstable. This condition is not uncommon in adolescents, and the knee and elbow are most often affected (Hixon and Gibbs, 2000); it has also been reported as a chronic occupational injury (During et al., 1994).

Nerve Injuries

The consequences of nerve injuries range from loss of sensation to loss of function. Usually the loss is temporary, but muscle atrophy may result, and if the nerve loss is prolonged or permanent, the bones will display signs of disuse atrophy as well. This sequel would be most likely in archaeological cases of neurotmesis, which is the complete division of a nerve through severing or severe scarring and the most serious type of nerve injury. In addition, nerve damage may mean that the individual will not feel pain and hence may continue to use the broken bone, impairing healing. Fracture of the vertebral column may result in damage to the spinal cord or spinal nerves, with paralysis below the level of the injury a possible outcome. Skull fractures must always be examined with respect to possible brain injury.

Hematoma Ossification

Posttraumatic ossification of a hematoma results when absorption of the hematoma is prevented by excessive stress placed on the periosteum. A smooth mass of bone is macroscopically visible after two months, with calcification being visible radiologically a few weeks after the injury. Although usually benign, movement may be restricted if there is joint involvement. Subperiosteal hematoma caused by trauma can be distinguished from inflammatory reaction because the former is found only external to the original bone surface, whereas the latter always involves the original bone surface as well (Schultz, 2001).

Arthritis

Traumatic arthritis is caused by trauma that bruises or tears the articular cartilage. Premature deterioration of articular cartilage and subsequent deterioration of subchondral bone are common complications of fractures affecting the joint surface itself, since cartilage repair is a very slow process. Such fractures also can result in ankylosis (fusion) of the joint. Traumatic arthritis also may result if joint function is affected by traumatic injury, even if the joint itself is not involved. A joint above or below the injured bone might be affected, but it is also possible that arthritis on the joint surfaces of an uninjured limb can develop; this may occur if weight bearing or other repeated activity was shifted onto an uninjured limb or body part in order to favor an injured area. Septic arthritis is an infection of the joint. Although there are several causes of septic arthritis, in the case of trauma it can result from the introduction of bacteria through a penetrating injury or an open fracture. The knee is the joint most commonly affected by septic arthritis (Goldenberg, 1998).

Nonunion

Nonunion occurs when the fracture fragments fail to unite and the marrow cavity seals. Initially, bone density at the ends of the fragments increases, which is a process known as sclerosis that is visible radiologically. Eventually the fragments take on a rounded appearance at their ends, which are connected by fibrous tissue. Nonunion may result from inadequate bone healing from infection, inadequate blood supply, insufficiency of
vitamin D or C or of calcium, excessive move-
ment between bone fragments during healing,
soft tissue being caught between the fragment
ends, poor contact between the fragments, pre-
sence of foreign material, or from the destruc-
tion of bone from preexisting disease or the
injury itself (Altner et al., 1975; Brinker,
2003; Karlstrom and Olerud, 1974; Sevitt,
1981; Stewart, 1974; Urist et al., 1954;
Yamagashi and Yoshimura, 1955). If there is
persistent movement between the ununited
ends, a pseudarthrosis (false joint) may form,
although this complication is relatively rare;
arceological populations reveal an
average frequency of 2% as do alloprimates
(Lovell, 1997). Nonunion and delayed union
of fractures in children is rare (Hensinger,
2003).

Malunion

A malunion consists of a fracture that heals
leaving a deformity. Malunion occurs com-
monly in children; most often it is associated
with a supracondylar fracture of the humerus.
According to clinical observation, most chil-
dren do not suffer impaired function but may
have a significant cosmetic deformity (Jones,
2003). In long bones of both children and
adults, the common deformities are shortening;
lengthening; and angular, horizontal, or
rotational displacement. Shortening is caused
by substantial angulation, crushing, gross bone
loss, or when muscular forces cause the frag-
ments to overlap. It typically occurs when
broken bones have not been reduced, often
because of severe pain or muscle spasm, or if
a fracture reduction failed because of instability.
Injuries to growing bone that lead to premature
fusion of the cartilaginous growth plate (in
long bones, this is the region between the dia-
ophys and the epiphysis) may result in shorten-
ing, as may bone ischemia such as can result
when sickled blood cells (in sickle cell
disease) obstruct the normal blood supply. The
presence of a shortened bone is most detrimen-
tal to the lower, weight-bearing limbs and can
lead to backache from pelvis tilting and from
lateral and rotational spinal deviation. Minimal
deformity in bones that are likely to be severely
affected when fractured has been interpreted as
evidence for immobilization of the injured part
and for possible medical treatment (Grauer
and Roberts, 1996), and this interpretation is
especially enhanced when the historical
context is supportive (e.g., Mitchell, 2006).

Bones can also be lengthened (distracted), by
the separation of bone fragments, which may
be from muscular forces, but bones themselves
may distract a fracture, such as when an intact
ulna pulls apart the fragment ends of a fractured
radius or when a fractured tibia is associated
with an intact fibula. Distraction also may
be caused when tissue is caught between
fragment ends.

Horizontal displacement between the frac-
tured bone refers to shifting of the distal frag-
ment in relation to the proximal end without
angulation (Lovell, 1997). If the bone is
viewed in anatomical position, a medial or
lateral shift may be observed; if viewed in a
lateral position, anterior or posterior displace-
ment may be observed.

Angulation refers to angular displacement
from the bone’s longitudinal axis of the
distal fragment of the bone relative to the
proximal fragment. Displacement can be
medial, lateral, anterior, posterior, or a combi-
nation of these. In the lateral view, anterior
angulation refers to the distal portion of the
distal fragment moving anteriorly so that the
fracture site appears posteriorly bowed. Posterior angulation refers to the distal
portion of the distal fragment moving posteri-
orly; the fracture site appears anteriorly
bowed. Rotation occurs when the distal frag-
ment of bone has turned, internally or exter-
nally, relative to the proximal fragment. This
rotation is usually easily identifiable in dry
bone, especially if the affected bone can be
compared with the contralateral element.
Rotation may result in osteoarthritis or in
ankylosis of a joint if ligaments were torn in
the injury.
Normal variation, burial effects, or excavation damage can cause features that can be mistaken for traumatic lesions. A great deal of variation occurs in the size and shape of bones from age, sex, or ancestry, as can be observed in any review of normal skeletal variation (as illustrated, for example, in White, 2000). Sutures and fissures unique to juvenile skeletons that represent normal features could be mistaken for linear fractures, and biparietal thinning, a condition of uncertain etiology but occasionally observed in the crania of older adults (Barnes, 1994; Cederlund et al., 1982, Wilms et al., 1983), could be mistaken for a depressed fracture. Much variation occurs in the appearance of normal bony features such as foramina, muscle attachments, and vascular impressions. Epigenetic traits such as sutural ossicles and mandibular and maxillary tori are also part of normal variation and could be mistaken for evidence of trauma (for a review of epigenetic traits, see Hauser and DeStefano, 1989, and Saunders, 1989).

Pseudopathological lesions from burial effects can be caused by insect or rodent activity, roots, water, temperature, soil pH, and soil pressure. Bone displacement in juvenile crania, for example, may be from postmortem warping (Crist et al., 1997). Excavation damage may lead to lesions that could be mistaken for cuts or fractures. However, postmortem fractures from burial effects or excavation damage can be distinguished from antemortem fractures because fresh bone and dry bone have different fracture properties and will be altered differently in the burial environment (Buikstra and Ubelaker, 1994; Maples, 1986; Mann and Murphy, 1990; Ubelaker and Adams, 1995). Pseudofractures tend to be characterized by 1) small fragments; 2) nonuniform coloration of fracture edges and adjacent bone surfaces, especially fracture edges that are much lighter in color than is the rest of the bone; 3) squared fracture edges; and 4) the absence of fracture patterning because of the tendency of dry, brittle bone to shatter on impact.

Postmortem destruction may overlay pathological lesions since pathological processes may weaken bone and make it more susceptible to burial effects, but there are several rules of thumb for distinguishing traumatic lesions from pseudotraumatic lesions:

1. Abnormal formation of bone always results from pathological processes.
2. Abnormal destruction of bone may result from pseudopathological processes.
3. Abnormal density of bone may result from pseudopathological processes.
4. Most antemortem processes leave smooth or rounded edges that are evidence of osteoblastic repair.
5. Antemortem osteoclastic activity is unmistakable when observed microscopically.

It is also possible, although perhaps not common, in the archaeological record that secondary burial treatment of the skeleton could be mistaken for lesions caused by interpersonal conflict. Olsen and Shipman (1994) have itemized features that distinguish the pseudopathological lesions for evidence of conflict, such as the frequency and patterning of cutmarks.
variables in trauma causation and provides valuable aids for the interpretation of fractures in antiquity, particularly with regard to skeletal patterning and the contexts of injury. For example, studies have shown that skeletal injuries in children are caused most often by short distance falls that result in a single extremity injury, more commonly in the upper than lower extremity (Ogden, 2000; Wilkins and Aroojis, 2006). The radius is generally agreed to be the most common fracture once children begin walking (Jones, 2003; Ogden, 2000; Wilkins and Aroojis, 2006). Toddlers and children fall from furniture, windows, stairs, trees, and fences, and injuries occur during adventuresome play and while doing chores (Ogden, 2000; Rang et al., 2005). In some cultures today, and no doubt also in the past, children of all ages participate in formal and informal labor forces in both rural and urban settings and face associated trauma risks.

Children also suffer trauma through abuse, which occurs in both industrialized and developing countries. A recent United Nations report (Pinheiro, 2006) identified several factors influencing the physical abuse of children that could have existed in the past, including marriage of young girls to older men, alcohol consumption, and poverty, as well as the contexts in which abuse can take place, such as corporal punishment as discipline in the home or workplace, traditional practices including bride-price related violence, and activities related to male socialization and norms of masculinity. Therefore, paleopathologists should be aware of the fracture patterns that might indicate child abuse in the past. Although the identification of domestic violence (e.g., abuse of a child, intimate partner, or elderly person) in a past population is difficult, the most specific indicators of physical abuse in any age group are multiple injuries at different stages of healing, disruption of callus formation, and evidence of chronic subdural and subperiosteal hematoma (Campbell and Schrader, 2006; Maples, 1986; Walker et al., 1997; Wilkinson and Van Wagenen, 1993; Worlock et al., 1986), although in adults these features may be characteristic of professional warriors (Sutter and Cortez, 2005). Head and neck injuries predominate in cases of child abuse (Blondiaux et al., 2002; Campbell and Schrader, 2006; Worlock et al., 1986), although the age of the child is an important variable: Accidental trauma is considered unlikely if the child is not yet walking, whereas toddlers may suffer skull fractures caused by falls from experimental standing, walking, running, or climbing. Typical abuse injuries to the limbs of children are reviewed extensively by Campbell and Schrader (2006) and Cramer and Green (2003) and include spiral or oblique fractures from a twisting or gripping motion; those of the humerus are considered a dominant characteristic, especially if occurring bilaterally (King et al., 1988; Loder and Bookout, 1991; Worlock et al., 1986). Epiphyseal–metaphyseal fractures are much less common than shaft fractures but are much more specific for child abuse, and a single transverse long bone fracture in a non-walking infant would be highly suggestive of abuse (Cramer and Green, 2003). Fracture of the distal humerus can be from birth trauma, but in a child older than a neonate is more likely from abuse (Cramer and Green, 2003). Rib fractures are the third most common site of skeletal injury in abused children, especially in children younger than two years; posterior rib fractures are the most common and are caused when the child is grasped and shaken with side-to-side compression of the infant chest by an adult’s hands (Cramer and Green, 2003).

Clinically, one third of all fractures from abuse are in children younger than one year of age, whereas accidental fractures occur rarely in this age group (Cramer and Green, 2003; Krishnan et al., 1990; Worlock et al., 1986). The incidence of fractures from abuse decreases as age increases, whereas the incidence of fractures from accident increases with age up to about 12 years (Cramer and Green, 2003). When considering past populations, the absence of context makes it difficult to infer abuse, and paleopathologists must be sure to evaluate other possible causes of fractures and fracture-related lesions in children. For
example, periosteal new bone formation suggestive of abuse may be from other systemic conditions such as rickets, scurvy, or congenital syphilis, and osteogenesis imperfecta (a hereditary disease characterized by extremely brittle long bones) may lead to a fracture pattern that mimics abuse (Cramer and Green, 2003). Although it has been estimated that 30% to 50% of physically abused children sustain significant skeletal fractures or related problems (Cramer and Green, 2003), Walker (2001) has not identified the battered child syndrome in any of the many archaeological skeletal collections that he has examined. An explicit assessment of the evidence, or lack of evidence, of child abuse in skeletal collections therefore may provide us with indications of the cultural value of children and of cultural beliefs about the value of children in different sex and age groups.

Interpersonal violence also can be detected through patterns of skeletal trauma in adults. Fractures of the cranium, ribs, or hands are more likely to indicate trauma from interpersonal violence than are fractures to the forearm, even though the ulna and radius are fractured more commonly than are any other skeletal elements (Grauer and Roberts, 1996; Gupta et al., 1982; Smith, 1996; Walker, 1997). Injuries that have a high specificity for a clinical diagnosis of assault are fractures of the skull (especially the nasal and zygomatic bones and the mandible), posterior rib, vertebral spinous process, and of hand and foot bones, all of which can result from the direct trauma of punches or kicks (Walker, 1997). Fractured or dislocated jaws and broken or chipped teeth are additional head injuries and have been described in past situations that may have involved interpersonal violence (Lukacs, 2007). In addition, the palmar surfaces of the manual phalanges may exhibit healed or unhealed cutmarks, originating as defensive wounds incurred as a victim of a knife or sword attack.

Several examples from past populations indicate interpersonal violence and, depending on the context, may point to such causes as social tension resulting from access to resources, environmental deterioration, or social change (e.g., Alvrus, 1999; Bloom and Smith, 1991; Dawson et al., 2003; Domett and Tayles, 2006; Hutchinson, 1996; Judd 2004, 2006; Lessa and Mendonça de Souza, 2004; Mitchell, 2006; Payne et al., 2006; Patrick, 2006; Stirland, 1996; Sutter and Cortez, 2005; Torres-Rouff and Costa Junqueira, 2006). In the absence of contributory evidence, the cause of the violence, such as crime, punishment, ritual, or competition for resources, should be considered speculative (Jurmain, 1999; Jurmain and Bellifemine, 1997). Although it is difficult to identify the cause of an isolated case of apparent interpersonal violence, some causes may be ruled out by secondary evidence such as the type of burial (e.g., Dawson et al., 2003), and patterns of trauma within a skeletal sample, such as by sex, may be illustrative. For example, women-directed violence by men is the most common pattern in domestic contexts, whether in industrialized nations or in developing countries (Espinoza, 2005; Kantor, nd; Koenig et al., 2003, 2006a,b; Kuruppuaracchi and Wijeratne, 2005; Lalasz, 2004), although abuse of male partners by women, same-sex partner abuse, and abuse of the elderly\(^5\) also has been documented. Violence perpetrated by women against other women is very rarely reported and generally involves little injury, but it has been noted between cowives in polygamous households, rivals in sex and marriage, mothers and daughters, mothers- and daughters-in-law, women and their domestic servants or slaves, and between domestic servants (Ahsan and Islam, 2004; Barat, 2004; 5Elder abuse is medically defined as acts of commission or omission that result in harm or threatened harm to the health or welfare of an adult usually over the age of 65 years (Amer Med Assoc, 1992) and would be difficult to identify in archaeological contexts, not only because of the paucity of skeletons that can be recognized confidently as those of older adults but also because of the likely existence of confounding predisposing conditions such as osteoporosis in such individuals (Collins, 2006). In addition, much of the published literature on elder abuse suggests that bedridden or otherwise physically or mentally incapacitated elders are particularly susceptible to abuse. \(^5\)
Burbank, 1987; Dinnen, 1997; Elbedour et al., 2006; Saravanan, 2000). Because injuries sustained in domestic assaults are under-reported by victims, the prevalence of specific fractures is unclear; however, head and neck injuries account for about 50% of clinically observed injuries in women, followed in frequency by fractures to the thoracic region and shoulder (Boes, 1998; Goldberg et al., 2000; Huang et al., 1998; Le et al., 2001; Pigadas and Oliver, 2006; Spedding et al., 1999).

Interpretation is fraught with pitfalls, and some researchers have argued that identifying the behavior behind trauma in past populations is rarely possible without compelling contextual evidence, and even then it is problematic (e.g., Jurmain, 1999; Jurmain and Bellefemine, 1997; Stirland, 1996; Wakely, 1996). Often the most compelling evidence for interpersonal violence lies in the presence of artifacts or other physical evidence of violence. Projectile points embedded in bone or recovered from the abdominal cavity provide unmistakable evidence of interpersonal violence and may point logically to the cause of other injuries as well (e.g., Bennike, 1985; Jurmain, 1991; Jurmain and Bellifemine, 1997; Smith, 1996; Walker, 1989; and several papers in Owsley and Jantz, 1994), although an isolated case of such an injury could alternatively suggest a hunting accident; thus, context is important for interpretation. Similarly, inflammatory lesions of the cranial vault consistent with scalping provide contributory evidence of violence (Owsley et al., 1994; Stewart and Quade, 1969).

Historical records and an understanding of culture historical context also can benefit the interpretation of interpersonal violence, such as the identification of strangulation from hyoid fractures and of hanging and decapitation from vertebral injuries (Anderson, 2001; Angel and Caldwell, 1984; Morimoto, 1987; Morimoto and Hirata, 1992; Spence et al., 1999; Ubelaker, 1992; Waldron, 1996) and the comprehensive analysis of trauma etiology in historical populations (Brickley and Smith, 2006; Williamson et al., 2003), in aboriginal populations of the Canadian northwest coast (Cybulski, 1992) and Arctic (Melbye and Fairgrieve, 1994), and in prehistoric populations of the American midwest (Milner, 1995). Historical and archaeological accounts also have provided context for the interpretation of battlefield trauma among soldiers of the fourteenth-century battle of Visby (Ingelmark, 1939; Knowles, 1983), the massacre at Fort William Henry (Liston and Baker, 1996), the War of the Roses (Novak, 2000), and the Battle of the Little Bighorn (Willey and Scott, 1996).

Despite the wealth of published reports and discussions of archaeological examples of interpersonal violence that have appeared in the past ten years, most traumatic injury in both clinical and archaeological contexts is from accidents. Features of the physical environment, such as irregular terrain, influence the frequency and nature of trauma (Alvrus, 1999; Jimenez-Brobeil et al., 2007; Judd, 2006; Kilgore et al., 1997), as does climate. Inclement weather and reduced winter daylight hours in northern latitudes increase fracture risk from mishaps caused by slippery surfaces and limited visibility, whereas decreased sunlight may impair calcium absorption and lead to fractures secondary to osteoporosis or rickets, and dietary inadequacies of vitamin C or calcium may reduce the quality and quantity of bone and hence increase the risk of fractures. Clinical studies show that children suffer higher rates of fracture in warm climates and warm seasons of the year, presumably because they spend more time in outdoor activities that carry a higher fracture risk (Glencross and Stuart-Macadam, 2000; Jimenez-Brobeil et al., 2007; Wilkins and Aroojis, 2006).

Features of the sociocultural environment similarly are factors in trauma. For example, differences in fracture rates have been linked to differences in subsistence activities, technology, and occupation (Domett and Tayles, 2006; Judd and Roberts, 1999; Judd, 2006), and patterns of trauma by age and sex also have sociocultural determinants. Traditionally, most fractures in females occur in the home, whereas most fractures in males occur outside
the home, although this trend varies according to country, income level, occupation, and age (reviewed by Judd and Roberts, 1999). High fracture risks exist in occupations that in the past likely were restricted to men, such as mining, logging, and construction, and household work such as carrying water and firewood would have posed fracture risks for women and children. Animal herding, milking, and harvesting may have been engaged in by both sexes and by both adults and children and would have associated fracture risks ranging from kicks from animals to falls from ladders or wagons. Both mechanized (e.g., bicycles and carts) and unmechanized transportation (e.g., horses, donkeys, and camels), also carry fracture risks.

In summary, this review of trauma and its causes provides descriptive protocols that can be used to improve paleopathological analysis and interpretation. These protocols include identification of the skeletal element(s) involved, the location of the injury, its appearance, and any evidence for complications of the injury. These descriptions then serve as a basis for inferences about the mechanism of injury. In some cases, particularly with well-remodeled fractures, it may not be possible to identify the type of fracture, but one or more types may be ruled out. Paleopathology reports often include a discussion of “differential diagnosis”; that is, the pros and cons of alternative diagnoses are discussed in an attempt to argue why one alternative may be more probable than another. Ultimately, the interpretation of the cause of trauma in antiquity benefits from the information derivable from the injury itself, from biological variables, and from the sociocultural and environmental context.

ACKNOWLEDGMENTS

I am indebted to the many students who have taken my course in paleopathology, whose enthusiasm has inspired me and whose challenging questions have led me to refine my approach to the analysis and interpretation of skeletal trauma. I thank Michele Buzon, Margaret Judd, Robert Jurmain, Lynne Kilgore, and Phil Walker, from whose knowledge and insights I have benefited greatly, and Debra Komar, who provided me with references and descriptions related to forensic applications. I also thank Anne Katzenberg and Shelley Saunders who provided suggestions for improving this chapter. I acknowledge, with thanks, permission to photograph specimens in the Physical Anthropology Collections, Department of Anthropology, University of Alberta (Figs. 11.10, 11.11, and 11.12). Some of the research referred to in this chapter was supported by grants from the office of the Vice President (Research) and the Faculty of Arts at the University of Alberta and the Social Sciences and Humanities Research Council of Canada.

REFERENCES


REFERENCES


Kantor P. nd. Domestic violence against women: a global issue. sss.ucis.unc.edu/Abuse/Abuse.html.


REFERENCES


Moholkar KD, Ziran BH. 2006. Local complications. In: Bucholz RW, Heckman JD,


Robinson CM. 2006. Fractures of the distal humerus. In: Bucholz RW, Heckman JD, Court-Brown C,


CHAPTER 12

LIGHT AND BROKEN BONES: EXAMINING AND INTERPRETING BONE LOSS AND OSTEOPOROSIS IN PAST POPULATIONS

SABRINA C. AGARWAL

INTRODUCTION

The study of health and disease in past human populations is well grounded in the field of anthropology. Anthropological studies contribute a holistic perspective, emphasizing both evolutionary and biocultural approaches to understanding the natural history of disease. Biological anthropologists are particularly interested in records of health and disease that can be read from ancient human skeletal remains. For example, the maintenance and fragility of bone tissue, and the associated skeletal disease osteoporosis, have received steady interest by biological anthropologists who have sought to reveal the patterns and prevalence of this disease in the past. In part the interest in examining osteoporosis in the past stems from the fact that it is a growing health concern in the aging populations of developed countries. Over 75 million people are affected in Europe, the United States, and Japan alone (EFFO, NOF, 1997). In the United States, it is reported that osteoporosis is a major health threat for 44 million Americans, with one out of every two women and one in four men over 50 years of age expected to suffer from an osteoporotic-related fracture in their lifetime (NIH/ORBD, 2006). These figures are expected to jump considerably higher over the next several decades. Osteoporotic-related fractures are associated with significant morbidity and mortality (Cauley et al., 2000; Center et al., 1999; Kanis et al., 2003), carry a high economic impact for treatment (Johnell, 1997; Johnell et al., 2004; Johnell and Kanis, 2005, 2006), and cause a significant reduction in quality of life (Cockerill et al., 2004). As such, it is not surprising that there is interest to better understand the natural history of bone loss and its prevalence in past societies. Several studies have examined bone loss in past populations with a wide variety of methodologies (see Table 12.1). Although bone loss and some types of typical fragility-related fractures have been observed in past human populations, the age- and sex-related patterns of bone maintenance and fragility often differ as compared with modern populations. This chapter will explore what the patterns of bone loss in the past might reflect and some of the interesting questions that can continue to be asked from studies of bone maintenance in archaeological samples. The unique methodological and paleodemographic challenges of assessing and interpreting bone loss in archaeological populations

will also be discussed. Finally, future research directions, and the importance of bioarchaeological studies in understanding bone loss in both the past and the present, will be highlighted.

**WHAT IS BONE LOSS AND OSTEOPOROSIS?**

Although we often think of bone as part of a dry and brittle skeleton, during life, bone is actually

<table>
<thead>
<tr>
<th>Method</th>
<th>What it Examines</th>
<th>Potential Problems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visual examination of whole bones</td>
<td>Gross examination of fractures</td>
<td>Difficult to assess unhealed perimortem fractures or postmortem breakage&lt;br&gt;Difficult to distinguish traumatic and fragility-related fractures&lt;br&gt;Uncertain when healed fractures may have occurred during life</td>
</tr>
<tr>
<td>Cortical histomorphometry</td>
<td>Quantitative examination of cortical micromorphological structures, bone remodeling, and turnover (static histomorphometry)</td>
<td>Invasive sampling required. Newer 3D microCT methods may soon be possible for noninvasive histomorphometry and are currently possible for analysis cortical canal network, although current limited gantry size accommodates only small-size tissue samples. Specialized processing, microscopy or CT equipment, and expertise required&lt;br&gt;Difficult to assess the many factors that can affect bone remodeling parameters</td>
</tr>
<tr>
<td>Metacarpal radiogrammetry</td>
<td>Measurement of cortical thickness and cortical index (percentage of bone width taken up by cortex) typically from radiographs</td>
<td>Does not measure bone loss in an area typically affected by fragility-related fracture</td>
</tr>
<tr>
<td>Cross-sectional geometry in long bones</td>
<td>Measurement of bone cross-sectional geometry (shape) and cortical thickness typically from X-rays or noninvasive CT imaging</td>
<td>Difficult to assess the biomechanical, ontogenetic, and life history influences on bone cross-sectional shape</td>
</tr>
<tr>
<td>Examination of trabecular architecture</td>
<td>Visual and quantitative examination of trabecular microarchitecture, connectivity, and deterioration</td>
<td>Specialized processing, microscopy equipment, and expertise required</td>
</tr>
<tr>
<td>Dual-energy x-ray absorptiometry (DEXA)</td>
<td>Absorptiometric measurement of bone mineral content and density that uses an X-ray source</td>
<td>Reliability of measurements are often questioned by possible chemical bone diagenesis that can alter mineral content in archaeological bone</td>
</tr>
</tbody>
</table>

(For extensive reviews of paleopathological studies of bone loss studies and methodology, see Agarwal and Grynpas, 1996; Brickley and Agarwal, 2003; and Nelson et al., 2003).
a highly dynamic tissue that allows the skeleton to act more like a bodily organ. The mammalian skeleton has a remarkable ability to grow, maintain, and renew itself through the coordinated efforts of its three major types of bone cells (osteocytes, osteoblasts, and osteoclasts). In the adult skeleton, the primary mode of maintenance and replacement for mechanically incompetent bone tissue is called remodeling, where bone is systematically resorbed and then replaced in the same area (Parfitt, 2003). In the immature skeleton, the primary mode of bone growth is called modeling, where bone is formed and then soon resorbed in different locations (Parfitt, 2003). However, modeling does not only occur in the juvenile skeleton, and remodeling does begin in childhood (Parfitt, 2003). Although modeling typically leads to an overall gain in bone mass, remodeling usually is associated with an overall loss of bone. Why bone remodels depends primarily on the function of the bone tissue that is being replaced (Frost, 2003; Parfitt, 2003). Bone tissue that serves a metabolic function will be remodeled as an important part of its role in calcium homeostasis and the production of blood cells in the body. However, the main function of bone is to resist mechanical loads, and like other materials that bear loads, over time the material accumulates fatigue damage. As such, bone tissue that serves mainly a structural function will be remodeled to maintain or repair the bone tissue material (Currey, 2003; Martin, 2003; Parfitt, 2003). Bone also remodels in response to unique traumatic situations, like healing fractures and also in times of intense physiologic demands for calcium such as growth, pregnancy, or lactation (Frost, 2003; Parfitt, 2003). Although the metabolic or physiologic role of the skeleton is both vital and intricate, the role of the skeleton as a structural and sensitive homeostatic mechanical organ is now recognized as central to our understanding of bone biology (Martin, 2003). Skeletal structure can change through modeling and remodeling in response to mechanical stimulation or stress in order to tolerate the effects of activity at the tissue level (Martin, 2003). The cellular mechanisms of how bone cells work together to remodel and model are beyond the scope of this chapter (for more information on bone remodeling, modeling, and loading, see Frost, 2003; Pearson and Lieberman, 2004; and also chapters by Robling and Stout and Ruff, this volume); however, the bone literally can “adapt” to mechanical loading sensed through its living bone cells (Martin, 2003). In the study of skeletal remains from archaeological populations, this means that the mechanisms of bone biology can record many aspects of life history at the bone tissue level. Significant lifetime changes during growth, nutrition, activity, or disease can conceivably be observed through the study of morphology, microstructure, and markers of bone turnover that are recorded in the skeleton. In the human skeleton, 20% of bone is trabecular bone and 80% is cortical bone. However, 80% of bone turnover occurs in the smaller amount of trabecular bone. As such, different aspects of bone adaptation during life are better studied in certain parts of the skeleton and bone tissue. A loss of bone or the differences in bone (micro) morphology within and between human populations, including past populations, is what is used to try to reconstruct the possible factors that may have contributed to bone biology during life.

Bone loss, and more typically the assessment of low bone mass, is what is most commonly familiar as the focus in clinical management and prevention of osteoporosis. During growth, the skeleton accrues bone to grow in length, breadth, mass, and volumetric density (Cooper et al., 2006; Parfitt, 1994). The exact age of peak bone mass differs at various skeletal sites, but peak bone mass generally occurs a few years after the fusion of the long bones and thus at different times in male and females. Bone mass and quality in adults is therefore dependent not just on how much bone is lost later in life but also on how much peak bone was gained during growth. At the bone remodeling level in the adult skeleton, bone loss is generally caused by a period in which bone resorption exceeds bone formation.
Osteopenia is the traditional term for loss in the amount of bone. Osteoporosis refers to a systemic skeletal disease that is characterized by such a significant reduction in bone mass and/or a deterioration of the microstructure of bone tissue that there is a consequent increase in bone fragility and susceptibility to fracture. There are two main types of osteoporosis: primary and secondary. Primary osteoporosis is first observed in women at menopause and in both sexes with age as bone loss becomes accelerated. Secondary osteoporosis results from a variety of identifiable conditions, some of which are listed in Table 12.2. Although many of the secondary causes of bone loss are less relevant to the study of osteoporosis in archaeological populations, many causes, particularly those related to metabolic disturbance, are important to consider when interpreting bone loss in paleopathological case studies. There are two kinds of primary osteoporosis: postmenopausal (type I) and age-related (type II) osteoporosis. Postmenopausal osteoporosis occurs after menopause because of the loss of estrogen resulting in an increase in bone turnover, and it tends to affect primarily trabecular bone in locations such as the vertebrae or distal forearm, with associated fractures in these locations (Riggs et al., 1991). Age-related osteoporosis is observed in both sexes typically after the age of 70 years, although women are more likely to suffer from age-related osteoporosis likely in part from the compounded effect with earlier postmenopausal bone loss. Type II osteoporosis results when the process of resorption and formation of bone are no longer coordinated and affects both trabecular and cortical bone resulting in fractures at the femoral neck, vertebrae, proximal humerus, and proximal tibia (Riggs et al., 1991). It may result from age-related reduction in vitamin D synthesis or resistance to vitamin D activity (possibly mediated by decreased or unresponsive vitamin D receptors in some patients) (Grynpas, 2003; Kanis, 1994).

The World Health Organization (WHO) defines osteopenia or low bone mass as between 1 and 2.5 standard deviations below the young adult reference mean and osteoporosis as greater than 2.5 SD below the young adult mean (Kanis, 1994; WHO, 1994). As such, measurement of bone mineral density (BMD) typically by dual-energy X-ray absorptiometry (DEXA) is the primary clinical method used to assess fracture. However, although bone mass does correlate with risk of fracture, several additional factors significantly contribute to bone strength that are independent of bone mass (Burr, 2004; Heaney, 1992; Turner, 2002; Watts, 2002). The mechanical properties of bone are affected by both the structure of the bone as a whole as well as by factors that contribute to the mechanical properties of the bony tissue itself. These factors are typically grouped as bone quality and include factors such as bone geometry, trabecular and cortical

### TABLE 12.2 Some Secondary Causes of Osteoporosis (Account for <5% of Osteoporosis Cases)

<table>
<thead>
<tr>
<th>Diseases that affect the endocrine system, e.g.:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>glucocorticoid excess</td>
<td></td>
</tr>
<tr>
<td>hyperparathyroidism</td>
<td></td>
</tr>
<tr>
<td>hyperthyroidism</td>
<td></td>
</tr>
<tr>
<td>hypogonadism</td>
<td></td>
</tr>
<tr>
<td>hyperprolactinemia</td>
<td></td>
</tr>
<tr>
<td>diabetes mellitus</td>
<td></td>
</tr>
<tr>
<td>Drugs, e.g.:</td>
<td></td>
</tr>
<tr>
<td>glucocorticosteroids</td>
<td></td>
</tr>
<tr>
<td>ethanol</td>
<td></td>
</tr>
<tr>
<td>dilantin</td>
<td></td>
</tr>
<tr>
<td>tobacco</td>
<td></td>
</tr>
<tr>
<td>barbiturates</td>
<td></td>
</tr>
<tr>
<td>heparin</td>
<td></td>
</tr>
<tr>
<td>Other miscellaneous conditions, e.g.:</td>
<td></td>
</tr>
<tr>
<td>immobilization</td>
<td></td>
</tr>
<tr>
<td>prolonged weightlessness as found in space</td>
<td></td>
</tr>
<tr>
<td>flight</td>
<td></td>
</tr>
<tr>
<td>chronic renal failure</td>
<td></td>
</tr>
<tr>
<td>liver disease</td>
<td></td>
</tr>
<tr>
<td>malabsorption syndromes (often associated with gastrointestinal disorders)</td>
<td></td>
</tr>
<tr>
<td>chronic obstructive lung disease</td>
<td></td>
</tr>
<tr>
<td>rheumatoid arthritis (RA)</td>
<td></td>
</tr>
<tr>
<td>sarcoidosis</td>
<td></td>
</tr>
<tr>
<td>malignancy</td>
<td></td>
</tr>
<tr>
<td>amenorrhea</td>
<td></td>
</tr>
</tbody>
</table>

(NIH/ORBD 2006; Riggs et al., 1991).
microarchitecture, and bone material properties (such as bone mineralization and mechanical properties) (see Fig. 12.1 and Box 12.1). Although the focus in the clinical setting is primarily on the risk of osteoporosis with reduced bone mass, considerable evidence exists that bone quality plays a more significant role in bone fragility with aging and disease (Burr and Turner, 1999; Cooper, 1993; Grynpas, 2003; Heaney, 1992; Watts, 2002), and as such, it is becoming increasingly realized that bone quality needs to be addressed in paleopathological studies of bone loss and fragility (Agarwal and Grynpas, 1996; Agarwal et al., 2004; Brickley and Agarwal, 2003). Although bone loss is clearly an age-related and post-menopausal phenomena, osteoporosis has a multifactorial etiology, and the risk of developing osteoporosis is mediated greatly by factors that are independent of the menopause-induced drop in estrogen levels in females (Table 12.3). For example, factors such as genetics, ethnicity, nutrition, physical activity, parity, and lactation also play an important role in bone maintenance (Nelson and Villa, 2003; Ralston, 2005; Sowers and Galuska, 1993; Stevenson et al., 1989; Ward et al., 1995). Perhaps most importantly, current experimental and epidemiological studies clearly demonstrate that bone loss and fragility accompanying old age is tied to the influences on bone maintenance early in growth and development (Cooper et al., 2006), and recent life-history approaches have emphasized the importance of understanding the interrelationships and joint cumulative contributions of different factors (e.g., genetics, diet, exercise, and reproduction) to bone development, maintenance and loss over the entire lifecycle (Agarwal and Stuart-Macadam, 2003; Agarwal et al., 2004; Fausto-Sterling, 2005; Weaver, 1998). Studies of skeletal populations have attempted to consider many of these influences on bone loss and fragility, along with age and menopause. As will be discussed, the role of some factors on bone maintenance in past populations seems more complicated and challenging to understand than as first suggested.

Figure 12.1  Aspects of bone quantity and quality that are known to affect bone fragility and could contribute to risk of osteoporotic fracture. Bone mass or BMD is the typical measure of bone quantity. Reduced bone mass is referred to as osteopenia and contributes to risk of fracture. Note that osteopenia may actually be a consequence after or exacerbated by fracture (Heaney 1992). More recently, several additional factors grouped as contributors to bone quality, including bone geometry, cortical bone histomorphology, trabecular architecture, and bone material properties (such as collagen amount or organization, bone mineralization, or fatigue damage) have been recognized as contributing greatly to bone strength and fragility.
BOX 12.1. WHAT IS BONE QUALITY?

The mechanical properties of bone, and thus the fragility of bone that occurs with age and disease, are affected by both the structure of the bone as a whole as well as the factors that contribute to the mechanical properties of the bony tissue. These factors are typically grouped as bone quality and include factors such as bone geometry, bone material properties, and trabecular and cortical microarchitecture. The structural properties are determined by the amount of bone that is present, but also they depend on the location of the bone with respect to some loading axis (Burr and Turner, 1999). Bone that has a large diameter may be stronger than a bone with a smaller diameter, and changes in bone shape geometry can reflect the mechanical environment that has shaped them and bone fragility (Burr and Turner, 1999; Ruff et al., 2006, see also Ruff, this volume). The material properties of bone depend on several factors. For example, the mineralization of the tissue which is determined by the rate of bone turnover, can affect the strength of the bone tissue. Hypermineralized bone, seen in older and pathological bone, is weaker because of increased brittleness (Burr and Turner, 1999; Grynpas, 2003, Meunier and Boivin, 1997). The organization and nature of the collagen proteins also affect the mechanical properties of bone. The orientation of the collagen fibers in bone tissue affect bone strength (Bromage et al., 2003), but also the actual collagen content and the intrafibrillar cross-links in the collagen correlate with bone toughness and its ultimate fragility (Burr, 2002). Another factor that affects the material properties of bone is the accumulation of microdamage. Bone from older individuals shows more accumulation of microcrack damage that increases bone fragility. Finally, changes in bone microstructure such as cortical porosity and trabecular architecture also contribute to bone quality. Change and loss in both the structure and the connectivity of the trabecular elements are known with age and disease (Mosekilde, 1993; Parfitt, 1992), and several studies have confirmed that microarchitecture does play a role, independent of mass, in bone strength (Compston, 1994; Goldstein et al., 1993; Mosekilde, 1989, 1993; Parfitt, 1992). Although the focus in the clinical setting is primarily on the risk of osteoporosis with reduced bone mass, considerable evidence exists that these aspects of bone quality play a significant role in bone fragility with aging and disease (Burr and Turner, 1999; Cooper, 1993; Grynpas, 2003; Heaney, 1992; Watts, 2002), and as such, it is becoming increasingly realized that bone quality needs to be addressed in paleopathological studies of bone loss and fragility (Agarwal, 1996; Agarwal et al., 2004; Brickley and Agarwal, 2003).

Images of trabecular microarchitecture from eighteenth-century archaeological human lumbar vertebrae. Left: from a younger individual. Right: from an older individual. Note the overall loss of trabecular elements in the older individual, with a preferential loss of horizontally oriented trabeculae, reducing the mechanical strength of the remaining vertical elements. The anisotropy of the trabecular architecture change occurs throughout the lifecycle and contributes to the overall fragility of the bone. Samples courtesy of the Museum of London Archaeological Service. Images courtesy of Sabrina Agarwal.
SORTING OUT THE INFLUENCES ON BONE MAINTENANCE IN THE PAST

The Role of Diet and Nutrition

Although it is recognized that several factors were likely influential in bone maintenance in past populations, many paleopathological studies have focused on the relationship between bone mass and nutrition. For example, dietary calcium has been discussed traditionally in the literature as an important influence on the evolution and maintenance of the human skeleton. It has been argued that as early as the expansion of anatomically modern Homo sapiens 2–1.5 million years ago into the northern climates of the Pleistocene, the consumption of dietary calcium would have dropped significantly with the increasing consumption of hunted meat (Eaton and Nelson, 1991, Eaton et al., 1988). Evolutionary reconstructions estimate that the earliest human ancestor 6–7 million years ago would have had a more varied diet of scavenged meat protein, roots, fruits, seeds, and wild plants. This diet of wild plants and meat is estimated to have provided as much as a 10-fold higher calcium content as compared with a diet heavy in hunted wild game (Eaton and Nelson, 1991; Eaton et al., 1988; Nelson et al., 2003). This shift to big-game hunting is thought to have had a significant impact on bone mineral metabolism (Nelson et al., 2003). This decline in dietary calcium content would have accelerated as Homo sapiens moved out of tropical Africa into northern climates and relied more heavily on meat consumption. Furthermore, the transition to subarctic life and big-game hunting on bone maintenance may have been exacerbated even more by decreasing exposure to ultraviolet radiation in northern climates and with the adaptation of warmer clothing (Nelson et al., 2003). The second significant temporal shift in the human diet that is hypothesized to have had a profound effect on many aspects of human health, including bone metabolism, is the transition from hunting and gathering lifestyles to the full-scale domestication of plants and animals. Although the full reliance on domesticated foods varied in different regions, the transition to food production (around the Neolithic period 12,000–10,000 years ago) in many areas is thought to have resulted in a reduction in the consumption of animal protein and instead a reliance on foods lacking in nutritional variety (Larsen, 2003; Nelson et al., 2003). For example, a reduction in foods high in calcium and iron is thought to have had an impact on bone metabolism, whereas the reliance on grains with an increase in phytate (which binds to calcium and reduces its availability in the body) would have a similarly negative impact (Larsen, 2003; Nelson et al., 2003).

The hypothesis that decreased calcium content in the diet of our Pleistocene hunting ancestors would have had a detrimental effect...
on bone mass and maintenance, has been compared with the large meat component diet of the recent living Inuit (Arctic) populations. Studies of bone maintenance in living Arctic groups have shown some links between high animal protein diet and bone loss (Nelson et al., 2003; Pfeiffer and Lazenby, 1994). Using differing methods, greater bone loss has been observed in Canadian and Alaskan Inuit populations as compared with age-matched U.S. Caucasians (Harper et al., 1984; Mazess and Mather, 1975; Pawson, 1974) which is attributed to the high-protein and acidotic affect of the Inuit diet. Skeletal calcium may be resorbed to buffer acid loads imposed by animal proteins (Orwoll, 1992), and calcium may be bound by the kidneys through sulphates and phosphorus produced in protein metabolism (Schuette et al., 1981). The association between protein intake and bone mass has also been investigated in archaeological populations. For example, bone core studies on different archaeological Inuit skeletal samples have found thicker cortices, lower bone mineral content, and increased secondary osteonal remodeling as compared with U.S. Caucasians, resulting in increased intracortical porosity and bone loss in these Inuit groups (Thompson and Gunnness-Hey, 1981; Thompson et al., 1981, 1983).

However, the influence of protein intake on bone mineral and metabolism is complex. The increase in urinary calcium excretion from high dietary protein essentially uses the skeleton to neutralize acid generated from the high protein diet (Kerstetter et al., 2003). Indeed, cross-sectional studies of modern populations have shown an association between high protein intake of dietary animal protein and a higher prevalence of hip fracture (Abelow et al., 1992). However, more recent studies have also indicated that dietary protein can have a positive effect on bone metabolism. Dietary protein increased circulating levels of Insulin-Like Growth Factor-1 (IGF-1), which affects skeletal development through increasing bone formation (Bonjour, 2005; Dawson-Hughes, 2003). Clinical studies have shown increases in serum IGF-1 levels with supplements in both elderly osteoporotics and healthy subjects (Cadogan et al., 1997; Heaney 1999; Schurch et al., 1998), and animal model studies have also shown dietary protein deficiency induces bone loss and reduced bone formation (Bourrin et al., 2000, Higashi et al., 1996). Dawson-Hughes (2003) has suggested that the effects of protein on bone may depend on its potential interaction with calcium intake. A higher intake of calcium would simply give more absorbed calcium to offset urine losses, and by lowering bone turnover even more, it could also reduce the effects of mild acidosis or amplify the positive effects of IGF-1 on bone maintenance (Dawson-Hughes, 2003). In sum, the effects of protein on bone seem to be dose dependent, and low intake can be deleterious to bone particularly in groups such as the elderly. High intakes are known to be equally harmful and depend on calcium intake and possibly also vitamin D intake to influence the impact of dietary protein on bone (Dawson-Hughes, 2003; Kerstetter et al., 2003). Although studies of high protein diet in archaeological populations are of interest, it is important to note that dietary protein can have both positive and negative effects on bone metabolism.

The change in diet with the transition to food production and its effect on bone loss has been investigated more widely in archaeological populations. Several paleopathological studies have reported low bone mass in early agricultural populations (for reviews, see Agarwal and Grynpas, 1996; Nelson et al., 2003; Pfeiffer and Lazenby, 1994). Morbidity and mortality are thought to have increased in farming societies in large part because of the increase in prevalence of infectious disease with the unsanitary living conditions of early, large sedentary populations, together with poor nutrition and periods of starvation (Nelson et al., 2003). Extensive studies of bone mass from archaeological samples dating between 350 B.C. and 1400 A.D. from Sudanese Nubia using various methodologies such as
cortical microscopy (Martin, 1981; Martin and Armelagos, 1979, 1985) and measurement of cortical thickness (Dewey et al., 1969a, 1969b) have been interpreted classically as reflecting chronic malnutrition. Nutritional hypotheses for bone loss in the past have also been suggested widely in studies of bone mass in Eastern and Southwestern North American prehistoric and contact period populations. For example, Ericksen (1976) suggested that nutrition was an important determinant of bone loss in her comparative study of Eskimo, Pueblo, and Arikara archaeological populations. She found cortical thinning of the humerus and femur bones of the Pueblo samples that had primarily a cereal-based diet. Additional studies of histological parameters of bone remodeling in these populations also suggested differences in bone turnover related to dietary differences between the groups, which again were interpreted to reflect the high-protein diet of the Eskimo and the low-protein diet of the sedentary Pueblo (Ericksen, 1980; Richman et al., 1979). Several bioarchaeological studies of Native American agricultural populations following Ericksen also noted bone loss suggestive of nutritional stress (Cassidy, 1984; Nelson, 1984; Pfeiffer and King, 1981). Although there have been almost no recent extensive studies of bone loss in North American archaeological populations, likely in part from recent and ongoing repatriation of North American skeletal collections, a focus on nutritional hypotheses remains to explain bone loss in prehistoric and historic skeletal populations. For example, several recent studies of low cortical and trabecular bone mass as compared with modern populations assessed with histomorphometry, quantitative computed tomography (QCT), and X-ray absorptiometry (or DEXA) in pre-Hispanic archaeological skeletal remains from Tenerife, Canary Islands have been explained as related to episodic starvation and low-protein and calcium intake (Gonzalez-Reimers et al., 2002, 2004, 2007; Velasco-Vazquez et al., 1999). However, as these archaeological populations are represented by mass burials, age at death estimations could not be made (and in some studies, sex could not be determined) for most of the samples, and as such interpretations of bone loss and nutrition in the populations are very limited.

Clearly the sources and amounts of dietary calcium intake have changed dramatically with the transition to agriculture. Eaton and Nelson (1991) estimate that cultivated food grains have lower calcium content as compared with uncultivated plant grains, with on average cereal grains having 29 mg of calcium per 100 g compared with almost 133 mg/100 g in uncultivated plant sources. However, it should be kept in mind that calcium is only one nutrient that would have affected skeletal health in past populations along with several other lifestyle and environmental factors. Furthermore, the focus in paleopathology on the influence of calcium intake is at odds with much of the recent biomedical literature. Recent clinical and epidemiological studies have raised controversy as to the effects of calcium intake on bone loss and fragility. Studies of calcium supplementation have been mixed, with some studies demonstrating supplementation to be insufficient in preventing bone loss in perimenopausal and older women, although responsiveness seems dependent on menopausal age and skeletal location of bone loss (Dawson-Hughes, 1991; Elders et al., 1994). Furthermore, calcium supplementation has been shown to have no effect in the prevention of osteoporotic fractures (Cumming and Klineberg, 1994; Feskanich et al., 1994), and more dramatic is the observation that hip fracture is correlated with calcium intake (Abelow et al., 1992; Hegstead, 1986; Kanis and Passmore, 1989). It may seem paradoxical, given the popular information on the benefits of calcium and bone health, that countries with the highest calcium dietary intake have the highest incidence of fracture. Explanations for this finding could be related to poor vitamin D status in North American and European countries as compared with Asian and African...
countries as well as the overall increase in food consumption and sedentary lifestyle in the past decades in these former countries (Kanis and Passmore, 1989; Nelson et al., 2003). However, what is clear, is that the incidence of fracture is rising at a rate that cannot be accounted for by either the aging of populations or by calcium intake alone (Kanis and Passmore, 1989).

Archaeological populations would have experienced diverse environments both temporally and geographically, and it is likely that calcium intakes need to be considered alongside other nutrients. The influence of phosphorus on calcium absorption and its role in bone maintenance has also been considered in the study of early agriculturalists. Cultivated cereal-based diets may have had a high ratio of phosphorus to calcium, which has been hypothesized to inhibit calcium absorption (Cohen, 1989; Nelson, 1984). Both calcium and phosphorus have a relationship with parathyroid hormone (PTH), and studies have suggested that increased phosphorus intake could adversely affect bone metabolism (with increased bone resorption and less bone formation) through increased PTH secretion when specifically combined with low calcium intake (Kemi et al., 2006; Palacios, 2006). However, the relationship between phosphorus and calcium is unclear (Slemenda and Johnston, 1990), and it seems the ratio of phosphorus to calcium intake is key rather than the absolute levels of phosphorus intake alone (Palacios, 2006). Clinical studies have not demonstrated a link between high phosphorus intake and bone loss in healthy humans with normal calcium intakes (Calvo, 1994; Palacios, 2006). The relationship between calcium intake and bone maintenance is also complex, and calcium absorption in the gut and excretion can contribute potentially more to calcium balance than can intake alone (Nordin et al., 1995). Perhaps a more influential nutrient on bone metabolism in past populations could have been vitamin D. Osteoporosis may be a consequence of subclinical mild vitamin D deficiency (Heaney, 1999; Parfitt, 1990; Vieth, 2003).

Vitamin D deficiency has been considered in interpretations of bone loss in historical populations (Agarwal et al., 2004; Mays, 1996), but mild vitamin deficiency and its effect on bone maintenance in past populations has not been examined widely. The human diet has varied throughout prehistoric and historic populations; however, the influence of calcium, protein, vitamin D, and other nutrients on skeletal health is complex. Future nutritional hypotheses of bone loss in past populations must consider the current biomedical research and the synergistic effects of other biological environmental influences on bone maintenance.

The Role of Mechanical Usage and Activity Patterns

The role of physical activity on bone maintenance and fragility has also been discussed in the interpretation of bone loss in archaeological populations. Although the subsistence shift to food production saw a significant change in the human diet, the other significant change of the Neolithic revolution was an increasing sedentary lifestyle. Many aspects of lifestyle, particularly workload and physical activity, are thought to have changed dramatically. Comparisons of agricultural and hunter-gatherer skeletal populations show a decline in indicators of activity, such as articular joint diseases or osteoarthritis, that suggest a decline in workload with domestication (Larsen, 2003). Evidence from the measurement of bone geometry in archaeological populations has also indicated generally a decline in bone strength with a sedentary agricultural lifestyle (Larsen, 2003; Ruff et al., 2006; see also Ruff, this volume). An early study of change in femoral cross-sectional geometry in an Amerindian sample from the Georgia Coast spanning the 4000 years from a hunting and gathering lifestyle to agriculture production found a decrease in cross-sectional size with time, which is suggestive of a less biomechanically demanding lifestyle in the agricultural population (Ruff et al., 1984). Another study of cross-sectional geometry from an agricultural archaeological
sample from the Pecos Pueblo, New Mexico, found the shape of the femoral cortex to be similar to the Georgia Coast early agriculturalists (Ruff and Hayes, 1983a, 1983b). Interestingly, the size of the femoral cortex in this population was more similar to the Georgia Coast hunter-gatherer group, leading the authors to suggest that although the type of loads in the Pecos may have been similar to the other agricultural group, the level of mechanical load and activity may have been similar to pre-agricultural levels (Ruff and Hayes 1983a, 1983b). More recent study of skeletal geometric properties in early agricultural populations from the South Eastern U.S. Atlantic Coast has revealed similar patterns (Larsen, 2001). However, another study of an archaeological population from Northwestern Alabama found cross-sectional strength to be greater in both sexes as compared with hunter-gatherers, which has been interpreted as indicating a more physically demanding lifestyle in this agriculturalist group (Bridges, 1991). These results emphasize that workload was likely still variable in agriculturalists depending on region and local terrain (Larsen, 2003; Nelson et al., 2003). Furthermore, it is uncertain how these adaptations to mechanical loading observed in cross-sectional geometry may have translated into bone loss or fragility in early agriculturalists, and additional studies of biomechanical properties together with other measures of bone loss and fragility in early sedentary agriculturalists need to be made. For example, for the Pecos archaeological sample, Burr et al. (1990) did examine age- and sex-related patterns of cortical histomorphology, finding small Haversian canal size in Pecos females and greater osteonal population density in males as compared with modern populations, which is suggestive of an active lifestyle. The authors also suggest that although cortical endosteal bone loss was observed in both sexes, this could have been compensated for geometrically in overall shape and in osteon dimensions, so that structural strength and fatigue properties of the tissue were maintained even in the presence of bone loss (Burr et al., 1990).

Although sedentary agricultural lifestyles began a change in physical activity levels in human evolution, studies of bone mass in historical European archaeological populations have suggested that physical activity in historical groups may still have been appreciably higher as compared with modern populations, thereby preventing bone loss and fragility. For example, in a study of femoral bone mineral density in female archaeological remains from Spitalfields, England dated between 1729 and 1852, Lees et al. (1993) found no evidence of premenopausal bone loss and less severe postmenopausal loss as compared with modern females, which they suggest to be the result of physical activity and possibly unidentified environmental factors. A subsequent study of cortical bone loss in the Spitalfields sample with the use of metacarpal radiogrammetry found both males and females to show a net loss of cortical bone with age that was similar to a comparative modern population, with a higher rate of endosteal metacarpal resorption than periosteal apposition (Mays, 2000, 2001). The reason for the differing results in the Spitalfield studies is unclear, but it could reflect the differing methods of analysis and patterns of bone loss that are often observed in different skeletal regions and tissues (trabecular vs. cortical). What is interesting, though, is that no typical fragility fractures (hip, wrist, or vertebrae) are observed in the Spitalfields sample (Lees et al., 1993; Mays, 2000). It is possible that the changes in cortical bone thickness at the metacarpal reflect metabolic stress (perhaps nutritional) during the growth period and accumulation of peak bone mass (Mays, 2000, 2001), but that the method is not an ideal predictor of bone strength or fragility in the overall skeleton. Another study of bone mineral density by Ekenman et al. (1995) of medieval skeletons from Stockholm dated between 1300 and 1530 A.D. found an absence of low bone density in older age groups, and higher diaphyseal bone density in the lower extremities as compared with modern reference values, which they also suggest could be the result of physical demands of the lower limbs...
in walking and standing. However, studies of bone loss in other medieval skeletal populations from other regions have not found similar results. For example, studies of cortical bone mineral density in the femur (Mays et al., 1998) and radius (McEwan et al., 2004) with absorptiometric methods and bone mass in the metacarpal with radiogrammetry (Mays, 1996) in the British medieval archaeological sample, Wharram Percy, found age-related bone loss in both sexes. These authors have suggested that the patterns of bone loss suggest that lifestyle factors such as rigorous agricultural activity in this rural medieval population were not sufficient in preventing bone loss. However, a recent study of trabecular microarchitecture in the Wharram Percy sample found that although loss of trabecular structure or connectivity was observed between the young and the middle aged, no change in trabecular structure was observed in old age in either sex. The reasons for these different patterns of bone loss at Wharram Percy are unclear, but they could reflect again differences in trabecular versus cortical tissue response and skeletal site (Agarwal et al., 2004). It is interesting that this archaeological sample does not show a significant number of typical fragility-related fractures, and it is possible that physical activity could have been significant enough to prevent fracture despite some bone loss (Agarwal et al., 2004; McEwan et al., 2004).

Although it is known that bone tissue responds to mechanical loading, the biomedical literature is unclear on what type and level of physical activity or exercise is needed to affect bone mass and more importantly bone strength. Although substantial evidence exists that exercise can increase bone mineral density during growth and development, particularly adolescence, exercise seems to have less long-lasting impact on the adult skeleton (Pearson and Lieberman, 2004; Rittweger, 2006). It is becoming increasingly understood that age plays a significant role in how the skeleton remodels and responds to mechanical loading and this changes throughout the lifecycle. Although there may be an ideal “window of opportunity” for the skeleton to grow big and robust bones during the acquisition of peak bone mass (Pearson and Lieberman, 2004), some high strain stress activity may still be effective at older ages (Rittweger, 2006). Furthermore, although the role of developmental age with exercise and activity has been studied widely with cortical bone remodeling, bone mass, and mineral density, less is known about the changing role of developmental age and mechanical loading on aspects of bone quality (such as trabecular architecture, bone geometry, or tissue-level material properties).

The Role of Reproductive Factors

Another factor that has been considered in the interpretation of bone loss in archaeological populations, especially in understanding the patterns of bone loss observed in females, is reproduction. Early observations of bone loss in the archaeological skeletons of Nubian females were thought to reflect pregnancy and lactation stress (Martin and Armelagos; 1979, 1985).

A recent study by Poulsen et al. (2001) also suggests that significant deficits in bone mineral density in young Danish medieval skeletons could be the result of pregnancy and lactation stress. The authors suggest that the physiologic demands associated with pregnancy and breastfeeding may have increased mortality in young medieval women (Poulsen et al., 2001). A similar study of bone mineral density in a medieval population from Norway also found loss of BMD at an early age in females that was explained by insufficient nutrition together with pregnancy and lactation stress (Mays et al., 2006; Turner-Walker et al., 2001). In contrast, Vogel et al. (1990) suggest that parity may have played a role in better trabecular connectivity as compared with modern populations in female skeletons from European historical populations.

Although pregnancy and lactation are known to be associated with high bone turnover rates, the long-term effect of pregnancy and lactation...
on bone loss and fragility is not understood clearly. Studies of the effects of pregnancy on bone have found conflicting results (Cross et al., 1995; Drinkwater and Chesnut, 1991; Kent et al., 1993; Naylor et al., 2000; Sowers et al., 1991); however, epidemiological evidence suggests that parity may decrease fracture risk and could increase bone density (Fox et al., 1993; Murphy et al., 1994; Sowers et al., 1992). Although longitudinal studies indicate that bone loss can occur during initial lactation (Affinito et al., 1996; Chan et al., 1982; Drinkwater and Chesnut, 1991; Hayslip et al., 1989; Kent et al., 1993; Lamke et al., 1977; Lopez et al., 1996; Sowers, 1996; Sowers et al., 1993; Sowers et al., 1995), substantial evidence exists that recovery of bone occurs with extended lactation and during weaning (Affinito et al., 1996; Kent et al., 1993; Lopez et al., 1996; Pearson et al., 2004; Sowers, 1996; Sowers et al., 1993; Sowers et al., 1995).

It seems likely that reproductive factors would have indeed played a role in bone maintenance in prehistoric and historic females. Bone loss observed in young age archaeological female skeletons makes sense, as we are likely observing skeletons of premenopausal women of reproductive age who were pregnant or lactating at the time of death (Agarwal et al., 2004). However, the long-term effects of pregnancy and lactation on maternal skeletal health and fragility in the past have not been addressed. Agarwal et al. (2004) have suggested that the unusual patterns of trabecular bone loss observed at Wharram Percy and other archaeological populations, such as the little loss observed between middle and old age, lack of significant sex difference in loss, and few fragility fractures, may be related to the long-term, perhaps protective, effects of reproductive factors. Historic practices of high parity and prolonged lactation would have created a very different hormonal milieu as compared with modern Western females in archaeological populations such as Wharram Percy. Although historic females of reproductive age could have suffered from transient bone loss, losses would not have affected long-term postmenopausal bone fragility (Agarwal and Stuart-Macadam, 2003; Agarwal et al., 2004). The less dramatic change in bone maintenance after menopause may also reflect the overall lower levels of steroid exposure in historical females caused by repeated pregnancies and prolonged periods of breastfeeding. The pattern of bone loss in modern females may be related to the sudden downregulation of bone forming osteoblast cells that are increased with chronically high levels of estrogen and other hormones (Agarwal et al., 2004; Weaver, 1998).

Although sorting out the influences on bone maintenance is complicated, it is clear that current explorations of the role of reproductive factors in bone maintenance and fragility in the past must consider the recent biomedical literature on the long-term contribution of parity and lactation to the female skeleton, while recognizing the role of changing reproductive cultural practices in the past.

**Unique Challenges to Diagnosis in Archaeological Populations**

**Where Are The Fragility Fractures?**

Although a great deal of interest exists in investigating bone loss and maintenance in past populations as an indicator of overall metabolic health and stress in the skeleton, it is the patterns of bone fracture and fragility in the past that need to be investigated to understand the prevalence of osteoporosis in the past. Typical fragility-related fractures are observed in the proximal femur (hip), distal radius (Colles’s fracture of the wrist), and vertebral bodies (spine). Each of these sites is rich in trabecular tissue, and the higher bone turnover in trabecular tissue likely predisposes these areas to osteoporotic fracture, although loss of cortical tissue at these sites does contribute to fracture risk. Although reports of fragility fractures have been made in the archaeological record (Brickley, 2002; Brickley and Agarwal, 2003), it is unclear how common these fractures were or whether age and sex patterns of fracture in archaeological populations are similar to modern populations. Several reported cases
of fracture exist from archaeological bone assemblages that may be from age or postmenopausal-related bone loss (for example, Dequeker et al., 1997; Foldes et al., 1995; Frigo and Lang, 1995; Mays et al., 2006; Sambrook et al., 1988). Vertebral compression fractures are the most common fragility fracture and the most widely reported fracture in the archaeological record (Brickley, 2002). Few reports of Colles’s fracture have been made in the archaeological record, although recent observations are being made (Brickley, 2002). A recent focused case study on Colles’s fracture and disability made by Mays (2006) in a third–fourth-century British population gives a good anatomical description of fracture at this site. However, it is still uncertain in paleopathological studies such as this one what overall prevalence rates are for these fractures for all age groups and for both sexes in the archaeological population. Without detailed study of age- and sex-related patterns in fragility fracture in past populations over the lifecycle, observations of osteoporotic fracture in the archaeological record remain isolated case studies. Although cases of spine and wrist fracture have been observed in the archaeological record, very few fragility-related hip fractures have been reported, and several authors have discussed the seemingly low number of typical fragility fractures in the past in general (Agarwal and Grynpas, 1996; Brickley, 2002; Brickley and Agarwal, 2003; Mays et al., 2006; Pfeiffer, 2000; Weaver, 1998). Several explanations exist for the paucity of fractures that should be considered when dealing with mortality samples. The lack of hip fracture in skeletal populations may simply reflect heterogeneity in archaeological samples, whereby those that lived to a greater age in the past could have been fitter biologically than elderly people today that have benefited from modern medicine. As such, healthier elderly individuals in past populations may simply not have suffered from many hip fractures (Agarwal et al., 2004). Hip fractures are also associated with a high level of morbidity (Johnell and Kanis, 2006), and many individuals in the past would have died from such a fracture with unhealed bones before burial (Brickley and Agarwal, 2003). Great difficulty exists in determining whether bones that are fractured when a skeleton is excavated are perimortem or postmortem (Pfeiffer, 2000). It has also been suggested that fracture risk may have been lower in some early populations (such as rural or prehistoric populations) that had environments where falls would be on softer ground and yielding surfaces that would have absorbed energy on impact as compared with softer ground and yielding surfaces that have absorbed energy on impact as compared with modern and urban populations (Mays et al., 1998, 2006).

One question is whether past populations lived long enough to sustain a high prevalence rate of fragility fracture, and it has been suggested that perhaps few individuals in the past would have reached extreme old age to suffer some types of fragility fracture, specifically age-related hip fracture (Mays, 1996, 2000). However, it should be noted that low life expectancy in the past is also related to high infant mortality. Surviving infancy, there was a good possibility of living to an old enough age to suffer some types of fragility fracture, particularly those associated with postmenopausal osteoporosis such as spine and wrist fracture (Agarwal et al., 2004). In the case of females, life expectancy is related not only to infant mortality but also to risks associated with childbirth. Furthermore, there is no reason to believe that human longevity has changed over time, and evidence exists that people did indeed live into old age (Jackes, 2000). Jackes (2000) cites estimates that a 10% survival beyond age 60 would actually be conservative, highlighting the demographic data of Russell (1985) who notes that several individuals were expected to have lived beyond 60 years across Europe and North Africa in the first 1500 years A.D., and the work of Sjovold (1978) who notes a significant number of deaths between the ages of 70 and 80 years in an Austrian village in the 250 years before 1852. Perhaps more importantly, even in modern times, fracture risk is not tied exclusively to life expectancy. Today, a secular trend exists whereby the increment in the
population over the age of 80 years has and will continue to rise exponentially as compared with overall population growth (Kanis, 1994). However, the change in demographics does not account entirely for the current increasing incidence of several types of fragility fracture. For example, Kanis (1994) notes that hip fracture incidence in Oxford, England, doubled in the 27 years after the 1950s, and similar increases have been documented in other parts of the world. It is clear that life expectancy is not the only factor involved in the increasing incidence of osteoporosis.

Of the fracture cases that have been reported in the archaeological record, it is not clear how many are truly (osteoporotic) fragility fractures as opposed to traumatic fractures. Certainly the focus in the archaeological record should be primarily on fractures that are known to be typically observed in osteoporosis today (i.e., hip, wrist, and spine). Some recent studies have attempted to correlate observations of low bone mass together with fracture prevalence. For example, in their study of bone mineral density in a medieval population from Norway, Mays et al. (2006) found the few individuals with fragility fractures did have lower bone mineral density (BMD) than those with other fracture types, which supports their classification of the fractures as truly osteoporotic. Additional correlation studies such as these in past populations need to also examine prevalence difference between the sexes. Also, many pathological conditions can cause a decrease in bone mass or mineral (ostopenia) and, in severe cases, can lead to a pathological fracture (Brickley and Agarwal, 2003). Careful differential diagnosis needs to be considered in cases of fracture in archaeological skeletons, such as age and sex of the individual, location of fracture, and any other pathological changes present on the skeleton (Brickley and Agarwal, 2003). Lastly, although healed fractures are becoming increasing examined in studies of bone loss and osteoporosis, it is impossible to be exactly sure when a fracture occurred in an individual’s life. Although an individual with typical fragility fracture may be identified as an older adult, it is possible that traumatic injury occurred at a younger age and was not related to age or postmenopausal bone loss (Brickley and Agarwal, 2003). This issue emphasizes the need to look not only at fracture prevalence in older adults but also in younger individuals to assess lifelong fracture patterns in a given archaeological populations.

What The Methods Can and Cannot Reveal in Past Populations. Although many methods have been used to examine bone loss and maintenance in archaeological bone (see Table 12.1), each method has its own advantages and disadvantages. One concern in the paleopathological literature has been the comparability of methods (Agarwal and Grynpas, 1996). Many methods measure different tissue types (trabecular versus cortical) and/or different skeletal sites. Trabecular tissue has higher bone turnover and could reflect different metabolic histories as compared with cortical tissue. Furthermore different skeletal sites show variability in bone mass, specifically between bones of the upper and lower limbs (Peck and Stout, 2007). As such, bone mass in skeletal elements is still very much tied to specific mechanical loading environments (Peck and Stout, 2007; Robling and Stout, 2004), and the use of bone mass in isolated skeletal elements (such as the metacarpal) to infer fragility at distant biomechanically sensitive sites or overall skeletal fragility needs to be interpreted cautiously.

Another significant concern in the case of methods that measure bone mineral content in archaeological bone is diagenesis (see Fig. 12.2). Absorptiometric methods such as DEXA offer noninvasive diagnosis in archeological specimens; the method is accepted generally as the gold standard for the assessment of bone loss in the clinical setting (Bates et al., 2002; Cummings et al., 2002). However, several investigators have questioned the reliability of bone mineral density measurements in archeological bone (Agarwal and Grynpas, 1996; Kneissel et al., 1994; Mays, 1996). Researchers
have assumed that if there are no highly unusual differential patterns of bone density among bones that are scanned (Bennike and Bohr, 1990) or if patterns of bone loss match what is expected from observations of modern populations (Mays et al., 1998, 2006), then diagenetic change is unlikely. However, diagenetic changes in bone mineral can occur in random bones and just as likely differentially modify entire samples. Although radiographic assessment of bone specimens before scanning can exclude samples with obvious inclusions, it is impossible to determine the extent of diagenetic alteration without invasive or chemical analysis of all or most samples (Brickley and Agarwal, 2003).

Finally, most analyses of bone loss in the past have focused on the examination of bone mass or BMD. This focus is in large part from the availability and comparability of well-established clinical methods such as metacarpal radiogrammetry or DEXA scan of BMD. However, it is now acknowledged that aspects of bone quality (such as bone material properties, architecture, and organization) contribute significantly and perhaps more to bone strength and fragility than bone mass and quantity (see Box 12.1). Nevertheless, there have been very few studies of bone quality in past populations. Material properties such as bone mineralization or microcracks cannot be examined in archaeological bone because of decomposition and burial; however, one aspect of bone quality that can be examined in archaeological bone is trabecular architecture. Few studies have been undertaken of trabecular architecture in past populations (Agarwal et al., 2004; Brickley and Howell, 1999; Kneissel et al., 1994, 1997; Vogel et al., 1990). It is interesting that although these studies have all shown age-related loss of trabecular structure or connectivity in archaeological populations, the age-and sex-related patterns of trabecular architecture differ as compared with modern populations. As mentioned, Agarwal et al. (2004) found age-related change in trabecular microarchitecture in the Wharram Percy sample to occur in young and middle-age groups with no change observed in old age and no significant sex difference. This is similar to Vogel’s et al. (1990) early study that reported less-pronounced age-related loss of bone volume and trabecular connectivity as compared with modern samples. Subsequent studies of Nubian vertebrae by Kneissel et al. (1997) did find significant loss of trabecular structure and connectivity in early age. However, they found no significant loss in the older age groups encompassing individuals 30–60 years of age (Kneissel et al., 1997). A study by Brickley and Howell (1999) of mean trabecular length in a British postmedieval sample also note

Figure 12.2 Scanning electron microscope image (in backscattered electron imaging mode) of human rib sections. Left: modern rib. Right: archaeological rib sample. Note the modern rib shows varying gray levels that correlate with normal differing areas of remodeled bone tissue and subsequent mineralization ages. The archaeological rib shows significant diagenetic destruction of both microstructure and chemical alteration of mineralization levels. (Images courtesy of Sabrina C. Agarwal.)
age-related loss of structure to occur primarily in young-age females and loss to be less marked in the oldest female age group. These patterns may reflect the differential response observed in trabecular tissue as compared with the many studies of age-related bone loss that have focused on cortical bone, and they highlight the need for additional studies of bone quality. It is possible that the low prevalence and unusual patterns of osteoporotic fracture in archaeological populations as compared with modern populations may be an accurate record of bone fragility in the past that is better revealed with measures of bone quality. Although we cannot perform direct measurement of biomechanical strength from dried archaeological bone samples, measures of bone quality such as trabecular architecture, bone geometric properties, or qualitative analysis of cortical histomorphology can offer a more complete picture of factors that could have influenced bone fragility more strongly than bone quantity alone.

CONCLUSIONS: WHAT DO THE PATTERNS OF BONE LOSS IN THE PAST MEAN?

Although recent studies have found age-related bone loss in past populations, the suggestion that the patterns and prevalence of osteoporosis were the same as in modern populations (Mays et al., 1998; Rosen, 1999; Tuner-Walker et al., 2001) is not entirely correct. Studies that have observed bone loss in past populations often observe loss in young-age individuals (for example, Agarwal et al., 2004; Brickley, 2002; Ekenman et al., 1995; Holck, 2007; Mays et al., 2006). It is unclear whether this finding is from the nature of mortality sample demographics such as the heterogeneity of older age groups, problems with age at death determination, and/or or loss in young-age females reflecting transient reproductive stress. However, bone loss is also observed in young-aged males as well, which is an observation not typical in modern populations (Agarwal et al., 2004). Furthermore, the amount of bone loss is often observed to be similar in both sexes at various ages, including old age (for example, Agarwal et al., 2004; Brickley 2002; Ekenman et al., 1995; Holck, 2007). This finding contrasts dramatically from patterns in modern populations where females typically demonstrate greater bone loss and higher rates of fracture as compared with males. As discussed, the patterns of fragility fracture in past populations are also unusual. Often a low prevalence or absence of some fragility fractures is found in the archaeological record (Agarwal and Grynpas, 1996; Brickley, 2002; Brickley and Agarwal, 2003; Mays et al., 2006; Pfeiffer, 2000; Weaver, 1998), and it remains unclear in many populations when fragility fractures are reported (Mays, 2006) if prevalence rates and patterns of fracture are actually similar or different between the sexes. An a priori expectation seems to exist that the same patterns of age-related bone loss and fragility fracture observed in modern populations should be found in archaeological populations (Agarwal et al., 2004; Weaver, 1998). However, as observed in modern human populations, population- and environmental-specific conditions simply may have preserved bone mass or contributed to a low prevalence of fragility fracture in some earlier human groups. We know that the indicators of bone turnover and maintenance that are studied by paleopathologists (whether indicators of bone quantity or quality) are influenced by a host of factors, including growth, biomechanics, genetics, and lifestyle. Bone maintenance and fragility in living human populations is clearly not tied to aging alone, as demonstrated by the widely increasing and different prevalence rates of osteoporosis around the world. As such it should not be surprising that past populations show varying patterns of bone loss and fragility that will not always follow expected temporal or regional expectations. Although researchers interested in bone loss in archaeological populations recognize the multifactorial etiology of bone maintenance, the expectation still exists that changes in bone (micro)structure or morphology should be associated most tightly
with aging (for example, Macho et al., 2005). When investigating osteoporosis in women in the past, it must also be kept in mind that bone maintenance and fragility in older women is the compound product of both age-related and postmenopausal osteoporosis. Furthermore, the long-term maintenance and strength of the older female skeleton is the product of more than just the menopausal loss of estrogen, but the life history of the skeleton, including aspects of growth, nutrition, activity, and reproduction.

So how might we better understand what the unique patterns of bone loss and fragility mean in prehistoric and historic populations? It is clear from current biomedical and clinical research that multiple influences on bone maintenance need to be considered simultaneously if we are to interpret bone loss in the past. It is interesting to note that early bioarchaeological studies tended to interpret patterns of bone loss primarily with nutritional hypotheses, whereas more recent studies in the past two decades have observed a shift toward an emphasis on postmenopausal estrogen loss and the influence of biomechanical stress on bone maintenance. This temporal trend matches that in biomedicine and what we have learned about bone biology in recent years. However, future paleopathological studies may be better informed by integrating information on the influences on bone maintenance to begin to recognize the perhaps synergistic influence of multiple factors. The use of multiple methods also seems to be increasingly necessary to obtain a full picture of bone maintenance and fragility in the past. Skeletal sites clearly differ in their tissue response to bone turnover, and measures of bone morphology need to take this into consideration. Furthermore, the unique results that have been gained from studies of bone quality in archaeological samples so far emphasize the need to examine other aspects of bone quality (such as trabecular architecture and bone geometric properties) along with traditional measures of bone quantity in past populations. Finally, although the emphasis is on osteoporosis bone loss in old age, we are beginning to realize that our understanding of long-term fragility, in both past and present populations, can be better achieved through the understanding of bone maintenance throughout the lifecycle (Agarwal and Stuart Macadam, 2003; Fausto-Sterling, 2005; Weaver, 1998). To do this, we need to make the shift to thinking about and investigating bone maintenance in whole paleopopulations—all sexes, genders, and age groups. With these developing directions, the study of bone loss and fragility remains an exciting and rapidly growing field in bioarchaeology that can reveal much about the life histories of ancient populations.

REFERENCES


REFERENCES


PART IV

CHEMICAL AND GENETIC ANALYSES
OF HARD TISSUES
CHAPTER 13

STABLE ISOTOPE ANALYSIS: A TOOL FOR STUDYING PAST DIET, DEMOGRAPHY, AND LIFE HISTORY

M. ANNE KATZENBERG

INTRODUCTION

This chapter is about chemical variation in the various components that make up bones and teeth and its application to studies of diet, demography, and life history. Bones and teeth provide direct evidence of past diets, including infant diets. Knowledge gained from bone chemistry relates to other sources of evidence for diet and, in turn, the interaction of diet, nutrition, and disease. Understanding infant diets and the duration of nursing relates to demographic variables such as birth spacing and population growth. Chemical variation in bones and teeth can also be linked to chemical variation in the environment and, thus, reveals information about place of residence and migration. Although stable isotope analysis may be viewed as a fairly technical research specialty, the results of such analyses make a significant contribution to the reconstruction of past human life.

Uses of Stable Isotope Analysis in Skeletal Biology

The routine use of stable carbon isotopes in current skeletal studies is very different from the excitement generated by the first applications of stable carbon isotopes to human paleodiet reconstruction. The idea that one could determine whether prehistoric peoples of North America consumed corn (maize) by performing chemical tests on their bones seemed like science fiction to most archaeologists and physical anthropologists in the mid-1970s when it was first attempted (Vogel and van der Merwe, 1977). Today, stable carbon isotope analysis is part of a suite of technical specialties performed on remains from archaeological sites. In addition to studying stable carbon and nitrogen isotopes in preserved protein, it is now possible to study stable isotopes of carbon, oxygen, and strontium from the mineral portion of bones and teeth. This chapter provides some background to the use of stable isotopes in bioarchaeological studies. It includes technical information on how such analyses are performed, with a sampling of applications, problems, and finally, promises for the future. This chapter is not intended to be a review of the now vast literature on methods and applications of stable isotope analysis in studies of past peoples. Reviews of the earlier literature are available elsewhere (Katzenberg and Harrison, 1997; Pate, 1994;
Developments in Stable Isotope Analysis

From the perspective of the archaeologist, the first stable isotope studies of past diet were carried out in the 1970s (DeNiro and Epstein, 1978, 1981; van der Merwe and Vogel, 1978; Vogel and van der Merwe, 1977). However, studies of stable isotopes began in the early years of the twentieth century in the laboratories of chemists and physicists. After the discovery of stable isotopes in 1913, improvements in instrumentation and intensive study resulted in the identification of most stable isotopes by the mid-1930s. The first commercial mass spectrometer was used to analyze petroleum in 1942 (Gross, 1979). Throughout the 1950s and 1960s mass spectrometers and the applications of stable isotope studies advanced rapidly in chemistry, biology, and geochemistry. Efforts were directed toward understanding variation in the relative abundances of stable isotopes of the various elements. For example, geochemists explored oxygen isotope variation and its potential for studies of past climate (reviewed by Luz and Kolodny, 1989). Major advances in understanding stable isotope variation in the biosphere and geosphere occurred during the 1950s and 1960s. Botanists and geochemists explored stable carbon isotope variation in plants (Craig, 1954; Smith and Epstein, 1971), and researchers in radiocarbon dating laboratories shared their interest (Bender, 1968; Hall, 1967; Lowdon, 1969) since this variation is relevant to radiocarbon dating methods.

Along with advances in understanding the processes that cause variation in stable isotope abundance ratios in different substances, major advances have occurred in instrumentation. Although improvements in resolution, detection, and overall design of mass spectrometers occurred throughout the twentieth century, advances in the last 15 years have had an enormous impact on the use of stable isotope methods because of the ability to run more samples at a much lower cost. Stable isotope analysis used to be a time-consuming and, therefore, expensive method. Preparation of samples, isolation of gasses containing specific elements of interest, and the actual analysis and corrections to standards were very laborious processes. Only a small number of samples per day could be analyzed, and analyses required constant attention by lab personnel. In the late 1980s new instrumentation was developed that simplifies sample preparation, automates introduction of samples into the system, requires much smaller samples, is much faster, and therefore is much less expensive. This development has opened up many more applications of stable isotope analysis, most notably in ecological research (Barrie et al., 1989; Fry, 2006; Griffiths, 1998; Rundel et al., 1989). It has also made it possible for many more samples to be analyzed in archaeological studies. Instead of selecting a few human bones for analysis, as was true of many studies carried out before 1990, researchers now routinely analyze nonhuman faunal bone and either prehistoric or modern plants to compare potential foods with human stable isotope data. These newer methods have allowed researchers to analyze many more human samples, revealing previously unknown variation within populations, including age differences and individuals who are outliers with respect to their chemical signatures. Stable isotope methods have now been applied to studies of demography, residence patterns, and disease in addition to studies of diet. Although newer instruments simplify analyses, it is imperative to have
trained personnel in reputable laboratories running the instruments to ensure good accuracy and precision.

**History of Applications to Analysis of Past Peoples**

The realization that stable isotopes of carbon could be used to investigate past diets can be traced to two different but related fields of study (reviewed by van der Merwe, 1982). Scientists working to determine $^{14}$C dates on ancient organic remains noted variation in dates derived from some human skeletal remains. Also, maize from archaeological sites gave anomalous dates relative to wood charcoal (Bender, 1968; Hall, 1967). These findings coincided with research on different biochemical pathways of photosynthesis among plants (Smith and Epstein, 1971). Maize fixes carbon by a different pathway and, as a result, contains more $^{13}$C relative to $^{12}$C than most other plants from temperate regions. Since samples for radiocarbon dating were assumed to have the same stable carbon isotope composition as the standard, those samples containing relatively more $^{13}$C, such as maize cobs and kernels, gave erroneous dates. In addition to its use in differentiating human consumption of plants with different photosynthetic pathways, carbon isotopes have been shown to differentiate marine from terrestrial-based diets in humans (Chisholm et al., 1982; Tauber, 1981).

Carbon was the first element for which stable isotope variation was used in archaeology, which follows from archaeologists’ familiarity with radiocarbon. Once the potential of studying stable carbon isotopes in preserved protein was understood, interest in other elements such as nitrogen, oxygen and sulfur flourished. Each of these elements and their stable isotopes have been studied extensively in geological and ecological systems as well. In fact, archaeologists are relative latecomers to the study of stable isotope variation of the elements, with the exception of carbon.

The second element to be used in paleodiet research was nitrogen. DeNiro and Epstein (1978, 1981) carried out controlled feeding experiments on several species to study the relationship between the stable carbon and stable nitrogen isotope ratios in diet and in animal tissues. Shortly after that, DeNiro, working with two postdoctoral researchers, explored trophic level and regional variation in nitrogen isotopes (Schoeninger and DeNiro, 1984) and trophic level variation and dietary differences in east Africa (Ambrose and DeNiro, 1987).

Carbon and nitrogen stable isotopes are the isotopes most commonly studied in human remains. More recently, stable oxygen isotopes and strontium isotopes have been studied in bone and in tooth enamel. Stable isotopes of sulphur have been studied in hair and to a lesser degree, bone, since hair keratin contains more sulphur than bone collagen, which contains very little sulphur.

**Basic Concepts of Stable Isotope Variation**

Isotopes are atoms of the same element with the same number of protons but different numbers of neutrons. Since the atomic mass is determined by the number of protons and neutrons, isotopes of an element vary in their masses. Table 13.1 shows some of the chemical elements that have several isotopes and the abundances of those isotopes. In contrast to unstable (radioactive) isotopes, stable isotopes do not decay over time. For example, $^{14}$C in a dead organism decays to $^{14}$N, whereas the amounts of $^{12}$C and $^{13}$C in the same organism will remain constant.

In chemical reactions, such as the conversion of atmospheric CO$_2$ into glucose by plants, the relative amounts of $^{12}$C and $^{13}$C differ in plant tissue relative to that of atmospheric CO$_2$. This variation is due to the fact that isotopes vary in mass and therefore have slightly different chemical and physical properties. Isotopes with higher mass (heavier isotopes) such as $^{13}$C usually react slightly more slowly than lighter isotopes such as $^{12}$C. Physical phenomena that occur during chemical reactions as a
result of the mass differences in isotopes are referred to as “isotope effects.” The resulting difference in the isotope ratio of the carbon in the plant tissues as compared with the carbon in atmospheric CO$_2$ caused by isotope effects is termed “fractionation.” Fractionation is the basis for stable isotope variation in biological and geochemical systems, and gaining an understanding of the chemical reactions that result in stable isotope variation allows the biological anthropologist, archaeologist, geochemist, or ecologist to put stable isotope analysis to work to solve a wide range of interesting problems. More detailed discussions of isotope effects and fractionation can be found in textbooks such as that by Hoefs (1997) and Fry (2006), with the latter providing an entire chapter on fractionation.

**Tissues Used in Stable Isotope Studies**

The first tissue to be used in archaeological stable isotope studies of human paleodiet was the collagen of bone. Methods for isolating collagen had already been developed in radiocarbon dating labs since collagen was used in dating. Information on isolating collagen from bones and teeth is provided by Ambrose (1990), who critically reviews the various methods available.

Bone is composed of an organic matrix of the structural protein, collagen, which is studded with crystals of calcium phosphate, largely in the form of hydroxyapatite. Dry bone is approximately 70% inorganic and 30% organic by weight. Most of the organic portion (85–90%) is collagen. The remainder includes noncollagenous proteins, proteoglycans, and lipids (Triffit, 1980). Because of the intimate structural relationship between collagen and hydroxyapatite, collagen may survive for thousands of years (Tuross et al., 1980), and protein that is probably degraded collagen has even been recovered from dinosaur fossils (Wyckoff, 1980). Collagen contains approximately 35% carbon and 11–16% nitrogen by weight (van Klinken, 1999), and it is the tissue of choice for stable carbon and nitrogen isotope analysis. The detection of postmortem degradation of collagen has been an active area of research (e.g., Child, 1995; Collins et al., 1995, 2002; DeNiro, 1985; Hedges, 2002; Schoeninger et al., 1989; Tuross, 2002).

Because collagen does degrade over time and at varying rates depending on the burial environment, researchers have sought other sources of carbon that are representative of lifetime carbon intake. Another biological source of carbon in bones and teeth is in the form of carbonate (CO$_3$), which occurs in the mineral portion of bone. Bone mineral is largely composed of hydroxyapatite, Ca$_{10}$(PO$_4$)$_6$OH$_2$. However, several ions can substitute for the constituent ions of the hydroxyapatite crystals (see Burton, this volume). Well known from the trace element literature are the substitutions of Sr$^{++}$ (strontium) or Pb$^{++}$ (lead) for Ca$^{++}$ (calcium). Another common substitution is CO$_3$$^-$$^-$ for PO$_4$$^{3-}$ (phosphate) (LeGeros et al., 1967). Sullivan and Krueger (1981) proposed using the carbon in bone mineral for stable carbon isotope studies in fossil bone, when

<table>
<thead>
<tr>
<th>Element</th>
<th>Isotope</th>
<th>Abundance (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hydrogen</td>
<td>$^1$H</td>
<td>99.985</td>
</tr>
<tr>
<td></td>
<td>$^2$H</td>
<td>0.015</td>
</tr>
<tr>
<td>Carbon</td>
<td>$^{12}$C</td>
<td>98.89</td>
</tr>
<tr>
<td></td>
<td>$^{13}$C</td>
<td>1.11</td>
</tr>
<tr>
<td>Nitrogen</td>
<td>$^{14}$N</td>
<td>99.63</td>
</tr>
<tr>
<td></td>
<td>$^{15}$N</td>
<td>0.37</td>
</tr>
<tr>
<td>Oxygen</td>
<td>$^{16}$O</td>
<td>99.759</td>
</tr>
<tr>
<td></td>
<td>$^{17}$O</td>
<td>0.037</td>
</tr>
<tr>
<td></td>
<td>$^{18}$O</td>
<td>0.204</td>
</tr>
<tr>
<td>Sulfur</td>
<td>$^{32}$S</td>
<td>95.00</td>
</tr>
<tr>
<td></td>
<td>$^{33}$S</td>
<td>0.76</td>
</tr>
<tr>
<td></td>
<td>$^{34}$S</td>
<td>4.22</td>
</tr>
<tr>
<td></td>
<td>$^{36}$S</td>
<td>0.014</td>
</tr>
<tr>
<td>Strontium</td>
<td>$^{84}$Sr</td>
<td>0.56</td>
</tr>
<tr>
<td></td>
<td>$^{86}$Sr</td>
<td>9.86</td>
</tr>
<tr>
<td></td>
<td>$^{87}$Sr</td>
<td>7.02</td>
</tr>
<tr>
<td></td>
<td>$^{88}$Sr</td>
<td>82.56</td>
</tr>
</tbody>
</table>

Extracted from Table 1.1, Ehleringer and Rundel, 1989.
collagen was too badly degraded to be used. This proposal was challenged by Schoeninger and DeNiro (1982), and a debate in the literature followed (reviewed by Krueger, 1991 and Lee-Thorp and van der Merwe, 1991). The challenge centered on whether bone carbonate was altered in the postmortem environment by exchange between constituents of buried bone and carbonates in sediments such that the carbon isotope ratios would not reflect lifetime carbon deposition. These concerns were addressed by Lee-Thorp (1989), who developed preparation methods to remove the more soluble carbonates, which are those most likely to be diagenetic in origin. More recent debate has centered on the use of carbonate in tooth enamel versus dentin and bone (Koch et al., 1997). There are two compelling reasons to pursue the use of carbonate in biological apatite as a source of carbon isotope ratios. It allows stable isotope studies to be applied to much older materials where it is no longer possible to isolate collagen (e.g., Lee-Thorp et al., 1989; Lee-Thorp and Sponheimer, 2006; Sponheimer and Lee-Thorp, 1999), and carbon from biological apatite records slightly different dietary information than does collagen.

The idea that the carbon in the carbonate of bones and teeth comes from different dietary components than the carbon in collagen was first proposed by Krueger and Sullivan (1984), Ambrose and Norr (1993) and Tieszen and Fagre (1993), in two separate controlled feeding experiments, demonstrated that Krueger and Sullivan were correct in suggesting that collagen carbon comes mainly from ingested protein in the diet, whereas the carbon in biological apatite reflects whole diet. The reason is that collagen is composed of a mixture of essential and nonessential amino acids. The essential amino acids come from ingested protein. The non-essential amino acids may come from ingested protein, or they may be formed from other dietary sources and breakdown products within the body. Carbonate in bone is formed from dissolved bicarbonate in the blood, and this comes from dietary carbohydrate, lipid, and protein. Therefore, the carbon in biological apatite provides a picture of the total diet, whereas collagen is more reflective of dietary protein.

The mineral portion of bone is also the source of oxygen and strontium used in isotope studies. Oxygen isotopes are most often isolated from PO₄ (Luz and Kolodny, 1989; Stuart-Williams and Schwarz, 1997) and have been used in paleoclimate and, more recently, paleodemographic studies. The oxygen in bone and tooth carbonate has also been analyzed for stable oxygen isotopes (Koch et al., 1997), and although it is a technically less demanding procedure relative to isolating phosphate, carbonate is more likely to be altered by diagenetic processes, more so for bone than for tooth enamel carbonate (see Garvie-Lok et al., 2004; Nielsen-Marsh and Hedges, 2000a, 2000b; for more detailed discussions of diagenesis and bone carbonate). Strontium is a common trace element in bone where it substitutes for calcium. Strontium isotopes have been used in paleodiet and residence studies (e.g., Bentley et al., 2005; Ericson, 1989; Ericson et al., 1989; Ezzo et al., 1997; Price et al., 1994a, 1994b; Sealy et al., 1991, 1995).

Methods for Isolating Specific Components for Stable Isotope Analysis

Specific instructions for isolating various components of bone can be found in the sources cited throughout this section. It is important to understand the chemical principals of each method. A wide range of variation exists in postmortem environments, duration of interment, and therefore the preservation of hard tissues. It is sometimes necessary to vary the methods for poorly preserved bone samples. For example, by diluting the acid solution, the process of dissolving the bone mineral is less harsh and proceeds more slowly so that partially degraded collagen may be recovered. Tropical environments often have poor preservation of bone with extensive collagen degradation; however, teeth may be better preserved, allowing extraction of collagen from dentine. Sometimes
postmortem alteration is so extensive that specific analyses are not possible.

**Collagen.** Most researchers use one of three methods for isolating collagen from bones and tooth dentine. Sealy (1986) describes a simple method in which small chunks of bone (1–3 g in total) are decalcified in a hydrochloric acid solution (between 1% and 5%, depending on the density of the sample and its outward appearance). An additional soak in sodium hydroxide (0.1 molar) may follow to remove decayed organic matter from the burial environment. The remaining collagen is freeze-dried.

In another method, (Bocherens et al., 1995; Tuross et al., 1988) small chunks of bone are demineralized in EDTA (ethylenediaminetetraacetic acid), a sodium salt, which separates collagen from bone mineral. A third method was originally developed by Longin (1971) and later modified by Schoeninger and DeNiro (1984) and by Brown et al. (1988). In this method, powdered bone is demineralized in 8% hydrochloric acid for a short period of time (around 18 minutes). This process is followed by a slow hydrolysis in weakly acidic hot water (pH 3). This method is preferable for poorly preserved bone; however, the risk is that one may obtain other organic matter in addition to collagen. Thus, an additional soak in sodium hydroxide (0.1 molar) to remove decayed organic matter is usually recommended.

Several researchers have compared methods and their yields (Chisholm et al., 1983; Schoeninger et al., 1989). Boutton et al. (1984) and Katzenberg (1989; Katzenberg et al., 1995) have demonstrated that some collagen is lost when demineralized bone is soaked in sodium hydroxide, but that the other material removed in the soak contains humic contaminants (decayed organic matter) that may skew $\delta^{13}C$ values. For example, Katzenberg et al. (1995) demonstrated that the residue removed during the sodium hydroxide soak of prehistoric human bones from southern Ontario had a much lighter $\delta^{13}C$ than the collagen, indicating that the residue contained decayed C$_3$ plant remains.

It is essential to demonstrate that the material being analyzed is collagen, and various criteria have been used for this purpose. DeNiro (1985) proposed that extracted material with a carbon-to-nitrogen (C/N) ratio in the range of 2.9 to 3.6 should preserve reasonable stable isotope ratios to those from the lifetime of the organism. Since the ratio of carbon to nitrogen in modern (never buried) bone collagen is 3.2, this range was suggested to take into consideration analytical error and some slight alteration over long periods of time. These figures are based on the atomic ratio of carbon to nitrogen in collagen. Modern light isotope mass spectrometers are usually interfaced with gas analyzers that provide data on the content of carbon and nitrogen in samples, and they also calculate carbon-to-nitrogen ratios. However, these are weight ratios and will be slightly lower (by a factor of 1.16667) than the atomic ratios given by DeNiro (1985, and see Ambrose, 1990). Recent publications often include data on percentage carbon and nitrogen in samples as well as the C/N ratios. These data are important for evaluating the validity of stable isotope results. For example, recent samples that contain lipids will have more carbon and erroneous stable carbon isotope ratios since lipid is less enriched in the heavier isotope than is collagen. Samples in which collagen is degraded often have low nitrogen content and may provide results for stable carbon isotopes but not stable nitrogen isotopes, since collagen contains much more carbon than nitrogen (thus the expected 3.2 ratio). In all of these procedures, the objective is to isolate collagen from bone mineral and any organic matter introduced in the postmortem environment.

**Compound-Specific Analyses.** With advances in the field, researchers have focused their attention on isolating individual amino acids, initially for accelerator radiocarbon dating (Stafford et al., 1991). Stable isotope values vary among the different amino acids, so preferential loss of certain amino acids from diagenesis can alter the overall $\delta^{13}C$ of a protein such as collagen (Hare and Estep, 1982).
Interest also exists in isolating the indispensable (essential) amino acids (those amino acids that must be obtained from the diet) from collagen for stable isotope analysis for dietary reconstruction (Hare et al., 1991). Since these amino acids come from dietary protein and are incorporated into human proteins such as collagen, they provide a more direct tracer than bulk collagen, which is made up of both dispensable amino acids (those that may be synthesized by the organism) and indispensable amino acids. Methods used for such study are much more complex than those for simply isolating collagen and are described by Stafford et al. (1991) and Hare et al. (1991). There is considerable promise in pursuing these methods, particularly with the development of GC/C/IRMS (gas chromatography/combustion/isotope ratio mass spectrometry) (Lichtfouse, 2000; Macko et al., 1997). This method allows the researcher to isolate specific organic compounds and then to introduce them into the mass spectrometer. Evershed pioneered this research in paleodiet studies by focusing on lipids and demonstrating the presence of specific substances in residues from prehistoric ceramic vessels (Evershed, 1993; Stott and Evershed, 1996), including dairy fats (Copley et al., 2005a, 2005b). Corr et al. (2005) have demonstrated that compound-specific stable carbon isotope analysis of collagen amino acids allows one to differentiate diets with marine protein and C4 plants in arid regions where \( \delta^{15}N \) values are unpredictable. Students interested in graduate study in this area are well advised to become familiar with the principles and methods of organic chemistry and biochemistry.

**Biological Apatite.** A method for isolating the carbonate fraction of bone mineral was developed by Lee-Thorp (1989; Lee-Thorp and van der Merwe, 1991; Lee-Thorp et al., 1989). Ground bone is soaked in sodium hypochlorite to remove organic material. Carbonate adsorbed from the burial environment is removed with one molar acetic acid. Samples are then reacted with phosphoric acid to release the structural carbonate. \( \text{CO}_2 \) is collected by cryogenic distillation. Specific steps are described in the references cited above and by Tieszen and Fagre (1993) and Ambrose and Norr (1993).

More recently, carbon in the carbonate fraction of tooth enamel has been analyzed using laser ablation stable isotope analysis (Sponheimer et al., 2006). The use of the laser allows sampling of such small quantities that sampling is nearly imperceptible for purposes of long-term curation, and seasonal variation can be explored, as was done by Sponheimer et al. (2006) on teeth from *Paranthropus robustus*.

Oxygen isotope measurements in bone usually make use of the oxygen in phosphate, which is less affected by diagenetic processes than carbonate. Stuart-Williams (1996; Stuart-Williams and Schwarcz, 1995, 1997) developed a method of isolating organic phosphate from bone that is simpler and safer than previous procedures. Bryant et al. (1996) compared the oxygen isotope results from phosphate and carbonate of tooth enamel and demonstrated that it is possible to estimate one from the other.

**Hydroxyapatite.** Strontium substitutes for calcium in the hydroxyapatite crystals of bone mineral. Two different methods of isolating bone mineral for analysis of strontium isotopes have been described in the recent literature. The method used by Sealy et al. (1995) and by Sillen et al. (1998) makes use of Sillen’s 1986 solubility profile method. Bone or tooth powder is washed in acetic acid and sodium acetate buffer solution repeatedly, saving each wash. The various washes are analyzed by ICP (inductively coupled plasma emission spectrometry). The first few washes presumably contain recently deposited contaminants from the burial environment and show variation in trace element concentrations. Later washes tend to show less variation in the concentration of strontium and other trace elements. It is these later washes, which are thought to contain the biologically deposited strontium, that are used for strontium isotope analysis by mass spectrometry.

A second method of preparation for strontium isotope analysis, described by Price and colleagues (1994a, 1994b) begins with
mechanical cleaning of the outer surface of bone, followed by an overnight soak in one normal acetic acid. The acid removes soluble carbonates and the portion of bone most likely to contain elements from the burial environment. The residue is then wet-ashed in nitric acid in preparation for mass spectrometry. Both methods attempt to isolate bone mineral that has not been diagenetically altered, and although Sillen’s method is more conservative, it is also more labor intensive. Sillen and Sealy (1995) have demonstrated that the solubility profile method does not result in any recrystallization of bone mineral, unlike methods that employ dry ashing (heating to high temperature to destroy the organic component).

**Lipids.** Methods for treating lipids from bone samples depend on whether one intends to remove the lipids so that only protein or carbonate is analyzed, or whether one wants to determine the δ13C of lipid. Lipids are less enriched in the heavier isotope (i.e., δ13C is more negative) than bone collagen. Therefore it is necessary to remove lipids from bone samples before analysis. This process is particularly important when analyzing bones of recent origin for comparative purposes. Liden et al. (1995) discuss methods of removing lipids. The most commonly used method involves soaking the bone in a mixture of chloroform and methanol (Bligh and Dyer, 1959; Folch et al., 1957) after demineralization. The residue must be rinsed carefully since these are organic solvents, which may contaminate the sample. Bone may also be soaked in diethyl ether before demineralization. The normal sodium hydroxide soak then follows demineralization (Ambrose and Norr, 1993). These methods will effectively eliminate lipids from collagen preparations. Bligh and Dyer (1959) describe a method for extracting and purifying lipid from biological materials. Recent work by Evershed and colleagues (Evershed, 1993; Stott and Evershed, 1996) on characterization of lipid extracts from ancient materials, such as ceramic vessels, has added another source of evidence for paleodiet studies by allowing δ13C determinations of lipids and, more recently, of specific compounds within lipids (see the previous section on compound-specific analyses).

**Mass Spectrometry**

Stable isotope abundance ratios are measured in isotope ratio mass spectrometers (IRMS), which should not be confused with organic mass spectrometers that are used to characterize complex organic molecules. Isotope ratio mass spectrometers have four components: an inlet system, an ion source, a mass analyzer, and a series of ion detectors (Fig. 13.1). For most elements of interest (H, O, N, C), the sample is introduced to the mass spectrometer as a gas (H2, CO2, N2, and CO2, respectively). Until recently, most stable isotope work performed with collagen or carbonate required combustion of the sample in sealed tubes. After combustion, the resultant CO2 and H2O were separated offline before the CO2 was let into the mass spectrometer. Modern instruments now interface combustion furnaces and gas analyzers with mass spectrometers to simplify and ease the conversion of the sample into the requisite gaseous form. In such a setup, collagen is weighed into tin sample holders, which are then placed into an automated sample tray. The revolving tray drops samples into the furnace where N2, CO2, and H2O are produced. These gases, carried by helium carrier gas, are separated before being swept into the mass spectrometer. (For a detailed presentation of continuous-flow stable isotope analysis, see Barrie et al., 1989 and Barrie and Prosser, 1996.)

Once in the mass spectrometer, the gas of interest is let into the second component of the mass spectrometer, the ion source. In the ion source, some gas molecules are ionized by electron bombardment, which allows them to be controlled and focused into a beam. The ion beam is then directed, via a flight tube, into the mass analyzer zone of the mass spectrometer. As the name suggests, the mass analyzer separates the ion beam into several smaller beams by passing it between the poles of a magnet. This process is directly analogous to the separation
of white light into its constituent wavelengths through a prism. The separation of one ion beam into several beams according to mass results in the desired “mass spectrum.” The beam intensities of the respective ion beams can then be measured in the ion collector section of the instrument. The relative intensities of the individual isotope ion beams are then reported as isotope ratios, for example, \(^{13}\text{CO}_2:\^{12}\text{CO}_2\). To report a meaningful value, the mass spectrometer alternately analyzes aliquots of the unknown sample and a known standard gas, thereby providing a ratio of the stable isotopes in the sample relative to that same ratio in the standard.

Because the element of interest cannot always be converted into an easily handled gas, it is sometimes necessary to introduce the sample in solid form. For strontium isotope analysis, the sample is deposited directly on a filament near the ion source where it is heated to evaporation and ionized under vacuum. This process is referred to as thermal ionization mass spectrometry (TIMS) and requires a different ion source (solid source) for analysis. Many labs will have separate instruments for analyzing gasses and solids. More recently, laser ablation ICP-MS (inductively coupled plasma–mass spectrometry) has been used for pinpoint sampling to analyze isotopes of heavier elements such as strontium (Latkoczy et al., 2001; Prohaska et al., 2002).

For compound-specific analyses, a gas chromatograph (GC) can be interfaced with an IRMS resulting in GC-C-IRMS (gas chromatography combustion isotope ratio mass spectrometry). This configuration allows one to determine the isotope ratio of a specific compound, such as cholesterol, that has been isolated using gas chromatography (Stott et al., 1999; Tripp and Hedges, 2004).

It is helpful to know enough about mass spectrometry to be able to discuss one’s needs with various laboratory personnel at the beginning of a research project. It is important to be able to understand problems that may develop when the data have been collected. Considerations include the general composition of the sample and the range of the expected results. For example, departures from the expected range of carbon to nitrogen in bones and teeth may signal poor preservation or contamination. It is necessary to have some idea of the range of isotope compositions expected since the standards used should bracket the expected values.

Mass spectrometers are analytical instruments that have many uses in chemistry, biochemistry, geochemistry, and ecology. Instruments vary in their setup depending on the needs of the researcher. A laboratory may be set up to perform analyses on certain types of samples and may specialize in certain elements. Laboratories carrying out ecological research are most likely able to accommodate most analyses of interest to the archaeologist and physical anthropologist interested in light isotopes (hydrogen, carbon, nitrogen, oxygen,
and sulphur), whereas geochemists are more likely to have TIMS and laser ablation setups. Organic chemists and biochemists may have GC-C-IRMS capabilities, and enterprising biological anthropologists and archaeologists may have their own instruments. Given the high cost of mass spectrometers and the need for full-time technical support, most institutions have shared facilities that serve the needs of multiple users.

**Standards, Precision, and Accuracy**

Stable isotope abundance ratios are determined relative to the ratios of those same isotopes in standard materials. The mass spectrometer compares the stable isotope abundance ratio in the sample with the stable isotope ratio in a standard. Thus, the reported value uses the following notation:

\[
\delta \text{ in } \%e = \frac{R_{(\text{sample})} - R_{(\text{standard})}}{R_{(\text{standard})}} \times 1000
\]

where \( R \) is the ratio of the number of heavier to lighter isotopes, so that for carbon isotopes, the equation is

\[
\delta^{13}C_{\%e} \text{ PDB} = \left[ \frac{^{13}C/^{12}C_{\text{sample}} - ^{13}C/^{12}C_{\text{standard}}}{^{13}C/^{12}C_{\text{standard}}} \right] \times 1000
\]

The \( \%e \) (permil sign) means “per thousand” since the ratio is multiplied times 1000.\(^1\) This calculation is done to amplify the difference between the ratio of the stable isotopes in the sample and the ratio of stable isotopes in the standard, which is usually a very small number.

\(^1\)The term “permil,” meaning per thousand, is similar to the term “percent,” meaning per hundred. In the isotope literature, using the delta notation, values are reported as some number permil (\( \%e \)). In the trace element literature, elements may be measured in parts per million (ppm). This measure is very different in that it is not a ratio and it refers to a much smaller proportion (out of one million). It is best to avoid using the phrase “parts permil,” which is sometimes heard from anthropologists. Although not incorrect, it is awkward and incongruous, since people do not say parts per cent. Standard terminology is “permil.”

International standards are available through the National Bureau of Standards (NBS) and the International Atomic Energy Agency (IAEA), Vienna. The circulation of these standards among laboratories allows comparison of results from different researchers working in different laboratories. However, because of the high cost, individual laboratories normally also have internal standards. These standards are substances whose isotopic ratio is well characterized relative to an international standard and that are run routinely with batches of unknowns to check for consistency in the instrument. Internal and reference standards are reported relative to a primary reference standard for a particular element, which by definition has a \( \delta \) value of 0. Absolute isotope abundances of some primary reference standards are available in textbooks such as that of Hoefs (1997) and Fry (2006). Primary and other reference standards for elements discussed in this chapter are listed in Table 13.2.

The sensitivity of mass spectrometers varies, and it is important to know the precision of the instrument used before making interpretations from the data. Most light isotope mass spectrometers can measure \( \delta^{13}C \) values with a precision of \( \pm 0.1\%e \) and \( \delta^{15}N \) values with a precision of \( \pm 0.2\%e \). Newer models have improved sensitivity, although some newer continuous flow systems that simultaneously measure more than one element in a sample sacrifice precision for speed and economy of sample size. Precision should be determined for individual instruments using multiple analyses of samples with similar composition to those of interest. Data should be reported in a manner that is consistent with the precision of the instrument. If the precision is \( \pm 0.2\%e \), then it is incorrect to report results to 0.01\%e.

**APPLICATION OF STABLE ISOTOPE ANALYSIS TO SELECTED PROBLEMS IN SKELETAL BIOLOGY**

During the 1990s, with the growing use of stable isotopes to address problems in biological
anthropology and archaeology, several review articles were published. Katzenberg and Harrison (1997) discuss developments and review the literature since 1989. Other reviews, which also cover basic concepts in stable isotope studies, include Schwarcz and Schoeninger (1991), Schoeninger and Moore (1992), and Pate (1994). A series of conferences entitled “Advanced Seminars in Paleodiet” has resulted in several publications (edited books as well as guest-edited issues of journals) of papers presented at those conferences (Price, 1989; Sillen and Armelagos, 1991; Lambert and Grupe, 1993; Bocherens et al., 1999; Ambrose and Katzenberg, 2001; Koch and Burton, 2003). A symposium in honor of the work of Dr. Harold W. Krueger, held at the 2001 meetings of the Society for American Archaeology, was also published as a special issue of a journal (Ambrose and Krigbaum, 2003). Collectively, these volumes provide a detailed look at the history and developments in the use of stable isotopes for studying archaeological human remains. For a similar review of developments in research on diagenesis, a series of conferences have been held on bone diagenesis, and these are also published as special issues of journals (Schwarcz et al., 1989; Hedges and van Klinken, 1995; Bocherens and Denys, 1997; Fernandez-Jalvo et al., 2002).

Initially, most applications of stable isotope analysis to human remains were concerned with reconstructing diet. Subsequently other research questions have been addressed with stable isotope methods. These questions include determining the duration of breastfeeding, effects of disease processes, and determination of residence and migration patterns. The following sections highlight some of these applications in addition to more traditional approaches to paleodiet studies.

Paleodiet

**C$_3$ and C$_4$ plants.** Maize is one of several tropical grasses that fixes carbon by a different photosynthetic pathway (referred to as the Hatch–Slack or C$_4$ pathway) than most plants found in temperate regions. C$_4$ plants, which also include sorghum, millet, and sugar cane, adapt to heat and aridity by minimizing the amount of time that the leaf pores (stomata) are open, thereby minimizing water loss. These plants discriminate less against the heavier isotope, $^{13}$C, than do temperate plant species, which use the C$_3$ (Calvin) photosynthetic pathway. Atmospheric CO$_2$ has a $\delta^{13}C$ value of $-8\%$ today, but before the widespread burning of fossil fuels, the value was around $-7\%$. $^2$ C$_4$ plants range from $-9$ to $-14\%$, whereas C$_3$ plants range from $-20$ to $-35\%$ (Deines, 1980). The non-overlapping ranges of C$_3$ and C$_4$ plants provide the basis for using stable isotopes of carbon in preserved human tissue for revealing diet.

Several studies have been carried out to determine the difference between the $\delta^{13}$C of the diet and the $\delta^{13}$C value of various body

---

**TABLE 13.2 Primary (International) and Reference Standards for Selected Elements**

<table>
<thead>
<tr>
<th>Element</th>
<th>Primary Standard</th>
<th>Other Reference Standards</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hydrogen</td>
<td>Vienna Standard Mean Ocean Water (VSMOW)</td>
<td>V-GISP, V-SLAP, NBS-30</td>
</tr>
<tr>
<td>Oxygen</td>
<td>Standard Mean Ocean Water (VSMOW)</td>
<td>NBS 19, 20, 18, 28, 30, V-GISP, V-SLAP</td>
</tr>
<tr>
<td>Carbon</td>
<td>PeeDee Belemnite (VPDB)</td>
<td>NBS 18, 19, 20, 21</td>
</tr>
<tr>
<td>Nitrogen</td>
<td>Atmospheric Nitrogen (Air)</td>
<td></td>
</tr>
<tr>
<td>Sulfur</td>
<td>Canyon Diablo meteorite troilite (VCDT)</td>
<td></td>
</tr>
</tbody>
</table>


---

$^2$Fossil fuels are depleted in $^{13}$C, and their use has resulted in a decrease in atmospheric $\delta^{13}$C of approximately 1%. The value of $-7\%$ is used in archaeological research to reflect the conditions present during the lifetime of the individuals being studied.
tissues (Ambrose and Norr, 1993; Lyon and Baxter, 1978; Tieszen and Fagre, 1993; Vogel, 1978). Bone collagen δ^{13}C is approximately 5‰ greater than δ^{13}C of the diet, which is the basis for the expression, “you are what you eat +5‰.” Interestingly, this number was first suggested by van der Merwe and Vogel (1978) based on measurements of free-ranging large mammals and their diets and then was confirmed recently by Ambrose and Norr (1993). In controlled feeding studies of rats, the diet to collagen spacing was only 5‰ when the dietary protein, carbohydrate, and fats were all from similar sources (C_3 or C_4). In situations where the protein source differed from that of the carbohydrates and fats, the spacing value varied (Ambrose and Butler, 1997). Such a situation in humans might occur if people were consuming the meat of C_3 browsers such as deer and C_4 plants such as maize. In the controlled feeding experiments on rats, such a situation results in a diet to collagen spacing of less than 5‰. Thus, although the spacing factor of 5‰ may be used as a guide, it should not be taken as an absolute value.

After the demonstration that stable carbon isotopes in bone collagen could be used to document the consumption of C_4 plants such as maize against a background of C_3 plants (Vogel and van der Merwe, 1977; van der Merwe and Vogel, 1978), several other researchers applied these same principles to other regions where maize was the major introduced cultigen (e.g., Buikstra and Milner, 1991; Katzenberg et al., 1995; Larsen et al.; Schurr and Redmond, 1991; Schwarz et al., 1985). The method works very nicely in eastern North America where maize is the predominant and, in some places, the only C_4 plant consumed in any quantity. Follow-up studies took into consideration the fact that if the animals exploited by human groups consumed C_4 plants, then their tissues would be enriched in the heavier isotope and this would show up in human bone collagen carbon (Katzenberg, 1989). Subsequent work on the differential routing of protein and nonprotein nutrients to the synthesis of collagen reinforces the importance of understanding the diets of the animals consumed by people (Ambrose and Norr, 1993; Tieszen and Fagre, 1993).

Several regions of archaeological interest exist where stable isotope analysis was not attempted because it did not seem that such analysis could provide any additional information. If all potential food sources have similar stable isotope ratios, then little can be learned, or if the central research question focuses on a C_4 plant, such as maize, but there are other C_4 plants in the region, it will not be possible to distinguish when maize was first introduced. The application of stable isotope methods came somewhat later to the American Southwest, where maize made an early appearance, but there are also several indigenous C_4 and CAM plants (a third photosynthetic pathway with values intermediate to C_3 and C_4 plants) in addition to human exploitation of animals that consume C_4 plants. Nevertheless, researchers did tackle this more complex region (Benson et al., 2006; Coltrain et al., 2006; Decker and Tieszen, 1989; Katzenberg and Kelley, 1991; Matson and Chisholm, 1991; Spielmann et al., 1990) with useful results in terms of documenting the intensity of maize use.

Many recent studies include stable carbon isotope data from both bone collagen and bone carbonate or, in cases where collagen is not preserved, just bone carbonate. Since bone carbonate is thought to better reflect the whole diet (Ambrose and Norr, 1993; Tieszen and Fagre, 1993) and collagen preferentially reflects dietary protein, analyses of both components provide additional information. For example, Harrison and Katzenberg (2003) used stable carbon isotope data from bone apatite to detect small amounts of maize in the diet that do not show up in collagen, when it was first introduced into southern Ontario. Ambrose et al. (2003) could better differentiate the diets of high and low status burials from Cahokia (in Illinois) using stable carbon isotope ratios from both collagen and carbonate and stable nitrogen isotope ratios from collagen.
Marine Versus Terrestrial Based Diets. Stable carbon isotopes are also useful in studies of people inhabiting or exploiting coastal areas where it is possible to test hypotheses about the relative importance of marine and terrestrial foods in the diet (Blake et al., 1992; Chisholm et al., 1982; Hayden et al., 1987; Keegan and DeNiro, 1988; Lubell et al., 1994; Norr, 1991; Walker and DeNiro, 1986; Tauber, 1981). Such studies now span the globe and include research in the Arctic (Coltrain et al., 2004), Tierra del Fuego (Yesner et al., 2003), and the Pacific (Ambrose and Butler, 1997; Pate et al., 2001). The main source of carbon for marine organisms is dissolved carbonate, which has a $\delta^{13}C$ value of 0‰, whereas the main source of carbon for terrestrial organisms is atmospheric CO$_2$, which had a $\delta^{13}C$ value of 27‰ in pre-industrial times. Tauber (1981) and Chisholm and colleagues (1982, 1983) demonstrated that this 7‰ difference is reflected in mammals, including humans, feeding from these two different ecosystems. Potential dietary change in northern Europe from the Mesolithic Period to the Neolithic Period has been investigated using stable carbon and nitrogen isotopes. Richards et al. (2003) have suggested that use of marine foods stopped when domesticated plants and animals were introduced into Britain, based on $\delta^{13}C$ in bone collagen. Others working in northern Europe have argued that such an abrupt transition did not occur and that the exploitation of marine foods continued into the Neolithic Period (Liden et al., 2004; Milner et al., 2004). This debate provides a good illustration of both the promises and the limitations of stable isotope analysis (Hedges, 2004), and it emphasizes the need to use the results of stable isotope data in the larger context of other archaeological evidence (Milner et al., 2004).

Nitrogen Isotopes and Diet

Trophic Level Distinctions. At the same time that DeNiro and Epstein (1981) demonstrated that carbon isotope ratios of diet are reflected in the tissues of an animal, they carried out a study of the relationship of diet and tissues for stable isotopes of nitrogen. Nitrogen isotopes vary depending on trophic level. Atmospheric nitrogen (N$_2$) is the primary standard, and its value is set at 0‰. Some plants (legumes) have a symbiotic relationship with bacteria of the genus Rhizobium. The bacteria live in the roots and can fix nitrogen (combine it with other elements such as hydrogen or oxygen), thereby making it available to the plant (Brill, 1977). Other plants must get their nitrogen from decomposed organic matter, which breaks down to compounds such as ammonia (NH$_3$) or nitrate (NO$_3$). Legumes have $\delta^{15}N$ values closer to that of atmospheric nitrogen, whereas nonleguminous plants are more enriched in $^{15}N$ and therefore have higher $\delta^{15}N$ values. Herbivore $\delta^{15}N$ values are approximately 3‰ higher than the $\delta^{15}N$ of their diet; thus, herbivores consuming legumes will have lower $\delta^{15}N$ values than those consuming non-leguminous plants. Carnivore tissues are again enriched in the heavier isotope resulting in $\delta^{15}N$ values approximately 3‰ higher than their diet. This principal of enrichment through successively higher trophic levels, which was first pointed out by Minagawa and Wada (1984) and Schoeninger and DeNiro (1984), provides the basis for using stable nitrogen isotopes to infer trophic level. Variation occurs in the magnitude of the trophic-level effect in stable nitrogen isotopes both among different tissues within the same organism and among different taxa (Vanderklift and Ponsard, 2003). It is important to understand this variation when working with tissues other than bone collagen and with fauna other than mammals.

Ideally, a range of animals and plants from the environment under study is analyzed and humans are viewed relative to the other organisms in their environment. An example is provided in Fig. 13.2, which shows carbon and nitrogen stable isotope ratios for humans from several prehistoric sites in New Mexico relative to other mammals from the region (from Katzenberg and Kelley, 1991). Humans have
the highest $\delta^{15}$N and $\delta^{13}$C values. In this region, humans consumed maize as well as animals that fed on C$_4$ plants. The figure indicates that both jackrabbits and bison consumed some C$_4$ plants. Humans are approximately 3‰ higher than deer, antelope, and bison for $\delta^{15}$N. The elevated $\delta^{15}$N for deer and antelope relative to cottontail may reflect consumption of some legumes by cottontail, but it may also be related to habitat. Cottontail and deer prefer forested, moister habitats, whereas antelope and jackrabbit are adapted to open range habitat. Heaton et al. (1986) and Ambrose (1991) have demonstrated that $\delta^{15}$N is sensitive to climate and is elevated in arid regions. For this reason, Ambrose (1991) suggests that species from different ecosystems cannot be compared directly without considering the isotopic composition of the local food web. This caution has been verified in subsequent studies.

**Freshwater Resources**

Initially, it was assumed by archaeologists that freshwater fish had $\delta^{13}$C values similar to those of terrestrial C$_3$-consuming organisms. Little was known about stable nitrogen isotope ratios in freshwater systems. In 1989, Katzenberg explored this question by analyzing bones of freshwater fish from archaeological sites around the Great Lakes. It was discovered that freshwater fish exhibit a trophic level effect, which results in higher $\delta^{15}$N and slightly increased $\delta^{13}$C in carnivorous fish. Thus, it is possible to estimate reliance on fish in regions like the Great Lakes where they are an abundant resource. This reliance is particularly significant since fish are frequently underrepresented in the zooarchaeological record because of cultural practices and, in earlier excavations, because of methods of recovery. Cultural practices, such as filleting and drying fish where they are caught and then transporting the dried fillets back to the village, will result in underrepresentation of fish bones relative to their dietary importance. Excavations carried out without fine screening will miss bones of fish as well as other small skeletal elements.

Freshwater fish also exhibit more variation in $\delta^{13}$C than had been assumed. Several studies of freshwater ecosystems by ecologists have provided the background to understanding the sources of this variation (France, 1995; Hecky and Hesslein, 1995; Kiyashko et al., 1991;
Freshwater plants have numerous sources of carbon unlike terrestrial plants whose source is atmospheric CO2. In freshwater ecosystems, carbon comes from atmospheric CO2, CO2 in the water, bicarbonate and carbonate from rocks and soils, and organic carbon as waste and decomposition products from plants and animals living in the water (Zohary et al., 1994). The result is that fish living in different habitats within freshwater lakes display widely varying δ13C. This is illustrated in Fig. 13.3, which shows a reconstruction of the stable isotope ecology of the region around Lake Baikal, Siberia (from Katzenberg and Weber, 1999). The δ13C of fish bones ranges from −14.2 to −24.6‰ with higher δ13C in species inhabiting the shallow waters and lower δ13C for fish inhabiting the deeper, open waters of the lake (Katzenberg and Weber, 1999). The higher δ13C for some fish explains the observed variation in human bone collagen δ13C from the region. This finding is important since there are no C4 plants in the area and yet some human samples analyzed showed evidence of food that was enriched in the heavier isotope. In other words, stable isotope evidence from humans and their presumed food sources can indicate when something has been missed.

Figure 13.3 also illustrates the trophic-level effect in δ15N in both terrestrial and freshwater organisms. Large terrestrial herbivores have δ15N around 4‰ to 5‰. Humans from a large number of sites in the region range from 10.1‰ to 14.4‰. Fish and the freshwater seals of Lake Baikal vary according to their trophic position. Carp (Caras a.) are bottom-feeders and have the lightest δ15N. Lenok (Branchimystax l.) are highest in the littoral food web and have higher δ15N. Seals occupy the highest position in the freshwater food web with δ15N around 14‰.

Nitrogen Isotopes and Water Stress

Environmental variation in δ15N of plants has been demonstrated in coastal versus inland regions and in arid versus wetter regions.
(Heaton, 1987; Shearer et al., 1983; Virginia and Delwiche, 1982). Heaton et al. (1986) and Sealy et al. (1987) have also shown that $\delta^{15}N$ varies in animals of the same species from arid versus wetter regions. Sealy et al. (1987) point out the importance of recognizing reasons for elevated $\delta^{15}N$ in arid regions when those regions are also in close proximity to coastal resources, as is the case on the southern cape of Africa. Incorrect dietary interpretations for human samples can result if higher $\delta^{15}N$ is attributed to the use of marine resources without recognizing that the same $\delta^{15}N$ might result from consumption of water-stressed terrestrial animals. Ambrose (1991) has explored the physiologic basis for $\delta^{15}N$ variation in mammals living in arid regions. He and DeNiro (1986) proposed a model based on varying nitrogen loss in urea, which is excreted in urine. Urea is depleted in $^{15}N$ relative to diet. Under conditions of water stress, more urea is excreted relative to the total volume of urine, and therefore, more of the lighter isotope, $^{14}N$, is lost. Therefore, more $^{15}N$ is retained in the body where it is available for tissue synthesis. The result is that tissue $\delta^{15}N$ will increase under prolonged water stress conditions.

**Nitrogen Isotopes and Protein Stress**

Another cause of elevated $\delta^{15}N$, relative to expectations from diet, is protein stress. This stress is related to the model above in that insufficient protein intake results in the breakdown and reutilization of existing tissues in the body that are already enriched in $^{15}N$ because of preferential excretion of $^{14}N$. Hobson and colleagues, in two different studies on birds (Hobson and Clark, 1992; Hobson et al., 1993), have found that under conditions of nutritional stress, new protein is synthesized from the products of catabolism of existing protein. Katzenberg and Lovell (1999) have presented evidence that suggests this may be detected in humans. They compared $\delta^{13}C$ and $\delta^{15}N$ in normal versus pathological segments of bones from individuals of known medical history. One individual, who died of AIDS and experienced some new bone deposition because of osteomyelitis, showed elevated $\delta^{15}N$ in the diseased segment relative to the two unaffected segments of bone. This case suggests that the recently deposited bone collagen was synthesized, at least in part, from amino acids liberated from the catabolism of existing proteins in the body. White and Armelagos (1997) reported elevated $\delta^{15}N$ in bones from individuals with osteoporosis from an arid region of Sudan. They discuss water stress as a possible factor to explain the nitrogen isotope data, but it is possible that protein stress was also a factor. The use of stable isotopes to address questions in paleopathology is a relatively new area of inquiry that promises to add to our understanding of disease processes as well as our understanding of stable isotope variation in metabolism.

**Infant Feeding and Weaning Studies**

One of the big questions that carries through many approaches to studying past peoples from their skeletal and cultural remains is the timing and rate of population growth. Buikstra et al. (1986) presented an interesting hypothesis that draws together the adoption of agriculture, the development of thin-walled ceramic vessels, and therefore the ability to prepare cereal pap as a weaning food. The result would be earlier weaning (defined here as the introduction of non-milk foods rather than the complete cessation of breastfeeding) and thus decreased inter-birth intervals, increased population growth, and larger population size. This result is based on the fact that nursing suppresses ovulation so that supplementing mother’s milk with cereal gruel will result in a shorter period of infertility after childbirth. The hypothesis that weaning occurs sooner in agricultural societies was tested by Fogel and colleagues (1989) in a study that included stable nitrogen isotopes from fingernails of nursing mothers and their infants. They found a trophic-level increase in $\delta^{15}N$ (+2.4‰ on
average) of infant protein beginning shortly after birth (around three months) and decreasing when supplemental foods were introduced. By three to five months after nursing ceased, fingernail $\delta^{15}$N for mothers and babies was the same. Fogel and colleagues (1989, 1997) then analyzed bone collagen from preagricultural and agricultural sites to see whether the latter children were weaned at an earlier age. In both groups, $\delta^{15}$N decreased at 18–20 months of age. This study demonstrated that the trophic-level effect of increased $\delta^{15}$N is reflected in protein, including bone collagen, and that the potential exists to apply this method to skeletal samples. In a more recent study, Schurr and Powell (2005) analyzed stable carbon and nitrogen isotopes from individuals from four archaeological sites in eastern North America, two preagricultural and two agricultural, and found no difference in the duration of nursing among the four samples.

Others have applied this method and have attempted to refine estimates of the duration of nursing in the past (reviewed by Katzenberg et al., 1996; Schurr, 1998). For example, Herring et al. (1998) studied the process of weaning (introduction of non-breast milk foods and the gradual cessation of breastfeeding) in a nineteenth-century skeletal sample from Ontario by using stable nitrogen isotopes, parish records, census data, and skeletal evidence. The sample is from an Anglican Church cemetery (see Saunders, this volume, for additional background information on the cemetery). Many skeletons of infants and young children were included in the sample, and parish records as well as the regional censuses could be consulted to obtain multiple sources of mortality data. Careful age determination of the individuals allowed the construction of a mortality profile, which could be compared with the parish records. Since mortality increases as infants are weaned, because of exposure to infectious agents and loss of passive immunity, it was possible to compare mortality with the nitrogen isotope evidence for weaning. The combined data show that non-breast milk foods were introduced around the age of 5 months, as indicated by increasing mortality, and that milk continued to be the major source of protein until around 14 months, as indicated by decreasing $\delta^{15}$N.

Fuller et al. (2006) carried out a longitudinal study of $\delta^{13}$C and $\delta^{15}$N in fingernails and hair of eight modern mother–infant pairs. Five infants were exclusively breastfed, whereas two were breastfed initially and then formula-fed and one was fed only formula. The results of this study provide confirmation for the use of stable nitrogen isotopes to document the duration of nursing and show the smaller trophic-level effect in stable carbon isotopes. Such carefully controlled studies are very important for providing a more precise understanding of the information conveyed through stable isotope analysis.

A potential problem with using stable isotopes of nitrogen to study past breastfeeding patterns is that the subjects under study died in infancy and early childhood of unknown causes. Since $\delta^{15}$N may be increased in situations of nutritional stress, it may not be possible to separate the trophic effect of breastfeeding from cause of death in some cases. This possibility has been examined, but it is not yet clear that it is a serious problem (Katzenberg, 1999).

One way to avoid the problem of analyzing bones from individuals who died in infancy or early childhood is to analyze tissues that were formed early in postnatal development that can be isolated from older children and adults. The permanent tooth crowns are formed between the ages of three months (incisors) and seven years (second premolars and second molars) (Hillson, 1996). The third molar crown forms later but is highly variable. Analysis of nitrogen isotopes in dentine collagen and oxygen and carbon isotopes in enamel apatite can provide information about nursing and weaning for individuals who survived into at least late childhood. This approach has been followed by Wright and Schwarcz (1998). They analyzed stable isotopes of carbon and oxygen from the carbonate in tooth enamel. Oxygen isotope
measurements are normally performed on the phosphate of bones and teeth since phosphate is less likely to undergo diagenetic change than is carbonate. Wright and Schwarz (1998) point out that the procedure for isolating phosphate is more complex than that for isolating carbonate, and that tooth enamel is less likely to be affected by diagenetic processes than is bone mineral. The principle of using stable isotopes of carbon and oxygen is as follows. Carbon isotopes reflect the introduction of weaning foods, which, in the Americas, are usually maize-based gruel. Therefore, $\delta^{13}C$ is expected to increase as this $C_4$-based weaning food is introduced into the diet. Oxygen isotopes reflect water source. Body water has a higher $\delta^{18}O$ than ingested water because more $^{16}O$ relative to $^{18}O$ is lost in expired water vapor. Since breast milk incorporates body water, it is enriched in the heavier isotope in comparison to other water sources for the infant. Therefore as the infant is weaned, $\delta^{18}O$ should decrease and breastfed infants should have higher $\delta^{18}O$ than infants who are not breastfed. Wright and Schwarz (1998) compared $\delta^{13}C$ and $\delta^{18}O$ in enamel carbonate from tooth crowns formed at different ages and showed that, as expected, $\delta^{13}C$ increases with age, whereas $\delta^{18}O$ decreases with age.

**RESIDENCE AND MIGRATION STUDIES**

From the preceding information, it is obvious that some principles regarding stable isotope variation may be used to indicate place of residence. Stable carbon isotope abundance ratios vary based on diet, whereas stable nitrogen isotope abundance ratios vary based on diet and habitat. Oxygen isotope abundance ratios vary based on climate and water source. In general, $\delta^{18}O$ decreases with increasing latitude, increasing distance from the coast, and increasing altitude, because more of the heavy isotope, $^{18}O$, falls in precipitation (Dansgaard, 1964). Other variables that affect the $\delta^{18}O$ of human bone phosphate include humidity and the plants and animals consumed (Luz and Kolodny, 1989). Comparisons of stable isotope ratios in tissues laid down early in life and those that turn over throughout life may be used to determine whether individuals have moved. An elegant example of this type of study is that of Sealy et al. (1995), who compared enamel, dentine, and bone from five individuals from the southern Cape of Africa. Two individuals are prehistoric Khoisan hunter-gatherers, and three others date to the historic period. Of the historic burials, two were thought to be males of European ancestry and one was thought to be a female who was brought to the Cape as a slave. Enamel, dentine, and bone were analyzed for stable isotopes of carbon, nitrogen, and strontium. The prehistoric individuals have similar isotope values for all tissues analyzed. However, two historic individuals show significant variation in stable isotope values for the different tissues, indicating that they moved between their childhood and several years before their death as adults.

The use of stable oxygen isotopes to identify the geographical origin of prehistoric peoples has been carried out on prehistoric human bones from Mexico (White et al., 1998, 2004) and on historic soldiers from northeastern North America (Schwarz and Schoeninger, 1991). The method becomes problematic in more recent peoples because our food and water comes from many sources (e.g., Perrier water from France is popular in North America, as is New Zealand lamb). However, in people who were unlikely to move frequently, and who had fairly monotonous diets, oxygen isotope analysis is a useful indicator of residence.

Strontium isotopes are also useful in residence and migration studies since their variation is tied to local geology. Strontium isotopes vary depending on the underlying bedrock that gives rise to soils, and they vary between people feeding on marine versus terrestrial resources (Sealy et al., 1991). In addition to the study described above by Sealy et al. (1995), strontium isotopes have been used in studies of
residence at the Grasshopper Pueblo in Arizona (Ezzo et al., 1997; Price et al., 1994b) and in a Bell Beaker sample from Bavaria (Price et al., 1994a). They have also been used to identify immigrants to large ceremonial sites such as the Mayan site of Tikal (Wright, 2005), and to investigate residential mobility in the Andes (Knudson et al., 2004). The potential of using strontium isotopes for residence studies was first pointed out by Ericson (1989). Because of the strontium isotope differences between marine and terrestrial foods, they can also be used in paleodiet studies (Sealy et al., 1991).

A DAY WITHOUT STABLE ISOTOPES: WHAT HAS THEIR USE ADDED TO OUR KNOWLEDGE?

The use of stable isotope methods to address archaeological questions now dates back over 30 years. Do we really need these sophisticated analyses to tell us what people ate, where they lived, and whether they migrated? Other sources of evidence exist in the archaeological record for all of these questions. The presence of a food item in an archaeological assemblage does not necessarily imply that it was locally produced or that it was eaten. Differential preservation of food remains may confuse interpretations of the relative importance of particular foods. Charred maize cobs and kernels may be well preserved in comparison with other plants, and this has resulted in situations where the importance of maize was overestimated, as shown by stable isotope analyses. The use of nitrogen isotopes has allowed researchers to have a clearer picture of the importance of freshwater fish in the diets of inland populations. Collectively the use of stable isotopes for dietary reconstruction refines estimates of the relative importance of various foods and therefore leads to more accurate interpretations of the effects of changing diet on health and demography. It is also the only means currently available of detecting sex and age differences in diet within past human groups. The use of nitrogen as well as carbon and oxygen isotopes to estimate the duration of breastfeeding also ties into demographic reconstruction. Residence and migration studies help explain population interaction and movement in the past.

With advances in instrumentation and the increasing use of stable isotopes in ecological studies, there has been a very fruitful exchange of information between bioarchaeologists and ecologists. This exchange is illustrated in the example of the Lake Baikal study by Katzenberg and Weber (1999), where much of the information on stable carbon isotope values for freshwater fish was obtained from ecological studies. The exchange has also been beneficial in the other direction in that ecological studies have benefited from the extensive stable isotope ecology reconstructions carried out by archaeologists such as Ambrose, working in East Africa (1986, 1991; Ambrose and DeNiro, 1986) and Sealy, working in Southern Africa (1986; Sealy and van der Merwe, 1988; Sealy et al., 1987).

There are still problems to solve, but many of the problems identified in 1989, in an important paper by Sillen et al. entitled “Chemistry and Paleodietary Research: No More Easy Answers” have seen significant progress. Controlled feeding studies such as those carried out by Ambrose and Norr (1993; Ambrose, 2001) and Tieszen and Fagre (1993) have addressed questions about $\delta^{13}C$ differences in collagen and bone carbonate. Others (Sponheimer et al., 2003) have focused on nitrogen isotopes in the diet and tissues of mammals. Research has become more sophisticated in terms of understanding the underlying biochemical processes that affect stable isotopes. Examples include Ambrose’s (2001) work on nitrogen isotopes and water stress, Fogel et al.’s (1997) insights into protein stress and amino acid metabolism, and Evershed’s (1993, Stott and Evershed 1996) research on stable isotopes in lipids. Research carried out on living subjects, using hair and fingernails as sources of protein for analysis, allows much more control over
diet, dietary change, and variation (Fuller et al., 2006; O’Connell and Hedges, 1999).

Ethical issues have an impact on all skeletal studies of past peoples (see Walker, this volume). In some situations, visual study is permitted but no destructive analyses are allowed. Even though analytical methods have increasingly been improved and refined so as to require milligram quantities of material, some legislation completely prohibits any destructive analyses of human remains. Even when destructive analyses are permitted on skeletal samples, one must be careful to consider the long-term integrity of collections. Are there other ways to obtain the information? Is the method likely to be successful and informative? Biological anthropologists must continue to work with concerned groups and to explain the value of such studies to everyone.

It is also important to be aware of limitations on the information that may be provided. This awareness is particularly important with paleodiet studies where consumption of several different combinations of foods with varying emphasis on each one may result in the same stable isotope ratios. Use of multiple elements and familiarity with other archaeological sources of dietary information can go a long way toward unraveling this problem. Recent developments using GC/IRMS in which specific amino acids, or other specific biomolecules, are isolated and analyzed can also help to solve this problem.

The future of stable isotope analysis will include many applications that have become routine over the last 30 years, but the field is also moving forward very quickly. It is increasingly important for individuals wishing to pursue this area of study to have training in chemistry and biochemistry, since the level of sophistication has increased substantially. It is true of both the methods for isolating specific components in bone samples as well as the interpretation of the results. Indeed many recent advances have been made by individuals with backgrounds in biochemistry, geochemistry, and ecology who have become interested in archaeological applications. At the same time, some of the most fruitful collaborations have been between individuals trained in the physical sciences and those trained in archaeology and physical anthropology.

**BOX 13.1**

When working with prehistoric skeletal remains, actual age at death is not known nor are actual familial relationships. However, when studying past diet in individuals buried in historic cemeteries, analyses may be carried out on individuals of known sex, age at death, and familial relationship since grave markers or coffin plates may be present. Such was the case with a small historic cemetery in southern Ontario (Katzenberg, 1991). Stable carbon and nitrogen$^3$ isotopes were analyzed in collagen from 15 individuals, of whom 9 were adults and 6 were children ranging in age from around the time of birth to six years. The oldest adults probably spent their earlier years in Great Britain before migrating to Canada. In some cases, this migration is also known since year of death is known (Saunders and Lazenby, 1991). Thus, it is possible in this small sample to view stable isotope data from migrants and those born in Canada, to assess diet, and to look for evidence of nursing and weaning in the bones of infants. In one particular case, a mother and child were buried together. Two presumed mother/infant pairs were also buried in the small cemetery.

$^3$Reanalysis of stable isotopes of nitrogen was carried out after the 1991 publication to obtain a complete set of data. The more recent results are provided herein.
The data are presented in the table below:

<table>
<thead>
<tr>
<th>Burial Number</th>
<th>Sex</th>
<th>Age</th>
<th>$\delta^{13}C$</th>
<th>$\delta^{15}N$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Female</td>
<td>25 years</td>
<td>−18.3</td>
<td>12.0</td>
</tr>
<tr>
<td>1a</td>
<td>Female</td>
<td>1 year, 6 months</td>
<td>−18.9</td>
<td>14.1</td>
</tr>
<tr>
<td>2</td>
<td>Male</td>
<td>1 year</td>
<td>−18.7</td>
<td>13.1</td>
</tr>
<tr>
<td>3</td>
<td>Male</td>
<td>31 years</td>
<td>−18.6</td>
<td>12.0</td>
</tr>
<tr>
<td>4</td>
<td>Male</td>
<td>71 years</td>
<td>−18.9</td>
<td>11.2</td>
</tr>
<tr>
<td>5</td>
<td>neonate</td>
<td></td>
<td>−16.5</td>
<td>12.0</td>
</tr>
<tr>
<td>6</td>
<td>Male</td>
<td>71 years</td>
<td>−19.5</td>
<td>11.5</td>
</tr>
<tr>
<td>7</td>
<td>Female</td>
<td>98 years</td>
<td>−17.8</td>
<td>12.2</td>
</tr>
<tr>
<td>8</td>
<td>Male</td>
<td>71 years</td>
<td>−20.9</td>
<td>10.7</td>
</tr>
<tr>
<td>9</td>
<td>1 year</td>
<td></td>
<td>−18.9</td>
<td>12.7</td>
</tr>
<tr>
<td>11</td>
<td>Female</td>
<td>34 years</td>
<td>−18.7</td>
<td>11.1</td>
</tr>
<tr>
<td>12</td>
<td>Female</td>
<td>6 years</td>
<td>−19.5</td>
<td>10.3</td>
</tr>
<tr>
<td>13</td>
<td>Female</td>
<td>31 years</td>
<td>−18.4</td>
<td>11.8</td>
</tr>
<tr>
<td>14</td>
<td>Male</td>
<td>57 years</td>
<td>−17.7</td>
<td>11.7</td>
</tr>
<tr>
<td>15</td>
<td>neonate</td>
<td></td>
<td>−18.9</td>
<td>12.0</td>
</tr>
</tbody>
</table>

The known mother and child are burials 1 and 1a, respectively. Notice that the $\delta^{15}N$ of the child, aged one year and six months at death, is 2.1‰ greater than that of the mother, indicating that the child nursed and may have still been nursing at the time of death. There is some time lag between the cessation of nursing and the deposition of new collagen that no longer reflects the trophic-level effect of nursing. In other studies, a small trophic-level shift in $\delta^{13}C$ has been observed, but it is not present here. The $\delta^{13}C$ of the child is 0.6‰ less enriched than that of the mother.

One of the presumed mother–infant pairs is burial 13 and burial 15. In this case, the infant died within a month of birth and the difference in $\delta^{15}N$ between mother and infant is only 0.2‰. Given that the precision of the analysis is ± 0.2‰, this difference is not significant, and repeat analyses may result in the numbers being the same or in the mother having a 0.1% or 0.2% difference from the infant in the other direction. The third presumed mother–infant pair (and the least certain, based on other sources of evidence) includes burials 5 and 11. Once again, the infant died within the first month of life. Here the $\delta^{15}N$ of mother and infant differ by 0.9‰. If we compare the $\delta^{13}C$ for these mother and infant pairs, we see that they are similar for burials 13 and 15 (difference of 0.5‰) but that they differ by 2.2‰ for burials 5 and 11. One would expect a newborn baby to have similar isotope ratios to its mother, so we might suggest that this is not a mother–infant pair, based on the stable isotope evidence. (In a matched study of fingernails from living mother–infant pairs, Fuller et al. (2006) found that $\delta^{13}C$ did not differ by more than 1‰ between mothers and their infants.)

The diet of people from Great Britain is based exclusively on C3 plants such as wheat, oats, and barley and on domesticated animals that also consumed C3 plants. The average $\delta^{13}C$ is around −20‰ to −21‰ (van Klinken et al., 2001). Settlers to southern Ontario brought that diet with them and characteristically have $\delta^{13}C$ values around −19‰ to −18‰, as seen here (Katzenberg et al., 2001). The slight enrichment is undoubtedly from the incorporation of the New World domesticate, maize, which is a C4 plant, that was used in bread. Another C4 plant product, cane sugar, was used in baked goods (Katzenberg et al., 2001).

The people buried in the cemetery came from Scotland and moved first to New York in 1810 and then to Ontario in 1817. Is it possible to see any dietary differences between the oldest adults, who would have spent some time in Scotland, and the younger individuals? At
In the first glance, only one individual stands out and that is burial 8, a 71-year-old man. His $\delta^{13}C$ value is $-20.9‰$, the most depleted in the heavier isotope of carbon. Burial 8 died in 1825, so he would have spent the first 56 years of his life in Scotland and only the last 15 years in North America. Collagen turnover is estimated to be about 10–20 years so it is possible that his bones still contained some collagen that formed during his many years in Scotland. Burial 7 is the oldest individual in the cemetery, and her $\delta^{13}C$ value is $-17.8‰$, which is different from burial 8 and more in line with a North American diet. However, she died in 1894 so she spent almost her entire life (the last 84 years) in North America and only the first 14 in Scotland. Therefore, her bone collagen reflects her North American diet.

This example illustrates some of the diverse information that may be obtained from stable isotope data. Obviously, some information in a historic sample will not be available in other situations; however, the certainty that is gained through such analyses is important in strengthening interpretations in less well-documented samples.

ACKNOWLEDGMENTS

For the first edition of this chapter, which appeared in 2000, I wish to express my appreciation to the Killam Foundation for granting me a Killam Resident Fellowship at the University of Calgary during the fall term 1998 for the purpose of preparing this chapter. I am also grateful to Steve Taylor and Roy Krouse, Department of Physics, University of Calgary, for providing useful comments on some sections of the chapter, and to my colleagues Shelley Saunders and Susan Pfeiffer for reading and commenting on the complete chapter. Thanks also to my graduate students, Sandra Garvie-Lok, Tamara Varney, and Roman Harrison who read and commented on the chapter. For the second edition, special thanks to Andrea Waters-Rist for research assistance and editorial comments, to Kristine Hепple for assistance with references, and to Kate Faccia and Shelley Saunders for editorial suggestions.

REFERENCES


Ambrose SH, Norr L. 1993. Experimental evidence for the relationship of the carbon isotope ratios of whole diet and dietary protein to those of...


REFERENCES

Sealy JC, van der Merwe NJ, Sillen AKFJ, Krueger HW. 1991. $^{87}$Sr/$^{86}$Sr as a dietary indicator in


REFERENCES

441


CHAPTER 14

BONE CHEMISTRY AND TRACE ELEMENT ANALYSIS

JAMES BURTON

BONE

A discussion of “bone composition” requires a prefatory clarification of exactly what one means by “bone.” As has been well described elsewhere in this book, a physical, real-world sample of bone is a highly complex, intimately intermixed collection of bone cells and other components (marrow, blood, serum, ...). The bone cells themselves necessarily have their full compliment of nucleic acids, proteins, and many other organic chemicals and inorganic salts necessary for cellular functions. Additionally, in the case of archaeological bones, nonbiological processes can alter composition both by physical contamination and by chemical reactions, processes collectively called “diagenesis.”

As a necessary simplification, bone chemistry literature commonly presents “bone” as a relatively simple and more or less homogeneous, binary material, one third of which by weight is protein (collagen) and two thirds are inorganic mineral salts, mainly calcium phosphates. Because analyses of the organic fraction focus mainly on the ratios of stable isotopes (e.g., $^{13/12}\text{C}$ and $^{15/14}\text{N}$) and are discussed elsewhere (see Katzenberg, this volume), the following discussion of bone emphasizes the elemental composition of the mineral fraction.

The human skeleton also contains two other calcified tissues of significance to compositional studies, i.e., dental enamel and dentine. Enamel, in contrast to bone, is not living tissue but consists almost entirely of well-crystallized hydroxyapatite with less than 1% organic material. Dentine is more like bone in that it is a living tissue but has more calcium phosphate (75%) than bone and less collagen (18%), with most of the remainder being water (Hillson, 1986).

THE MINERAL FRACTION OF BONE

Hydroxylapatite, or hydroxy-calcium phosphate, is the principal mineral component of bone. Because minerals by definition have a specific structure as well as composition, and the structure of hydroxylapatite requires two calcium phosphate molecules, the formula for hydroxylapatite is formally expressed as $\text{Ca}_10(\text{PO}_4)_6(\text{OH})_2$ rather than by the formula for the unitary molecule, $\text{Ca}_5(\text{PO}_4)_3(\text{OH})$. From either formula one may calculate that ideal hydroxyapatite contains 39.8% calcium and 18.5% phosphorus ($\text{Ca}/\text{P} = 2.15$).
X-ray diffraction studies to determine the size of hydroxylapatite crystals in fresh bone reveal that most bone “mineral” particles have a size on the order of tens of nanometers ($10^{-9}$ m). This is far too small to be seen by optical microscopy and is comparable in scale with the minimal “unit-cell” size of the hydroxylapatite crystal itself. These particles are too small to be a well-formed crystal lattice of many molecules and thus are more amorphous (poorly crystalline) than true hydroxylapatite.

The extremely small size of the mineral particles greatly enhances chemical reactivity through increased solubility. Also, because the biological particles are almost molecular in size, a large fraction of each molecule’s atoms are close to the particle’s surface and thus are available for ion exchange. In living bone tissue, an indefinite fraction of the calcium phosphate is in a reactive flux, which can contain a variety of transient, metastable mineral phases among which brushite ($\text{CaHPO}_4 \cdot 2\text{H}_2\text{O}$), octocalcium phosphate ($\text{Ca}_8\text{H}_2\text{(PO}_4\text{)}_6 \cdot 5\text{H}_2\text{O}$), whitlockite ($\text{Ca}_3(\text{PO}_4)_2$), defective apatite, and other phases have been reported. Small amounts of sodium, magnesium, potassium, and other elements are also present in this flux. Nonetheless, under physiologic conditions, hydroxylapatite is the only truly stable mineral and represents a good approximation to actual bone composition: 38% calcium, 18% phosphorous ($\text{Ca}/\text{P} = 2.13$) with traces of sodium (0.6%), and magnesium (0.6%) and smaller amounts of other elements.

**ELEMENTS IN THE MINERAL FRACTION OF BONE**

Because the amounts of calcium and phosphorus in bone mineral are fixed within relatively invariant limits by the formula constraints, or “stoichiometry,” of hydroxylapatite, analysts measure them mainly as an indicator of the quality of the sample. For example, a bone badly contaminated by calcium carbonate would have high Ca but lower phosphorus, yielding $\text{Ca}/\text{P}$ higher than 2.13. Likewise, because whole bone is approximately one-third protein, it has only two thirds as much Ca and P as pure hydroxyapatite. Fossil bones can range from 26% calcium in samples with fully preserved collagen (2/3 x 38%) to 38% calcium in those with no remaining collagen. To adjust for variations in composition that are simply from varying amounts of collagen, fossil bone is commonly burned in a kiln before analysis to remove organic matter and reduce the sample to just the ash mineral component. The percent mineral versus percent organic fraction can then be assessed from the weight lost during this ashing.

The primary elements of interest to bone chemists are thus not calcium and phosphorous, the major elements of fixed composition, but minor or trace levels of other elements that can substitute for them. By far the most important of these is strontium, followed by barium and lead. Because these elements can easily replace calcium, and accumulate in bone, they are known as “bone-seeking.” Because they do not have metabolic functions, their biological levels are not regulated but tend to be proportional to their amount in the diet, leading to efforts to use them as indicators of prehistoric diets. Zinc has also received significant attention as a possible paleodietary indicator, followed by several other elements of more speculative utility, some with quite imaginative applicability.

Strontium is an alkaline-earth element; i.e., in the periodic chart it is immediately under calcium, which it chemically resembles. Strontium atoms are positively charged and divalent (+2), have an ionic size (200 pm) similar to that of calcium (180 pm), and thus have enough similarity to calcium that strontium can replace calcium atoms in hydroxylapatite. Barium, another alkaline-earth under strontium in the periodic table, is likewise positively charged, divalent, only slightly larger (215 pm) than strontium, and substitutes for calcium but with slightly less efficiency than strontium. Lead is not an alkaline earth element, but nonetheless it has a positive, divalent ion that has
the same size (180 pm) as calcium, which it can likewise replace.

Dietary calcium can be considered to be “dirty” or “contaminated” with these divalent ions (e.g., Sr⁺⁺, Ba⁺⁺, and Pb⁺⁺). These larger ions are, however, preferentially retained in the gut and excreted such that the percentage of these elements in biologically retained calcium is lower than their percentage in dietary calcium (Elias et al., 1982). “Biopurification” is the collection of processes that tend to preferentially remove these ions from calcium as it progresses through the food chain from lower to higher order consumers. Similar processes occur in plants; these other divalent ions “centripetally” accumulate in the lower parts of plants such that calcium farther from the root will be cleaner than calcium closer to the roots (Bowen and Dymond, 1955, 1956). Calcium can be further purified by other biological activities as well; e.g., additional purification occurs in the mammary gland such that milk Sr/Ca (and Ba/Ca, etc.) will be lower than that of the nursing mother (Blakely, 1989; Wasserman et al., 1958).

The end result of these processes is a much lower ratio of strontium to calcium—and likewise for barium and lead—in the bones of animals that are high in the food chain: Sr/Ca (soil) > Sr/Ca (plants) > Sr/Ca (herbivores) > Sr/Ca (carnivores) (Fig. 14.1).

Geologists initially tried to apply this knowledge of these processes to determine the diets of extinct animals, resolving grazers, browsers, and carnivores (Toots and Voorhies, 1965). Anthropologists and archaeologists soon followed this lead in adopting the technique to assess whether prehistoric humans were similarly mainly vegetarian or carnivorous (e.g., Brown, 1973, 1974; Kavanaugh, 1979; Price and Kavanagh, 1982; Schoeninger, 1982; Sillen, 1981, 1986). This offered the promise of being the first method in which one could directly assess prehistoric human diet in contrast to trying to infer diet from artifacts and other archaeological remains.

It was initially presumed by anthropologists that levels of strontium in human bone would be proportional to the amounts of plants, with relatively high strontium, in the diet compared with the amount of meat, with more biopurified calcium, i.e., with less strontium. A given strontium level, e.g., 200 ppm, would be meaningess by itself, but would imply more plants in the diet than an individual with 150 ppm. In efforts to make the method more quantitative, analysts compared human levels with those of pure herbivores and carnivores, assuming that intermediate strontium levels would be proportional to intermediate plant/meat ratios; e.g., an individual with a bone strontium level halfway between carnivore and herbivore...
levels would be consuming a diet with a 50/50 plant/meat ratio. Thus, it was believed anthropologists could begin to address, quantitatively, highly significant issues such as hunting versus gathering and the transitions to agriculture as well as status-related differential access to foods.

During this initial period of high enthusiasm, such applications were supplemented by empirical studies of modern food webs—studies that demonstrated the reduction of strontium, barium, and lead with increasing position in the food chain (e.g., Elias et al., 1982; Price et al., 1985)—and with controlled feeding studies, in which bone levels were shown to be proportional to dietary levels of these elements (Lambert and Weydert-Homeyer, 1993a, 1993b; Weydert, 1990).

Nonetheless, application of this principle to determine the percentage of meat in the diet has proved problematic. The most significant difficulty develops from the fact that bone strontium in pure herbivores differs from that in pure carnivores, but it is relatively insensitive to the plant/meat ratio of intermediate diets, i.e., chemical equivocality (Burton and Wright, 1995). Another major problem, which affects all aspects of bone composition and which will be discussed separately below, is the difficulty in assessing and removing postmortem, or diagenetic, changes to the original, biological composition of bone.

Probably the most confounding aspect of such dietary studies is that even though bone strontium levels demonstrably and consistently reflect dietary strontium levels, and modern bones of carnivores have consistently less strontium than bones of herbivores, intermediate diets do not yield proportionally intermediate strontium levels. That is to say, a diet of half meat and half vegetation will not—as has been traditionally assumed—produce bone strontium levels halfway between that of a pure carnivore and a pure herbivore.

It is significant that the bone level of strontium is not proportional simply to the amount of strontium in the diet but to the mean Sr/Ca ratio of the diet:

$$(\text{Sr/Ca})_{\text{bone}} = (0.2) (\text{Sr/Ca})_{\text{diet}}$$

It is the amount of strontium relative to the total amount of available bone-forming cations $(\text{Ca} + \text{Sr} + \text{Ba} + \text{Pb} \cong \text{Ca})$ that determines what percentage of the bone mineral is represented by strontium. Thus, carnivores with nearly single-component diets have relatively constant and predictable Sr/Ca ratios in their bones. If two bobcats consume different quantities of hares, and thus different amounts of strontium, their bone Sr/Ca ratios will nonetheless be the same because the amount of strontium relative to calcium is the same for both cats (i.e., about 20% of the Sr/Ca of the hares). It is for such single-component “end-member” cases that one obtains the well-known food chain, or “trophic,” response. For example, for the data of Elias et al. (1982), martens $(\text{Sr/Ca} = 0.00067)$ mostly eat voles $(\text{Sr/Ca} = 0.0042)$, which consume mainly the local sedges $(\text{Sr/Ca} = 0.0170)$ (Fig. 14.1).

To predict bone Sr/Ca for multicomponent diets, however, it is necessary to know both the total strontium consumption and the total calcium consumption. The conventional use of bone strontium as a measure of the dietary plant-to-meat ratio implicitly assumes that the Sr/Ca ratio of the diet is proportional to a sum by weight of the Sr/Ca ratios of the various dietary components:

$$\Sigma_i (\text{Sr/Ca}_i) \, (\% \text{ of } x_i)_i$$

where $i$ is added for the percent of each dietary component “$x_i$.”

The actual weighting factor in this dietary equation turns out not to be the weight-fraction of a dietary component, as had been assumed, but the relative calcium contribution of that component. For a simple two-component diet:

$$(\text{Sr/Ca})_{\text{diet}} = (\text{Sr from a} + \text{Sr from b})/ (\text{Ca from a} + \text{Ca from b})$$
\begin{align*}
\text{i.e.,} & \\
& = \frac{[(\text{Sr})_a + (\text{Sr})_b]}{[(\text{Ca})_a + (\text{Ca})_b]} \\
& = \frac{[(\text{Sr}/\text{Ca})_a(\text{Ca})_a + (\text{Sr}/\text{Ca})_b(\text{Ca})_b]}{[(\text{Ca})_a + (\text{Ca})_b]} \\
& = \frac{(\text{Sr}/\text{Ca})_a(\text{Ca})_a}{[(\text{Ca})_a + (\text{Ca})_b]} \\
& + \frac{(\text{Sr}/\text{Ca})_b(\text{Ca})_b}{[(\text{Ca})_a + (\text{Ca})_b]} \\
& = (\text{Sr}/\text{Ca})_a (\% \text{ of total Ca from a}) \\
& + (\text{Sr}/\text{Ca})_b (\% \text{ of total Ca from b})
\end{align*}

This is not equal to \((\text{Sr}/\text{Ca})_a (\% \text{ of a}) + (\text{Sr}/\text{Ca})_b (\% \text{ of b}).\)

By extension to multicomponent diets, Diet \(\text{Sr}/\text{Ca} = \sum_i (\text{Sr}/\text{Ca})_i (\% \text{ of total Ca})_i, \text{ not } \sum_i (\text{Sr}/\text{Ca})_i (\% \text{ of } x_i)_i, \) where \(i\) is added for each dietary component “\(x_i\).”

The significance of this can be understood graphically for a simple two-component diet (Fig. 14.2). The dietary \(\text{Sr}/\text{Ca}\) ratio will be mainly that of whichever item has the most calcium. The item with less calcium will not affect the \(\text{Sr}/\text{Ca}\) ratio of the diet unless it contributes most of the calcium, i.e., until it is overwhelmingly the greatest dietary component. In other words, dietary \(\text{Sr}/\text{Ca}\) does not reflect the relative amounts of various components, just the \(\text{Sr}/\text{Ca}\) of whatever component contributes the calcium. In an extraordinary case where two items have identical mineral concentrations, then, and only then, will diet \(\text{Sr}/\text{Ca}\) be proportional to the amount of each item in the diet. The primary consequence of this \(\text{Sr}/\text{Ca}\) dependency on calcium content is that foods high in calcium have a disproportionately large effect on bone \(\text{Sr}/\text{Ca}\), whereas low-calcium foods may be invisible or even produce a paradoxical response in bone \(\text{Sr}/\text{Ca}\) ratios. This dependence of bone \(\text{Sr}/\text{Ca}\) on the calcium content of diet, as well as on the strontium content, was explicitly recognized in several reports of Runia (1987a, 1987b, 1988) and quantitatively used in the study of white-tailed deer by Price et al. (1985), but few anthropological studies have considered the implications in assessing prehistoric diet. Bone strontium gives some idea of the degree of biopurification of the principal calcium source, but it does not provide a quantitative assessment of that source.

Recognition of the importance of the \(\text{Sr}/\text{Ca}\) ratio as the relevant parameter can explain the idiosyncratic results of the controlled-feeding experiment of Lambert and Weydert-Homeyer (1993a, 1993b). They paradoxically observed

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{Figure14.2.png}
\caption{Figure 14.2 Logarithm of \(\text{Sr}/\text{Ca}\) of a two-component diet as the diet varies from 100% plant (amaranth) to 100% meat. Notice that the ratio hardly differs for diets with less than 80% meat.}
\end{figure}
that their highest strontium diets, i.e., “meat” and “fish,” produced the lowest bone strontium levels, whereas plants, with relatively low strontium, yielded high bone strontium. This produced a negative correlation between the concentration of strontium in the diet with that in the bone, as well as an inverse “trophic” effect. However, the “meat” and “fish” diets also contained high levels of bone and, hence, the highest levels of calcium. Their highest strontium diets actually have the lowest Sr/Ca ratios and appropriately show a positive correlation with dietary Sr/Ca ratios. Lambert and Weydert-Homeyer similarly found an unexpected, negative correlation between dietary barium and bone barium that disappears when the calcium-weighted Ba/Ca ratio is considered.

It is also imperative to remember that high calcium sources extend beyond a simple plant/meat dichotomy and can include items such as dairy products, bone meal, salt, and even lime. The Sr/Ca of the diet would then match the Sr/Ca of such components regardless of what other items might be consumed. In general, bone strontium is an average of many different inputs. Knowing the amounts of different foods and their calcium and strontium levels permits one to estimate an average bone Sr/Ca, but generally one cannot estimate multiple components given merely this average. The same principles similarly affect barium and lead.

Barium is chemically similar enough to strontium and calcium that it behaves essentially like strontium except that, being a larger ion, it is biologically assimilated with less efficiency than strontium and thus has a greater degree of biopurification than strontium. Bone Ba/Ca tends to be about one tenth that of the dietary Ba/Ca, whereas bone Sr/Ca tends to be about one fifth that of the diet (Fig. 14.2). Thus it has been suggested that barium could be used in addition to strontium as a trophic indicator (Burton and Price, 1991), but barium also exhibits chemical equivocality (Burton and Wright, 1995). Thus it, too, reflects the degree of biopurification of the principal calcium source and is not a quantitative indicator of trophic position.

There is one significant difference between barium and strontium, i.e., the difference in the solubility of their sulfate salts. Strontium sulfate is soluble, but barium sulfate is not. Thus, in an environment that has abundant sulfate ions, the most important of which is the ocean, strontium is available in solution, whereas barium is not, being removed as insoluble barite (BaSO₄). Ba/Ca in marine foods is thus an order of magnitude lower than that in comparable terrestrial diets. Although this suggests the possibility of revealing marine diets (Burton and Price, 1990), it again must be understood that Ba/Ca reflects that of the principal calcium source and is not a quantitative measure of seafood any more than of trophic position.

A good example of how barium and strontium reflect diet but fail as quantitative measures are the studies by Larsen, Ezzo, and others (Ezzo et al., 1995; Larsen et al., 1992) of prehistoric populations on an island on the southeastern coast of the United States. Early populations had similar strontium but much higher barium than later populations. This decrease in barium but not in strontium could be interpreted traditionally as an increase in the consumption of seafood in the later populations. However, this study also included stable isotope measurements, which indicated an increase in maize but no change in meat or fish. This apparent paradox is resolved by remembering that Sr/Ca and Ba/Ca indicate the principal calcium source but not its amount. In diets in which there is some seafood component, a change from plants high in calcium to maize, which is extraordinarily low in calcium, means that more calcium is coming from the seafood component than before even though seafood remains quantitatively constant. Ba/Ca drops, whereas Sr/Ca remains the same because corn, not seafood, is increasing.

The classic biopurification study of Elias et al. (1982), cited in most trophic-level Sr and Ba studies, was actually part of a study of lead exposure and accumulation. Biologically lead
behaves much like strontium and barium; it replaces calcium in hydroxylapatite, accumulates in bone, and even has been shown to exhibit a pronounced trophic effect. However, in contrast to strontium and barium, neither of which have any known physiologic effect, lead is extraordinarily toxic, at levels as low as one part-per-billion, which affects all organ systems.

Fortunately, levels of lead in unpolluted biological systems are many orders of magnitude lower than those of strontium and barium. Calcified tissues no older than the Roman era, however, are likely to reflect exposure to anthropogenic sources such as culinary items containing lead (lead crystal and lead-glazed pottery), lead paints, and leaded gasoline. It is to this end that lead analyses are commonly applied, i.e., to assess lead exposure.

Bone samples are rarely available from living humans, so studies of lead exposure have focused mainly on lost deciduous teeth and postmortem and archaeological bone samples. To assess nonanthropogenic background levels of lead, anthropologists and others have relied on pre-Roman archaeological samples in Europe and preindustrial era samples for the Western hemisphere.

Although bone lead levels significantly above background (i.e., on the order of parts per billion) reveal exposure to one or more anthropogenic lead sources, bone lead levels, analogous to strontium and barium levels, respond not directly to Pb exposure but to the Pb/Ca ratio. Acute lead exposure can even be ameliorated by promptly giving large doses of calcium before the lead can be assimilated, hence lowering Pb/Ca.

Increased lead levels reveal exposure to potentially toxic anthropogenic sources, but they do not distinguish among sources when there are multiple possibilities. Lead, however, has four isotopes, one stable ($^{204}$Pb) and three radiogenic ($^{206}$Pb, $^{207}$Pb, $^{208}$Pb), the relative abundances of which vary among various ore deposits and are not altered biologically. Thus, by comparing the ratios of these isotopes in calcified tissues with those of possible anthropogenic sources, one can potentially identify the specific sources of lead exposure (e.g., Gulson, 1996; Gulson and Wilson, 1994; Gulson et al., 2004).

Early enthusiasm in the 1980s for these “bone-seeking” elements quickly spread to other elements, with zinc being both the most popular and most controversial. A thesis by Gilbert (1975) reported higher levels of zinc in bones that were larger than others and posited that zinc correlated with better dietary health and greater protein consumption in particular. Although this was presented somewhat speculatively, zinc was embraced rapidly as an indicator of meat in the diet. Subsequent controlled feeding experiments (Weydert, 1990; Lambert and Weydert-Homeyer, 1993a, 1993b) did not show an increase in bone zinc with high zinc diets, and studies of modern food webs (Burton et al., 1999) did not show higher zinc in carnivore bones than in herbivore bones, although numerous papers continued to interpret increased zinc as indicating more meat in the diet. Ezzo (1994a) critically addressed this use of zinc by tracking the citation chain for studies using zinc as a meat indicator back uniquely to the Gilbert thesis and proposed that zinc be dropped as a paleodietary indicator, arguing that zinc is not incorporated in hydroxylapatite but is metabolically controlled so that levels should be constant rather than dependent on diet. It is certainly with some irony that a metabolic study of dietary and bone zinc levels by Hunt and Johnson (1992) found that bone zinc is not proportional to the amount of zinc in the diet, but if dietary zinc levels are adequate, bone zinc does increase with a high protein diet. Because this is species specific (i.e., it is not a result of trophic-level biopurification), neither the food chain studies nor the feeding experiments with high zinc are inconsistent with Gilbert’s postulate that higher bone zinc could be caused by increased meat consumption.

After the apparent success of strontium and zinc as paleodietary indicators, anomalies were sought in levels of other elements in bone, which, when found, engendered a search for consumable items that are preferentially
enriched in that element and led to imaginative inferences ranging from goat-herding (Edward et al., 1984) to maggot cultivation being a prelude to agriculture (Arrhenius, 1990). The methodological misstep most responsible for this interpretive chaos began with the inclusion of elements other than strontium and barium. Although it required extensive research to reveal the bone–diet connection for strontium and barium, this was reduced, evidently for simplicity, and not altogether without validity, to the idea that plants are higher in strontium than is meat, and that bones reflect this dietary difference. Likewise with Gilbert’s thesis suggesting zinc in meat becomes zinc in bone, the concept as transcribed to archaeologists was that bones simply reflect the chemistry of the diet.

Ezzo (1994b), reviewing this “you are what you eat” concept that more of any element in the diet produces more in the bone, enumerated specific criteria that should be attributes of any inorganic paleodietary indicator. The element of interest should be biologically incorporated within bone mineral and should be relatively free from metabolic control (i.e., biological levels should not be homeostatic and thus independent of diet). Biological levels should also considerably exceed levels that can be anticipated from postdepositional processes—a problem even for strontium.

We can improve the current situation by simply omitting such elements from our analyses, but we would nonetheless forfeit useful information. Although elements that are not biologically incorporated in hydroxylapatite are not likely to inform us about prehistoric diet, the fact that they are not biologically incorporated within bone mineral makes them useful as diagenetic indicators (Ezzo, 1994b; Price, 1989; Price et al., 1992). The use of non-alkaline-earth elements such as zirconium to assess diagenetic contamination is actually one of the earliest applications of such elements (Katzenberg, 1984).

Several other elements appear in the mineral component of bone at minor (0.1–1.0%) or trace levels (<0.1%), even though they are not known to substitute for ions within crystalline hydroxylapatite. These elements probably occur in the more amorphous mineral matrix from which they can more readily exchange with other ions in the blood. Magnesium (Mg), sodium (Na), and potassium (K) are important examples of such elements that cannot be structurally incorporated into hydroxylapatite in significant amounts yet are certainly present in bone, in which they serve as a buffer to maintain blood levels of these elements within strict physiologic limits. Because levels of such essential elements are maintained within critical limits, bone levels are not likely to exhibit much variation and hence not likely to inform about diet. The extraordinary reactivity of these elements in the amorphous bone matrix and consequent susceptibility to diagenetic contamination probably preclude paleoanthropological applications, but nonetheless there may be undiscovered relationships with diet and exposure in modern populations.

Although most analyses have focused on the positively charged ions that replace calcium, there are also negatively charged ions that can substitute for phosphate \( \text{PO}_4^{3-} \) and hydroxyl \( \text{OH}^- \) ions in hydroxylapatite. Phosphorus is structurally well bound within the phosphate ion and is unavailable for exchange and substitution with other elements, but carbonate \( \text{CO}_3^{2-} \) may substitute in small amounts for phosphate as an entire unit. Carbonate \( \text{CO}_3^{2-} \) can also take the place of the hydroxyl ion in hydroxylapatite, as can fluorine \( \text{F}^- \) and chlorine \( \text{Cl}^- \). Although there is scant interest in the quantitative measurement of carbonate and phosphate levels, both are of major importance for stable isotope ratios in bone (i.e., \( \text{C}_{13}/\text{C}_{12} \) and \( \text{O}_{18}/\text{O}_{16} \)), as discussed in the chapter by Katzenberg in this volume.

Quantitative measurements of fluorine have received some attention because natural, nonbiological fluorapatite, \( \text{Ca}_5(\text{PO}_4)_3\text{F} \), is the fluorine analog of hydroxylapatite. Replacement of the hydroxyl ion by fluorine increases the stability of apatite, making it more resistant to attack by acids, prompting public health agencies to add fluorine to water supplies to reduce the prevalence of dental cavities. In a postmortem
environment in which bone is exposed to groundwater solutions, the less stable biological hydroxylapatite can acquire fluorine and convert, at least partly, to the more stable fluorapatite. Such a transformation was used to debunk the famous Piltdown hoax in which the jaw of an orangutan was combined cleverly with the cranium of a modern human to appear to be an early hominid skull hundreds of thousands of years old (Wiener et al., 1950). Analysis of fluorine in the fossils revealed that the mandible had far less than the cranium and that it was thus likely to be only a few decades old, and not from the same skeleton, i.e., that the fossil was fake.

Unfortunately others have tried to foster fluorine measurements as a method to relatively date skeletal assemblages (e.g., Ezzo, 1992; Schurr, 1989; Schurr and Gregory, 2002), but fluorine assimilation does not behave with the regularity of isotopic decay and is not a simple function of time. Different bone tissues do not have the same propensity to absorb fluorine, nor must similar bones in different burial contexts acquire fluorine at similar rates. Although a museum-curated sample such as the modern Piltdown mandible should have far less fluorine than a cranium that has been buried for years, more modest differences in fluorine could reflect easily random variation in burial contexts or cumulative biological uptake, which can exceed a few thousand parts per million in modern bone. Thus, bone fluoride should not be used to assess absolute burial times or even relative ages.

POST-MORTEM ALTERATION (DIAGENESIS)

To be able to quickly regulate blood levels of critical elements, bone has to be both highly porous as well as extraordinarily chemically reactive. In the postmortem environment, the high porosity aids physical contamination by salts deposited from soil solutions and by physical infiltration of the soil itself. The extraordinary reactivity, likewise, does not stop immediately at death but promotes more chemical reactions with elements available in the soil. Both kinds of contamination potentially affect all elements of interest and can be difficult to assess and difficult or impossible to reverse.

The recognition of the chemical alteration of bone was concurrent with the enthusiasm for measuring strontium to assess trophic level. Often, fossil carnivore bones did not have much lower strontium or, paradoxically, had even more strontium than those of herbivores. Bones of the same species that should have had similar strontium levels often had different strontium according to their depositional context, and bones that should have shown large differences likewise did not (e.g., Blitz, 1995). Expectation of biological values (e.g., a Ca/P ratio slightly larger than two for hydroxyapatite and a 30% weight loss on ignition from the removal of collagen) frequently were not met, and buried bones were often found to have large amounts of elements such as uranium and the rare-earth elements (REEs) that are normally absent in biological bones. Such elements also vary across sections of bone, being higher near the inner and outer surfaces of bone, which indicates diffusion into the bone from the external environment. Concentrations of nonbiological elements also sometimes showed correlations with strontium and barium, which suggests they too were diagenetic additions to the bone.

Extensive research thus began to focus on ways both to assess the degree of contamination of bones and to try to remove contaminants and recover biological information. Biological integrity can be assessed structurally at the cellular level with an optical microscope and at the subcellular level with infrared spectroscopy and X-ray diffraction. Through the analysis of microscopic thin sections and the use of a microprobe, Hassan and Ortner (Hassan, 1975; Hassan and Ortner, 1977) demonstrated the presence of several physical and chemical contaminants in fossil bone. They posited that contamination occurred through both precipitation from groundwater and the physical movement of materials into
the bone. Inclusions such as quartz went into bone as solid grains; fungal hyphae, rootlets, and fragments of charcoal were also observed within the bone structure as physical contaminants.

Examination of bone with a petrographic or polarizing light microscope can also reveal the postdepositional disruption of the bone structure by microorganisms, which can also alter the chemical or isotopic composition (Grupe and Piepenbrink, 1988; Schoeninger et al., 1989; White and Hannus, 1983). In modern, unaltered bone, an oriented optical pattern can be observed in thin sections of the bone using polarized light, whereas decomposition of the organic matrix produces a random optical pattern. Because apatite is more exposed to diagenetic alteration when the organic matrix decomposes, this loss of orientation, along with an increase in the crystallinity of bone mineral, can be correlated with chemical alteration.

Studies using X-ray diffraction and infrared spectroscopy (Blumenthal et al., 1975; Schoeninger et al., 1989; Sillen, 1989; Tuross et al., 1989) indicate that, when the organic material is lost, either through diagenesis or through ashing, the crystallinity measurably increases. Because much of calcium phosphate in living bone is in a metastable, amorphous state, X-ray diffraction patterns, which reveal mineral structure, have broadly diffuse bands. Postmortem apatite, on the other hand, is well crystallized with a well-defined atomic arrangement so the diffraction pattern has sharp peaks rather than diffuse bands. Likewise, infrared spectroscopy, which shows the absorption of light by molecular bonds (e.g., the hydroxyl oxygen–hydrogen bond), also has absorption peaks that are better defined in crystallized apatite than in amorphous bone. Unfortunately both the postmortem degradation of cellular texture and the increased crystallinity of bone mineral proceed regardless of contamination. They are observable in most fossil bone and have little power to inform about the degree of chemical contamination.

To assess more directly chemical contamination, researchers measure elements that are much more abundant in the burial environment than in uncontaminated bone. Among the most sensitive of these are uranium and the REEs, both of which may be considered “phosphorous-seeking.” Biological levels in bones tend to be below tens of parts-per-billion, whereas exposure to ambient levels in soils can cause postmortem levels to rise several orders of magnitude. Williams (1988; Williams and Marlow, 1987; Williams and Potts, 1988) made a set of uranium measurements across femoral cross sections that showed uranium enrichment in a U-shaped pattern, which was higher at both the external and the internal surfaces, and decreasing into the interior, which was characteristic of contamination by diffusion. Schoeninger et al. (2003) later studied dental enamel with cathodoluminescence to reveal that enamel older than approximately 10,000 exhibits a contaminated gradient decreasing from the surface. Neither method allows one to quantitatively assess the extent of contamination by other elements (e.g., strontium and barium), but because of their ubiquity in soil and their tendency to combine with phosphate, the absence of uranium and REE is positive evidence against chemical contamination.

Titanium, zirconium, manganese, and hafnium are some other elements that are much more abundant in the soil. These elements are not soluble and hence are not available in solution to react with bone. These elements contaminate bones simply through physical inclusion as soil, especially in highly porous trabecular tissues. Significant amounts of these elements in a sample suggest that elements such as strontium and barium might also be postdepositionally increased because of their similarly ubiquitous presence in soil (Katzenberg, 1984; Parker and Toots, 1970).

Quantitative analyses to test for nonbiological values may also include determination of the ratio of mineral to organic matter and measurement of the calcium and phosphorus ratio. Since unaltered bone contains approximately 70% mineral matter and 30% organic
matter, the amount of mineral remaining after removal of the organic component by ashing should yield about 70% of the original weight. Increased values indicate either addition of diagenetic minerals or decomposition of the organic component. The mineral component (bone ash) of modern bone contains about 38% by weight of calcium and 17% by weight of phosphorous, and thus, it has a Ca/P ratio slightly greater than two.

Unfortunately, because diagenesis can occur through many means, each of which affects composition differently, none of these measurements informs about other modes of contamination. A bone sample could “pass” one or more of these assessments and yet be substantially contaminated through other means. Likewise “failure” in methods such as measurement of crystallinity does not necessarily imply contamination. None of these methods can be used to apply a quantitative correction to measurements of elements of interest such as strontium. Nonetheless, evidence of severe contamination implies that any measurement will have an unknown amount of contamination and thus should not be used for biological data. It should also be remembered when using these tests that the absence of evidence for contamination is not the same as evidence of the absence of contamination.

Many, probably most, archaeological samples fail many of these tests of biological integrity. Thus, much subsequent research has focused on ways to remove diagenetic contamination, mainly through combinations of mechanical cleaning followed by chemical cleaning by washing with acid.

Initial efforts to reduce or eliminate contaminants involved the removal of the outermost cortex and the inside surface of the bone by physical abrasion (Lambert et al., 1989). This abrasive cleaning greatly reduced the amounts of potassium, iron, aluminum, and manganese, which are abundant in soil oxides and clays but are characteristically low (<1000 ppm) in fresh bone. Later studies (Henderson et al., 1983; Katzenberg, 1984; Pate and Hutton, 1988; Williams, 1988; Williams and Marlow, 1987; Williams and Potts, 1988) of the amounts and distributions of trace elements that are common in soil but are absent in biological bone (e.g., Y, U, Th, and rare earths) indicate that bones are not just contaminated in the surface regions but can be pervasively affected throughout by adsorption and cation-exchange. Abrasive cleaning removes the most strongly affected portions of the bone, along with adhering soil minerals, but it does not remove the more chemically mobile components, especially carbonate salts that might have penetrated the entire bone and can introduce high levels of strontium and barium. Initial efforts to remove this chemical contamination by Krueger and Sullivan (1984) involved soaking bones in dilute acetic acid to dissolve selectively and wash out the more soluble minerals. Although Ca/P values were often much greater than two even after acid washing, the Ca/P values were almost invariably lower in the washed bone.

To study the effectiveness of acid washing in more detail, Sillen (1986, 1989) sequentially washed bone powder with dilute, buffered acetic acid and analyzed the acid-extract rather than the bone mineral. Sillen discussed the differential solubility of carbonates, hydroxylapatites, and fluorapatites. In acetate buffer at pH 4.5, apatites are increasingly soluble as a function of carbonate content and decreasingly soluble as a function of fluoride content (Biltz and Pellegrino, 1977; Sillen, 1986). Sillen (1986) examined a series of 2-million-year-old bones of herbivores, carnivores, and omnivores from the Omo Basin in Ethiopia. Supernatants from a series of 24 washes of bone powder in an acetate buffer at pH 4.5 were analyzed for elemental content. Sillen found that his initial extracts contained a soluble component with excess calcium (i.e., Ca/P > 2.14). The initial extracts did not show the expected separation of strontium values by trophic level, but after repeatedly washing the bone with fresh aliquots of acid, the Ca/P values returned to the expected biological value and the proper relative trophic positions appeared in the strontium values in intermediate solutions in the acid-washing
profile. A later analysis of animal bone from the site of Mehrgahr by Sillen showed a similar pattern. Sillen interpreted this pattern as an initial removal of highly soluble contaminants followed by slower dissolution of relatively less-soluble biological hydroxyapatite.

Price et al. (1992) experimented with a similar sequential procedure and confirmed Sillen’s results. They used a continuous acid-elution in which a small amount of bone powder (0.2 g) was placed in a tapered glass tube loosely plugged with glass wool to retain the bone powder. Dilute acid was flushed continuously through this tube, and the effluent was saved in sequential aliquots for analysis. Using this procedure on modern, unaltered bone, elemental concentrations quickly reached a steady value and the Ca/P ratio replicated the biological value of two. Washes of archaeological material, however, almost invariably had Ca/P ratios that were highly increased in initial washes and fell slowly to a steady value that was commonly close to but often slightly above the biological ratio. Values for barium and strontium from the final extracts, for cases in which the solution profile was flat and Ca/P = 2, had the expected separation of herbivores and carnivores. They cautioned that, although Ca/P ratios recover to the expected, biological value, this shift does not necessarily indicate that values for other elements also recover to biological values.

Lambert et al. (1990) and subsequently Lee-Thorp and van der Merwe (1991) experimented with various cleaning procedures and recommended that unbuffered acetic acid not be used because the stronger acidity causes rapid recrystallization of bone, at least partly trapping contaminants in the newly recrystallized material. They recommend washing bone with dilute acid (0.1 M#) buffered to a pH of about 6.

ENAMEL

In sharp contrast to the propensity of bone for contamination, enamel seems to be extraordinarily stable. Although bone must be porous and reactive, enamel has to be chemically inert under a broad range of possible chemical environments, including exposure to acidic environments in which calcium phosphates are normally soluble. The properties of bone that make it highly susceptible to contamination, i.e., porosity and chemical reactivity, are absent in enamel. Enamel is deposited early in childhood as dense, well-crystallized hydroxyapatite that is nearly devoid of either pore spaces or organic material, which could decompose to acids that advance mineral decomposition (White and Hannus, 1983). Accordingly, enamel is far more robust against conditions that promote either physical or chemical contamination. Although any calcified tissue under appropriately severe burial conditions could be totally dissolved or completely replaced by nonbiological minerals, enamel often persists under conditions where bone cannot. Bones buried for only a few years commonly show signs of decomposition and contamination; tooth enamel thousands of years old commonly retains its biological attributes as well as an absence of nonbiological contaminants such as uranium. Beyond 10,000 years, however, even enamel can exhibit evidence of chemical contamination through diffusion from the surface (Schoeninger et al., 2003).

Because most enamel (excepting that of the third molar) is deposited relatively early in childhood and—in contrast to bone—does not chemically remodel, the chemistry of the enamel reflects that obtained through the diet of early childhood and thus has the potential to inform about changes in childhood diet (e.g., nursing and weaning) as well as pre- and postnatal processes (e.g., Dolphin et al., 2005; Lane and Peach, 1997). Dietary studies involving strontium and barium, even when the enamel is chemically pristine, are, however, still subject to the above caveats about chemical equivocality as well as to complications such as additional biopurification in the mammary gland (Wasserman et al., 1958) and reduced biopurification during the rapid, early development
of calcified tissues. On the other hand, the availability of deciduous teeth and even extracted adult teeth, in contrast to the accessibility of modern human bone, makes enamel valuable as a monitor in modern humans of exposure to bone-seeking toxins such as lead and cadmium and to examine correlations among trace-element levels and other health factors (e.g., Brown et al., 2002, 2004; Budd et al., 1998, 2004; Curzon and Losee, 1977; Dolphin et al., 2005; Losee et al., 1974).

Also, because enamel retains the elements of the diet during early childhood, it is most valuable for the study of human mobility and provenience. The amounts of strontium, barium, and lead can vary geographically and thus have some potential in enamel to distinguish places of human origin (Burton et al., 2003) but also vary, as discussed above, for many reasons besides geography. Lead and strontium, however, both have radiogenic and nonradiogenic isotopes. For each of these elements, the relative isotopic abundances vary geographically with geology, but they are unaffected by biological processes. Enamel retains these geographically specific ratios so that measurements of these ratios place strong constraints on the geographic origins of individuals (e.g., Bentley et al., 2003; Knudson et al., 2004; Montgomery et al., 2005; Price et al., 1998, 2000, 2002, 2004; Wright, 2005).

INSTRUMENTATION

Accuracy, Precision, and Sensitivity

Key factors for choosing instrumentation for elemental analysis of bone—or any other material—are precision, accuracy, and sensitivity. Precision is the ability of an instrument to replicate the same result, independent of how correct that measurement might be, i.e., the uncertainty in the magnitude of the measurement. Accuracy is the ability of the instrument to provide the correct answer. An instrument that gave readings of 68, 70, and 72 would be more precise than an instrument that gave readings of 80, 90, and 100, regardless of whatever the real value might be. If the true value in the above example were 100, then the second instrument would be more accurate, although less precise, than the first. An instrument measuring 98, 100, and 102 for the same sample would be both more accurate and precise. Sensitivity is the measure of the smallest amount that can be measured reliably.

An instrument that can measure microgram (0.000001 g) quantities is much more sensitive than one that can only measure milligrams (0.001 g) or more. Precision is an important limit to sensitivity; if the uncertainty, or imprecision, of a measurement is ±0.00050 g, then the instrument is incapable of reliable microgram measurements, but it can easily measure milligram quantities.

Although it might seem paradoxical, precision is generally the most important criterion of all three because sensitivity is limited to the level of precision. Moreover, if the measurements are precise, an analysis can be made of standard samples, of known composition, and inaccuracies can be thus quantitatively determined and often corrected. If the results were not precise to begin with, then such a correction factor would be limited by the imprecision. In cases where one needs to know the actual level in the sample, e.g., the amount of lead to assess exposure, then accuracy and precision are both important. Often in the measurement of bone composition, results are only for comparative purposes in which it has to be determined whether one sample set differs significantly from another, whereas the accuracy of the measurement itself is less important. In such cases, precision is more important.

Considerable variation exists among different instrumental techniques in their suitability for specific elements, but, ideally, any instrument designed for elemental analysis, with adequate sensitivity, precision, accuracy, and proper calibration, can be used to measure bone composition. Today the major types of elemental instrumentation include atomic absorption spectroscopy (AA), atomic emission spectroscopy (AES) including inductively
coupled plasma optical emission spectroscopy (ICP-OES or, more commonly, just ICP), inductively coupled plasma mass spectroscopy (ICP-MS), and neutron activation analyses (NAA or, sometimes, INAA for instrumental NAA).

Atomic absorption and emission spectroscopy (mainly inductively coupled plasma spectroscopy) are best for easily ionized atoms that absorb light in the visible wavelengths, e.g., those in the first columns of the periodic table, including some of major interest: Ca, Sr, and Ba, but they have lower sensitivity, being more suited for parts-per-million (e.g., for Sr, Ba and Zn) than parts-per-billion measurements (e.g., not for uranium and rare-earth elements). ICP-MS, currently one of the most popular tools for elemental analysis, has much higher sensitivity in the parts per trillion range, but suffers from molecular interferences in the low mass ranges (which include calcium and phosphorus). NAA likewise is best at the bottom of the periodic table for large atoms that are easy to “activate” by hitting them with neutrons (e.g., uranium and REE). A particular research issue might require an instrument with unusual quality in particular elements—for example, the need to determine uranium in dental enamel requires the high sensitivity of ICP-MS, but the measurement of strontium levels of hundreds of parts per million could be achieved easily by any of the above instruments.

One promising area of current research exploits the high sensitivity of ICP-MS by coupling a laser to the sample input of the ICP-MS (“LA-ICP-MS”). The high sensitivity of ICP-MS requires much smaller sample sizes so that an adequate amount can be obtained by having the laser ablate microscopic amounts of material from the surface of the sample. Thus, the method is minimally destructive. It does not require dissolution of the sample and allows multiple measurements to be made on a microscopic scale across the surface (e.g., across growth lamellae in a tooth; see Dolphin et al., 2005; Kang et al., 2004; Uryu et al., 2003).

Cost, Time, and Access

Excepting the exploratory LA-ICP-MS research, all of the above instruments have adequate data quality for most purposes. In practice, the choice of instrumentation rarely depends on data quality but on far more pragmatic grounds of availability. Key factors in choosing particular instrumentation tend to be the accessibility of the instrument, the “turnaround” time required for the analyses, and the costs of analysis. Obviously any one of these could be, and commonly is, prohibitive. If, for example, one can choose between an inexpensive but rarely available research instrument at a university and an expensive but immediately available commercial service, then funding and time priorities will be determining factors in the decision. The pragmatic issue of sample size is not likely to be significant where several gram quantities of clean bone tissue are easily available for destructive analyses, but for extremely small or extraordinarily precious samples such as teeth of early hominids, a technique requiring less than one milligram (e.g., ICP-MS) of sample might be selected, whereas a method requiring a half-gram of sample, or requiring that a flat section be made from the sample, would be unacceptable. Each laboratory will likely have its own preparation protocols depending on the particular requirements of each type of instrument and the types of samples that they normally analyze.

CURRENT PERSPECTIVE

Although most compositional studies of bone have used strontium and other elements to infer the amount of meat in prehistoric diets, we now know that such measurements indicate the degree of biopurification of the principal source of calcium rather than quantitative measures of dietary components. Herbivores have lower Sr/Ca ratios than the plants they eat, and carnivores have lower Sr/Ca ratios than the meat they eat, but multicomponent diets do not produce a simple correspondence
between bone Sr/Ca and the dietary abundance of meat and plants. That, along with the severe problems of diagenetic contamination, has led to the almost universal replacement of such elemental analyses by stable isotope analyses ($^{13}$C/$^{12}$C and $^{15}$N/$^{14}$N) of collagen and dental enamel to infer differences in diet. Nonetheless Sponheimer, Lee Thorp and colleagues (e.g., Sponheimer et al., 2005) have again picked up the trace-element gauntlet, using enamel to minimize diagenesis and creating empirically determined dietary categories instead of the traditional plant/meat dichotomy. Multielement studies of dental enamel, which has far less problems with diagenesis, are also currently being studied for their potential to inform about modern diet and health (e.g., Dolphin et al., 2005; Lane and Peach, 1997) as well as chemical and isotopic records of childhood residence. The exploration of new isotope systems, e.g., $^{44}$Ca/$^{40}$Ca (Clementz et al., 2003; Skulan and DePaolo, 1999; Skulan et al., 1997) and $^{34}$S/$^{32}$S (Richards et al., 2003), as indicators of trophic position, are also especially promising areas of current interest.

Despite of the disappointment, after more than a decade of research, in quantitatively applying strontium, other elements are present in bone such as magnesium and zinc, which have not been studied adequately from the perspective of anthropological applications. Although it would be premature to state that such applications exist, a greater understanding of the variations in these elements and recognition of the causes of such variations are issues worthy of additional investigation.

REFERENCES


INTRODUCTION

Molecular archaeology, the study of DNA from archaeological remains, is a relatively new and exciting field that uses techniques from molecular biology to address anthropological questions. As a result, researchers who work on ancient DNA need to be informed about population genetics, human genetics, and archaeology. This knowledge is particularly important for the design of a research project so that the data obtained are informative from an anthropological perspective, statistically robust, and comparable with data from modern populations.

The analysis of DNA from ancient bone and tissue is feasible because of advances in molecular genetics. The first experiments to determine whether DNA survived in ancient material used dried tissue, such as skin from a 2400-year-old Egyptian mummy (Pääbo, 1985a). A few years later, DNA was extracted from human bone (Hagelberg et al., 1989). The initial research was aimed primarily at successfully extracting the DNA, examining its state of preservation, and demonstrating its authenticity. To date, the results of ancient DNA investigations have been applied to questions in physical anthropology, archaeology, evolutionary biology, forensic science, and paleopathology. These data have shed some light on the sex of individuals, relationships between individuals within a cemetery, origin of migrant populations, history of animal and plant domestication, diseases in the past, and phylogenetic relationships between modern and extinct species, including Neandertals and modern humans. Although the results generated thus far are limited, ancient DNA research holds great promise in addressing archaeological questions; however, the technical difficulties and problems with contamination make it slow, and at times frustrating, work.

Most ancient DNA research has targeted mitochondrial DNA (mtDNA); however, several studies have examined nuclear DNA loci, including Y chromosome sequences for sex identification and short tandem repeat (STR) loci for determining relatedness between individuals (see BOX 15.1). The locus or loci chosen for analysis should depend on the question of interest. MtDNA and Y chromosome polymorphisms have been examined extensively in living populations (e.g., Bolnick et al., 2006; Comas et al., 1998;; Hammer et al., 1997; Knight et al., 2003; Tarazona-Santos et al., 2001.
MITOCHONDRIAL DNA

The mitochondria are located in the cytoplasm and are the energy-producing organelles of the cell. Each cell typically contains several hundred mitochondria, and each mitochondrion has several copies of its own DNA. MtDNA is circular and approximately 16,500 base pairs (bp) in length. Most of the mtDNA genome is a coding sequence except for the displacement loop (or D-loop) where replication is initiated and except for some short sequences between genes (Fig. 15.1). Because of the high copy number of mtDNA per cell, it is more likely to survive over time than nuclear DNA. As a result, it has been the subject of most ancient DNA analyses. In addition, mtDNA is maternally inherited, does not recombine, and has a higher mutation rate. These features make it useful for population genetic analyses. Specifically, the lack of recombination makes it easier to analyze since it is not shuffled each generation, and the higher mutation rate makes it useful for understanding recent evolutionary history. Finally, although both men and women have mtDNA, only women can pass it to their offspring, and therefore, it is used to examine female population history (including migration rates and effective population sizes).

MtDNA analyses have focused on the two hypervariable segments of the control region, a non-coding “spacer” region where one of two origins of replication is located, on restriction sites located throughout the mitochondrial genome, and on a 9-bp deletion found in some populations in a small noncoding segment between the cytochrome oxidase subunit II gene and a gene coding for 13 proteins (or subunits of proteins), 2 ribosomal RNA subunits, and 23 transfer RNAs (necessary because the mtDNA genetic code is slightly different from the mtDNA genetic code). Different genes are found on the heavy (H) and light (L) strands of the DNA. These are transcribed in opposite directions. (Reprinted with permission from Page and Holmes, 1998.)

Figure 15.1  The mtDNA genome codes for 13 proteins (or subunits of proteins), 2 ribosomal RNA subunits, and 23 transfer RNAs (necessary because the mtDNA genetic code is slightly different from the mtDNA genetic code). Different genes are found on the heavy (H) and light (L) strands of the DNA. These are transcribed in opposite directions. (Reprinted with permission from Page and Holmes, 1998.)
for a lysine tRNA. Characteristic polymorphisms have been used to define mtDNA haplogroups (a group of similar mtDNA lineages). One limitation of mtDNA is that its lack of recombination means that it should be treated as one locus (since everything is linked and passed from mother to child as a unit) in population genetic analyses. This reduces the statistical power of conclusions about population history that only use this locus.

NUCLEAR DNA

Although nuclear DNA is more difficult to recover from skeletal remains, the large number of different loci that can be examined provides the opportunity to examine relationships among populations, and kinship relationships, identify the sex of an individual, and distinguish disease-causing alleles. Nuclear DNA is present in the nucleus of the cell. In humans, nuclear DNA (Fig. 15.2) consists of 22 pairs of autosomes and 1 pair of sex chromosomes (the X chromosome and the Y chromosome). Nuclear DNA is passed down from both parents (i.e., you receive half from each parent). Nuclear DNA recombines each generation, and thus, it represents a mosaic of DNA from all ancestors. The one exception is the nonrecombining portion of the Y chromosome (the tips of the Y chromosome, known as the pseudoautosomal region, do recombine with the X chromosome during meiosis). Because nuclear DNA is only present in one copy in each cell, it is more difficult to recover from ancient remains. However, because many independent markers can be examined, it offers the opportunity to rigorously test hypotheses with multiple independent loci. Such markers can be coding or noncoding and can include single nucleotide polymorphisms (SNPs), insertions or deletions of DNA, and short tandem repeats (STRs or microsatellites).

Figure 15.2 A normal karyotype for a human female. (Reprinted with permission from Strachen and Read, 1999.)
Y chromosome

Most the Y chromosome does not recombine, and thus, it can be examined to understand male population history. As in mtDNA, the lack of recombination makes it easier to analyze since it is not shuffled each generation but limits its statistical power for conclusions about population history. Although DNA on the Y chromosome does not mutate as rapidly as mtDNA, many polymorphisms have been identified. Y chromosome sequences can also be used in the identification of the sex of an individual.

Xue et al., 2006), and thus, there are many comparative datasets for addressing questions about maternal and paternal population history in a given region of the world. In addition, each dataset provides information about sex-specific migration and relationships that can illuminate patterns of organization or admixture in a society (Merriwether et al., 1997; Oota et al., 2001; Seielstad et al., 1998; Zerjal et al., 2003). Finally, a few studies have investigated bacterial DNA or genetic disease loci from individuals with lesions suspected to result from a particular disease.

METHODS

Ancient DNA can be extracted from cellular remains such as tissue (preserved in water, frozen or dried), bone, tooth roots, coprolites, seeds, and other plant materials. Typically, the first step in the extraction procedure is to prepare the sample by removing any surface contamination from previous handling of the material. This removal is particularly important for human remains, which may have been handled by numerous excavators, archaeologists, and osteologists as well as by the laboratory workers. Surface contamination can be removed by cutting or grinding away the exposed layers, irradiating the surface with ultraviolet (UV) light, or soaking the material in a hydrochloric acid or bleach solution. After a bone or tooth root sample is cleaned, it is ground to dust in a bone mill or other grinder, whereas soft tissue samples are cut into fine pieces. This process increases the surface area of the material, enhancing the release of the DNA from the material during the extraction.

Two principle techniques have been used to extract DNA from ancient remains. The first technique is a proteinase K digestion followed by a phenol/chloroform extraction (Hagelberg and Clegg, 1991). The proteinase K digestion works to break up the proteins in the tissue to release the DNA, whereas the extraction separates the DNA from the proteins in the solution. This method is standard in biology and results in a high yield of DNA. One problem, however, with the phenol/chloroform protocol for ancient DNA work has been the coextraction of inhibitors that make it difficult to copy and analyze the DNA. The other primary method, the silica/guanidine isothiocyanate technique (Höss and Pääbo, 1993), has the advantage of removing the inhibitors. In addition, it is a fairly simple and fast method, but the DNA yield may not be as high. Other methods for extracting DNA from ancient remains include using silica-based spin columns (Cano and Poinar, 1993; Yang et al., 1998), cetyltrimethylammonium (CTAB) buffer (Yang et al., 1997), chelex (Faerman et al., 1995; Poinar et al., 1993), or a combination of the two primary methods (Krings et al., 1997).

DNA Preservation

The original environment and the treatment of samples after they are discovered affect the preservation of ancient DNA. In general, samples recovered from environments with cooler temperatures, neutral or slightly alkaline pH, and dry conditions are best for DNA preservation, although samples found in wet anoxic conditions
(Hagelberg and Clegg, 1991; Lawlor et al., 1991; Pääbo et al., 1988) or frozen in permafrost (Hagelberg et al., 1994; Höss et al., 1994) have also yielded DNA. Differing microenvironments, even within the same burial or excavation site, may cause varying success for DNA analysis (Hagelberg and Clegg, 1991; Stone and Stoneking, 1999). To date, the oldest samples that have yielded reliable DNA results are from cave or permafrost environments (Hagelberg et al., 1994; Hänni et al., 1994; Höss et al., 1996a; Höss et al., 1994; Krings et al., 1997; Yang et al., 1996). The source of the material used in the extraction can also affect the amount of DNA recovered. Typically, the quality and quantity of DNA isolated from an individual is better from tooth roots than bone and better from bone than from soft tissue (O’Rourke et al., 1996). In mummified soft tissue, DNA is usually best preserved in peripheral tissues that are more likely to desiccate rapidly, thus escaping extensive degradation from lytic enzymes (Pääbo, 1985b; Pääbo et al., 1989). Other environmental factors include substances that can coextract with the DNA to inhibit PCR and make DNA analysis difficult. Fulvic acids, which are breakdown products of organic soils, and Maillard reaction products, which are produced during the initial decay of organic matter (Poinar et al., 1998; Tuross, 1994), fall into this category.

The methods used to preserve a sample after excavation can also influence the amount of DNA recovered (Cooper, 1993; Thuesen and Engberg, 1990), with some preservatives such as formaldehyde, gamma radiation, and tanning agents destroying or degrading the remaining DNA. In addition, surface coating of glue or varnish may introduce contaminants from the preparer, from previous samples that were treated with the varnish, or from the preservative itself (if it was of animal or plant origin) (Cooper, 1993; Krings et al., 1997). Solutions to these problems include using bovine serum albumin (BSA) in the PCR to bind some interfering molecules, removing surface treatments, modifying the extraction method as noted above, and withholding preservative treatment during or after excavation for samples that will be subject to DNA analysis.

Depending on the conditions of preservation both in and out of the ground, and after pushing the limit on molecular technology, the success rate of retrieving DNA from ancient remains at a given site ranges from 0% to 90%. Success also depends on the genetic locus examined.

**DNA Analysis**

After extraction, the polymerase chain reaction (PCR) is used to copy the DNA fragment of interest millions of times so that there is sufficient DNA for analysis (Box 15.2). Damage to the DNA and the low number of starting molecules (often less than 1% of the original amount) make PCR amplification of ancient DNA difficult (Handt et al., 1994b; Höss et al., 1996b; Pääbo, 1989).

After PCR, the DNA can be analyzed by cutting the DNA with restriction enzymes that cleave the DNA at sequence specific sites (to investigate restriction fragment length polymorphisms or RFLPs) or by sequencing the DNA to determine the order of nucleotide bases. Sequencing can be done in two ways. Direct sequencing is where one sequence is derived from the PCR product, and this sequence is essentially a summary of everything amplified by the primers. This technique is standard with modern DNA; however, PCR from ancient materials may amplify multiple products, including authentic DNA, damaged authentic DNA, or contaminating DNA. Thus, for ancient DNA research, sequences are often determined using cloning. Cloning takes a single DNA molecule from the PCR product, inserts it into a bacterium that then grows clonally in a colony copying the insert as it reproduces. After a sufficient number of bacteria exist, the insert is extracted and sequenced. Typically, many clones representing many independent molecules from the PCR product are examined. If all clones (or most) are consistent, it suggests that the DNA is well preserved and not contaminated.
BOX 15.2. PCR AND ELECTROPHORESIS

PCR copies, or amplifies, the fragment using a three-step process of DNA denaturation, primer annealing, and DNA extension (Fig 15.3a). First, the DNA is heated to separate (or denature) the double-helix structure into single strands. The temperature is then decreased to let primers stick (or anneal) to the DNA. Primers are short, single-stranded fragments of DNA (usually 15 to 25 base pairs long) that are complementary to a portion of the sequence of interest and that define the segment of DNA to be amplified. Last, the temperature is increased to approximately 72°C to enhance the activity of the Taq polymerase, an enzyme that sits down next to the primer and begins adding complementary bases to extend the DNA molecule. This cycle of denaturation, annealing, and extension is then repeated 25 to 40 times during PCR. Although initially only a small amount (maybe only a few molecules) of DNA is present, PCR increases the amount of a specific DNA fragment exponentially so that after PCR, millions of copies are present. After PCR, a few microliters of the PCR products, as well as a size standard, are loaded into an agarose gel (Fig 15.3b). An electric current is then applied, causing the slightly negatively charged DNA to migrate toward the positive pole. As the DNA migrates through the gel, large fragments move slowly, whereas smaller fragments move more quickly; thus, the DNA is separated by size. The DNA is then visualized with ethidium bromide or some other type of stain on a UV light table.

Regardless of the protocols used to extract and analyze ancient human DNA, or the locus examined, the greatest concern is contamination. Precautions must be taken to assure the authenticity of the results (Cooper and Poinar, 2000; Handt et al., 1994a; Handt et al., 1996; Richards et al., 1995; Stoneking, 1995) (Box 15.3). The care necessary to work with ancient DNA and the need for multiple independent analyses of materials makes this research slow, expensive, and sometimes very frustrating. Thus, the questions asked and the probability of actually getting sufficient data to answer the questions should be carefully considered before undertaking a project.
BOX 15.3. CRITERIA OF AUTHENTICITY

**Physically isolated work area:** Ancient DNA work should be performed in a dedicated laboratory that is completely separate from the main laboratory where modern DNA and post-PCR products are analyzed and stored to avoid contamination. This laboratory should not share an air supply or materials with other laboratories.

**Control amplifications:** The reagents must also be tested for contamination by including “blank” tubes without any DNA in each extraction and PCR. Positive controls should be avoided since these could contaminate the samples.

**PCR product size:** PCR product yield should be related inversely to product size. Authentic ancient DNA usually only yields PCR products that are smaller than 300bp. Thus, to examine larger sequences, multiple primer sets that amplify smaller regions that overlap are used (the sequences of the overlapping fragments can also be compared to confirm results).

**Cloning:** Direct PCR sequences should be verified by cloning PCR products to identify the percentages of authentic and contaminating sequences and damage-induced errors.

**Preservation:** The preservation of DNA in a sample can be assessed by examining the extent of degradation in amino acids or other biomolecules. In addition, the extraction of DNA from associated faunal remains can be used to gauge the preservation of DNA in human remains. This can also provide a means of testing whether the remains from a site are contaminated by modern human sequences since these should not be found in faunal remains.

**Reproducibility:** All results must be confirmed by multiple independent extractions and analyses. Ideally, separate samples of a specimen should be extracted and sequenced in independent laboratories to confirm results. The results should also make phylogenetic sense (i.e. the sequences should not match the investigator or belong to a cow when you are expecting a horse sequence).

**DNA quantification:** The copy number of starting molecules can be determined using a competitive or labeled PCR. If the starting number of template DNA molecules is low (<1000), the results may not be reproducible.

---

**Analyses**

The data obtained from ancient DNA are typically examined using standard population genetic analyses. Ancient DNA data can be used to investigate many different questions of anthropological interest, and these questions can focus on an individual or on an entire community or region. Specific analyses are chosen based on the hypothesis that is being tested. For population-level analyses that address hypotheses about the relationship of an ancient population to another group (either modern or ancient), large datasets (from ≥25 individuals) from each group as well as data from multiple loci are preferable (Jobling et al., 2004; Nei, 1987). Larger sample sizes prevent erroneous conclusions from sampling error. For example, Stone and Stoneking (1993) found that one mtDNA haplogroup at the Norris Farms site was only present in males, suggesting that these individuals could represent maternally related males who immigrated into the community, possibly to offset losses from violent conflict. Subsequent research, however, found this haplogroup in females as well (Stone, 1996; Stone and Stoneking, 1998). Larger sample sizes are also helpful for investigating the relationships between individuals within an archaeological site. If two people share a genetic polymorphism, or allele, that is rare in
that population, they are more likely to be closely related than two individuals who share a common polymorphism.

Just as few samples can lead to misleading results, the examination of only one locus may not adequately represent the history of a population. For example, if the maternal and paternal histories of a population are different, then only analyzing one type of marker could lead to conclusions that are only partially correct. This problem may occur in cases of directional mating where males remain in the village but marry females from outside the population; the levels of diversity and ancestry reflected in different loci may tell different stories (Merriwether et al., 1997; Oota et al., 2001). Multiple markers are also needed to test the hypothesis of a close relationship such as a parent–child or sibling relationship between two individuals that is statistically supported. Such analyses are similar to those used in genetic tests of forensic samples.

Sample size and the analysis of multiple markers may not be as important in some cases. In a single individual, DNA can identify biological sex or the cause of a skeletal disorder and, in some special cases such as with Neanderthal or Cro-Magnon remains, can provide evolutionary information about our species. DNA can also be used to identify the species classification of plant and animal remains for dietary information or to determine the identity of a person recovered in a forensic investigation. The following studies provide examples of the different kinds of questions addressed using analyses of ancient DNA.

CASE EXAMPLES

What Can an Individual Sample Tell You?

In special cases, individual samples can provide answers to archaeological questions. Stone and Stoneking (1996) analyzed mtDNA sequences from an 8000-year-old skeleton found in a high-altitude cave in Colorado. The results indicated that haplogroup B, thought by some to have been brought to the New World by late migrants from the north or from Polynesia, was indeed present in early Native Americans. On the other hand, the mtDNA sequence of the Iceman, a man found high in the Tyrolean Alps and dating to the Copper Age, belongs to a haplogroup that is common in modern Europeans (Handt et al., 1994b). This result was not particularly surprising.

Individual samples can also be used to address phylogenetic questions. Although most of human evolutionary history is beyond the scope of ancient DNA analyses because of age, DNA was extracted successfully from the Neandertal-type specimen, a skeleton found in 1856 in the Neander Valley of Germany (Krings et al., 1997, 1999). This individual has classic Neandertal morphological features, and radiocarbon dating indicates an age of approximately 40,000 years, making it one of the oldest hominin skeletons discovered with preserved DNA (Schmitz et al., 2002).

DNA was obtained in multiple extractions from the humerus of Neandertal 1. To examine the preservation of DNA in the bone, amino acid racimization as well as quantification assays were performed (Krings et al., 1997). These analyses indicated, first, that DNA should be recoverable since the amino acid preservation was in the range compatible with DNA survival (Poinar et al., 1996) and, second, that PCR began from approximately 50 molecules. For results to be consistent, PCR should begin with 1000 molecules or more (Handt et al., 1996). As a result, it was necessary to reproduce results by using multiple PCRs and by sequencing multiple clones to rule out sequence changes caused by nucleotide misincorporation during early stages of amplification or by sporadic contamination. By amplifying and cloning small overlapping segments, the mtDNA sequence for a total of 357 bp of the HVI and 340 bp of HVII was obtained.

The Neandertal sequence was then compared with almost 700 modern human sequences from around the world. This comparison showed an average of about 35.3 differences between
modern humans and the Neandertal sequence. The average number of differences between any two modern humans was about 10. European sequences were not found to be more closely related to the Neandertal sequence than those of other modern humans. Using population genetic methods, the age of the common ancestor of modern human and Neandertal mtDNA sequences was estimated to be 465,000 years (confidence interval of 317,000–741,000 years), almost three times the age of the common ancestor of modern human mtDNA sequences (163,000 years). This result points to a long period of time during which Neandertals were distinct from the lineage leading to modern humans. Since the Neandertal sequence seems to fall outside the range of modern human variation and since European mtDNA types did not seem to be more closely related to the Neandertal sequence, Krings et al. concluded that Neandertals became extinct without contributing mtDNA to modern humans (Krings et al., 1997, 1999).

More recently, mtDNA HVI sequences from three additional Neandertals were recovered (Krings et al., 2000; Ovchinnikov et al., 2000; Schmitz et al., 2002). The sequences from these individuals show similar distances from modern human sequences as the Neanderthal-type specimen. The mean difference among the four Neandertal sequences is 1.7%, suggesting that Neandertal diversity was relatively low compared with that found in great apes but similar to that found in modern humans (1.8%) (Schmitz et al., 2002). Sequences obtained to date from early modern humans in Europe resemble those of modern humans (Caramelli et al., 2003; Serre et al., 2004). Thus, these subsequent results also support the hypothesis that Neandertals became extinct with little or no admixture with anatomically modern humans.

**How Can We Test for Population Replacement or Continuity?**

From an archaeological perspective, a migration into a previously uninhabited area can be fairly easy to document, whereas the movement of new peoples into a previously occupied area may be difficult to distinguish from the movement of new trade items or ideas. In both cases, the source of the migrant population and the numbers involved are of interest and can be estimated using ancient DNA data.

mtDNA diversity in ancient Native American populations of the Great Basin has been examined to investigate population change. In this region, the Fremont culture disappears from the archaeological record after about 1300 A.D. A new tradition identified as Paiute–Shoshoni then emerges. This chain of events has been hypothesized to indicate a replacement of the Fremont by Numic-speaking migrants ancestral to modern Utes, Paiutes, and Shoshoni during the fourteenth century. To test this hypothesis, mtDNA markers characteristic of four haplogroups (A, B, C, and D) common in Native Americans were examined in 47 prehistoric Fremont skeletons from Utah (Carlyle et al., 2000; Parr et al., 1996). The remains, dating from 252 A.D. to 1296 A.D., were excavated from the Great Salt Lake wetland. Of the 47 people sampled, the haplogroup could be identified for 32 (68%), whereas DNA from two individuals did not fall into one of the four major haplogroups. Haplogroup A lineages were not present in the sample, while group B lineages were most common (Table 15.1). In modern Numic-speaking Shoshone and Paiute populations, the frequency of the 9-bp deletion (group B) is lower (Lorenz and Smith, 1994, 1996), and Parr et al. (1996) suggested that this provides evidence that Numic speakers did replace the Fremont populations. Carlyle et al. (2000) found that the Fremont seem more closely related to modern Southwestern populations such as Jemez Pueblo.

Analysis of mtDNA in remains excavated from sites in western Nevada provides additional evidence that Numic-speakers replaced ancient Great Basin populations (Kaestle, 1997; Kaestle and Smith, 2001; Kaestle et al., 1999). Twenty individuals from Pyramid Lake (860 to 5905 years B.P.) and 27 individuals
from Stillwater Marsh (290 to 3290 years B.P.) were tested for five mtDNA haplogroup (A, B, C, D, and X) markers common in Native Americans. Of these markers, haplogroups were identified in 18 (90%) samples from Pyramid Lake and 21 (78%) from Stillwater Marsh. Haplogroups B and D were most common at Pyramid Lake and Stillwater Marsh (Table 15.1), whereas neither haplogroup C nor X was present. Haplogroup A was present in only 11% of the samples from Pyramid Lake and 5% of the samples from Stillwater Marsh. Kaestle and Smith (2001) concluded that the ancient people from western Nevada are most closely related to modern California Penutians and significantly different from modern Numic-speaking populations as well as from the Washo, a population living in the Great Basin and speaking a language that has not been clearly linked to other languages. Analyses also showed significant differences in mtDNA haplogroup frequencies between the Great Salt Lake wetland Fremont sample and both samples from western Nevada. Thus, although all three ancient samples appear to be unrelated to modern Numic speakers in the Great Basin, those from western Nevada seem to be more closely related to California populations and those from Utah seem to be more closely related to Puebloan groups. The analysis of Y chromosome and autosomal loci and the modeling of evolutionary forces such as genetic drift through time would provide additional support for the hypothesis of replacement.

### How Can We Use DNA to Help Reconstruct Social Organization at a Prehistoric Site?

DNA data that are used to identify the sex of individuals buried in a cemetery and understand the relationships between individuals can be included with archaeological and osteological data in mortuary analyses to illuminate the social organization of past peoples. The population history of a prehistoric population and the association between DNA types and cemetery organization were investigated in a 700-year-old Native American community from central Illinois (Stone and Stoneking, 1993, 1998). From a burial population of approximately 270 individuals, 152 were selected for DNA analysis. Four markers that define the four major Native American mtDNA haplogroups (groups A–D) were examined. In the Norris Farms population, the mtDNA haplogroup was identified successfully in 108 people (71%) using these markers, and all four major groups of lineages were found. In addition, six individuals had mtDNA types that did not seem to belong to one of these haplogroups. Sequence analysis of a 353-bp segment of HVI was also performed in 52 (34%) of the 108 individuals for whom the
mtDNA haplogroup was defined successfully. These individuals included members of the four primary clusters as well as those who did not seem to fall into one of these groups, in roughly proportional numbers to the presence of these groups in the cemetery. Twenty-three lineages were retrieved (Stone and Stoneking, 1998) with characteristic HVI mutations that were concordant with the haplogroup designation in each individual. Two additional lineages were recovered that were determined to be the result of contamination from modern sources and thus not included in additional analyses.

In addition to mtDNA analyses, the sex of 46 individuals was investigated using the sequence-specific probe method of Stone et al. (1996), which targets the amelogenin gene (see Box 15.4). The sex of 21 (47%) individuals was identified, whereas PCR amplification was not successful in 25 samples (Stone and Stoneking, 1999; Stone et al., 1996). In adults with clear morphological indicators of sex, the DNA results were in agreement with the morphology with the exception of one individual where the difficulty of PCR amplification indicated insufficient DNA for successful molecular analysis (Stone et al., 1996). Unfortunately many of the samples that were not successful for amelogenin analysis were juveniles, which may be from poorer preservation and density of subadult bones.

The DNA data were used to examine pre-Columbian mtDNA diversity, to investigate

Figure 15.4 Spatial patterning of mtDNA lineages in the Norris Farms Cemetery from the initial study (Stone and Stoneking, 1993). Outlines indicate the position of burials, and the arrow indicates a burial found with two individuals, each with the Hinc II -nt. 13259 marker (haplogroup C).
the genetic history of the peopling of the New World, and to detect patterns of spatial organization in the Norris Farms cemetery. The spatial patterning of mtDNA lineages in the Norris Farms cemetery investigated to determine whether maternally related individuals were more likely to be interred near each other (Stone, 1996; Stone and Stoneking, 1993). Such a pattern was not found, although several graves contained multiple individuals with the same mtDNA haplogroup (Fig 15.4). Initial results from Norris Farms showed that one haplogroup was only present in males, suggesting that these individuals could represent maternally related males who immigrated into the community, possibly to offset losses from violent conflict (Stone and Stoneking, 1993). Subsequent research, however, found this haplogroup in females as well (Stone, 1996; Stone and Stoneking, 1998).

The genetic data were combined with data on grave form and location, burial goods, age, morphologically determined sex, skeletal and dental anomalies, and circumstances of death (Stone, 1996). The results indicate that this was a egalitarian community where people were primarily differentiated, at least upon their deaths, by age and sex. In addition, the mtDNA data suggest women were no more fixed than men in terms of where they lived after marriage and point to possible relationships between a few high ranking men as well as among people buried in several graves containing more than one person (Stone, in preparation).

**BOX 15.4. SEX IDENTIFICATION**

Molecular sex identification tests have been used in anthropological research to indicate the sex of juveniles and fragmentary remains and to compare molecular results with morphological and archaeological indicators of sex. Molecular methods identifying the sex of an individual are particularly useful for juvenile and fragmentary remains where sex determination is difficult by traditional morphological techniques. Methods of sex identification using DNA focus on either repetitive DNA sequences that are chromosome specific (Kogan et al., 1987; Witt and Erickson, 1989) or single-copy genes found on both the X and the Y chromosomes (Aasen and Medrano, 1990; Cui et al., 1994; Ebensperger et al., 1989; Nakahori et al., 1991a). Analysis of ancient DNA for sex identification (see Brown, 1998 for another review) has adapted some of these methods for degraded DNA and primarily has focused on repetitive sequences (Honda et al., 1990; Hummel and Herrmann, 1991; Ovchinnikov et al., 1998) or the amelogenin gene (Faerman et al., 1995; Gill et al., 1994; Stone et al., 1996).

Repetitive sequences on the Y chromosome that have been examined include the DYZ1 and DYZ2 repeat sequences, which are found at the ends of the chromosome, and alphoid satellite sequences, which are long repetitive DNA sequences that are found in the region around the centromere. DXZ3 repeat sequences from the X chromosome have also been examined (Ovchinnikov et al., 1998). The sequences of these repeats are chromosome specific, and each is present in hundreds of copies. Because of the high copy number, repeat sequences should be much easier to amplify in ancient remains than single copy sequences. However, one major flaw in this system exists: Although the presence of a product after PCR amplification of a Y-specific repeat sequence indicates that an individual is male, the absence of a product does not necessarily indicate that an individual is female. PCR amplification failure may simply be the result of low-quantity or low-quality DNA or of PCR inhibition by other components in the sample.

The amplification of both X and Y chromosome sequences with one set of primers solves the dilemma revealed by repeat sequence analysis for sex identification. Examining these sequences with one set of primers is accomplished by amplifying a fragment of a gene found on both
chromosomes that have slight sequence differences on the X and Y chromosomes that allow the source of the fragment to be identified. One such gene is the amelogenin gene that is involved in enamel formation and has sequences on the X and Y chromosome that are approximately 90% homologous (Nakahori et al., 1991b). The most common method employed to identify sex using this gene amplifies a region containing a size difference between the X and Y fragments (Sullivan et al., 1993). The PCR products are then easily distinguished using gel electrophoresis (Fig. 15.5).

Although the amelogenin systems may not be as sensitive to degraded, low-quantity DNA as using the X or Y chromosome repeats, it is easily apparent whether the reaction fails or works. Several researchers have noted, however, that the small amount of ancient DNA that is the starting material for PCR may only include a few molecules (or even a single molecule!) of DNA from the region of interest, and thus, sometimes only the X or only the Y copy may be amplified during a particular PCR (Faerman et al., 1995; Lassen et al., 1996; Stone et al., 1996). This problem is referred to as allele dropout and is solved by performing multiple PCR experiments. For example, assuming that both X and Y copies are equally likely to be amplified, the Y chromosome copy of the fragment should be detected after four successful PCR attempts with a 94% probability if it is present. This assumption may not be true where the copies differ by size since smaller fragments are more easily amplified from ancient DNA (Handt et al., 1994a).

How Do We Test for Family Relationships Between Individuals in the Same Tomb or Cemetery?

Kurosaki et al. (1993) used ten STR loci (see Box 15.5) to examine kinship in two sets of remains from the Hirohata and Hanaura sites on the southern main island of Japan. The Hirohata site is approximately 1500 years old, and 26 individuals were excavated from 16 chamber tombs. At the Hanaura site, which is approximately 2000 years old, 29 burials were recovered. These people were individually buried in earthenware jars. The individuals chosen for ancient DNA analysis at each site were thought to be related based on archaeological data. At Hirohata, two males, one juvenile, and one adult buried in the same tomb were sampled. Two individuals from Hanaura were also sampled. These individuals are an adult and a juvenile, both thought to be female, buried side by side about halfway up the hill.

Figure 15.5 Results of PCR using the Sullivan et al. (1993) method. Lane 1: 1-kb ladder (size standard). Lanes 2–4: One band is present, indicating that these individuals are female. Lanes 5–6: Two bands are present indicating that these individuals are male. Lane 7: PCR blank.
DNA was extracted from both tooth and bone samples for each individual. Eight dinucleotide STRs, one trinucleotide STR, and one pentanucleotide STR were examined using PCR. In the males from Hirohata, nine loci (all but one of the dinucleotide STRs) could be amplified. The results are consistent with kinship among these individuals. Although their results seem to corroborate the archaeological data, Kurosaki et al. (1993) are cautious in defining the degree of kinship since the frequencies of these STR alleles are unknown in this ancient population.

At the Hanaura site, two females found buried side-by-side in a hillside were hypothesized to be mother and daughter. Alleles in three of the seven STR loci examined in these individuals do not match, thus ruling out a parent–child relationship. A comparison of their mtDNA sequences also precluded a kinship relationship on the maternal side, such as maternal aunt and niece. Their results do not rule out kinship on the paternal side.

**BOX 15.5. SHORT TANDEM REPEATS**

The biological relationship of two individuals can be established with a high degree of certainty using multiple, highly polymorphic loci. STR sequences are commonly used for this purpose in modern forensic cases. STRs consist of sequences of two to five bases that are tandemly repeated, sometimes hundreds of times, in a region of nuclear DNA (Fig. 15.6). For individual identification or the establishment of relationships between individuals, typically 10–20 independent STR loci are necessary.

In some ancient samples with low DNA quantity, only one allele may be amplified during one PCR, apparently indicating homozygosity. This process is referred to as allelic dropout. In subsequent PCRs, two alleles may be found, indicating heterozygosity, or there may be homozygosity for a different allele, making the results difficult to interpret. Problems with allelic dropout underscore the need for DNA quantification to identify problematic samples as well as for multiple PCR tests of samples to ensure accurate results. Zierdt et al. (1996) noted that such allelic dropout may be the cause of the apparent deficiency of heterozygotes in a population.

**How Do We Learn About Disease in the Past?**

Disease in the past can be placed into two general categories, infectious and genetic, although some “genetic” diseases such as cancer can be influenced by the environment and even by infections (e.g., cervical cancer and papilloma virus). Several genetic disorders, such as thalassemia, dwarfism, and Gaucher’s disease, result in skeletal changes that can be identified in ancient populations (Ortner and Putschar, 1981). The genetic mutations that cause many of these disorders have been discovered and can be examined in past peoples. Filon et al. (1995) examined DNA from the skeletal remains of a child excavated from a Phoenician cemetery in Israel approximately 3800 years old. The child died at the age of 8 years with skeletal pathology indicating severe anemia. Anemia can be caused by genetic and environmental conditions, including malnutrition, sickle cell anemia, and thalassemia. A 232-bp segment of the β-globin gene, a component of hemoglobin, was examined, and the child was found to be homozygous for...
a mutation that causes a null phenotype (i.e., no β-globin to be produced). This mutation is found at a frequency of 2–10% in modern populations living in the eastern Mediterranean region. Most people with this particular form of thalassaemia require blood transfusions from early childhood (when fetal hemoglobin is replaced by adult hemoglobin) to survive. The ancient child, however, lived to the age of 8 years, and haplotype analysis of the DNA sequence suggests that the child had increased levels of fetal hemoglobin that enhanced survival (Filon et al., 1995).

Using DNA to diagnose pathogenic diseases in prehistoric skeletons is intriguing because it provides the opportunity to look at the evolution of a disease-causing agent, to investigate the history of the spread of a disease around the world, and to identify the causative agent in cases where several agents may result in similar morphological changes in the skeleton. Sequence information from ancient pathogens can help in the understanding of why some strains of a virus or bacteria are more virulent than others. Pathogens such as the flu, tuberculosis, and the plague have had large effects on human populations, and ancient DNA analysis can add insight into what may have been significant forces of natural selection in humans.

Ancient DNA techniques also allow researchers to reexamine old archived medical specimens that may reveal that a newly discovered pathogen was the cause of disease in someone in the past and, thus, that this pathogen has been affecting people for some time. For example, pathogenic DNA was first recovered from archived paraffin-embedded tissues approximately 40 years old. The results indicated that the samples were positive for human papilloma virus (Shibata et al., 1988). More recently, sequences from the 1918 flu virus, which was particularly virulent, were obtained from preserved samples from a U.S. serviceman who died during the epidemic (Taubenberger et al., 1997). Additional samples were then obtained from the lung tissue of an Alaskan victim of the pandemic who was buried in permafrost, and the complete sequence of the virus (eight genes in total) was recovered (Basler et al., 2001; Reid et al., 1999, 2000, 2002, 2004; Taubenberger et al., 1997, 2005). These analyses revealed that the 1918 flu was an avian virus that had adapted to humans and identified 10 amino acid changes important for this adaptation.

DNA from pathogenic agents is probably the most difficult type to recover from ancient skeletal and mummified remains since the total amount of viral or bacterial DNA present is likely to be small even in a heavily infected individual. However, the opportunity to answer questions about disease evolution and history has led to many attempts to examine ancient pathogen DNA. For example, tuberculosis is a reemerging infectious disease with an ancient history in humans. A long-standing controversy in physical anthropology concerns the origin of TB, particularly in the pre-Columbian Americas. Twentieth-century researchers such as Hrdlicka (1909), Morse (1961), and Cockburn (1962) did not recognize the extrapolmonary spinal TB (Pott’s disease) present in the ancient New World, and thus, they proposed an Old World origin for the disease. This idea was largely based on the notion that infectious diseases only developed with the advent of agriculture, animal domestication, and relatively large, permanent settlements of humans (Buikstra, 1976, 1981). The high morbidity and mortality from TB in Native American populations after European contact bolstered the Old World origin scenario, with New World groups considered “virgin soil” for the organism.

By the mid-twentieth century, however, pre-Columbian skeletons with classic Pott’s disease had been described from several areas in North America, with the earliest dated around 1000 A.D. In South America, skeletons and mummies confirmed the presence of TB in even earlier times (certainly by 700 A.D.) (Roberts and Buikstra, 2003). Although many researchers adhered to the Old World origin scenario, others explored the possibility that M. tuberculosis was indeed present in the pre-Columbian New World (e.g., Buikstra, 1976,
1981, 1999; Daniel, 2000 [and references therein])

Salo et al. (1994) were able to recover tuberculosis DNA from a lung lesion in a 1000-year-old mummy from Peru. A small (123bp) fragment of an insertion element called IS\textit{6110} present in several copies in the genomes of members of the tuberculosis complex (which includes \textit{M. tuberculosis}, \textit{M. bovis}, \textit{M. canettii}, \textit{M. africanum}, and \textit{M. microti}) was amplified. These results indicate that tuberculosis was present in the New World before European contact. Molecular analysis was subsequently used to confirm the presence of tuberculosis in a 12-year-old girl from Chile who died approximately 900 years ago (Arriaza et al., 1995), and Braun et al. (1998) also amplified IS\textit{6110} from fifteenth century Canadian and eleventh century Middle Mississippian bone samples to establish further the presence of \textit{M. tuberculosis}.

The presence of the IS\textit{6110} element, however, only points to the presence of bacteria of the tuberculosis complex and does not provide information about the species or the specific strain. Thus, it was suggested that the New World pathogen was \textit{M. bovis} rather than \textit{M. tuberculosis} (Stead, 1997; Stead et al., 1995), based on the hypothesis (Cockburn, 1962; Rich, 1944) that human TB developed from the bovine variety in the Old World. Were this indeed the case, it would have been impossible for \textit{M. tuberculosis} to be present in the New World before European contact. However, genetic research of modern strains indicates that \textit{M. tuberculosis} and \textit{M. bovis} developed much earlier from a common ancestor (Baker et al., 2004; Brosch et al., 2002; Gutierrez et al., 2005), and that \textit{M. bovis} was not ancestral to \textit{M. tuberculosis}.

Rothschild et al. (2001) attempted to find out more about the type of tuberculosis present in the Americas before European contact through the analysis of ancient DNA. They examined a metacarpal, dated to approximately 17,000 years ago, that had pathological changes suggestive of tuberculosis from an extinct form of bison. After DNA extraction, the IS\textit{6110} element was amplified and found to be present. Rothschild and colleagues also examined the DR locus. This locus contains polymorphic, short, direct repeats interspersed with nonrepetitive spacers, and in addition to variation in the number of direct repeats, strains also vary in the presence and absence of particular spacers (Kamerbeek et al., 1997). Analysis of the DR locus uses a method known as spacer-oligonucleotide typing (spoligotyping). In spoligotyping, PCR primers target the direct repeats to amplify multiple fragments from this locus. Then multiple oligonucleotides that hybridize to the different spacer sequences are used to discern patterns of variation. The results from the ancient sample did not match any pattern present in the database, but it was closest to patterns from \textit{M. africanum} and \textit{M. tuberculosis}. Spoligotyping, however, is problematic both from an ancient DNA methodological perspective and a phylogenetic perspective. Some studies of ancient remains have found that different patterns may result from independent replications, perhaps because the various segments within the locus may amplify differentially resulting in allele dropout (i.e., fragments that do not amplify because of low quantity) (Mays et al., 2001). In ancient DNA research, absence of evidence is not evidence of absence, which makes interpretation of the hybridization patterns problematic, particularly when results are not consistent. In addition, spoligotyping has been found to be less informative for phylogenetic analyses than DNA sequencing or IS\textit{6110} RFLP pattern analysis because the evolution of the region is complex, involving deletion events, strand slippage during replication leading to duplication, point mutations, and IS\textit{6110}-mediated mutation, and they may result in lineages that are identical by state but not by descent (Filliol et al., 2006; Gori et al., 2005; van Embden et al., 2000; Warren et al., 2002). Thus, if results of Rothschild et al. (2001) are valid, then ancient \textit{M. tuberculosis} DNA was present in an extinct long-horned bison from Wyoming some 17,870 years ago. Additional analyses focusing on
phylogenetically informative SNPs would assist in better understanding the evolution and history of tuberculosis in the Americas.

As noted, ancient DNA can contribute more information about the history and evolution of disease agents in humans (but see Barnes and Thomas, 2006; Bouwman and Brown, 2005 for a discussion of some problems). Questions about the origins of diseases such as syphilis could also be addressed. New genome sequences as well as phylogenetic analyses of modern strains provide the necessary comparative data for understanding the relationships between modern and ancient pathogens. Future analyses of ancient pathogens would also benefit from the use of more samples with historical documentation, such as old samples from medical school collections, to confirm the ability to amplify a specific pathogen before ancient samples from individuals who died of unknown causes are tested.

How Do We Learn About Diet Using Ancient DNA?

In addition to contributing data about the people themselves, ancient DNA can provide information about their environment. Ancient DNA has been used to examine plants found with the Iceman (Rollo et al., 1994); to learn about the history of wheat, corn, and cattle domestication (Bailey et al., 1996; Brown and Brown, 1992; Goloubinoff et al., 1993); and to identify the source of paints and parchment (Reese et al., 1996). DNA was also extracted from three 2000-year-old human coprolites excavated from Hinds Cave, Texas (Poinar et al., 2001). Human mtDNA sequences were amplified as well as segments of the 12S and 16S rRNA genes for the identification of animals and fragments of the chloroplast rbcL gene for the identification of plants in the diets of these individuals. The mtDNA analyses found that one individual had characteristic mtDNA markers and HVI sequences consistent with haplogroup B, whereas two individuals had mtDNA that belonged to haplogroup C with their HVI sequences differing by one mutation. Examination of the rbcL sequences revealed that the diets of these individuals included plant foods from the order Liliales (such as Agave and Yucca), and the families Asteraceae (sunflowers), Ulmaceae (includes hackberry), Fabaceae (the oak family), Solanaceae (such as Physalis), Fabaceae (legumes), Fouquieriaceae (ocotillos), and Rhamnaceae (the buckthorn family). Of these families, Liliales, Fabaceae, and Ulmaceae were confirmed microscopically in the feces. Most fragments of the 12S and 16S rRNA genes that were cloned and sequenced were of human origin; however, some sequences consistent with bighorn sheep, pronghorn antelope, and cottontail rabbit were also identified.

CONCLUSIONS AND FUTURE PROSPECTS

Ancient DNA research has moved past the initial stage of “look! we got DNA!” and has been applied successfully to many questions of anthropological interest. Several problems, however, continue to plague the field. Of these problems, contamination is the greatest one, requiring great care in prevention as well as time and expense to repeat experiments to ensure authenticity. In particular, the repetition of experiments using independent extractions is necessary to verify previous results and, for nuclear loci, to ensure that allele dropout has not occurred. These steps and others outlined in the criteria of authenticity are required to ensure that ancient DNA results can be taken seriously.

Experimental design also warrants greater attention. The significant challenges of working with ancient DNA, its destructive nature, and the expense of such work require careful thought about the hypotheses to be examined and whether ancient DNA data can test those hypotheses realistically. For example, if claims are made about such issues as familial kinship between individuals in a cemetery, the relatedness of two populations, separated in time or space, or the relationship between a disease
agent and a modern pathogen, several questions come to mind: Are a sufficient number of loci being examined in a large enough sample? Are these loci highly variable in modern populations, and thus, could they be expected to be powerful in statistical analyses? Could genetic drift or natural selection have large effects on these loci? Do sequence data from the loci investigated in disease agents provide enough information to distinguish different strains and place them in a phylogenetic context with modern strains? DNA research clearly must be done in combination with other analyses, whether of material culture, paleopathology, osteology, archaeology, or of other fields, and it must be informed by population genetic research of modern groups, in terms of the loci examined, understanding the extent of polymorphism at those loci, and the types of statistical analyses used to investigate particular questions. Under these circumstances, ancient DNA analysis can be a powerful tool, and “molecular archaeology” promises to continue to be a lively and informative field.

REFERENCES


REFERENCES


Lorenz JG, Smith DG. 1996. Distribution of four founding mtDNA haplogroups among Native


PART V

QUANTITATIVE METHODS AND POPULATIONS STUDIES
CHAPTER 16

METRIC ANALYSIS OF SKELETAL REMAINS: METHODS AND APPLICATIONS

MICHAEL PIETRUSEWSKY

BACKGROUND

Measurement in Anthropology

Human skeletal and dental remains continue to provide the most direct evidence for examining the biology of past populations (Larsen, 1997, 2002). Quantitative (i.e., metric) as well as more qualitative (i.e., nonmetric or discrete) aspects of human biological variation remain central to studies in physical anthropology since its inception. Measurement of living (anthropometry) and skeletal remains (osteometry and craniometry), and the methods for analyzing this category of variation represent one of the discipline’s most notable contributions to science.

Much of this early attention to measurement and description in physical anthropology focused on investigations of human population structure and past biological relationships, including the assignment of unknown specimens to known reference groups. As was typical of the discipline as a whole during this formative period, these early studies of population history were associated frequently with the now outdated concept of racial typology and classification. After this formative period, there have been major refinements in biological evolutionary theory and the methods for analyzing quantitative data, refinements that now provide a solid foundation for understanding the history of human groups and biological relatedness among groups, past and present.

Biological Distance Studies

Although the focus of interest in skeletal biology has now greatly expanded to include studies of health and well-being, diet, physical activity, paleodemography, violence, and trauma of past populations, interests in determining biological relatedness or ancestry, today referred to as biological distance (or bi-distances) (Buikstra et al., 1990), continue to attract attention in physical anthropology, bioarchaeology, and forensic anthropology.

Since metric (and nonmetric) aspects of skeletal and dental form and structure (i.e., morphology) have a genetic basis, these categories of variation provide valuable information about past genetic relationships. Groups that share more metric and nonmetric features in common are considered to be more closely related than groups not sharing these same features (Larsen, 2002). Analysis of morphological and, in particular, quantitative traits thus provides an indirect reflection of genetic variation within and between human groups. Several sophisticated models have
been developed specifically that allow the genetic analysis of quantitative traits (e.g., Konigsberg, 1990, 2000; Konigsberg and Blangero, 1993; Relethford, 1994, 2001, 2002; Relethford and Blangero, 1990; Relethford and Harpending, 1994, 1995; Rogers, 1986; Rogers and Harpending, 1983; William-Blangero and Blangero, 1992).

Although biodistance analyses have typically used both metric (i.e., continuous linear measurements and indices) and nonmetric (i.e., discontinuous or quasi-continuous traits) categories of data, the multivariate techniques discussed in this chapter are more amenable to the analysis of continuous or metric data rather than nonmetric data. (For a discussion of nonmetric traits, see Saunders and Rainey, this volume). Furthermore, because of the popularity of the skull in investigating population structure and past relationships, the applications selected to illustrate these methods will be based on multivariate analyses of craniometric data. In addition to studies that focus on methodological issues and the genetic analysis of craniometric data, the literature provides many examples of analytical applications of biodistance studies (e.g., Hemphill, 1998; Howells, 1973, 1989, 1995; Jantz and Owsley, 2001; Pietrusewsky, 1990a, 1994, 1997, 2004, 2005; Powell and Neves, 1999; Relethford and Lees, 1982; Sciulli, 1990; Sokal et al., 1987; Steele and Powell, 1993; Stefan and Chapman, 2003; Stefan, 2004, Varela and Cocilovo 2002; and many others).

**Landmark Measurements and Alternative Data**

Traditionally, the variables used in metric analyses have been simple distances and angles (linear dimensions) defined by craniometric landmarks. These variables are measured directly on the specimen using handheld calipers. Recent alternatives to landmark measurements include coordinate points (Dean, 1995), surface contouring (e.g., moire fringe patterns), holography, methods that capture images through the use of digitizing pads, video-based systems, and three-dimensional imaging techniques such as those used in geometric morphometric studies (Adams et al., 2002; Bookstein, 1991; Rohlf, 1998). This chapter will focus exclusively on the application of multivariate statistical procedures using traditional craniometric landmark data. Richtsmeier et al. (2002) have recently reviewed comparable methods for analyzing three-dimensional landmark coordinate data in morphometric analyses. Although these technical developments in recording three-dimensional objects are becoming more common, traditional measurement methods are still practical for geographically dispersed study samples and provide the foundation for student instruction.

**Justification for Continued Use of Metric Data**

Several convincing arguments have been made supporting the use of measurements, especially craniometry, as sources of taxonomic information (see, e.g., Howells, 1973, 1989, 1995; Van Vark and Howells, 1984). The precision and repeatability of measurements, the conservative nature of continuous variation, the direct link with the past, and the demonstration of a heritability component for this category of biological variation (e.g., Devor, 1987; Sjøvold, 1984) provide the theoretical foundation for metric analysis. Isolation by distance models indicates a strong correlation between genetic and craniometric results that also suggests that craniometric variation (or at least the multivariate patterns of among-group variation they produce) is, on average, selectively neutral (González-José et al., 2004; Relethford, 2002; Roseman, 2004; Roseman and Weaver, 2004). Perhaps the most important reason, from a statistical and mathematical point of view, is the continuous (and correlated) nature of measurements, which makes them well suited for applying multivariate statistical procedures.

Despite the development of mathematical theory, which underlies multivariate statistical methods, in the early decades of this century,
the analysis of metric data using these procedures was slow to gain widespread usage. Much of the initial reluctance can be attributed to the often extensive and tedious computations involved. A variety of historical, political, and scientific reasons, in the mid-twentieth century, diverted attention away from the study of skulls, including the use of measurements, to other areas of interest in the field (Lahr, 1996:xv). General applications of these methods had to wait for the appearance of mainframe computers, commencing in the late 1960s and early 1970s, and then personal computers that have become available only since then. The use of these methods continues to the present.

The focus of this chapter will be the analysis of metric, or quantitative, data using multivariate statistical procedures, methods that have emerged as the most appropriate and advanced for investigating metric data. Although the methods to be discussed are based on complex and sometimes intimidating mathematical statistical theory, this discussion will be largely nontechnical and devoid of mathematical formulas and detailed statistical discourse. Although familiarity with such theory facilitates the use of these methods, this chapter will focus on the application and interpretation of results primarily generated by computer statistical packages, a strategy that will hopefully encourage students and others to use these methods. The use of multivariate statistics in skeletal biology, including biodistance analyses, has been reviewed by Buikstra et al. (1990). More recently, Larsen (1997, 2002) and Wright and Yoder (2003) have provided summaries and syntheses of specific studies in skeletal biology that make extensive use of biodistance measures using metric as well as nonmetric data. Several publications further summarize the use of multivariate statistics and biodistance studies in biological anthropology (e.g., Feldesman, 1997; Larsen, 1997; Van Vark, 1985, 1995; Van Vark and Howells 1984; Van Vark and Schaafsma, 1992).

**METRIC ANALYSIS: UNIVARIATE METHODS**

The use of measurements in physical anthropology and skeletal biology has a long history. Cranial morphology, including craniometry, has been an integral part of this long history. The ubiquitous cranial index and the attendant labels, brachycranic, dolichocranic, and mesocranic, which were used to express cranial vault shape, are synonymous with early physical anthropology and its attempt to reconstruct human racial history. As W.W. Howells (in Lahr, 1996:viii) has stated:

> When Anders Retzius, a century and a half ago, invented the cranial index, he gave us an answer for which there was no question.

These early, often very industrious, efforts in physical anthropology, now regarded as exercises in typological racial classification, resulted in massive compilations of descriptive information on human populations, which including their use as classificatory devices are now of uncertain value. Individuals, or their skeletal remains, were assigned (typed) to groups on the basis of several observable qualitative and/or quantitative features. The approach was largely visual where combinations of observed features determined group membership. More mathematically based approaches to this problem followed a good deal later.

Given this early concern with quantification, it is not surprising that the fields of physical anthropology and mathematical statistics have shared, at least in their formative stages, a mutually beneficial relationship (Howells, 1984). Early statisticians, such as Karl Pearson, used measurements recorded in skulls and people to develop new statistical procedures. These earliest analytical methods were predominantly restricted to descriptive, or univariate, statistics such as measures of central tendency, dispersion, and variance, the statistics of measurements, not individuals, or populations (Howells, 1969:312). Although comparisons between populations
might proceed one measurement at a time, or possibly two at a time as in the case of an index or bivariate plot, the statistics of populations and the treatment of individual specimens in the context of their parent population had to await the introduction of multivariate statistical procedures by Hotelling (1933), Fisher (1936), Mahalanobis (1936), Rao (1948, 1952), and Mahalanobis et al. (1949), among others, beginning in the third decade of the twentieth century.

**MULTIVARIATE STATISTICAL PROCEDURES**

Multivariate statistical procedures comprise a family of related mathematical procedures that allow the simultaneous analysis of multiple variables (i.e., measurements or nonmetric traits) recorded in individuals or objects from one or more groups. More rigorous definitions require that all variables must be random and interrelated and that their different effects cannot be interpreted individually in a meaningful way (Hair et al., 2006:4). Multivariate statistical procedures are exceptionally well suited for investigating interrelationships among the variables, examining group differences, and making other inferences of the variables and groups selected.

By definition, measurements, which represent distance or quantity, have continuous distributions and phenotypes. Statistically, metric traits (e.g., maximum cranial length) can take any value within a range of a scale whose values change smoothly or continuously rather than abruptly. By contrast, nonmetric or discrete traits have discontinuous (qualitative) phenotypes that are expressed by a certain finite, usually small, number of values (or character states) that do not transition smoothly from one value or character state to the next. The distinction between these two kinds of data, continuous and discontinuous, is important in selecting the appropriate multivariate statistical procedure.

Another important concept in multivariate analysis is the variate, here defined as a linear combination of weighted variables represented by a single value. These variates are generated by a specific multivariate statistical procedure such as discriminant analysis or factor analysis. Some multivariate techniques may be considered extensions of univariate techniques; others, like factor analysis, discriminant analysis, and generalized distance, require more complicated statistical procedures that resolve the entire data matrix requiring the use of matrix algebra. The data for a single group usually take the form of a data matrix consisting of \( N \) rows of observations (number of individuals) and \( P \) variables (number of measurements), which are arranged in columns. Reading the rows, an individual becomes a vector of its scores on \( P \) traits, whereas the vector of measurements is read by examining the scores for \( N \) individuals in the columns. The mean vector (or centroid) is the mean of each variable. Univariate statistical techniques are concerned with mean vectors and variance, whereas multivariate statistics must consider the entire matrix of numbers, which is a procedure that requires a consideration of covariance (and correlation) of each possible pair of variables in the matrix.

The primary purpose of multivariate analysis is to investigate relationships among the transformed variables. By reducing the information contained in the original measurements to a smaller number of uncorrelated (orthogonal) variables, or scales, multivariate procedures overcame one of the greatest obstacles posed in earlier attempts to devise a distance statistic [e.g., Pearson’s Coefficient of Racial Likeness (Pearson, 1926) and Penrose’s Size and Shape statistic (Penrose, 1954)], namely that of correcting for the correlation among measurements and the repeated influence of size in individual measurements. Finally, these secondary, transformed, variables allowed individuals and/or populations to be located in multivariate space. The mathematical basis of multivariate statistical methods relies on the matrices of variation and covariation. These methods allow
individual specimens to become vectors of their measurements that, in turn, allow them to be located in multivariate space defined by the newly created transformed variables. This appropriateness of multivariate analyses in handling populations is succinctly stated by Howells (1973:3–4):

Methods of multivariate analysis...allow a skull to be treated as a unit, i.e., as a configuration of the information contained in all its measurements. Next, they allow populations to be treated as configurations of such units, taking account of their variation in shape because they in turn are handled as whole configurations of individual dimensions. Finally, the relations and differences between all the populations being considered are set forth in terms of their several individual multivariate ranges of variation. Thus it is possible to see the range of the whole species in such complete and objective informational terms. That is the importance of multivariate statistics: they fit the model of populations looked on not as centroids or means, but as swarms of the varying individuals who compose them; and the differentiation of these swarms from one another constitutes a statement of the degree and nature of the difference between the populations. Although the information is ultimately limited by the measurements selected to describe the skull, their relationships and their relative taxonomic significance are not otherwise biased by the worker.

These refinements in technique have made multivariate methods theoretically the most soundly based for analyzing metric variation. The procedures adopted depend on the questions being asked, but traditionally, differences (or distances) between human groups and classification (grouping analysis) have been the two principal concerns in physical anthropology. Skeletal biologists have addressed similar issues, usually from an archaeological and/or bioarchaeological perspective, including studies that investigate the processes of evolution (selection, drift, gene flow, and the effects of geography), and whether differences in population structure can be attributed to internal versus external (i.e., introduced) influences. Other studies in skeletal biology have been designed to examine more specific issues such as the identification of population (ethnic) boundaries, postmarital residence patterns, familial and kin relationships, cemetery structure, evidence for social stratification, and the presence of intrusive individuals and/or evidence of admixture with different groups (e.g., Buikstra et al., 1990; Konigsberg and Buikstra, 1995; Relethford and Blangero, 1990; Steadman, 2001; Williams-Blangero et al., 1990). Other potential uses of this family of statistical procedures include the identification of individual crania for repatriation claims and forensic applications (Ousley and Jantz, 2005; Powell and Rose, 1999; Snow et al., 1979).

ASSUMPTIONS OF MULTIVARIATE DATA

In recent years, several researchers have cautioned against the inappropriate and “blind” use of multivariate statistical procedures (e.g., Kowalski, 1972; Rhoads, 1984; Van Vark, 1976; Van Vark and van der Sman, 1982). In an attempt to prevent possible misuses of these procedures, careful scrutiny of the data and the assumptions underlying multivariate data are recommended.

All multivariate statistical procedures have a number of underlying statistical and conceptual assumptions that require evaluation if statistical inferences are to be made (e.g., Campbell, 1978; Corruccini, 1975; Hair et al., 2006; Kowalski, 1972; Krzanowski 2000; Tabachnick and Fidell, 2001; Van Vark and Pasveer, 1994). Among the assumptions, multivariate normality and equality of covariances (or within-group variances) are perhaps the most critical. Adherence to the mathematical conditions of multivariate analyses, including adequate sample sizes, however, helps alleviate most of the assumption violations.

Multivariate procedures require sufficiently large sample sizes. As a general rule, some
(e.g., Corruccini, 1975; McHenry and Corruccini, 1975) have suggested that the sample size should exceed the number of variables used. Others (e.g., Lachenbruch and Goldstein, 1979:70) suggest that there should be at least three times as many individuals for each sample as there are measurements (variables). Aside from these and other more general guidelines, there is lack of unanimous agreement on what constitutes a “sufficiently” large sample. Howells (1973, 1989, 1995), who maintains that samples should be relatively large and comparable in size, ultimately selected 50–55 specimens of each sex in his multivariate analyses. Opinions vary on the effects of unequal sample size especially with regard to the covariance matrix. As a general precaution, however, it is advised that sample sizes be kept uniform. The techniques of statistical power analysis and sample size estimation, now part of most statistical software packages (e.g., SAS), can be used to determine how large a sample is needed for statistical judgments to be accurate and reliable.

Multivariate normality of the data is the assumption that each variable, and all linear combinations of the variables are normally distributed (Tabachnick and Fidell, 2001:72). Ideally, the distribution should be normal, but opinions vary as to the adversity of the effects of non-normal distributions in influencing the results of multivariate analyses. Maintaining equal and sufficiently large samples would seem to satisfy this condition. Reyment (1990) discusses some of the more robust techniques that control for non-normality. There are also distribution-free inferential techniques such as jackknife and bootstrap analyses, based on resampling, that do not require normally distributed data.

Perhaps the most critical assumption of the multivariate conditions is homogeneity of the covariance (variance–covariance) matrix (Campbell, 1978), which is an assumption that may never be completely satisfied (Van Vark and Schaafsma, 1992:236). If the covariance matrices of individual groups (these represent the deviation between one variable and its mean times the deviation between a second variable and its mean, etc.) are unequal, the results of multivariate analyses can be affected adversely. Maintaining large and approximately equal sample sizes should, at least on mathematical and statistical grounds, reduce the likelihood of violating the homogeneity covariance assumption.

However daunting these assumptions may appear, potential users, including those with little or no background in mathematics and/or statistics, should not be unduly dissuaded from applying these methods. Statistical tests for normality and equivalence of covariances, and their remedies, are available in most multivariate statistical packages. Choosing samples with care, maintaining equal and large sample sizes, careful scrutiny of the data, and the possible pre-selection of variables using discriminant analysis as prescribed by some (Van Vark and Schaafsma, 1992), help to avoid violating these underlying assumptions.

CLASSIC MULTIVARIATE STATISTICAL METHODS

The statistical procedures most commonly used by skeletal biologists and physical anthropologists include factor analysis, principal components analysis, discriminant function analysis, and generalized distance. The latter two multivariate procedures are designed to handle two or more groups, whereas principal components analysis, factor analysis, and related techniques are for investigating underlying patterns in a single group. Various clustering methods, and multidimensional scaling techniques, provide a useful means of visualizing the results of multivariate procedures. Some of the newest approaches include the use of digitizing nodes to produce three-dimensional images as well as other uses of coordinate geometry that generate a multitude of measurements in place of the traditional caliper-generated ones.
**Principal Components and Factor Analysis**

Factor analysis and related procedures such as principal components analysis (PCA), by focusing on the interrelationship (covariation) among a large number of variables of a single sample, seek to identify common underlying patterns of variation through an inspection of their shared underlying factors, or axes. Unlike discriminant function analysis, PCA does not employ any criterion for maximizing differences among the groups. Individual specimens can be scored or located on these new axes or factors. Examples that use factor analysis and PCA in physical anthropology include studies by Howells (1957, 1972, 1973) and Brown (1973).

**Discriminant Function (Canonical) Analysis**

The major purpose of discriminant analysis is to maximize differences between two groups. The mathematical basis (Goldstein and Dillon, 1978) for this procedure is to weight and combine, in a linear manner, two or more discriminating variables in such a way that the intercorrelations of the variables are considered and the ratio of between-group variance to the within-group (total) variance is maximized (Tatsuoka, 1970). This concept can then be extended to include more than two groups; in which case, the procedure is commonly referred to as multiple discriminant function, or canonical variate, analysis. The resulting transformed variables, known as discriminant functions, or canonical variates, possess the important property of being orthogonal (uncorrelated, or independent). Individuals and/or groups can then be placed in a multidimensional space, thus providing a means of visualizing these interrelationships. The total number of functions is one less than the number of groups entered or one less than the original number of measurements if that is less. Typically, the first few transformed variables account for the preponderance of the variation among groups. The remaining functions, usually ranked in decreasing importance, are responsible for the residual variation. Although originally designed to assign an unknown specimen to one or more groups, discriminant analysis is now widely used as a measure of group separation (Campbell, 1978).

The original measurements selected in computing the linear classification functions can be chosen in a stepwise manner (stepwise discriminant function analysis) such that, at each step, the measurement that adds the most to the separation of the groups is entered into the discriminant function in advance of the others (Dixon and Brown, 1979:711). The technique also identifies which of the measurements (variables) are most responsible for the observed differentiation. Another utilitarian aspect of the latter procedure is that it allows for the selection of a subset of measurements for use in subsequent distance analyses (Heathcote, 1994; Rightmire, 1970). In the applications presented in this chapter, the interpretation of the discriminant functions and the patterns of group separation is based on an inspection of standardized canonical, or discriminant, coefficients.

At the end of the stepping process, each specimen is classified into one of the original groups based on the several discriminant scores it receives. The probability of group membership can be evaluated mathematically through the calculation of posterior probabilities and/or typicality probabilities. Posterior probabilities assume that the unknown specimens belongs to one of the groups included in the function, whereas typicality probabilities evaluate how likely the unknown belongs to any, or none, of the groups based on the average variability of all the groups in the analysis (Tatsuoka, 1971:228–232; Van Vark and Schaafsma, 1992:244–246). The results are presented in the form of a classification matrix. The classification results provide an additional check on how well the groups are, or are not, differentiated from one another as well as provide a general guide for assessing the homogeneity or heterogeneity of each group. Most major computer packages provide cross-validation procedures (e.g., jackknife methods) to check
discriminant results and the probability that an individual belonging to a specified group has been misclassified. To ensure that the results are externally as well as internally valid, the final stage of discriminant analysis typically involves the validation of the discriminant function results. One such technique, as will be demonstrated in the examples in this chapter, is the jackknife classification method in which the discriminant function is reestimated on multiple subsets of the original samples. In these examples, each case is reclassified into a group (its own or any other one included in the analysis) according to the classification functions computed from all data except the individual case being classified.

Some of the earliest and best known examples that use discriminant function analysis in skeletal biology are concerned with assigning an unknown individual specimen to a given reference group for determining ancestry (Giles and Elliot, 1962) and sex (Giles and Elliot, 1963). In these early examples, discriminant functions (or equations with weights for a number of measurements of an individual which are multiplied to provide a single score) are computed from the measurements of two defined groups (e.g., male and female) such that the differences between the two groups are maximized (F-ratios are maximized), whereas the deviation of individual cases from their respective means remains minimal. Once a discriminant function has been computed between two groups, an unknown specimen can be assigned to one of these groups. There have been many other applications, before and since, including, most notably, the assignment of unknown human or nonhuman primate fossils (e.g., Albrecht, 1992; Campbell, 1984; Howells, 1966; Kammainga and Wright, 1988; Powell and Neves, 1999; Rightmire, 1979; Van Vark, 1995; Wright, 1992a).

Several statistical packages, with a variety of multivariate procedures designed for mainframe and personal computers such as BMDP (Dixon, 1990a, 1990b, 1992; Dixon and Brown, 1979), SAS (1990a, 1990b), SSPS (1993, 1999a, 1999b, 1999c), SYSTAT (1992), and NTSYS (Rohl, 1993) are now available. Other, more specialized programs written for personal computers, such as CRANID2/CRANID3 (Wright, 1992a, 1992b, 2005) and FORDISC3 (Ousley and Jantz, 2005), are also available. Complementing these are software programs that may be downloaded from websites like that of L.W. Konigsberg at http://konig.la.utk.edu/.

Similar procedures, although not accessible to the general public, have been introduced by Van der Sluis et al. (1985) and Howells (1995). The former method, POSCON, used by Van Vark and Schaafsma (1992), is similar to CRANID3 in that both use principal components analysis rather than discriminant function analysis, and the database provided by Howells (1989), to assign an unknown skull. POSCON uses Euclidean distances (explained in the next section) between the unknown skull and the group centroids for classification, whereas CRANID3 examines the Euclidean distances (space) of all the individuals of the reference groups and determines the identity of the unknown skull by reading the list of the 50 nearest individuals. Howells (1995) uses methods similar to both of these (DISPOP and POPKIN), but these are based on discriminant function analysis.

**Mahalanobis's Generalized Distance and Euclidean Distance**

Although several different distance measures are available, the most commonly used one is Euclidean distance, which can be represented geometrically (hence, the name) as the length of the hypotenuse of a right angle that is calculated by the formula: \[ \sqrt{(X_2-X_1)^2+(Y_2-Y_1)^2}, \] where \(X_1Y_1\) and \(X_2Y_2\) represent the respective coordinates of two points plotted on two variables, \(X\) and \(Y\). This concept may then be extended to more than two variables. Mahalanobis’s distance (\(D^2\)) uses the squared (i.e., the sum of squared distances without taking the square root in the above formula) Euclidean distance.

Mahalanobis’s generalized distance, or \(D^2\) (Mahalanobis, 1930, 1936; Mahalanobis et al., 1949; Rao, 1948, 1952), remains the classic, if
only realistic (Reyment et al., 1984), measure of biological distance for analyzing metric data. The immense popularity of the $D^2$ statistic as a measure of distance stems from its theoretical soundness. Generalized distance is computed by maximizing the difference between pairs of groups by maximizing the between-group variance to the pooled within-group variance. This process involves an inversion of the pooled correlation (within-group variance–covariance) matrix. The usual assumptions of equivalence of covariance matrices, normal distribution of variables, and large sample sizes are usually satisfied by pooling sample covariance matrices and avoiding small and extremely uneven sample sizes. Gower (1972), as well as Sneath and Sokal (1973), provides the mathematical basis for computing $D^2$.

Through this procedure, the original variables are transformed into a new set of variables whose correlation with the remaining variables has been removed. $D^2$ represents the summed squared difference between the transformed mean values of any two groups compared. The failure to correct for this correlation was a major flaw with earlier proposed distance statistics such as Pearson’s CRL (Pearson, 1926) and Penrose’s Size and Shape statistic (Penrose, 1947, 1954). Another attraction of generalized distance is that its values do not change if the number and kinds of measurements differ (Van Vark and Shaafsma, 1992:238).

Statistical testing of the significance of the derived distances was first introduced by Hotelling (1933). The method described in Rao (1952:245), one used by Talbot and Mulhall (1962), and later reiterated by Buranarugsa and Leach (1993:17), is the one used in the applications presented in this chapter: The quantity, $(n_i/n_j + n_j) D^2_{ij}$, is distributed as chi-square with $p$ degrees of freedom ($n_i$ = sample size of group $i$; $n_j$ = sample size of group $j$; $D^2_{ij}$ = square of the generalized distance between groups $i$ and $j$; and $p$ = number of variables employed. The nonsignificance of $D^2$ generally indicates that the differences are too small for detection of group differences and/or that the sample sizes are too small.

Because a single quantitative value, which measures dissimilarity between pairs of groups is obtained, another attraction of the generalized distance statistic is that various methods for clustering groups based on these values can be applied. Most computer statistical packages routinely generate Mahalanobis’s generalized distances as output in discriminant function analysis programs.

Other Multivariate Methods (Q-mode and R-mode Analyses)

Other ways of portraying relationships between individual specimens or groups include canonical plots, metric multidimensional scaling (Torgerson, 1952), and principal coordinate analysis (Gower, 1966). The latter two procedures are most often used in association with factor analysis and Q-mode principal components analysis.

Q-mode and R-mode analyses represent two related multivariate techniques, which represent the inverse of principal components analysis and related procedures. R-mode analysis focuses on relationships between variables, whereas Q-mode analysis focuses on relationships between individuals. The principal aim of Q-mode analysis is a graphical visualization of the inter-relationships between individuals of a sample and the identification of clusters (Reyment, 1990:125). Examples that apply Q-mode analysis for analyzing skeletal metric data include those of Howells (1989) and Hanihara (1996).

Alternatives to these methods include the use of Chernoff “faces” (Chernoff, 1973, 1978), Fourier transformations, and boxplots, which some researchers have adopted (e.g., Brown, 1996; Wilson, 1984). Most applications in skeletal biology, however, use distance matrices or some measure of similarity, or dissimilarity.

Finally, group means, or centroids, can be plotted for the first few canonical variates, or functions, in multiple discriminant function analysis or canonical variate analysis to
represent intergroup relationships. Although the procedures are mathematically unrelated, substantial agreement exists between tree construction procedures, which are discussed in the next section, and these other methods of representing multivariate results.

**CLUSTER ANALYSIS AND CLUSTERING ALGORITHMS**

Cluster analyses are often regarded as one group of related multivariate procedures, which differ from the other procedures discussed in this chapter in that they do not actually estimate a variate but rather use one that has already been specified (e.g., $D^2$). Again, these techniques involve the simultaneous utilization of many variables whose purpose is to group individuals (or objects) on the basis of characteristics they possess (Hair et al., 2006). The results of cluster analysis are typically depicted as a tree-like structure, or dendrogram, which has become a popular and convenient way to graphically illustrate and summarize multivariate data (Everitt and Dunn, 1992). Wilmink and Uytterschaut (1984) provide the historical and theoretical background for the use of cluster analysis in physical anthropology. Gower (1967, 1972) provides good comparisons of several methods of cluster analysis.

Although clustering is not phylogeny (Howells, 1984), dendrograms like those based on Mahalanobis’s distances, have implications for the latter and are routinely used in interpreting past relationships. The theoretical and methodological background for most of these techniques derives from classic numerical taxonomy (Sneath and Sokal, 1973). Numerical taxonomy includes quantitative (numeric) methods for grouping usually based on overall morphological (phenetic) similarity where each character, when assigned a specified numerical value, is considered to have equal weight. These methods typically allow the selection of the largest number of traits possible. Numerical, or phenetic, taxonomy represents one of the two major approaches to systematic biological classification (Lewin and Foley, 2004). The other major approach to classification is phylogenetic systematics, or cladistics, in which evolutionary history is inferred from branching patterns of phylogeny that require the careful weighting of only a few traits.

The most commonly used clustering methods comprise a family of procedures known as the hierarchical (either agglomerative or divisive) clustering techniques (Sokal and Rohlf, 1995; Sneath and Sokal, 1973). Agglomerative clustering techniques commence by placing each object (or group) in a single cluster. In subsequent steps, the two most similar clusters are combined into a new (aggregate) cluster, which is a process that continues until all groups are combined into a single cluster. In the opposite clustering procedure, hierarchical divisive methods, the process begins by placing all groups into a single cluster that is then divided into two clusters that contain the most dissimilar groups. Since clusters at any stage are obtained by the combination (or division) of two clusters from the previous stage, these methods lead to a hierarchical structure for the diagram. Different options are available (e.g., single lineage, complete linkage, or group average) depending on how distance, or similarity, of the clusters is measured.

One of the most commonly used agglomerative clustering techniques is UPGMA, or the unweighted pair-group method, arithmetic average algorithm, which measures similarity as the average distance between all cases in one cluster to all cases in another. That is, the average distance between all cases in the resulting cluster is as small as possible and the distance between two clusters is taken as the average between all possible pairs of cases in the cluster. For those interested in phylogenetic tree reconstruction, this method assumes a constant rate of evolution.

An alternative clustering algorithm, the neighbor-joining (NJ) method (Saitou and Nei, 1987; Saitou et al., 1991), which was used initially to construct trees from genetic data, has also been applied to distances.
derived from skeletal data. This algorithm, also known as the method of minimum evolution, is conceptually related to cluster analysis but differs from UPGMA in that it does not assume that all the lineages have diverged equal amounts, thus removing the assumption of constant rates of evolution. Recent computer simulations (Saitou and Imanishi, 1989; Saitou and Nei, 1987) have suggested that the NJ method yields more accurate trees than does the UPGMA method and thus produces a truer phylogenetic tree, at least when gene frequency data are used. The NTSYS-pc computer software program by Rohlf (1993) provides one of the most comprehensive selections of clustering algorithms of interest to skeletal biologists.

RESEARCH DESIGN

Regardless of the intrinsic problems in defining human groups, biologically and statistically, selection and definition of groups is a necessary preliminary step to data analysis. Most of the multivariate statistical procedures reviewed in this chapter require that individual specimens be assigned to groups a priori. The skeletal series encountered most frequently in skeletal biology represent aggregates of individuals found in a specified area at a given time or some subgrouping of individuals below the species level. As is often the case in skeletal biology, archaeological human remains may easily represent individuals who lived and died at different times and whose ancestral–descendant relationships are unknown. In many instances, the “skeletal series” represents a collection of skulls housed in museums documented as having been collected from a specific village, island, or region. In ideal circumstances, exact provenance of the specimens may be available, but often in these situations, completeness, preservation, and the number of specimens available is frequently problematic.

Two basic underlying theoretical assumptions in distance studies using morphological data are that morphometric similarity implies genetic similarity and that the more similar two groups are, the more closely related they must be relative to groups that exhibit greater differences. Despite the demonstration of significant heritability components for both metric and nonmetric traits, morphological traits (including measurements) are also subject to nongenetic influences. Environment and/or allometry rather than gene flow, migration, genetic drift, and other isolating factors may be responsible, or at least contribute, to the results obtained in biodistance analysis. Despite these reservations, general consensus exists (see, e.g., Van Vark and Schaafsma, 1992:241) that biological relatedness as measured by biological distances based on metric data reflects genetic similarity. Konigsberg and Ousley’s (1995) finding of a correlation between anthropometric and quantitative genetic analyses strengthens this assertion.

Of the two basic approaches to the analysis of metric data in studies of population structure outlined by Relethford and Lees (1982), model-bound and model-free, the latter approach remains the one most frequently adopted in studies of physical anthropology and skeletal biology. Multivariate statistical methods, such as measures of biological distances and discriminant functions, have frequently been used in anthropology as exploratory tools for summarizing data on patterns of variability and overall similarities among groups (often accompanied by comparisons with other types of data) and for reconstructing the evolutionary histories of these groups, regardless of cause. Model-bound approaches, although methodologically appealing because they incorporate measures of population similarity directly into models of population structure to estimate one or more parameters (e.g., admixture, genetic drift, gene flow, or effective population size) like the methods used to analyze living populations, have been used less frequently in metric studies of past populations.
The approach used in the example in this chapter is model free. Mahalanobis’s generalized distance and discriminant function analysis are used to investigate patterns of craniometric variation, which, in turn, are used to reconstruct the evolutionary histories and possible origins of the groups, regardless of cause.

**AN APPLICATION**

The example of an application of some methods discussed in this chapter focuses on the population history and biological relatedness among recent and near recent peoples who occupy eastern Asia and the Pacific. Two multivariate statistical procedures, stepwise discriminant function analysis and Mahalanobis’s generalized distance, are applied to 27 cranial measurements recorded in a total of 2805 male crania representing 63 cranial series representing modern and near-modern inhabitants of Oceania, Australia, Southeast Asia, and East Asia. The data derive from earlier studies (e.g., Pietrusewsky, 1990a, 1994, 2004, 2005).

**Preparation of Data**

**Variable Selection.** Although the need in craniometric analysis to improvise new measurements continues (see, e.g., Howells, 1973), many measurements currently used by skeletal biologists can be traced to early international conventions (e.g., Frankfort Agreement of 1882) and attempts to standardize the technique to ensure the comparability of measurements (e.g., Broca, 1875; Martin and Saller, 1957; Vallois, 1965). These traditional measurements require detailed definitions, including landmark definitions and instrumentation (e.g., spreading, sliding, or coordinate calipers), which is information that is available in several recent publications (Bräuer, 1988; Buijkstra and Ubelaker, 1994; Moore-Jansen et al., 1994). Although there are exceptions (e.g., Van Vark and Pasveer, 1994:233), for most multivariate analyses, it is generally preferable to begin with as many variables as possible from which subsets of variables may be selected. Which measurements are ultimately selected for analysis depends on the research questions that are being addressed. The 27 standard cranial measurements included in this application are defined in Table 16.2 later on in this chapter. The number of measurements represents the greatest number of variables comparable with all the series.

**Errors.** The reliability (i.e., the extent to which a metric variable is reproducible over time) of metric data hinges on precision (freedom from measurement error at the individual level) as well as on the dependability of the variable or its freedom from short-term random influences (Marks et al., 1989). The latter, which are generally beyond the control of the individual observer, probably are of minor concern compared with error introduced at the observer level. The recognition of climatic influences on cranial measurements (Utermhole et al., 1983), however, would seem to argue for proper storage of specimens.

Certainly, the precision of measurements can be enhanced greatly with the standardization of the technique and with calibration of the measuring equipment. Other sources of possible error (and imprecision) may be attributed to reading, recording, and/or data entry errors. If more than a single observer is involved, the possible sources of error may be compounded greatly. In the case of a single observer, the approach is to focus on intraobserver, or within-observer, reliability and replication. For intraobserver error, access to the original material offers the best solution. Interobserver (more than one observer) error is more of a concern in reliability studies. A variety of statistical techniques (from basic descriptive statistics to the analysis of variance and correlation coefficients, and so on) are available for assessing the degree of error (Utermhole and Zegura, 1982). Although the potential for interobserver error would seem to argue against the combining of data from different observers, a recent study (Willis, 1999) that uses data recorded by several independent researchers...
spanning a considerable period of time concluded that interobserver error, at least among those measurements found to be common, was not a major concern. This finding should be of some comfort to future researchers contemplating metric analyses but who are otherwise denied access to the original skeletal remains.

**Missing Data.** Most multivariate procedures require complete sets of data, which often means that missing measurements, a frequent occurrence in studies involving archaeological human remains, must be supplied by estimation. A common solution to this problem is to replace the missing observation with its regressed value (Howells, 1973; Van Vark and Pasveer, 1994). Various statistical packages such as BMDP (Dixon and Brown, 1979; Dixon, 1992) contain suitable regression analysis programs for such purposes. These procedures, however, should only be used when a few of the measurements are missing either per individual and/or per variable, which means that only complete or nearly complete specimens can be used ultimately in multivariate statistical analyses. More details on missing observations and general guidelines for dealing with them are described by Howells (1973:33–35), Van Vark (1985, 1995), Van Vark and Shaafsma 1992, Van Vark and Pasveer (1994), and Schafer (1997, 1999).

In this application, missing measurements, which were minimal, were replaced with regressed values obtained through stepwise regression analysis using the computer program, PAM, of the BMDP statistical package (Dixon and Brown, 1979).

**Removal of the Size-Based Component (Z-Scores and C-Scores).** A fairly common concern in biodistance analyses that use multivariate statistical procedures is to determine the relative contribution of size and shape in distance measures. Size is defined as the magnitude of a vector of measurements on an organism, whereas shape is a function of relative proportion normalized by size (Corruccini, 1987:289, 290). Many investigators regard shape, rather than size, as being of greater importance when the taxonomic units in question are above the deme or subspecies level and, thus, warrant the removal of the size-based component (Corruccini, 1973:743). Several researchers (e.g., Brace and Hunt, 1990; Brace et al., 1990; Brace and Tracer, 1992; Howells, 1989) have advocated the use of C-scores as a way to compensate, at least partially, for the unequal influence that size differences may exert on the patterns of variation.

However, although the use of C-scores theoretically compensates for size differences and hence their unequal influence on the patterns of variation, others (e.g., Green, 1990; Pietrusewsky, 1994, 1995) have demonstrated that removal of this size-based component has little or no effect on the interpretation of patterns of craniometric variation. Accordingly, C-score measures are not used in this application.

**Cranial Series**

The approximate locations of the 63 male cranial series used in the application, representing modern and near-modern inhabitants of the Pacific, Australia, Southeast Asia, and East Asia, are shown in Fig. 16.1. Additional information for these comparative series is given in Table 16.1.

**Results of Applying Multivariate Analysis**

**Stepwise Discriminant Function Analysis.** Stepwise discriminant function analysis was applied to 27 cranial measurements recorded in 63 male cranial series using the computer program, BMDP-7M (Dixon, 1992), which was written for the mainframe computer.

A summary of the measurements, ranked according to the F-values [tests of equality of group means using classic one-way analysis of variance] received in the final step of discriminant function analysis, is given in Table 16.2. The order of the measurements in
this table provides an indication of the discriminatory power of the original variables. In this analysis, three breadth measurements (maximum cranial breadth, biorbital breadth, and minimum cranial breadth), basion–nasion length, and nasion–alveolare length received the highest ranking.

Eigenvalues, also referred to as latent roots, represent the amount of variance accounted for by each function or canonical variate, which is here expressed as the percentage of total dispersion; the level of significance tested by Bartlett’s criterion (Rao, 1952:373) for the 27 canonical variates are presented in Table 16.3. The maximum number of canonical variates (and computed eigenvalues) is equal to the number of original variables entered into the analysis. Eigenvalues identify those variates that contribute the most to the discrimination; generally values greater than 1 are considered significant, whereas those less than 1 are insignificant. This procedure thus reduces the original number of variates while retaining as much of the information in the original variables (i.e., measurements). The values under % Dispersion indicate percentage of dispersion accounted for by each corresponding transformed, or canonical, variate. In this analysis, the first three canonical variates account for 63.7% of the total variation. All eigenvalues are significant at the 1% level, which indicates significant heterogeneity for these canonical variates, or axes.

Canonical coefficients, which represent those values by which an individual’s measurements may be multiplied to obtain its score, for 27 measurements, for the first three canonical variates are given in Table 16.4.

Figure 16.1  Map showing the approximate locations of the 63 cranial series used in the application. The numbers correspond to the cranial series in Table 16.1.
<table>
<thead>
<tr>
<th>Series Map No. &amp; Name (abbrev.)</th>
<th>No. of Crania</th>
<th>Locationa and Number of Crania</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Polynesia</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Tonga-Samoa (TOG)</td>
<td>19</td>
<td>BER-3; AMS-2; DRE-1; PAR-1 BPB-4; AIM-2; AUK-5; SIM-1</td>
<td>Fourteen specimens are from Tonga and five are from Samoa. Included in the Tongan series are three skulls from Pongaimotu excavated by McKern in 1920; two from To-At-1 &amp; To-At-2 excavated by J. Davidson in 1965; and five from To-At-36 excavated by D. Spennemann in 1985–1986. The remaining specimens are from museums in Berlin, Paris, and Sydney.</td>
</tr>
<tr>
<td>2. Rapa Nui (RAP)</td>
<td>50</td>
<td>BER-5; DRE-9; PAR-36</td>
<td>Pinart collected most crania in Paris in 1887 at Vaihu and La Perouse Bay, Rapa Nui (Easter Island).</td>
</tr>
<tr>
<td>3. Hawaii (HAW)</td>
<td>60</td>
<td>BPB-20; HON-20; SIM-20</td>
<td>An equal number of specimens have been randomly chosen from three different skeletal series: Mokapu (Oahu), Honokahua (Maui), and Kauai. All specimens are presumed to be prehistoric.</td>
</tr>
<tr>
<td>4. Marquesas (MRQ)</td>
<td>63</td>
<td>PAR-49; LEP-1; BLU-1; BPB-12</td>
<td>Crania are from four islands, Fatu Hiva, Tahuata, Nuku Hiva and Hiva Oa.</td>
</tr>
<tr>
<td>5. New Zealand (NZ)</td>
<td>50</td>
<td>BRE-3; PAR-21; SAM-1; AIM-13; GOT-1; ZUR-5; DRE-6</td>
<td>A representative sample of New Zealand Maori crania from the North and South Islands of New Zealand.</td>
</tr>
<tr>
<td>6. Chatham Is. (CHT)</td>
<td>45</td>
<td>DUN-8; OTM-2 WEL-4; CAN-10 AIM-3; DRE-5 AMS-2; DAS-3 GOT-4; PAR-4</td>
<td>Moriori crania from the Chatham Islands, New Zealand.</td>
</tr>
<tr>
<td>7. Society Is. (SOC)</td>
<td>44</td>
<td>PAR-33; BPB-11</td>
<td>Crania are from the island of Tahiti, Society Islands.</td>
</tr>
<tr>
<td>8. Tuamotu Arch. (TUA)</td>
<td>18</td>
<td>PAR-18</td>
<td>Most specimens are from Makatea in the Tuamotu Archipelago. Dumoutier collected most of these crania from an abandoned cemetery on Magareva Islands, Gambier Islands, French Polynesia, circa 1874.</td>
</tr>
<tr>
<td>9. Gambier Is. (GAM)</td>
<td>7</td>
<td>PAR-7</td>
<td></td>
</tr>
<tr>
<td>Island Melanesia</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Fiji (FIJ)</td>
<td>42</td>
<td>BER-1; SAM-3; QMB-1; DRE-4 FRE-3; CHA-1; BPB-11; PAR-7 AMS-3; DUN-6; SIM-2</td>
<td>Crania are from all major islands, including the Lau Group in the Fiji Islands.</td>
</tr>
<tr>
<td>11. Vanuatu (VAN)</td>
<td>47</td>
<td>BAS-47</td>
<td>F. Speiser collected most specimens in 1912 from Malo, Pentecost and Espiritu Santo Islands, Vanuatu.</td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Series Map No. &amp; Name (abbrev.)</th>
<th>No. of Crania</th>
<th>Location(^a) and Number of Crania</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>12. Loyalty Is. (LOY)</td>
<td>50</td>
<td>BAS-43; PAR-7</td>
<td>Crania are from Mare, Lifou, and Ouvea Island Groups, Loyalty Islands.</td>
</tr>
<tr>
<td>13. New Caledonia (NCL)</td>
<td>50</td>
<td>BAS-34; PAR-16</td>
<td>Crania are from several coastal and inland locations on New Caledonia. Most of these specimens were collected in the late nineteenth century.</td>
</tr>
<tr>
<td>14. Santa Cruz Is. (SCR)</td>
<td>46</td>
<td>SAM-4; AMS-2; BAS-40</td>
<td>Felix Speiser collected the crania located in Basel in 1912 (Speiser, 1928).</td>
</tr>
<tr>
<td>15. Solomon Is. (SOL)</td>
<td>49</td>
<td>DRE-4; BER-1; NMV-1; QMB-3; AMS-16; DAS-10; BAS-14; GOT-1</td>
<td>Crania are from Buka Island (1), New Georgia (5), Guadalcanal (9), San Cristobal Island (7), and other locations in the Solomon Islands.</td>
</tr>
<tr>
<td>16. New Britain (NBR)</td>
<td>50</td>
<td>CHA-20; DRE-30</td>
<td>The specimens from New Britain in Dresden were collected by A. Baessler in 1900, and those in Berlin were collected by R. Parkinson in 1911. These specimens were collected from trading posts near Rabul in the Gazelle Peninsula and most likely represent Tolai crania (see Pietrusewsky, 1990b:236–237; Howells, 1973:24–25).</td>
</tr>
<tr>
<td>17. New Ireland (NIR)</td>
<td>53</td>
<td>AMS-4; BER-2; BLU-6; DRE-18; GOT-15; QMB-1; SAM-6; TUB-1</td>
<td>Pöhl collected most of the crania in Dresden in 1887–1888 from the northern end of the island; the specimens in Göttingen were collected during the Südsee Expedition in 1908.</td>
</tr>
<tr>
<td>18. Admiralty Is. (ADR)</td>
<td>50</td>
<td>DRE-20; GOT-9; CHA-6; TUB-15</td>
<td>Specimens from Hermit, Kaniet, and Manus Islands, Admiralty Islands.</td>
</tr>
<tr>
<td>New Guinea</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19. Sepik R. (SEP)</td>
<td>50</td>
<td>DRE-33; GOT-10; TUB-7</td>
<td>Otto Schlaginhaufen collected the specimens in Dresden in 1909 from various locations along the Sepik River, Papua New Guinea.</td>
</tr>
<tr>
<td>20. Biak Island (BIK)</td>
<td>48</td>
<td>DRE-48</td>
<td>Most (45) of the specimens were collected by A.B. Meyer in 1873 on Biak Island (Mysore), Geelvink Bay, Irian Jaya.</td>
</tr>
<tr>
<td>21. Fly R. (FLY)</td>
<td>42</td>
<td>DRE-35; QMB-7</td>
<td>Most of the skulls in Dresden were collected by Webster in 1902 along the Fly R. of Papua New Guinea. Many of the crania are decorated and have engraved frontal bones (see Pietrusewsky, 1990b:235–236 for further details).</td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Series Map No. &amp; Name (abbrev.)</th>
<th>No. of Crania</th>
<th>Location(^a) and Number of Crania</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>22. Purari Delta (PUR)</td>
<td>50</td>
<td>DRE-50</td>
<td>Decorated (engraved) skulls obtained by Messrs. Gerrard and Webster between 1900 and 1902 are from along the Purari River and Purari Delta regions, Papua New Guinea.</td>
</tr>
<tr>
<td>23. D’Entrecasteaux Is. (DTX)</td>
<td>26</td>
<td>FRE-21; DRE-4; QMB-1</td>
<td>Crania are from Fergusson (16) and Normanby (10) Islands of the D’Entrecasteaux Islands.</td>
</tr>
<tr>
<td>24. Dawson Strait Is. (DAW)</td>
<td>48</td>
<td>ROM-48</td>
<td>Crania are from the islands of the Dawson Straits (between Normanby and Fergusson Islands of the D’Entrecasteaux Islands), which were collected by L. Loria on a voyage to Papua New Guinea between 1889 and 1890.</td>
</tr>
<tr>
<td>Australia/Tasmania</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25. Murray R. (MRB)</td>
<td>50</td>
<td>AIA-39; DAM-11</td>
<td>Australian Aboriginal crania were collected by G.M. Black along the Murray River (Chowilla to Coobool) in New South Wales between 1929 and 1950.</td>
</tr>
<tr>
<td>26. New South Wales (NSW)</td>
<td>62</td>
<td>AMS-21; DAS-41</td>
<td>Australian Aboriginal crania from the coastal locations in New South Wales.</td>
</tr>
<tr>
<td>27. Queensland (QLD)</td>
<td>54</td>
<td>AMS-21; DAS-3; QMB-30</td>
<td>Australian Aboriginal crania from the southeastern and middle-eastern regions of Queensland.</td>
</tr>
<tr>
<td>28. Northern Territory (NT)</td>
<td>50</td>
<td>AIA-4; AMS-3; MMS-1; NMV-38; QMB-1; SAM-3</td>
<td>Australian Aboriginal crania from Port Darwin (39) and Arnhemland (36) in the Northern Territory, Australia.</td>
</tr>
<tr>
<td>30. Western Australia (WA)</td>
<td>47</td>
<td>WAM-47</td>
<td>Australian Aboriginal crania from the central (20), eastern (4), northern (14), and southern (9) regions of Western Australia.</td>
</tr>
<tr>
<td>31. Tasmania (TAS)</td>
<td>26</td>
<td>THM-22; CHA-1; SAM-2; NMV-1</td>
<td>The crania represent Tasmanian Aborigines.</td>
</tr>
</tbody>
</table>

(Continued)
### TABLE 16.1 Continued

<table>
<thead>
<tr>
<th>Series Map No. &amp; Name (abbrev.)</th>
<th>No. of Crania</th>
<th>Location and Number of Crania</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Micronesia</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>32. Guam (GUA)</td>
<td>46</td>
<td>BPB-42; PAR-4</td>
<td>Pre-Spanish Chamorro crania associated with late structures collected in the 1920s by Hans Hornbostel along Tumon Beach, Tumon Bay, Guam.</td>
</tr>
<tr>
<td>33. Caroline Is. (CAR)</td>
<td>24</td>
<td>TRO-7; DRE-9; PAR-4; GOT-3; AMS-1</td>
<td>The crania are from Kosrae Island (1), Pohnpei (16), and Chuuk (7) Islands of the central and eastern Caroline Islands, Federated States of Micronesia.</td>
</tr>
<tr>
<td>34. Marshall/ Kiribati Is. (MSK)</td>
<td>13</td>
<td>PAR-6; GOT-3; FRE-3; BER-1</td>
<td>Crania are from the Marshall (7) and Kiribati (6) Islands of eastern Micronesia.</td>
</tr>
<tr>
<td><strong>Island Southeast Asia</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>35. Sumatra (SUM)</td>
<td>39</td>
<td>BER-1; BRE-1; DRE-5; LEP-4; PAR-3; ZUR-25</td>
<td>The specimens in Zurich are designated “Battak.” Specific locations within the island of Sumatra are not known.</td>
</tr>
<tr>
<td>36. Java (JAV)</td>
<td>50</td>
<td>BER-1; BLU-8; CHA-9; DRE-1; LEP-24; PAR-7</td>
<td>Crania were collected from several different localities in Java.</td>
</tr>
<tr>
<td>37. Borneo (BOR)</td>
<td>34</td>
<td>BER-2; BRE-2; DRE-6; FRE-4; LEP-8; PAR-12</td>
<td>A great many of the specimens are indicated as representing Dayak tribes; some have elaborate decorations.</td>
</tr>
<tr>
<td>38. Sulawesi (SLW)</td>
<td>41</td>
<td>BAS-7; BER-10; DRE-4; FRE-7; LEP-5; PAR-8</td>
<td>An exact location is known for many of these specimens.</td>
</tr>
<tr>
<td>39. Lesser Sunda Is. (LSN)</td>
<td>61</td>
<td>BAS-5; BER-6; BLU-2; CHA-1; DRE-17; LEP-1; PAR-6; ZUR-7</td>
<td>Crania from Bali (13), Flores (9), Sumba (1), Lombok (2), Alor (2), Timor (11), Wetar (2), Leti (4), Barbar (1), Tanimbar (13), Kai (2), and Aru (1) Islands of the Lesser Sunda Islands.</td>
</tr>
<tr>
<td>40. S. Molucca Is. (SML)</td>
<td>65</td>
<td>FRE-47; DRE-17</td>
<td>Crania are from the Seram (48) and Buru (17) Islands in the Southern Molucca Islands.</td>
</tr>
<tr>
<td>41. Sulu (SUL)</td>
<td>38</td>
<td>LEP-1; PAR-37</td>
<td>The specimens in Paris were collected by Montano-Rey circa 1900.</td>
</tr>
<tr>
<td>42. Philippines (PHL)</td>
<td>28</td>
<td>BER-9; DRE-19</td>
<td>Most specimens are from Luzon Island.</td>
</tr>
<tr>
<td><strong>Mainland Southeast Asia</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>43. Viet Nam (VTN)</td>
<td>49</td>
<td>HCM-49</td>
<td>Specimens are from Hanoi (Van Dien Cemetery) and Ho Chi Minh City.</td>
</tr>
<tr>
<td>44. Bachuc Village (BAC)</td>
<td>51</td>
<td>BAC-51</td>
<td>Victims of the 1978 Khmer Rouge massacre in Bachuc Village in western Angiang Province, Viet Nam.</td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Series Map No. &amp; Name (abbrev.)</th>
<th>No. of Crania</th>
<th>Location(^a) and Number of Crania</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>45. Cambodia &amp; Laos (CML)</td>
<td>40</td>
<td>PAR-40</td>
<td>A combined sample of crania from various locations in Cambodia and Laos collected between 1877 and 1920.</td>
</tr>
<tr>
<td>46. Thailand (THI)</td>
<td>50</td>
<td>SIR-50</td>
<td>Most specimens represent dissecting room cases from Bangkok.</td>
</tr>
<tr>
<td>47. Burma (BUR)</td>
<td>16</td>
<td>ZUR-16</td>
<td>The crania in Zurich are from a series (Cat. Nos. 93-125) of skulls collected in Mandalay, Myanmar (Burma), described in a catalogue dated circa 1900.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>East Asia: Japan/Ryukyu Is.</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>48. Kanto (KAN)</td>
<td>50</td>
<td>CHB-50</td>
<td>A dissecting room population of modern Japanese from the Kanto District of eastern Honshu. Most individuals were born during the Meiji period (1868–1911) and died well before 1940.</td>
</tr>
<tr>
<td>49. Tohoku (TOH)</td>
<td>53</td>
<td>SEN-53</td>
<td>Dissecting room specimens of modern Japanese from the Tohoku District in northern Honshu Island.</td>
</tr>
<tr>
<td>50. Kyushu (KYU)</td>
<td>51</td>
<td>KYU-51</td>
<td>Modern Japanese that derive mostly from Fukuoka Prefecture in Kyushu Island. Other specimens are from Yamaguchi, Saga, Nagasaki, and adjoining prefectures.</td>
</tr>
<tr>
<td>51. Ainu (AIN)</td>
<td>50</td>
<td>SAP-18; TKM-5; TKO-27</td>
<td>Yoshikiyo Koganei collected these skeletons in 1888–1889 from abandoned Ainu cemeteries in Hokkaido (Koganei, 1893–1894).</td>
</tr>
<tr>
<td>52. Ryukyu Is. (RYU)</td>
<td>64</td>
<td>KYO-18; KAN-21; RYU-8; KYU-5; TKO-8</td>
<td>Eighteen crania are from Tokunoshima Island, one of the Amami Islands located north of the Okinawa Group in the central Ryukyu Islands; 21 specimens are from two different locations on Kume Island, an island located west of Okinawa Island; Yattchi (17) and Hiyajo (4); and 24 specimens are from five separate islands in the Sakishima Group of the southern Ryukyu Islands: Hateruma (2); Miyako (4); Iriomote (2); Ishigaki (1), and Yonaguni Islands (12).</td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Series Map No. &amp; Name (abbrev.)</th>
<th>No. of Crania</th>
<th>Location and Number of Crania</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>China/E. &amp; N.E. Asia</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>53. Shanghai (SHA)</td>
<td>50</td>
<td>SHA-50</td>
<td>The specimens are mostly from post-Qing cemeteries in Shanghai.</td>
</tr>
<tr>
<td>54. Hangzhou (HAN)</td>
<td>50</td>
<td>SHA-50</td>
<td>Ancient skeletal remains exhumed in the modern city of Hangzhou, Zhejiang Province, in eastern China.</td>
</tr>
<tr>
<td>55. Nanjing (NAJ)</td>
<td>49</td>
<td>SHA-49</td>
<td>Ancient remains exhumed from the modern city of Nanjing, Jiangsu Province, in eastern China.</td>
</tr>
<tr>
<td>56. Chengdu (CHD)</td>
<td>53</td>
<td>SHA-10; CHE-43</td>
<td>Most of these specimens date to the Ch’ên Dynasty (1796–1908 A.D.) and are from Chengdu, Sichuan Province, in western China. Ten crania are from Leshan, Lizhong County, Sichuan Province.</td>
</tr>
<tr>
<td>57. Hong Kong (HK)</td>
<td>50</td>
<td>HKU-50</td>
<td>Specimens represent individuals who died in Hong Kong between 1978 and 1979.</td>
</tr>
<tr>
<td>58. Taiwan (TAI)</td>
<td>47</td>
<td>TPE-47</td>
<td>Modern Chinese living in Taiwan who trace their immediate origins to the Fujian and Guangdong provinces on the mainland of China.</td>
</tr>
<tr>
<td>59. Hainan Island (HAI)</td>
<td>47</td>
<td>TPE-47</td>
<td>Chinese immigrants originally from the Canton region of China who began arriving around 200 B.C. (Howells, 1989:108). This material was excavated by T. Kanaseki in Haikou City on Hainan Island.</td>
</tr>
<tr>
<td>60. Manchuria (MAN)</td>
<td>50</td>
<td>TKO-50</td>
<td>Many specimens are from northeastern China or the region formerly referred to as “Manchuria,” which today includes the Heilongjiang and Jilin provinces and adjacent northern Korea. A great many of these specimens are identified as soldiers or cavalrymen who died in battle in the late nineteenth century.</td>
</tr>
<tr>
<td>61. Korea (KOR)</td>
<td>32</td>
<td>KYO-7; SEN-3; TKM-2; TKO-20</td>
<td>Specific locations in Korea are known for most of these specimens.</td>
</tr>
<tr>
<td>62. Mongolia (MOG)</td>
<td>50</td>
<td>SIM-50</td>
<td>The skulls are identified as coming from Ulaanbaatar (Urga), Mongolia, and were purchased by A. Hrdlička in 1912.</td>
</tr>
</tbody>
</table>
Biorbital breadth, nasion–alveolar height, nasal height, basion–prosthion length, and nasio–occipital length (those variables with the highest coefficients regardless of sign) are the most important variables in producing group separation in the first canonical variate. This first variate may, therefore, be defined as a biorbital breadth, facial and nasal height, and cranial base and vault length discriminator.

Minimum cranial breadth, orbital height, alveolar breadth, nasal height, and nasal breadth are most responsible for group separation produced in the second canonical variate. Maximum cranial length, orbital breadth, nasal height, and bijugal breadth are primarily responsible for the discrimination produced in the third canonical variate. Often, such procedures are used as a means of selecting specific variables for additional multivariate analyses.

Table 16.5 contains the overall group classification results based on posterior probabilities, regular and jackknifed, for this analysis. A mathematical explanation of posterior probability is given in Van Vark and Schaafsma.
Jackknifed classification is a common cross-validation procedure used in multiple discriminant analysis, where cases are classified without using misclassified individuals in computing the classification function.

Mongolia, Swanport (Australia), Chatham Islands (New Zealand), Rapa Nui (Easter Island), Guam, Dawson Strait, Ainu, Tasmania, and Western Australia are among the series having the best jackknifed classification results (i.e., more than 57% of the cases are assigned correctly to each of these original groups). Groups with the poorest jackknifed classification results (less than 20% of the cases correctly classified to their original group) include the Solomon Islands, Lesser Sunda Island, New Ireland, Hangzhou, Sulawesi, Hainan Island, Nanjing, Borneo, Sumatra, Shanghai, and Fiji series; eight of the series with poor classification results represent Chinese and Southeast Asian island cranial series.

Table 16.2 provides a summary ranking of cranial measurements according to F-values received in the final step of discriminant function analysis (63 male groups, 27 measurements).

<table>
<thead>
<tr>
<th>Step No.</th>
<th>Measurementa,b</th>
<th>F-Value</th>
<th>d.f.B/d.f.Wc</th>
<th>p d</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Maximum cranial breadth (M-8)</td>
<td>47.6</td>
<td>62/2742</td>
<td>*</td>
</tr>
<tr>
<td>2</td>
<td>Biorbital breadth (H-EKB)</td>
<td>20.7</td>
<td>62/2741</td>
<td>*</td>
</tr>
<tr>
<td>3</td>
<td>Minimum cranial breadth (M-14)</td>
<td>26.2</td>
<td>62/2740</td>
<td>*</td>
</tr>
<tr>
<td>4</td>
<td>Basion–nasion length (M-5)</td>
<td>17.6</td>
<td>62/2739</td>
<td>*</td>
</tr>
<tr>
<td>5</td>
<td>Nasion–alveolare (M-48)</td>
<td>14.4</td>
<td>62/2738</td>
<td>*</td>
</tr>
<tr>
<td>6</td>
<td>Maximum cranial length (M-1)</td>
<td>12.1</td>
<td>62/2737</td>
<td>*</td>
</tr>
<tr>
<td>7</td>
<td>Basion–bregma height (M-17)</td>
<td>12.5</td>
<td>62/2736</td>
<td>*</td>
</tr>
<tr>
<td>8</td>
<td>Biauricular breadth (M-11b)</td>
<td>12.0</td>
<td>62/2735</td>
<td>*</td>
</tr>
<tr>
<td>9</td>
<td>Basion–prosthion (M-40)</td>
<td>10.9</td>
<td>62/2734</td>
<td>*</td>
</tr>
<tr>
<td>10</td>
<td>Nasal height (H-NLH)</td>
<td>10.4</td>
<td>62/2733</td>
<td>*</td>
</tr>
<tr>
<td>11</td>
<td>Nasion–occipital length (M-1d)</td>
<td>8.3</td>
<td>62/2732</td>
<td>*</td>
</tr>
<tr>
<td>12</td>
<td>Nasal breadth (M-54)</td>
<td>8.1</td>
<td>62/2730</td>
<td>*</td>
</tr>
<tr>
<td>13</td>
<td>Bipugal breadth [M-45(1)]</td>
<td>8.0</td>
<td>62/2731</td>
<td>*</td>
</tr>
<tr>
<td>14</td>
<td>Bifrontal breadth (M-43)</td>
<td>7.3</td>
<td>62/2729</td>
<td>*</td>
</tr>
<tr>
<td>15</td>
<td>Alveolar breadth (M-61)</td>
<td>7.1</td>
<td>62/2728</td>
<td>*</td>
</tr>
<tr>
<td>16</td>
<td>Mastoid height (H-MDL)</td>
<td>6.9</td>
<td>62/2727</td>
<td>*</td>
</tr>
<tr>
<td>17</td>
<td>Cheek height (H-WMH)</td>
<td>6.6</td>
<td>62/2726</td>
<td>*</td>
</tr>
<tr>
<td>18</td>
<td>Nasion–bregma chord (M-29)</td>
<td>5.7</td>
<td>62/2725</td>
<td>*</td>
</tr>
<tr>
<td>19</td>
<td>Orbital height, left (M-52)</td>
<td>5.3</td>
<td>62/2724</td>
<td>*</td>
</tr>
<tr>
<td>20</td>
<td>Bimaxillary breadth (M-46)</td>
<td>4.9</td>
<td>62/2723</td>
<td>*</td>
</tr>
<tr>
<td>21</td>
<td>Orbital breadth, left (M-51a)</td>
<td>4.7</td>
<td>62/2720</td>
<td>*</td>
</tr>
<tr>
<td>22</td>
<td>Bistephanic breadth (H-STB)</td>
<td>4.7</td>
<td>62/2722</td>
<td>*</td>
</tr>
<tr>
<td>23</td>
<td>Maximum frontal breadth (M-10)</td>
<td>5.5</td>
<td>62/2721</td>
<td>*</td>
</tr>
<tr>
<td>24</td>
<td>Minimum frontal breadth (M-9)</td>
<td>4.1</td>
<td>62/2719</td>
<td>*</td>
</tr>
<tr>
<td>25</td>
<td>Mastoid width (H-MDB)</td>
<td>3.8</td>
<td>62/2718</td>
<td>*</td>
</tr>
<tr>
<td>26</td>
<td>Bregma–lambda chord (M-30)</td>
<td>3.5</td>
<td>62/2717</td>
<td>*</td>
</tr>
<tr>
<td>27</td>
<td>Biasterionic breadth (M-12)</td>
<td>3.3</td>
<td>62/2716</td>
<td>*</td>
</tr>
</tbody>
</table>

a M = Martin and Saller (1957) [the numbers after M refer to the original numbered measurements in this standard].
b H = Howells (1973) [the letters after H refer to the designations used by Howells].
c d.f.B/d.f.W = degrees of freedom between/degrees of freedom within.
d p < .01; n.s. = not significant.

the 49 crania originally assigned to the Solomon Islands are reassigned to the Solomon Islands series; the remaining specimens originally assigned to this cranial series are “misclassified” to cranial series from the islands of Melanesia and the Caroline Islands. The New Ireland, Solomon Islands, Sumatra, and Southern Moluccas series each have reclassifications to 25 or more groups in this analysis!

The misclassifications of the Polynesian series are generally among other Polynesian groups, although a few are misclassified to an Southeast Asian (e.g., Southern Moluccas, Lesser Sunda Islands, or Java) island series, Fiji, or one of the cranial series from the islands of Melanesia. Five of the New Zealand Maori crania are assigned to Southern Moluccas, and seven others are reassigned to Loyalty, Solomon, or New Ireland.

Most of the misclassifications between the New Guinea and the Melanesian series are to a series from the same geographical region. As expected, most of the misclassifications among the Australian/Tasmanian series are to these and to some of the Melanesian series. Misclassifications for the cases originally

<table>
<thead>
<tr>
<th>Canonical Variate</th>
<th>Eigenvalue</th>
<th>% Dispersion</th>
<th>Cumulative % Dispersion</th>
<th>d.f.</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.43542</td>
<td>41.7</td>
<td>41.7</td>
<td>88</td>
<td>*</td>
</tr>
<tr>
<td>2</td>
<td>1.13325</td>
<td>13.7</td>
<td>55.4</td>
<td>86</td>
<td>*</td>
</tr>
<tr>
<td>3</td>
<td>0.68543</td>
<td>8.3</td>
<td>63.7</td>
<td>84</td>
<td>*</td>
</tr>
<tr>
<td>4</td>
<td>0.56173</td>
<td>6.8</td>
<td>70.5</td>
<td>82</td>
<td>*</td>
</tr>
<tr>
<td>5</td>
<td>0.37941</td>
<td>4.6</td>
<td>75.1</td>
<td>80</td>
<td>*</td>
</tr>
<tr>
<td>6</td>
<td>0.28659</td>
<td>3.5</td>
<td>78.6</td>
<td>78</td>
<td>*</td>
</tr>
<tr>
<td>7</td>
<td>0.24680</td>
<td>3.0</td>
<td>81.6</td>
<td>76</td>
<td>*</td>
</tr>
<tr>
<td>8</td>
<td>0.20748</td>
<td>2.5</td>
<td>84.1</td>
<td>74</td>
<td>*</td>
</tr>
<tr>
<td>9</td>
<td>0.17662</td>
<td>2.2</td>
<td>86.3</td>
<td>72</td>
<td>*</td>
</tr>
<tr>
<td>10</td>
<td>0.14432</td>
<td>1.7</td>
<td>88.0</td>
<td>70</td>
<td>*</td>
</tr>
<tr>
<td>11</td>
<td>0.13498</td>
<td>1.6</td>
<td>89.6</td>
<td>68</td>
<td>*</td>
</tr>
<tr>
<td>12</td>
<td>0.12389</td>
<td>1.5</td>
<td>91.1</td>
<td>66</td>
<td>*</td>
</tr>
<tr>
<td>13</td>
<td>0.11808</td>
<td>1.5</td>
<td>92.6</td>
<td>64</td>
<td>*</td>
</tr>
<tr>
<td>14</td>
<td>0.09761</td>
<td>1.2</td>
<td>93.8</td>
<td>62</td>
<td>*</td>
</tr>
<tr>
<td>15</td>
<td>0.08235</td>
<td>1.0</td>
<td>94.8</td>
<td>60</td>
<td>*</td>
</tr>
<tr>
<td>16</td>
<td>0.07295</td>
<td>0.8</td>
<td>95.6</td>
<td>58</td>
<td>*</td>
</tr>
<tr>
<td>17</td>
<td>0.06267</td>
<td>0.8</td>
<td>96.4</td>
<td>56</td>
<td>*</td>
</tr>
<tr>
<td>18</td>
<td>0.06016</td>
<td>0.7</td>
<td>97.1</td>
<td>54</td>
<td>*</td>
</tr>
<tr>
<td>19</td>
<td>0.04573</td>
<td>0.6</td>
<td>97.7</td>
<td>52</td>
<td>*</td>
</tr>
<tr>
<td>20</td>
<td>0.03864</td>
<td>0.5</td>
<td>98.2</td>
<td>50</td>
<td>*</td>
</tr>
<tr>
<td>21</td>
<td>0.03053</td>
<td>0.3</td>
<td>98.5</td>
<td>48</td>
<td>*</td>
</tr>
<tr>
<td>22</td>
<td>0.02999</td>
<td>0.4</td>
<td>98.9</td>
<td>46</td>
<td>*</td>
</tr>
<tr>
<td>23</td>
<td>0.02392</td>
<td>0.3</td>
<td>99.2</td>
<td>44</td>
<td>*</td>
</tr>
<tr>
<td>24</td>
<td>0.02213</td>
<td>0.3</td>
<td>99.5</td>
<td>42</td>
<td>*</td>
</tr>
<tr>
<td>25</td>
<td>0.01979</td>
<td>0.2</td>
<td>99.7</td>
<td>40</td>
<td>*</td>
</tr>
<tr>
<td>26</td>
<td>0.01491</td>
<td>0.2</td>
<td>99.9</td>
<td>38</td>
<td>*</td>
</tr>
<tr>
<td>27</td>
<td>0.01041</td>
<td>0.1</td>
<td>100.0</td>
<td>36</td>
<td>*</td>
</tr>
</tbody>
</table>

*a d.f. = degrees of freedom = \((p + q - 2), (p + q - 4)\).

*b p ≤ 0.01 when eigenvalues are tested for significance according to criterion \(N - \frac{1}{2} (p + q) \log_e (\lambda + 1)\), where \(N = \) total number of crania, \(p = \) number of variables, \(q = \) number of groups, and \(\lambda = \) eigenvalue, all of which are distributed approximately as chi-square (Rao, 1952:373).
assigned to the two Micronesian series are generally to Micronesian, Polynesian, and Melanesian series.

Four of the Southeast Asian island series, Lesser Sunda Islands, Sulawesi, Borneo, and Sumatra, also have high misclassifications, with most of these being to another cranial series from Southeast Asia. The crania from the Lesser Sunda Islands are misclassified to more groups (32 in all!) than any other group, including two Polynesian series, New Zealand and Tonga-Samoa. At least 14 of the Southern Molucca Islands specimens are reclassified to the Polynesian series (e.g., Marquesas, New Zealand, and Hawaii) and at least 17 more are reassigned to cranial series from Melanesia and New Guinea.

Most importantly, these classification results serve to highlight those regions (e.g., Solomon Islands, Lesser Sunda Islands, Southern Moluccas, Vietnam, Taiwan, and Hainan Island) that exhibit the greatest heterogeneity and possibly where contact with other regions was the most intense and/or sustained.

A plot showing the 63 group means on the first two canonical variates results in three separate clusters shown in Fig. 16.2. Cranial series from Australia, New Guinea, and geographical Melanesia are contained in one of these general clusters, with little overlap between the Australian and the Melanesian series. The Polynesian cranial series and those from Guam and the Marshall-Kirbati Islands form a second isolated constellation in this diagram.

### TABLE 16.4 Canonical Coefficients of 27 Cranial Measurements for the First Three Canonical Variates

<table>
<thead>
<tr>
<th>Cranial Measurement</th>
<th>Canonical Variate 1</th>
<th>Canonical Variate 2</th>
<th>Canonical Variate 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum cranial length (M-1)</td>
<td>0.08825</td>
<td>−0.08027</td>
<td>−0.15823</td>
</tr>
<tr>
<td>Orbital breadth, left (M-51a)</td>
<td>0.09176</td>
<td>0.03942</td>
<td>−0.12983</td>
</tr>
<tr>
<td>Nasal height (H-NLH)</td>
<td>0.10629</td>
<td>0.10212</td>
<td>0.09897</td>
</tr>
<tr>
<td>Bijugal breadth [M-45(1)]</td>
<td>−0.02228</td>
<td>0.09763</td>
<td>−0.09039</td>
</tr>
<tr>
<td>Basion–bregma height (M-17)</td>
<td>−0.02201</td>
<td>0.04550</td>
<td>0.08050</td>
</tr>
<tr>
<td>Biauricular breadth (M-11b)</td>
<td>−0.01973</td>
<td>0.06519</td>
<td>−0.07778</td>
</tr>
<tr>
<td>Nasion–alveolare (M-48)</td>
<td>−0.12798</td>
<td>−0.08643</td>
<td>−0.06925</td>
</tr>
<tr>
<td>Nasion–occipital length (M-1d)</td>
<td>−0.09897</td>
<td>0.08086</td>
<td>0.06442</td>
</tr>
<tr>
<td>Mastoid width (H-MDB)</td>
<td>−0.03215</td>
<td>0.02777</td>
<td>0.06404</td>
</tr>
<tr>
<td>Bistephanic breadth (H-STB)</td>
<td>−0.03680</td>
<td>0.05350</td>
<td>0.06035</td>
</tr>
<tr>
<td>Orbital height, left (M-52)</td>
<td>−0.03903</td>
<td>0.11651</td>
<td>−0.05403</td>
</tr>
<tr>
<td>Bifrontal breadth (M-43)</td>
<td>−0.03779</td>
<td>−0.08027</td>
<td>0.05216</td>
</tr>
<tr>
<td>Bimaxillary breadth (M-46)</td>
<td>−0.06497</td>
<td>−0.00025</td>
<td>0.04754</td>
</tr>
<tr>
<td>Minimum frontal breadth (M-9)</td>
<td>0.05735</td>
<td>−0.01817</td>
<td>−0.04493</td>
</tr>
<tr>
<td>Biorbital breadth (H-EKB)</td>
<td>0.14151</td>
<td>−0.02456</td>
<td>0.04411</td>
</tr>
<tr>
<td>Maximum frontal breadth (M-10)</td>
<td>−0.01108</td>
<td>−0.06485</td>
<td>−0.03890</td>
</tr>
<tr>
<td>Basion–nasion length (M-5)</td>
<td>−0.04514</td>
<td>0.03642</td>
<td>−0.03873</td>
</tr>
<tr>
<td>Bregma–lambda chord (M-30)</td>
<td>0.01679</td>
<td>−0.01977</td>
<td>0.02805</td>
</tr>
<tr>
<td>Basion–prosthion (M-40)</td>
<td>0.10021</td>
<td>0.05748</td>
<td>0.02460</td>
</tr>
<tr>
<td>Nasal breadth (M-54)</td>
<td>−0.01999</td>
<td>−0.09942</td>
<td>−0.02160</td>
</tr>
<tr>
<td>Maximum cranial breadth (M-8)</td>
<td>−0.05423</td>
<td>−0.00605</td>
<td>−0.01649</td>
</tr>
<tr>
<td>Biasterionic breadth (M-12)</td>
<td>0.01295</td>
<td>−0.03754</td>
<td>−0.01514</td>
</tr>
<tr>
<td>Mastoid height (H-MDL)</td>
<td>0.04276</td>
<td>0.07088</td>
<td>0.01043</td>
</tr>
<tr>
<td>Minimum cranial breadth (M-14)</td>
<td>−0.09832</td>
<td>−0.12304</td>
<td>0.00823</td>
</tr>
<tr>
<td>Nasion–bregma chord (M-29)</td>
<td>0.05184</td>
<td>0.03643</td>
<td>−0.00731</td>
</tr>
<tr>
<td>Alveolar breadth (H-M61)</td>
<td>0.06024</td>
<td>−0.10264</td>
<td>0.00553</td>
</tr>
<tr>
<td>Cheek height (H-WMH)</td>
<td>−0.07379</td>
<td>0.09241</td>
<td>0.00249</td>
</tr>
</tbody>
</table>
The remaining groups, representing cranial series from eastern and northern Asia and mainland and island Southeast Asia, form a relatively compact third grouping. The cranial series representing the Southern Moluccas and Caroline Islands are peripheral members of the greater Melanesian–Australian grouping.

The plot of the 63 group means on the first three canonical variates, which is separated into two diagrams for easier viewing, is presented in Figs. 16.3 and 16.4. The Atayal series is included in both representations to provide continuity in viewing these plots. The Polynesian and two Micronesian cranial series are well separated from the Australian and Melanesian samples in Fig. 16.3. The Admiralty Island, Marshall-Kiribati, and Caroline cranial series occupy intermediate positions between these two major groupings. The island and mainland Southeast Asian cranial series form a relatively distinct association in Fig. 16.4. The Chinese, Japanese,
TABLE 16.6 Some Jackknifed Classification Results Obtained From Stepwise Discriminant Function Analysis Showing the Cases Reclassified at the End of the Stepping Process (numbers in parentheses represent the number of crania originally assigned to each group)

<table>
<thead>
<tr>
<th>Tonga-Samoa (19)</th>
<th>Rapa Nui (50)</th>
<th>Hawaii (60)</th>
<th>Marquesas (63)</th>
<th>New Zealand (50)</th>
<th>Chatham Is. (45)</th>
<th>Society Is. (44)</th>
<th>Tuamotu Arch. (18)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOG 7</td>
<td>RAP 35</td>
<td>HAW 30</td>
<td>MRQ 14</td>
<td>NZ 19</td>
<td>CHT 32</td>
<td>SOC 12</td>
<td>TUA 9</td>
</tr>
<tr>
<td>HAW 3</td>
<td>GAM 6</td>
<td>TOG 3</td>
<td>TUA 9</td>
<td>SML 5</td>
<td>NZ 2</td>
<td>TUA 11</td>
<td>SOC 3</td>
</tr>
<tr>
<td>FIJ 2</td>
<td>HAW 2</td>
<td>RAP 3</td>
<td>SOC 6</td>
<td>MRQ 3</td>
<td>GUA 2</td>
<td>HAW 5</td>
<td>RAP 1</td>
</tr>
<tr>
<td>GAM 1</td>
<td>SOC 1</td>
<td>CHT 3</td>
<td>FIJ 5</td>
<td>TUA 3</td>
<td>HAW 1</td>
<td>CAR 3</td>
<td>NZ 1</td>
</tr>
<tr>
<td>GUA 1</td>
<td>TUA 1</td>
<td>SOC 3</td>
<td>NZ 4</td>
<td>LOY 3</td>
<td>SOC 1</td>
<td>MRQ 2</td>
<td>GAM 1</td>
</tr>
<tr>
<td>JAV 1</td>
<td>FIJ 1</td>
<td>GUA 3</td>
<td>HAW 3</td>
<td>CHT 2</td>
<td>GAM 1</td>
<td>CHT 2</td>
<td>CAR 1</td>
</tr>
<tr>
<td>CML 1</td>
<td>LOY 1</td>
<td>MRQ 2</td>
<td>RAP 2</td>
<td>SOL 2</td>
<td>FIJ 1</td>
<td>FIJ 2</td>
<td>SML 1</td>
</tr>
<tr>
<td>KAN 1</td>
<td>NIR 1</td>
<td>GAM 2</td>
<td>GAM 2</td>
<td>NIR 2</td>
<td>SEP 1</td>
<td>KYU 2</td>
<td>SUL 1</td>
</tr>
<tr>
<td>RYU 1</td>
<td>DAW 1</td>
<td>KAN 2</td>
<td>LSN 2</td>
<td>RAP 1</td>
<td>JAV 1</td>
<td>TOG 1</td>
<td></td>
</tr>
<tr>
<td>TAI 1</td>
<td>MSK 1</td>
<td>ADR 1</td>
<td>AIN 2</td>
<td>HAW 1</td>
<td>AIN 1</td>
<td>RAP 1</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gambier Is. (7)</th>
<th>Fiji (42)</th>
<th>Vanuatu (47)</th>
<th>Loyalty Is. (50)</th>
<th>New Caledonia (50)</th>
<th>Santa Cnzu Is. (46)</th>
<th>Solomon Is. (49)</th>
<th>New Britain (50)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RAP 2</td>
<td>FIJ 7</td>
<td>VAN 11</td>
<td>LOY 12</td>
<td>NCL 20</td>
<td>SCR 20</td>
<td>NBR 6</td>
<td>NBR 14</td>
</tr>
<tr>
<td>GAM 2</td>
<td>LOY 12</td>
<td>NBR 7</td>
<td>CHT 3</td>
<td>BIK 5</td>
<td>SCR 4</td>
<td>NBR 5</td>
<td>NBR 10</td>
</tr>
<tr>
<td>TUA 1</td>
<td>TOG 7</td>
<td>NSW 3</td>
<td>FIJ 3</td>
<td>NT 12</td>
<td>NBR 3</td>
<td>NBR 1</td>
<td>NBR 5</td>
</tr>
<tr>
<td>FIJ 1</td>
<td>NBR 2</td>
<td>TAS 3</td>
<td>QLD 3</td>
<td>VAN 2</td>
<td>NBR 2</td>
<td>NBR 3</td>
<td>NBR 3</td>
</tr>
<tr>
<td>GUA 1</td>
<td>LOY 2</td>
<td>SOL 2</td>
<td>BIK 2</td>
<td>SEP 23</td>
<td>SCR 2</td>
<td>SCR 2</td>
<td>SCR 2</td>
</tr>
<tr>
<td>DAW 2</td>
<td>NCL 2</td>
<td>NIR 2</td>
<td>RAP 2</td>
<td>CAT 2</td>
<td>TUA 2</td>
<td>CAR 2</td>
<td>CAR 2</td>
</tr>
<tr>
<td>SUL 2</td>
<td>NBR 2</td>
<td>CAR 2</td>
<td>PUR 2</td>
<td>DX 2</td>
<td>SCR 2</td>
<td>SCR 2</td>
<td>SCR 2</td>
</tr>
<tr>
<td>SOC 1</td>
<td>FLY 2</td>
<td>CAR 2</td>
<td>CAR 2</td>
<td>RAP 1</td>
<td>CAR 2</td>
<td>CAR 2</td>
<td>CAR 1</td>
</tr>
<tr>
<td>TUA 1</td>
<td>DAW 2</td>
<td>RAP 1</td>
<td>HAW 1</td>
<td>NCL 1</td>
<td>CAR 2</td>
<td>NCL 1</td>
<td>NCL 1</td>
</tr>
<tr>
<td>-----------------</td>
<td>------------------</td>
<td>--------------</td>
<td>--------------</td>
<td>-------------</td>
<td>------------------</td>
<td>--------------------------</td>
<td>------------------------</td>
</tr>
<tr>
<td>NBR 6</td>
<td>ADR 23</td>
<td>SEP 12</td>
<td>BIK 11</td>
<td>FLY 13</td>
<td>PUR 26</td>
<td>DTX 9</td>
<td>DAW 29</td>
</tr>
<tr>
<td>NIR 5</td>
<td>SEP 4</td>
<td>BIK 4</td>
<td>FJU 5</td>
<td>PUR 9</td>
<td>FLY 10</td>
<td>DAW 3</td>
<td>DTX 5</td>
</tr>
<tr>
<td>BIK 5</td>
<td>BIK 3</td>
<td>PUR 4</td>
<td>SEP 5</td>
<td>DTX 3</td>
<td>SCR 2</td>
<td>SCR 2</td>
<td>ATY 2</td>
</tr>
<tr>
<td>ADR 4</td>
<td>SML 3</td>
<td>DAW 4</td>
<td>VAN 4</td>
<td>VAN 2</td>
<td>SOL 2</td>
<td>SOL 2</td>
<td>WA 2</td>
</tr>
<tr>
<td>LOY 3</td>
<td>BOR 2</td>
<td>MSK 4</td>
<td>NIR 3</td>
<td>SCR 2</td>
<td>FLY 1</td>
<td>ADR 1</td>
<td>SML 2</td>
</tr>
<tr>
<td>NCL 3</td>
<td>RAP 1</td>
<td>DTX 3</td>
<td>PUR 3</td>
<td>WA 2</td>
<td>NCL 1</td>
<td>SEP 1</td>
<td>NBR 1</td>
</tr>
<tr>
<td>SEP 3</td>
<td>MRQ 1</td>
<td>CAR 3</td>
<td>NBR 2</td>
<td>JAV 2</td>
<td>NBR 1</td>
<td>BIK 1</td>
<td>SEP 1</td>
</tr>
<tr>
<td>FLY 2</td>
<td>SOC 1</td>
<td>SUL 3</td>
<td>FLY 2</td>
<td>FLY 1</td>
<td>NIR 1</td>
<td>FLY 1</td>
<td>BIK 1</td>
</tr>
<tr>
<td>CAR 2</td>
<td>FJ 1</td>
<td>SOL 2</td>
<td>DTX 2</td>
<td>SOL 1</td>
<td>ADR 1</td>
<td>NT 1</td>
<td>FLY 1</td>
</tr>
<tr>
<td>SML 2</td>
<td>VAN 1</td>
<td>ADR 2</td>
<td>RAP 1</td>
<td>SEP 1</td>
<td>BIK 1</td>
<td>TAS 1</td>
<td>QLD 1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>MRB 20</td>
<td>NSW 19</td>
<td>QLD 13</td>
<td>NT 15</td>
<td>SAS 27</td>
<td>WA 27</td>
<td>TAS 15</td>
<td>GUA 28</td>
</tr>
<tr>
<td>WA 9</td>
<td>QLD 14</td>
<td>NSW 10</td>
<td>MRB 7</td>
<td>QLD 3</td>
<td>TAS 4</td>
<td>NSW 2</td>
<td>HAW 4</td>
</tr>
<tr>
<td>NSW 5</td>
<td>NT 5</td>
<td>MRB 8</td>
<td>WA 6</td>
<td>MRB 2</td>
<td>NT 3</td>
<td>NZ 1</td>
<td>TOG 3</td>
</tr>
<tr>
<td>QLD 4</td>
<td>SAS 5</td>
<td>NT 7</td>
<td>NSW 5</td>
<td>VAN 1</td>
<td>NSW 2</td>
<td>VAN 1</td>
<td>MRQ 3</td>
</tr>
<tr>
<td>NT 4</td>
<td>MRB 3</td>
<td>WA 5</td>
<td>QLD 4</td>
<td>NBR 1</td>
<td>SAS 2</td>
<td>ADR 1</td>
<td>CML 2</td>
</tr>
<tr>
<td>VAN 2</td>
<td>SCR 2</td>
<td>VAN 2</td>
<td>SCR 3</td>
<td>NSW 1</td>
<td>VAN 1</td>
<td>SEP 1</td>
<td>NZ 1</td>
</tr>
<tr>
<td>LOY 1</td>
<td>BIK 2</td>
<td>SOL 2</td>
<td>FLY 2</td>
<td>AIN 1</td>
<td>LOY 1</td>
<td>MRB 1</td>
<td>NBR 1</td>
</tr>
<tr>
<td>NCL 1</td>
<td>WA 2</td>
<td>LOY 1</td>
<td>SOL 2</td>
<td>SOL 1</td>
<td>WA 1</td>
<td>SUM 1</td>
<td></td>
</tr>
<tr>
<td>SEP 1</td>
<td>NZ 1</td>
<td>NBR 1</td>
<td>VAN 1</td>
<td>NBR 1</td>
<td>BOR 1</td>
<td>JAV 1</td>
<td></td>
</tr>
<tr>
<td>PUR 1</td>
<td>LOY 1</td>
<td>FLY 1</td>
<td>NCL 1</td>
<td>DAW 1</td>
<td>KAN 1</td>
<td>SUL 1</td>
<td></td>
</tr>
</tbody>
</table>

(Continued)
TABLE 16.6  Continued

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>CAR</td>
<td>5</td>
<td>MSK</td>
<td>3</td>
<td>SUM</td>
<td>6</td>
<td>BOR</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>NBR</td>
<td>4</td>
<td>NZ</td>
<td>2</td>
<td>BOR</td>
<td>3</td>
<td>CML</td>
</tr>
<tr>
<td>RAP</td>
<td>1</td>
<td>NIR</td>
<td>2</td>
<td>JAV</td>
<td>2</td>
<td>SLW</td>
<td>4</td>
</tr>
<tr>
<td>HAW</td>
<td>1</td>
<td>MRQ</td>
<td>1</td>
<td>SUL</td>
<td>2</td>
<td>LSN</td>
<td>3</td>
</tr>
<tr>
<td>MRQ</td>
<td>1</td>
<td>SOC</td>
<td>1</td>
<td>BUR</td>
<td>2</td>
<td>SUM</td>
<td>3</td>
</tr>
<tr>
<td>SOC</td>
<td>1</td>
<td>TUA</td>
<td>1</td>
<td>TOH</td>
<td>2</td>
<td>BAC</td>
<td>3</td>
</tr>
<tr>
<td>FIJ</td>
<td>1</td>
<td>ADR</td>
<td>1</td>
<td>SHA</td>
<td>2</td>
<td>THI</td>
<td>3</td>
</tr>
<tr>
<td>VAN</td>
<td>1</td>
<td>BIK</td>
<td>1</td>
<td>RAP</td>
<td>1</td>
<td>TOG</td>
<td>1</td>
</tr>
<tr>
<td>NCL</td>
<td>1</td>
<td>CAR</td>
<td>1</td>
<td>MRQ</td>
<td>1</td>
<td>NZ</td>
<td>1</td>
</tr>
<tr>
<td>SCR</td>
<td>1</td>
<td>FIJ</td>
<td>1</td>
<td>SOC</td>
<td>1</td>
<td>BIK</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sulu (38)</th>
<th>Philippines (28)</th>
<th>Viet Nam (49)</th>
<th>Bachue Village (51)</th>
<th>Cambodia &amp; Laos (40)</th>
<th>Thailand (50)</th>
<th>Burma (16)</th>
<th>Kanto (50)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SUL</td>
<td>10</td>
<td>PHL</td>
<td>6</td>
<td>VTN</td>
<td>13</td>
<td>BAC</td>
<td>21</td>
</tr>
<tr>
<td>CML</td>
<td>7</td>
<td>VTN</td>
<td>5</td>
<td>PHL</td>
<td>6</td>
<td>CML</td>
<td>18</td>
</tr>
<tr>
<td>JAV</td>
<td>3</td>
<td>CML</td>
<td>4</td>
<td>BAC</td>
<td>21</td>
<td>CML</td>
<td>18</td>
</tr>
<tr>
<td>BOR</td>
<td>3</td>
<td>SLW</td>
<td>2</td>
<td>THI</td>
<td>4</td>
<td>HAI</td>
<td>3</td>
</tr>
<tr>
<td>SLW</td>
<td>3</td>
<td>TOG</td>
<td>2</td>
<td>HAI</td>
<td>3</td>
<td>MSK</td>
<td>2</td>
</tr>
<tr>
<td>TOG</td>
<td>1</td>
<td>ADR</td>
<td>1</td>
<td>BAC</td>
<td>2</td>
<td>BUR</td>
<td>1</td>
</tr>
<tr>
<td>HAW</td>
<td>1</td>
<td>BUR</td>
<td>1</td>
<td>HAI</td>
<td>1</td>
<td>PHL</td>
<td>2</td>
</tr>
<tr>
<td>TUA</td>
<td>1</td>
<td>ADR</td>
<td>1</td>
<td>BAC</td>
<td>2</td>
<td>HAI</td>
<td>2</td>
</tr>
<tr>
<td>NIR</td>
<td>1</td>
<td>HAW</td>
<td>1</td>
<td>LSN</td>
<td>1</td>
<td>KOR</td>
<td>2</td>
</tr>
<tr>
<td>ADR</td>
<td>1</td>
<td>CAR</td>
<td>1</td>
<td>HAW</td>
<td>1</td>
<td>KOR</td>
<td>2</td>
</tr>
<tr>
<td>Tohoku (53)</td>
<td>Kyushu (51)</td>
<td>Ainu (50)</td>
<td>Ryukyu Is. (64)</td>
<td>Shanghai (50)</td>
<td>Hangzhou (50)</td>
<td>Nanjing (49)</td>
<td>Chengdu (53)</td>
</tr>
<tr>
<td>------------</td>
<td>-------------</td>
<td>-----------</td>
<td>-----------------</td>
<td>--------------</td>
<td>--------------</td>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>TOH 19</td>
<td>KYU 16</td>
<td>AIN 29</td>
<td>RYU 23</td>
<td>SHA 12</td>
<td>SHA 12</td>
<td>NAJ 7</td>
<td>CHD 26</td>
</tr>
<tr>
<td>KYU 6</td>
<td>KAN 8</td>
<td>KAN 5</td>
<td>KAN 5</td>
<td>HAN 8</td>
<td>HAN 6</td>
<td>MAN 7</td>
<td>NAJ 4</td>
</tr>
<tr>
<td>AIN 4</td>
<td>TOH 4</td>
<td>TOH 3</td>
<td>VTN 4</td>
<td>NAJ 7</td>
<td>NAJ 6</td>
<td>SHA 5</td>
<td>MOG 4</td>
</tr>
<tr>
<td>KOR 4</td>
<td>KOR 4</td>
<td>KYU 3</td>
<td>TOH 3</td>
<td>THI 4</td>
<td>CHD 5</td>
<td>HAN 5</td>
<td>VTN 3</td>
</tr>
<tr>
<td>ATY 3</td>
<td>RYU 3</td>
<td>RYU 2</td>
<td>TAI 3</td>
<td>KAN 4</td>
<td>HK 5</td>
<td>CHD 5</td>
<td>HK 3</td>
</tr>
<tr>
<td>MAN 3</td>
<td>HK 3</td>
<td>RAP 1</td>
<td>TOG 2</td>
<td>CHD 4</td>
<td>MOG 4</td>
<td>HK 3</td>
<td>SHA 2</td>
</tr>
<tr>
<td>LSN 2</td>
<td>ATY 2</td>
<td>HAW 1</td>
<td>ATY 2</td>
<td>HK 4</td>
<td>SUM 2</td>
<td>SUM 2</td>
<td>HAN 2</td>
</tr>
<tr>
<td>KAN 2</td>
<td>TAI 2</td>
<td>CHT 1</td>
<td>MSK 2</td>
<td>MAN 2</td>
<td>LSN 2</td>
<td>RYU 2</td>
<td>CHT 1</td>
</tr>
<tr>
<td>TAI 2</td>
<td>HAI 2</td>
<td>FIJ 1</td>
<td>KYU 2</td>
<td>MOG 2</td>
<td>MAN 2</td>
<td>KOR 2</td>
<td>LSN 1</td>
</tr>
<tr>
<td>CHT 1</td>
<td>MAN 2</td>
<td>CAR 1</td>
<td>HAI 2</td>
<td>TOG 1</td>
<td>ADR 1</td>
<td>SEP 1</td>
<td>BUR 1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Hong Kong (50)</th>
<th>Taiwan (47)</th>
<th>Hainan Is. (47)</th>
<th>Manchuria (50)</th>
<th>Korea (32)</th>
<th>Mongolia (50)</th>
<th>Atayal (36)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HK 23</td>
<td>TAI 24</td>
<td>HAI 6</td>
<td>MAN 18</td>
<td>KOR 8</td>
<td>MOG 40</td>
<td>ATY 15</td>
</tr>
<tr>
<td>HAN 5</td>
<td>TOH 4</td>
<td>KOR 6</td>
<td>RYU 4</td>
<td>HAI 5</td>
<td>SHA 2</td>
<td>DAW 4</td>
</tr>
<tr>
<td>MAN 4</td>
<td>HAI 3</td>
<td>BUR 4</td>
<td>CHD 4</td>
<td>TOH 3</td>
<td>CHD 2</td>
<td>KAN 3</td>
</tr>
<tr>
<td>KYU 3</td>
<td>KOR 3</td>
<td>RYU 4</td>
<td>HK 3</td>
<td>KYU 3</td>
<td>JAV 1</td>
<td>KYU 2</td>
</tr>
<tr>
<td>THI 2</td>
<td>BAC 2</td>
<td>TAI 4</td>
<td>KOR 3</td>
<td>HAW 2</td>
<td>THI 1</td>
<td>AIN 2</td>
</tr>
<tr>
<td>TAI 2</td>
<td>LSN 1</td>
<td>THI 3</td>
<td>ATY 2</td>
<td>NAJ 2</td>
<td>TOH 1</td>
<td>KOR 2</td>
</tr>
<tr>
<td>TOG 1</td>
<td>SML 1</td>
<td>GUA 2</td>
<td>KAN 2</td>
<td>ATY 1</td>
<td>RYU 1</td>
<td>ADR 1</td>
</tr>
<tr>
<td>CHT 1</td>
<td>PHL 1</td>
<td>PHL 2</td>
<td>TOH 2</td>
<td>SUM 1</td>
<td>HAN 1</td>
<td>SEP 1</td>
</tr>
<tr>
<td>ADR 1</td>
<td>VTN 1</td>
<td>CHD 2</td>
<td>KYU 2</td>
<td>BOR 1</td>
<td>HAI 1</td>
<td>PHL 1</td>
</tr>
<tr>
<td>ATY 1</td>
<td>CML 1</td>
<td>HK 2</td>
<td>SHA 2</td>
<td>PHL 1</td>
<td>TOH 1</td>
<td></td>
</tr>
</tbody>
</table>

See Table 16.1 for explanation of abbreviations.
Manchurian, and Korean cranial series, including the Ainu, form another separate grouping. The cranial series from Mongolia is the most isolated series in this representation.

To examine differences between individual groups more closely requires another multivariate procedure, namely Mahalanobis’s generalized distance.

Figure 16.2  A plot of 63 group means on the first two canonical variates that results from the application of stepwise discriminant function analysis to 27 cranial measurements.

Figure 16.3  A plot of 35 of the 63 groups on the first three canonical variates that results from the application of stepwise discriminant function analysis to 27 cranial measurements.
Mahalanobis’s Generalized Distance-$D^2$. Some of the results of applying Mahalanobis’s generalized distances to the same measurements analyzed by stepwise discriminant function analysis are given in Table 16.7. The results given in this table list the ten smallest distances for each of the 63 groups, which are groups that are most similar to each of these groups.

Although the cranial series from Polynesia and Micronesia are generally closest to other Polynesian/Micronesian series, several island Southeast Asian series (e.g., Java, Borneo, Sulawesi, Sulu, Lesser Sunda Islands, and Southern Moluccas) rank among the smallest distances for Tonga-Samoa, Hawaii, Marquesas New Zealand, Chatham Islands, and Tuamotu. The cranial series from the Southern Moluccas, followed by Marquesas, Marshall-Kiribati, Solomon Islands, Caroline Islands, New Ireland, and Biak Island, are among the groups closest to the New Zealand Maori series. A few island Melanesian cranial series (e.g., Fiji, Loyalty, and New Caledonia) are among the smallest distances to the Polynesian series. Among the distances closest to the Caroline Islands and Marshall/Kiribati Islands series are those associated with these two series, New Zealand, Marquesas, several Melanesian series, and the Southern Moluccas series.

Inspection of the smallest distances for the island Melanesian (e.g., Fiji, Vanuatu, Loyalty, or New Caledonia) and New Guinea series (e.g., Sepik, Fly River, or Purari Delta) reveals that most of these are associated with cranial series from the same geographical region as well as the Caroline, Marshall-Kiribati, and Southern Moluccas series. Without exception the distances closest to the Australian series are other Australian/Tasmanian and Melanesian (e.g., New Britain, Vanuatu, or Santa Cruz) series.

Interestingly, one series from geographical Melanesia, Admiralty Island, is found to be closest to the Southern Moluccas series and several other island Southeast Asian series, including Borneo, Sulu, and the Lesser Sunda Islands.

Most of the cranial series closest to the island and mainland Southeast Asian series are generally from Southeast Asia. Somewhat surprisingly, the cranial series closest to one of the island Southeast Asian series, the

Figure 16.4 A plot of the 28 of the 63 groups on the first three canonical variates that results from the application of stepwise discriminant function analysis to 27 cranial measurements.
### TABLE 16.7 The Smallest Mahalanobis’s Distances for 63 Groups (except where indicated, all variance ratios calculated for distances are significant at the 1% level)

<table>
<thead>
<tr>
<th></th>
<th>Tonga-Samoa</th>
<th>Rapa Nui</th>
<th>Hawaii</th>
<th>Marquesas Islands</th>
<th>New Zealand</th>
<th>Chatham Islands</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>HAW</strong></td>
<td>6.643</td>
<td>GAM 5.439b</td>
<td>TOG 6.643</td>
<td>SOC 3.583</td>
<td>SML 3.553</td>
<td>NZ 7.468</td>
</tr>
<tr>
<td><strong>SLW</strong></td>
<td>8.290a</td>
<td>MRQ 8.811</td>
<td>SUL 7.953</td>
<td>GAM 5.819b</td>
<td>MSK 5.473b</td>
<td>HAW 9.633</td>
</tr>
<tr>
<td><strong>CML</strong></td>
<td>8.557a</td>
<td>CAR 10.236</td>
<td>JAV 8.143</td>
<td>TUA 7.167</td>
<td>SOL 5.631</td>
<td>GAM 10.105b</td>
</tr>
<tr>
<td><strong>LSN</strong></td>
<td>8.805</td>
<td>MSK 11.045</td>
<td>GUA 8.898</td>
<td>MSK 7.681a</td>
<td>NIR 6.839</td>
<td>SOC 13.959</td>
</tr>
<tr>
<td><strong>GUA</strong></td>
<td>9.006</td>
<td>SOC 11.087</td>
<td>SLW 8.969</td>
<td>SML 8.116</td>
<td>BIK 7.071</td>
<td>SML 14.514</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Society Islands</strong></th>
<th><strong>Tuamotu Islands</strong></th>
<th><strong>Gambier Islands</strong></th>
<th><strong>Fiji</strong></th>
<th><strong>Vanuatu</strong></th>
<th><strong>Loyalty Islands</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MRQ</strong></td>
<td>3.583</td>
<td>SOC 4.690b</td>
<td>RAP 5.439b</td>
<td>BIK 4.232</td>
<td>NBR 1.742b</td>
</tr>
<tr>
<td><strong>TUA</strong></td>
<td>4.690b</td>
<td>MRQ 7.167</td>
<td>MRQ 5.819b</td>
<td>SOL 5.010</td>
<td>BIK 4.155</td>
</tr>
<tr>
<td><strong>MSK</strong></td>
<td>7.364b</td>
<td>MSK 10.102b</td>
<td>NZ 7.462b</td>
<td>LOY 5.185</td>
<td>NIR 4.533</td>
</tr>
<tr>
<td><strong>GAM</strong></td>
<td>8.551b</td>
<td>NZ 11.019</td>
<td>SOC 8.551b</td>
<td>NIR 5.884</td>
<td>SOL 4.695</td>
</tr>
<tr>
<td><strong>NZ</strong></td>
<td>8.796</td>
<td>GAM 12.724b</td>
<td>CHT 10.105b</td>
<td>CAR 6.091a</td>
<td>SCR 5.244</td>
</tr>
<tr>
<td><strong>CAR</strong></td>
<td>9.914</td>
<td>CAR 14.994b</td>
<td>MSK 11.845b</td>
<td>NCL 6.685</td>
<td>QLD 5.954</td>
</tr>
<tr>
<td><strong>TOG</strong></td>
<td>10.963</td>
<td>TOG 15.249b</td>
<td>TUA 12.724b</td>
<td>VAN 6.796</td>
<td>NT 5.954</td>
</tr>
<tr>
<td><strong>RAP</strong></td>
<td>11.087</td>
<td>RAP 15.386</td>
<td>FIJ 13.014b</td>
<td>SML 7.451</td>
<td>NCL 6.652</td>
</tr>
<tr>
<td><strong>FIJ</strong></td>
<td>11.632</td>
<td>SML 15.921</td>
<td>GUA 13.968b</td>
<td>NZ 7.711</td>
<td>SML 6.678</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>New Caledonia</strong></th>
<th><strong>Santa Cruz Islands</strong></th>
<th><strong>Solomon Islands</strong></th>
<th><strong>New Britain</strong></th>
<th><strong>New Ireland</strong></th>
<th><strong>Admiralty Islands</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>LOY</strong></td>
<td>3.313</td>
<td>NBR 4.747</td>
<td>NIR 2.389a</td>
<td>VAN 1.742b</td>
<td>SOL 2.389a</td>
</tr>
<tr>
<td><strong>NBR</strong></td>
<td>5.982</td>
<td>NT 5.089b</td>
<td>CAR 3.515b</td>
<td>NIR 3.485</td>
<td>NBR 3.485</td>
</tr>
<tr>
<td></td>
<td>SOL 6.270</td>
<td>VAN 5.244</td>
<td>SEP 3.873</td>
<td>SOL 4.170</td>
<td>CAR 3.663&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>-------</td>
<td>----------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
<td>------------------------</td>
</tr>
<tr>
<td>NIR</td>
<td>6.284</td>
<td>SOL 5.932</td>
<td>BIK 4.093</td>
<td>SCR 4.747</td>
<td>BIK 3.825</td>
</tr>
<tr>
<td>VAN</td>
<td>6.652</td>
<td>NCL 6.716</td>
<td>NBR 4.170</td>
<td>NCL 5.982</td>
<td>SEP 4.011</td>
</tr>
<tr>
<td>SCR</td>
<td>6.716</td>
<td>NIR 6.879</td>
<td>DTX 4.620&lt;sup&gt;a&lt;/sup&gt;</td>
<td>NT 6.102</td>
<td>VAN 4.533</td>
</tr>
<tr>
<td>CAR</td>
<td>7.174</td>
<td>SEP 6.886</td>
<td>VAN 4.695</td>
<td>NSW 7.013</td>
<td>FIJ 5.884</td>
</tr>
<tr>
<td>MSK</td>
<td>8.044&lt;sup&gt;b&lt;/sup&gt;</td>
<td>PUR 7.613</td>
<td>FLY 4.874</td>
<td>SEP 7.241</td>
<td>FLY 5.912</td>
</tr>
<tr>
<td>SEP</td>
<td>8.134</td>
<td>FLY 7.883</td>
<td>FIJ 5.010</td>
<td>LOY 7.461</td>
<td>DTX 6.070</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>D'Entrecasteaux</th>
<th>Biak Island</th>
<th>Fly R.</th>
<th>Purari Delta</th>
<th>Islands</th>
<th>Dawson Strait</th>
</tr>
</thead>
<tbody>
<tr>
<td>DTX</td>
<td>3.714&lt;sup&gt;b&lt;/sup&gt;</td>
<td>NIR 3.825</td>
<td>PUR 2.101&lt;sup&gt;b&lt;/sup&gt;</td>
<td>FLY 2.101&lt;sup&gt;b&lt;/sup&gt;</td>
<td>SE 3.714&lt;sup&gt;b&lt;/sup&gt;</td>
<td>DTX 3.746&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>SOL</td>
<td>3.873</td>
<td>SML 3.873</td>
<td>SEP 4.226</td>
<td>SEP 6.317</td>
<td>DAW 3.746&lt;sup&gt;b&lt;/sup&gt;</td>
<td>BIK 5.498</td>
</tr>
<tr>
<td>NIR</td>
<td>4.011</td>
<td>SEP 4.015</td>
<td>SOL 4.874</td>
<td>SOL 6.454</td>
<td>SOL 4.620&lt;sup&gt;a&lt;/sup&gt;</td>
<td>SEP 5.892</td>
</tr>
<tr>
<td>BIK</td>
<td>4.015</td>
<td>SOL 4.093</td>
<td>BIK 5.317</td>
<td>BIK 7.261</td>
<td>SML 5.489</td>
<td>NIR 6.103</td>
</tr>
<tr>
<td>CAR</td>
<td>5.297&lt;sup&gt;a&lt;/sup&gt;</td>
<td>FIJ 4.232</td>
<td>DTX 6.129&lt;sup&gt;a&lt;/sup&gt;</td>
<td>DTX 8.044</td>
<td>FLY 6.129&lt;sup&gt;a&lt;/sup&gt;</td>
<td>SOL 6.808</td>
</tr>
<tr>
<td>DAW</td>
<td>5.892</td>
<td>FLY 5.317</td>
<td>FIJ 6.488</td>
<td>NT 8.408</td>
<td>BIK 6.799</td>
<td>LSN 7.318</td>
</tr>
<tr>
<td>SML</td>
<td>6.083</td>
<td>DAW 5.498</td>
<td>VAN 7.659</td>
<td>NIR 8.695</td>
<td>SCR 7.946</td>
<td>FIJ 7.795</td>
</tr>
<tr>
<td>MSK</td>
<td>6.195</td>
<td>CAR 5.535&lt;sup&gt;a&lt;/sup&gt;</td>
<td>CAR 7.669</td>
<td>NBR 9.007</td>
<td>PUR 8.044</td>
<td>ATY 7.998</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Murray R.</th>
<th>New South Wales</th>
<th>Queensland</th>
<th>Northern Territory</th>
<th>Swanport, S.A.</th>
<th>Western Australia</th>
</tr>
</thead>
<tbody>
<tr>
<td>QLD</td>
<td>2.696&lt;sup&gt;a&lt;/sup&gt;</td>
<td>QLD 1.575&lt;sup&gt;b&lt;/sup&gt;</td>
<td>NSW 1.575&lt;sup&gt;b&lt;/sup&gt;</td>
<td>QLD 2.560&lt;sup&gt;a&lt;/sup&gt;</td>
<td>NSW 6.116</td>
</tr>
<tr>
<td>NSW</td>
<td>2.840</td>
<td>MRB 2.840</td>
<td>NT 2.560&lt;sup&gt;a&lt;/sup&gt;</td>
<td>MRB 2.850&lt;sup&gt;a&lt;/sup&gt;</td>
<td>MRB 7.167</td>
</tr>
<tr>
<td>NT</td>
<td>2.850&lt;sup&gt;a&lt;/sup&gt;</td>
<td>NT 3.238</td>
<td>MRB 2.696&lt;sup&gt;a&lt;/sup&gt;</td>
<td>NSW 3.238</td>
<td>QLD 8.069</td>
</tr>
<tr>
<td>VAN</td>
<td>6.738</td>
<td>VAN 5.344</td>
<td>VAN 5.954</td>
<td>SCR 5.089</td>
<td>VAN 8.371</td>
</tr>
<tr>
<td>WA</td>
<td>6.875</td>
<td>WA 5.943</td>
<td>WA 6.345</td>
<td>VAN 5.954</td>
<td>NT 9.780</td>
</tr>
</tbody>
</table>

(Continued)
### TABLE 16.7  Continued

<table>
<thead>
<tr>
<th>Murray R.</th>
<th>New South Wales</th>
<th>Queensland</th>
<th>Northern Territory</th>
<th>Swanport, S.A.</th>
<th>Western Australia</th>
</tr>
</thead>
<tbody>
<tr>
<td>TAS</td>
<td>8.733</td>
<td>NBR 7.013</td>
<td>SAS 8.069</td>
<td>WA 7.160</td>
<td>LOY 12.520</td>
</tr>
<tr>
<td>SCR</td>
<td>9.358</td>
<td>BIK 8.679</td>
<td>FIJ 8.320</td>
<td>BIK 7.820</td>
<td>TAS 12.822</td>
</tr>
<tr>
<td>SOL</td>
<td>11.175</td>
<td>TAS 8.893</td>
<td>TAS 8.442</td>
<td>FLY 8.231</td>
<td>FIJ 13.645</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tasmania</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>QLD</td>
<td>8.442</td>
<td>HAW 8.898</td>
<td>SOL 3.515</td>
<td>CAR 4.239</td>
<td></td>
</tr>
<tr>
<td>VAN</td>
<td>8.486</td>
<td>TOG 9.006</td>
<td>NIR 3.663</td>
<td>BOR 2.239</td>
<td></td>
</tr>
<tr>
<td>MRB</td>
<td>8.733</td>
<td>NZ 9.412</td>
<td>MSK 4.239</td>
<td>SLW 3.125</td>
<td></td>
</tr>
<tr>
<td>NSW</td>
<td>8.893</td>
<td>LSN 9.649</td>
<td>SEP 5.297</td>
<td>VTN 4.333</td>
<td></td>
</tr>
<tr>
<td>NBR</td>
<td>10.662</td>
<td>BOR 9.777</td>
<td>BIK 5.535</td>
<td>JAR 4.399</td>
<td></td>
</tr>
<tr>
<td>NT</td>
<td>12.085</td>
<td>SUL 10.286</td>
<td>FJ 6.091</td>
<td>THI 4.321</td>
<td></td>
</tr>
<tr>
<td>SAS</td>
<td>12.822</td>
<td>MRQ 10.714</td>
<td>SML 4.448</td>
<td>SLW 3.125</td>
<td></td>
</tr>
<tr>
<td>BIK</td>
<td>14.168</td>
<td>SLW 11.794</td>
<td>NCL 7.174</td>
<td>LSN 4.865</td>
<td></td>
</tr>
<tr>
<td>SML</td>
<td>14.700</td>
<td>CML 11.859</td>
<td>LOY 7.500</td>
<td>VTN 6.070</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marshall-Kiribati</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>QLD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VAN</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MRB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NSW</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NBR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SAS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BIK</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SML</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indonesia</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LSN</td>
<td>1.810</td>
<td>JAV 2.334</td>
<td>BOR 1.810</td>
<td>NZ 3.553</td>
<td></td>
</tr>
<tr>
<td>SUM</td>
<td>2.239</td>
<td>CML 2.607</td>
<td>SUM 2.622</td>
<td>BIK 3.873</td>
<td></td>
</tr>
<tr>
<td>SLW</td>
<td>2.658</td>
<td>BOR 2.658</td>
<td>SLW 2.737</td>
<td>NIR 4.255</td>
<td></td>
</tr>
<tr>
<td>SUL</td>
<td>4.053</td>
<td>LSN 2.737</td>
<td>JAV 3.611</td>
<td>SOL 4.373</td>
<td></td>
</tr>
<tr>
<td>JAV</td>
<td>4.434</td>
<td>SUL 2.853</td>
<td>PHL 3.706</td>
<td>LSN 4.865</td>
<td></td>
</tr>
<tr>
<td>PHL</td>
<td>4.746</td>
<td>SUM 3.125</td>
<td>VTN 4.608</td>
<td>BOR 5.328</td>
<td></td>
</tr>
<tr>
<td>VTN</td>
<td>4.871</td>
<td>PHL 3.426</td>
<td>SOL 4.852</td>
<td>LSN 4.852</td>
<td></td>
</tr>
<tr>
<td>CML</td>
<td>5.168</td>
<td>THI 3.952</td>
<td>SML 4.865</td>
<td>ADR 5.930</td>
<td></td>
</tr>
<tr>
<td>SML</td>
<td>5.328</td>
<td>VTN 5.408</td>
<td>SOL 5.430</td>
<td>SEP 6.083</td>
<td></td>
</tr>
<tr>
<td>SOL</td>
<td>7.003</td>
<td>BAC 5.918</td>
<td>RYU 5.980</td>
<td>DAW 6.185</td>
<td></td>
</tr>
</tbody>
</table>

### Table Notes
- The table continues with additional rows and columns, detailing various islands and their corresponding data points.
- The data points are likely to include measurements or values related to the regions mentioned.
- The table continues to list data for different regions, each with specific columns for data representation.
<table>
<thead>
<tr>
<th>Viet Nam</th>
<th>Bachuc Village</th>
<th>Cambodia &amp; Laos</th>
<th>Thailand</th>
<th>Burma</th>
<th>Kanto</th>
</tr>
</thead>
<tbody>
<tr>
<td>PHL</td>
<td>2.353</td>
<td>THI 3.169</td>
<td>SLW 2.607</td>
<td>BAC 3.169</td>
<td>SUL 6.202</td>
</tr>
<tr>
<td>RYU</td>
<td>4.247</td>
<td>HAI 3.780</td>
<td>SUL 2.742</td>
<td>SLW 3.952</td>
<td>SUM 6.442</td>
</tr>
<tr>
<td>THI</td>
<td>4.296</td>
<td>VTN 5.139</td>
<td>JAV 3.498</td>
<td>VTN 4.296</td>
<td>CML 7.013</td>
</tr>
<tr>
<td>SUM</td>
<td>4.333</td>
<td>CML 5.757</td>
<td>BOR 5.168</td>
<td>JAV 4.321</td>
<td>SLW 7.216</td>
</tr>
<tr>
<td>LSN</td>
<td>4.608</td>
<td>SLW 5.918</td>
<td>THI 5.495</td>
<td>PHL 4.821</td>
<td>JAV 7.324</td>
</tr>
<tr>
<td>BOR</td>
<td>4.871</td>
<td>JAV 6.036</td>
<td>BAC 5.757</td>
<td>CML 5.495</td>
<td>BAC 7.464</td>
</tr>
<tr>
<td>KOR</td>
<td>5.090</td>
<td>PHL 6.235</td>
<td>PHL 5.955</td>
<td>HAI 5.558</td>
<td>THI 7.756</td>
</tr>
<tr>
<td>BAC</td>
<td>5.139</td>
<td>KOR 6.358</td>
<td>LSN 5.991</td>
<td>KAN 5.933</td>
<td>VTN 9.208</td>
</tr>
<tr>
<td>HAI</td>
<td>5.191</td>
<td>HK 6.910</td>
<td>SUM 6.413</td>
<td>SUM 5.936</td>
<td>BOR 9.328</td>
</tr>
<tr>
<td>SLW</td>
<td>5.408</td>
<td>SUL 7.213</td>
<td>BUR 7.013</td>
<td>KOR 6.100</td>
<td>HAI 9.586</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tohoku</th>
<th>Kyushu</th>
<th>Ainu</th>
<th>Ryukyu Islands</th>
<th>Shanghai</th>
</tr>
</thead>
<tbody>
<tr>
<td>KYU</td>
<td>2.167</td>
<td>TOH 2.167</td>
<td>KYU 5.371</td>
<td>VTN 4.247</td>
</tr>
<tr>
<td>KOR</td>
<td>3.219</td>
<td>KAN 2.784</td>
<td>TOH 6.067</td>
<td>KYU 4.790</td>
</tr>
<tr>
<td>SLW</td>
<td>3.333</td>
<td>KOR 2.844</td>
<td>RYU 6.366</td>
<td>KOR 5.280</td>
</tr>
<tr>
<td>CML</td>
<td>4.659</td>
<td>RYU 4.790</td>
<td>KAN 6.402</td>
<td>HAI 5.638</td>
</tr>
<tr>
<td>SUL</td>
<td>5.438</td>
<td>HAI 4.876</td>
<td>SUM 6.840</td>
<td>TAI 5.842</td>
</tr>
<tr>
<td>AIN</td>
<td>6.067</td>
<td>ATY 5.011</td>
<td>LSN 8.503</td>
<td>LSN 5.980</td>
</tr>
<tr>
<td>NAJ</td>
<td>6.179</td>
<td>MAN 5.198</td>
<td>VTN 9.644</td>
<td>SUM 6.303</td>
</tr>
</tbody>
</table>

(Continued)
TABLE 16.7  Continued

<table>
<thead>
<tr>
<th>Hangzhou</th>
<th>Nanjing</th>
<th>Chengdu</th>
<th>Hong Kong</th>
<th>Taiwan</th>
</tr>
</thead>
<tbody>
<tr>
<td>SHA</td>
<td>0.549</td>
<td>HAN</td>
<td>1.398</td>
<td>NAJ</td>
</tr>
<tr>
<td>NAJ</td>
<td>1.398</td>
<td>SHA</td>
<td>1.985</td>
<td>HAN</td>
</tr>
<tr>
<td>HK</td>
<td>3.536</td>
<td>CHD</td>
<td>2.584a</td>
<td>SHA</td>
</tr>
<tr>
<td>CHD</td>
<td>3.767</td>
<td>MAN</td>
<td>3.302</td>
<td>MAN</td>
</tr>
<tr>
<td>MAN</td>
<td>4.994</td>
<td>HK</td>
<td>4.018</td>
<td>KOR</td>
</tr>
<tr>
<td>KOR</td>
<td>5.493</td>
<td>KOR</td>
<td>4.376</td>
<td>HK</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Hainan Island</th>
<th>Manchuria</th>
<th>Korea</th>
<th>Mongolia</th>
<th>Atayal</th>
</tr>
</thead>
<tbody>
<tr>
<td>KOR</td>
<td>2.415</td>
<td>NAJ</td>
<td>3.302</td>
<td>HAI</td>
</tr>
<tr>
<td>TAI</td>
<td>3.165a</td>
<td>KOR</td>
<td>3.645a</td>
<td>KYU</td>
</tr>
<tr>
<td>BAC</td>
<td>3.780</td>
<td>TOH</td>
<td>4.659</td>
<td>TOH</td>
</tr>
<tr>
<td>KYU</td>
<td>4.876</td>
<td>HAN</td>
<td>4.994</td>
<td>MAN</td>
</tr>
<tr>
<td>VTN</td>
<td>5.191</td>
<td>KYU</td>
<td>5.198</td>
<td>TAI</td>
</tr>
<tr>
<td>THI</td>
<td>5.558</td>
<td>CHD</td>
<td>5.678</td>
<td>KAN</td>
</tr>
<tr>
<td>RYU</td>
<td>5.638</td>
<td>HK</td>
<td>5.734</td>
<td>NAJ</td>
</tr>
<tr>
<td>MAN</td>
<td>5.804</td>
<td>HAI</td>
<td>5.804</td>
<td>VTN</td>
</tr>
<tr>
<td>ATY</td>
<td>6.271</td>
<td>SHA</td>
<td>5.894</td>
<td>RYU</td>
</tr>
</tbody>
</table>

*aVariance ratio significant at 5% level.
*bVariance ratio not significant at 5% level.
Southern Moluccas, is New Zealand, a Polynesian series. Seven of the smallest distances to the Southern Moluccas include seven series from New Guinea and the neighboring regions of island Melanesia, (i.e., New Ireland, Admiralty, Dawson, Admiralty Islands, and D’Entrecasteaux).

Several Southeast Asian series are among the ten smallest distances to the Japanese series. Korea generally ranks in the top ten closest groups to the Japanese and Chinese series. Although not significant, Korea is closest to the Hainan Island series. Other groups closest to Hainan Island include Taiwan Chinese, two Vietnamese series, Thailand, and the Ryukyu Islands.

The groups found to be closest to the Ainu cranial series include several modern Japanese

Figure 16.5  Dendrogram showing the relationship of 63 groups that results from a cluster analysis (UPGMA) of Mahalanobis’s distances using 27 cranial measurements.
followed by the Ryukyu Islanders and several island Southeast Asian series. Distances closest to the Ryukyu Islands include Vietnam, Kyushu, Korea, Hainan Island, Taiwan, Lesser Sunda Islands, and Sumatra. The series closest to the Atayal (Taiwan Aborigines) series include three Japanese series, Korea, Hainan Island, and Vietnam.

Applying the UPGMA clustering algorithm to these distances results in the dendrogram shown in Fig. 16.5. Two major divisions are evident in this diagram of relationships, the first includes all Asian (North, East and Southeast Asia) and Polynesian cranial series and the second includes all Australian, Tasmanian, New Guinea and Melanesian cranial series.

With the exception of New Zealand, the Polynesian (and one from Guam) cranial series form a distinct cluster that ultimately connects to one containing East Asian, North Asian, and Southeast Asian cranial series. The New Zealand Maori and Southern Moluccas cranial series group with three Melanesian (Fiji, Solomon Islands, and New Ireland) series, a single north coastal New Guinea series (Biak Island), and two Micronesian cranial series (Caroline Islands and Marshall-Kiribati Islands).

The cranial samples representing Southeast Asia occupy two separate clusters, representing primarily the mainland and island Southeast Asia series, respectively. The cranial series representing modern Japanese align with Taiwan and Hainan Island Chinese, Korea, and more remotely with the Ryukyu Islands and Atayal (Taiwan Aboriginal) series. The remaining series representing China and Manchuria occupy a separate branch. The Ainu and Mongolian series are the last two series to connect with this exclusively Asian division that comprises the cranial series from Northern, Eastern, and Southeast Asia.

**Discussion of Results.** The results of the multivariate craniometric analysis used in this application indicate the presence of two major divisions representing the inhabitants of the Pacific and neighboring regions of eastern Asia. One of these major divisions includes cranial series representing the indigenous inhabitants of Australia, Tasmania, and geographical Melanesia, whereas a second major division includes cranial series from East Asia, North Asia, Southeast Asia, and Remote Oceania. The sharpness of the distinction suggests separate origins for the indigenous inhabitants of these two regions, one that coincides with the initial peopling of Australia and Near Oceania and a second event that accounts for the earliest human colonization of Remote Oceania.

Likewise, in these results, the Polynesian cranial series occupy a separate branch of the greater East/Southeast Asian division, one that is well removed from the division that includes all the cranial series from Melanesia and Australia, which is a relationship that is more consistent with an ancestral homeland for Polynesians in Eastern and Southeast Asia rather than one within geographically adjacent Melanesia. The classification results (Table 16.6) and detailed inspection of Mahalanobis’s distances (Table 16.7) also indicate connections between several of the Polynesian series and cranial series from eastern Indonesia. The Micronesian series are variable; some (e.g., Guam) show Polynesian affinities, whereas others (e.g., Caroline Islands) reveal connections with Melanesia.

These multivariate craniometric results share much in common with archaeological, historical linguistic, and recent molecular genetic evidence (e.g., Blust, 1995; Kirch, 1997, 2000; Lum and Cann, 1998; Lum et al., 1998; Melton et al., 1995; Merriwether et al., 1999; Oppenheimer and Richards, 2001; Redd et al., 1995; Richards et al., 1998; Su et al., 2000) that favors a relatively rapid eastward migration and colonization of Remote Oceania by peoples and cultures (the so-called Lapita expansion) whose ancestors are traceable to somewhere in eastern island Southeast Asia.

In these same results, island and mainland Southeast Asian cranial series form two separate
branches well separated from the East and North Asian series, which is a distinction that implies long-term in situ development in both regions rather than displacement (see, e.g., Bellwood, 1997) to account for the current-day inhabitants of Southeast Asia, conclusions that are supported by dental (e.g., Turner, 1987, 1989, 1992) and other studies that use craniometric data (e.g., Hanihara, 1996).

Finally, contrary to others (e.g., Brace et al., 1990), there is no support for a close biological connection between the Ainu and the Polynesian series in these results. Rather, the Ainu are members (albeit marginal) of a greater East/North Asian division and do not connect directly with any of the Polynesian series, which is a conclusion supported by other skeletal evidence (see, e.g., Hanihara, 1993, 1996).

**CONCLUSIONS**

The application of analytical methods to metric data, which has characterized physical anthropology since its inception, remains a major focus of interest in biological anthropology. Likewise, determining biological relatedness, as in the application provided in this chapter, continues to attract considerable attention in skeletal biology and physical anthropology. Multivariate statistical procedures remain the most robust procedures available for analyzing both metric and non-metric variation. Discriminant function analysis and Mahalanobis’s generalized distance constitute two of the most popular multivariate procedures for determining biological relatedness and for classification of unknown specimens using metric variables. Various clustering procedures and other methods of ordination provide an important means of visualizing multivariate results.

The general reluctance among anthropologists and skeletal biologists to use multivariate statistical procedures, stems, in part, from the complex, often daunting, mathematical theory that underlies these methods. Refinements of the methodological and theoretical concerns, which have often resulted in mathematical adjustments or “corrections” to the existing procedures, have done little to advance the easy acceptance and use of these methods. However, given that many of these methods are now readily available in statistical packages designed for personal computers lessens the need for users to be highly skilled or trained in mathematics and/or statistics.

Coinciding with anthropologists’ earliest applications of multivariate statistical procedures to metric data, the principal concern has been determining past relationships and classification or allocation of individual (usually fossil) specimens. More recently, the controversy surrounding the discovery of the Kennewick skeleton (Chatters, 2001) and determining ancestral–descendant relationships of skeletal remains of indigenous peoples has attracted considerable attention. This trend, which primarily addresses biological relationships of past populations, is likely to continue well into the future. Likewise, the use of multivariate statistical procedures for classificatory purposes will most certainly continue to find applications in forensic anthropology and in cases involving repatriation claims and NAGPRA legislation. Although global appraisals of skeletal (especially cranial) variation will undoubtedly continue, studies of human skeletal remains having more restricted regional and temporal focus are likely to increase in the future.

With the recognition by bioarchaeologists of the importance of determining biological relatedness to contextualize issues pertaining to health, disease, nutrition, demography, and epidemiology in past populations, studies that apply multivariate statistical procedures to metric and nonmetric data are expected to increase in number.

Finally, given the statistical and mathematical underpinnings of the methods and the problems associated with interpreting the results
of multivariate statistical analyses, physical anthropologists and skeletal biologists, especially those highly tutored in quantitative methods, are expected to continue to refine the methodological and theoretical concerns associated with these methods. It is anticipated that these endeavors will make these methods both more accessible and relevant for analyzing and interpreting anthropological and skeletal biological data.

ACKNOWLEDGMENTS

My thanks to Billie Ikeda of the University of Hawaii Instructional Support Center for her assistance with producing the figures used in this chapter. I am grateful for the assistance of Scott Reinke and Rona Ikehara-Quebral in data analysis. As always, I am extremely grateful to Michele Toomay Douglas for her comments used in preparing this chapter. My thanks to RoutledgeCurzon (London) for permission to reproduce the five illustrations in this chapter that first appeared in Pietrusewsky (2005). Permission to examine the cranial series used in the example used in this chapter has been acknowledged elsewhere (Pietrusewsky, 1994, 1995, 2004; Pietrusewsky and Chang, 2003).

REFERENCES

Campbell NA. 1984. Some aspects of allocation and discrimination. In: Van Vark GN, Howells WW,


REFERENCES

529


SPSS, Inc. 1999a. SPSS Advanced Models 10.0. Chicago, Ill.: SPSS, Inc.

SPSS, Inc. 1999b. SPSS Base 10.0 Applications Guide. Chicago, Ill.: SPSS, Inc.

SPSS, Inc. 1999c. SYSTAT 9 Statistics I and II. Chicago, Ill.: SPSS, Inc.


the study of human populations from their skeletal remains. Homo 27:94–114.


Nonmetric traits are morphological variants of anatomy, typically of a feature or of an anatomical landmark. They can be found in any tissue, but the hard tissues, tooth and bone, are of main interest to the biological anthropologist. This chapter focuses on skeletal variants, whereas the chapter by Scott (this volume) focuses on dental nonmetric variation.

Skeletal nonmetric traits have sometimes been confused with anthroposcopic traits. Anthroposcopic traits are features of shape observable in all skeletons, such as a particular form of the palate or the position and height of the bridge of the nose. These slight variations in shape are often described as discontinuous or ranked traits, but often they can also be measured as continuous variables, for example, nasal height, which may be measured with calipers as well as ranked. In contrast, nonmetric traits are minor skeletal and dental variants that may or may not be present. When nonmetric traits were discovered in the 1800s, they were described as curious anatomical features, not present on all skeletal remains or groups of skeletons. To explain these occasional features, they were assumed to cluster in large or small population groups. Since their discovery, several hundred nonmetric traits have been reported for the skull and infra-cranial skeleton (Hauser and De Stefano, 1989; Ossenberg, 1976; Saunders, 1989). Figures 17.1 to 17.4 show some examples of nonmetric skeletal traits, which were chosen because they are well known in the literature, easy to identify, or well described.

The expression of nonmetric traits does not occur in a vacuum. Rather, hard-tissue development is influenced by the anatomy and development of the surrounding soft-tissue structures (cf. Hall, 1998). In many cases, these variations do not affect the intended function of the bony element. It is rarely possible to look at a living individual and identify the presence of a given trait, although in certain circumstances, some traits may produce clinical symptoms or the presence of the trait may be associated with more severe, genetically based syndromes or other conditions.

Nonmetric traits have several different aliases in the literature (Table 17.1) that often imply subtle differences in definition, particularly with regard to the nature or etiology of the traits. The term minor skeletal variants (e.g., Hauser and De Stefano, 1989) implies
that traits are unlikely to affect the necessary functions of the skeleton and associated soft-tissue. Because the "normal" function is not impaired, they are not considered pathological. This distinction can be useful at methodological and theoretical levels, but it can also be problematic, as discussed below. The terms nonmetric, nonmetrical, and discrete are similar in their meanings, implying that traits cannot be measured in incremental units, as are linear measures of skeletal features (such as cranial length). The term dichotomous means that traits are scored as present or absent, highlighting what is thought to be their discontinuous nature. It is rarely that simple. For example, osseous bridging of a vascular opening such as the supraorbital foramen on the superior orbital rim (Fig. 17.1a,C and D) may be present in differing degrees: absent, trace, partial, and complete. But these categories are not discrete. There is an underlying continuous distribution from absent to complete.

Figure 17.1  (a) Anterior view of skull: A. metopism, B. frontal grooves (bilateral), C. supraorbital foramina, D. supraorbital notch. E. trochlear spur, F. zygomatico-facial foramina, G. infraorbital suture medial from infraorbital rim, H. infraorbital suture from zygomaxillary suture, I. infraorbital foramina, J. Os japonicum (bipartite zygomatic bone) and K. mental foramina (b) Lateral view of skull: J. Os japonicum (bipartite zygomatic bone). K. mental foramina, L. auditory exostosis, M. squamosmastoid suture, N. mastoid foramina, O. occipitomastoid ossicle, P. ossicle at asterion, Q. sutura mendosa, R. parietal notch bone, S. bipartite parietal bone, T. coronal ossicle, U. epipterion bone and V. marginal tubercle (c) Posterior view of skull: O. occipitomastoid ossicle, P. ossicle at asterion, W. parietal foramina, X. sagittal ossicle, Y. ossicle at lambda, Z. lambdoid ossicles (wormian bones) and AA. Os inca.
Consequently, the term *quasi-continuous* was introduced by Grüneberg (1952) to highlight the underlying continuous distribution of the traits while still distinguishing them from metric traits. *Epigenetic variant* contrasts traits that exhibit a Mendelian pattern of inheritance and focuses instead on the polygenic nature of these traits as well as on the possibility of their modification during ontogeny (Hauser and De Stefano, 1989).

The term *atavism* refers more specifically to a throwback, which is a trait that has been retained from a last common ancestor and may have an ongoing evolutionary significance (Ossenberg, 1969). For example, the supratrochlear spur of the humerus (Fig. 17.2a,B), which is occasionally observed in humans, is the normal state in cats and some other carnivorous mammals so its rare presence in humans is thought by some to be controlled by conserved genetic information that becomes manifest in an individual for an unknown reason. Therefore, the supratrochlear spur is a shared ancestral trait among mammals, illustrating how atavistic traits can be used to analyze evolutionary processes. *Nonmetric*, which is a somewhat general
and neutral term is now widely accepted as the most appropriate when referring to groups of traits.

There are hundreds of recorded nonmetric traits, and researchers such as Ossenberg have classified them; however, no standardized system exists. Ossenberg (1969) developed a classification system for cranial nonmetric traits based on tissue development and form, which includes five categories (Table 17.2). Saunders (1989) adopted these categories for infracraniial traits and added the additional categories: spinal variants, prominent bony processes, and facet variations. This classification includes alternative categories for traits. For example, the third trochanter of the femur can be categorized as either “prominent bony processes” or “hyperostotic.” Schwartz’s (1995) system differs in that it is ordered hierarchically, providing subdivisions for the major categories defined by Ossenberg and Saunders (see Table 17.2 for a comparison of these approaches).

Historically, classification began as a description of form. More recently, trait classification has been changed to reflect the common morphogenesis of traits (Hall, 1998). Common to all classification systems are the categories of hyperostotic and hypostotic. Hyperostotic refers to traits with an excess of bone formation, but this definition can incorporate numerous different reasons for why the excess bone formed. The ossification is usually an excess over the nonanomalous condition and can include ossification of soft-tissue structures such as cartilage, ligaments, or dura and accelerated rates of fusion or closure (Ossenberg, 1969; Saunders, 1989). The supratrochlear spur is an example of a hyperostotic trait (Fig. 17.2a,B). Hypostotic traits are those exhibiting incomplete ossification, arrested development, or the retention of the immature or embryonic form (Ossenberg, 1969; Saunders, 1989). Apertures, or holes, can be examples of hypostotic nonmetric traits. The most well-known aperture traits are the septal aperture of the humerus.
(Fig. 17.2a) and the sternal aperture of the body of the sternum.

Trait classification is not necessarily explanatory. For example, extra-sutural ossicles such as wormian bones, lambdoidal ossicles, and others are supernumerary elements acting as additional centers of ossification, but Os Inca (Inca bone) is a hypostotic trait, which is defined as incomplete coalescence of elements (Schwartz, 1995) (Fig. 17.1b and 17.1c). It is therefore likely that the underlying genetic and developmental factors are different for Os Inca and sutural ossicles. Barnes (1994) classifies cervical and lumbar ribs as cases of border shifting, but Schwartz (1995) calls them supernumerary elements with aggressive pre-ossification differentiation. Schwartz (1995:259) classifies sacralization of the fifth lumbar vertebra as an accelerated closure or union trait and describes it as “a reduction in lumbar vertebrae via incorporation into the sacrum.” Alternatively, Barnes (1994) sees this as border shifting, where the borders of development fields have shifted and anatomical morphology typically observed in one field is present in a neighboring field, either cranially or caudally. These two different classifications imply different causes.

In the twentieth century, cranial, infracranial and dental nonmetric traits came to be recognized as having a genetic basis. For this reason, biological anthropologists studying human skeletons believed that the varying frequencies of traits in different samples could
be used to identify biological relationships among past populations. Biological distance refers to the measure of the divergence among populations based on genetic relationships (Buikstra et al., 1990). However, physical traits reflect the phenotype of the individual and not the genotype (they do not have simple, Mendelian patterns of inheritance). It is assumed that populations who share similar frequencies in the occurrence of traits (similar phenotypes) are more closely related genetically than populations with different frequencies. However, this assumes that genes largely control traits. In the late 1970s and early 1980s, it was found that many nonmetric traits could be influenced by other factors, and when some distance analyses produced unexpected patterns of biological relationships, nonmetric trait studies fell out of fashion (see the Historical Review section below). More recently there has been some resurgence in trait study as investigators discover new ways to characterize their distributions and new ways to integrate the study of traits with advances in developmental and evolutionary biology.

This chapter presents a historical approach to the rise and fall of skeletal nonmetric traits and important advancements in our knowledge about these traits.¹ We also discuss issues of how to record traits and the interaction between nonmetric traits and various physiological and external factors such as sex, age, activity, and pathological conditions. Finally, we offer some suggestions for additional research directions that focus on nonmetric traits and their relationship to developmental and population studies.

### TABLE 17.1 Names Used to Describe Nonmetric Traits

<table>
<thead>
<tr>
<th>Name</th>
<th>Reference</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minor skeletal variants</td>
<td>Berry and Searle, 1963</td>
<td></td>
</tr>
<tr>
<td>Nonmetric variants</td>
<td>Cheverud et al., 1979</td>
<td></td>
</tr>
<tr>
<td>Nonmetric traits</td>
<td>Sjovold, 1984</td>
<td></td>
</tr>
<tr>
<td>Nonmetrical variants</td>
<td>Prowse and Lovell, 1996</td>
<td></td>
</tr>
<tr>
<td>Quasi-continuous variants</td>
<td>Grüneberg, 1952</td>
<td></td>
</tr>
<tr>
<td>Discrete traits</td>
<td>Corruccini, 1974</td>
<td></td>
</tr>
<tr>
<td>Discreta</td>
<td>Rösing, 1982</td>
<td></td>
</tr>
<tr>
<td>Discontinuous traits</td>
<td>Ossenberg, 1969</td>
<td></td>
</tr>
<tr>
<td>Anomalies</td>
<td>LeDouble, 1903</td>
<td></td>
</tr>
<tr>
<td>Morphological variants</td>
<td>Hallgrímsson et al., 2004</td>
<td></td>
</tr>
<tr>
<td>Atavisms</td>
<td>LeDouble, 1903</td>
<td></td>
</tr>
<tr>
<td>Epigenetic</td>
<td>Berry, 1963</td>
<td></td>
</tr>
<tr>
<td>polymorphisms</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Epigenetic variants</td>
<td>Berry and Berry, 1967</td>
<td></td>
</tr>
<tr>
<td>Threshold characters</td>
<td>Falconer, 1965</td>
<td></td>
</tr>
</tbody>
</table>

¹For a more in-depth review prior to the 1970s, see Saunders (1978).

### HISTORICAL REVIEW

In the nineteenth and early twentieth centuries, interest in descriptive morphology and embryology flourished, leading to the discovery and description of many nonmetric skeletal traits. This interest was stimulated in part by two popular theories about biological evolution, which were later discredited when the “evolutionary synthesis” was developed. These theories were “the theory of recapitulation” and “the theory of acquired characteristics.” Originally proposed by Ernst Haeckel, the theory of recapitulation stated that during ontogeny, an organism literally passes through the adult forms of its ancestors. The most famous example refers to the apparent gill slits of human embryos that Haeckel interpreted as features of adult ancestral fish pushed back into the early developmental stages of humans by a universal acceleration of developmental rates (Gould, 1977; Mayr, 1982). Recapitulation was an attractive explanation for those human skeletal traits that appeared early in development and resembled constant features of ancestral or more “primitive” adult forms.

The theory of acquired characteristics, an ancient concept made famous by Lamarck, suggested that features acquired by use or disuse and the influence of environmental conditions could become inherited and thus passed on to subsequent generations (Balter, 2000; Landman, 1991). This was an appealing explanation for the presence of some skeletal
<table>
<thead>
<tr>
<th>Category</th>
<th>Reference</th>
<th>Traits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hyperostotic: an excess of ossification into structures that are normally composed of cartilage, ligaments, or dura</td>
<td>Ossenberg (1969)</td>
<td>Squamo-parietal synostosis, marginal foramen of the tympanic plate, pterygo-spinous bridge, pterygo-basal bridge, auditory exostosis (Fig. 17.1b,L) intermediate condylar canal</td>
</tr>
<tr>
<td></td>
<td>Saunders (1989)</td>
<td>Trochlear spur (Fig. 17.1a,E) pterygospinous bridge, clinoclinoid and caroticoclinoid bridging, atlas bridging (posterior and lateral), (Fig. 17.3a and 17.3b) supratrochlear spur of the humerus, (Fig. 17.2a,B) divided hypoglossal canal, ossified apical ligament, jugular canal bridging, cervical transverse foramen double, (Fig. 17.3c,G and 17.3c,H) third trochanter of the femur, peroneal trochlea of the calcaneus</td>
</tr>
<tr>
<td></td>
<td>Schwartz (1995)</td>
<td>1. Division/bridging of foramina/canals: optic canal, hypoglossal (anterior condylar) canal, infraorbital foramen, (Fig. 17.1a,I) jugular foramen, transverse foramen (vertebral), bridging of neural canal of first cervical vertebrae (lateral or posterior) (Fig. 17.3a and 17.3b)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Mineralization of connective tissue/deep fascia: complete supraorbital foramen, (Fig. 17.1a,C) extension of lacrimal hamulus, pterygo-spinous bridge, pterygobasal bridge, clinoid bridge</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. Response to soft tissue: trochlear spur, (Fig. 17.1a,E) extension of vaginal process, distension of maxillary tuberosity, enlarged hamulus, genial tubercles shape and size, precondylar tubercle, third trochanter (femur), peroneal tubercle (calcaneus)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Accelerated closure/union: reduction of lumbar vertebrae via incorporation into sacrum</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5. Excessive bone deposition of unknown origin/etiology: palatal torus, mandibular torus</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6. Excessive bone deposition of apparent (non-soft tissue) known origin/etiology: auditory exostoses (Fig. 17.1b,L)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>7. Supernumerary elements: additional centres of ossification: sutural ossicles (Fig. 17.1b and 17.1c)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8. Supernumerary elements: aggressive preossification differentiation: vertebrae, digits, cervical ribs</td>
</tr>
<tr>
<td>Hypostotic: incomplete ossification or arrested development reflecting the retention of an immature or embryonic stage</td>
<td>Ossenberg (1969)</td>
<td>Metopism, (Fig. 17.1a,A) Os Japonicum, (Fig. 17.1a,J) infraorbital suture, (Fig. 17.1a,G and 17.1a,H) tympanic dehiscence, foramen spinosum communicates with sphenoid petrous fissure</td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Category</th>
<th>Reference</th>
<th>Traits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Saunders (1989)</td>
<td></td>
<td>Tympanic dehiscence, trace of <em>Os Japonicum</em>, infraorbital suture,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Fig. 17.1a,G and 17.1a,H) septal aperture of the humerus, (Fig. 17.2a,A)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>sternal aperture, lumbosacral spina bifida (occulta)</td>
</tr>
<tr>
<td>Schwartz (1995)</td>
<td></td>
<td>1. Incomplete coalescence of elements: cranial sutures metopism, (Fig.17.1a,A) <em>Os Inca</em>, (Fig. 17.1c,AA) infraorbital suture, (Fig. 17.1a,G and 17.1a,H) squamomastoid suture, (Fig. 17.1b,M) <em>Os Japonicum</em> (Fig. 17.1b,J)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Incomplete coalescence of elements: postcranial double occipital condyle, (Fig. 17.3a,C) spina bifida (lumbosacral; occulta or severe), sacral segments (separate)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. Incomplete mineralization of soft tissue precursor: foramen of Huschke (tympanic dehiscence), septal aperture, (Fig. 17.2a,A) infraorbital groove length, enlarged foramen lacerum, incomplete carotid canal, emarginate/ (vastus) notched patella (Fig. 17.3d)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Incomplete preossification development: sternal aperture, pharyngeal fossa, anterior facial cleft</td>
</tr>
<tr>
<td>Supernumerary vault sutures</td>
<td>Ossenberg (1969)</td>
<td><em>Os Inca</em>, (Fig. 17.1c,AA) parietal notch bone, (Fig. 17.1b,R) lambdoid wormians, (Fig. 17.1c,Z) occipito-mastoid wormians, (Fig. 17.1c,O) epiteric bones (Fig. 17.1b,U)</td>
</tr>
<tr>
<td></td>
<td>Saunders (1989)</td>
<td>Lambdoid wormian bones, (Fig. 17.1c,Z) pterionic ossicle, parieta l notch ossicle, (Fig. 17.1b,R) asterionic ossicle (Fig. 17.1b,P)</td>
</tr>
<tr>
<td>Foramina, canals, and grooves</td>
<td>Ossenberg (1969)</td>
<td>Parietal foramen, (Fig. 17.1c,W) post condylar canal, foramen of Vesalius, parieta l process of temporal squama, accessory optic canal, supraorbital foramen, (Fig. 17.1a,C) supratrochlear foramen, frontal grooves, (Fig. 17.1a,B) foramen in the lateral pterygoid plate, accessory mental foramen (Fig. 17.1a,K)</td>
</tr>
<tr>
<td>blood vessels and nerves</td>
<td>Saunders (1989)</td>
<td>Supraorbital notch or foramen, (Fig. 17.1a,C and 17.1a,D) frontal grooves, (Fig. 17.1a,B) accessory optic canal, zygomatico facial foramen (absent or number present) (Fig. 17.1a,F) supraclavicular nerves pierce the clavicle</td>
</tr>
<tr>
<td></td>
<td>Schwartz (1995)</td>
<td>1. Absent: posterior ethmoid, mastoid, (Fig. 17.1a,N) foramen spinosum, zygomaticofacial (Fig. 17.1a,F)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Atypically present: parieta l, (Fig. 17.1c,O) zygomatico facial, (Fig. 17.1a,F) supraspinous (scapula), foramen in clavicle (for supraclavicular nerve)</td>
</tr>
</tbody>
</table>

(Continued)
features such as joint facet variations or accessory joint facets where it seemed that the variants might be produced during an individual’s lifetime and then passed on to the offspring. In addition, several of these early studies used traits to support notions of racial hierarchy. For example, some traits considered to be throwbacks were said to be more prevalent in the “lower” races (Finnegan and McGuire, 1979).

The rediscovery of Mendel’s principles of inheritance and of the human blood groups in the early twentieth century somewhat eclipsed the use of skeletal morphological traits to describe population variation, although blood group studies could only be done in living populations. In these early days of genetics, a controversy developed between the Mendelians and the biometricians. The Mendelian geneticists, recognizing the particulate nature of inheritance, stressed discontinuous variation as the important material of evolution, whereas the biometricians argued that continuous variation was the only relevant material of natural selection, but they assumed inheritance to be blending (Mayr, 1982). Later, advances in evolutionary genetics were able to show that there is no conflict among particulate inheritance, continuous variation, and natural selection, but...
until that time skeletal nonmetric traits were in a somewhat ambiguous position. The biometricians also used the human skeleton widely for their arguments for the importance of continuous variation, producing broad metrical descriptions of the human skull and long bones (cf. Pearson and Bell, 1919).

In the early twentieth century, LeDouble, an anatomist, published a series of papers in which he described in great detail the form and variation (a description of traits) found on the cranium, face, and spine (LeDouble, 1903, 1906, 1912). In the 1930s and 1940s, several other researchers produced a series of lengthy articles on either individual traits found on specific bones or general morphological variation of the scapula, humerus, femur, spine, and cranium (Akabori, 1934; Hrdlička, 1916, 1932, 1934, 1935, 1937, 1942a, 1942b, 1942c; Wood-Jones, 1930a, 1930b, 1930c, 1933; Odgers, 1931; Trotter, 1934; 1936/1937; 1937). Recognition that traits could vary across populations was limited, although some of these authors emphasized that traits could be used for race identification. Two other important studies appeared in this formative period that, among other investigations, reported the frequencies of a variety of traits within a population rather than taking an anatomical approach, emphasizing variation in specific parts of the skeleton. The first of these was the influential study by Earnest Hooton (1930) of the skeletal remains of people from the Pecos Pueblo site in New Mexico, and the second was that by Charles E. Snow (1948), who examined the large sample of burials from the Indian Knoll Archaic period mound in Kentucky, which spanned several millennia. Hooton observed many morphological traits of both the cranial and the infracranial skeleton, although he combined anthroposcopic and nonmetric traits. He was one of the first to sort his observations by degree of expression, sex, and time period and provide detailed statistical tables of the results.

Until the 1950s, biological anthropology consisted mainly of collecting, categorizing, and comparing data to fit into fixed typological classifications. Avenues of research were limited by the lack of variety in theory. In 1951, Sherwood Washburn presented “The New Physical Anthropology,” which is a contribution that would influence the field for the next 20 years. Reflecting parallel developments in evolutionary biology, Washburn stressed that researchers needed to focus on hypothesis testing and on explanations of the processes and mechanisms involved in evolutionary change and the appearance of adaptations.

The incorporation of genetics (i.e., an understanding of the mechanisms of evolution) into the study of modern human variation and population relationships necessarily led to the rejection of (but not complete dissolution) of traditional typological classifications, particularly with regard to human races (Armelagos et al., 1982). The division of a breeding population into a series of fixed racial types was unjustified and irreconcilable with evolutionary genetics and population history. Since typological classification focuses on the average and ignores variation, it was completely counter to the evolutionary perspective that focuses on variation and sees the average as a nonexistent abstraction (Mayr, 1959).

Washburn’s theoretical foreshadowing helped prepare the way for interest on the part of biological anthropologists in the genetic studies of mice in the 1950s and 1960s. The mouse studies were a series of articles published in The Journal of Genetics aimed at searching for evidence of genetic inheritance of skeletal and dental anomalies in convenient laboratory subjects. In 1952, Grünberg established the link between absence of the third molar in mice with the size of the tooth germ, the underlying genetic constitution. He showed that there is no simple genetic interpretation to explain nonmetric traits. The distribution of traits in populations is described as quasi-continuous; they have an underlying continuous distribution that is rendered discontinuous by a limiting threshold effect. The threshold can be altered by developmental or environmental factors
that influence the continuous distribution, thus altering the liability of the population to show the trait. For example, mice with large litters will show higher frequencies of missing third molars since the offspring compete for womb space and are smaller in size. In this case, the tooth germs in the smallest mice cannot reach the critical threshold for tooth development (Falconer, 1965; Grünberg, 1952). Thus, womb space and litter size are physiologic factors that will influence the position of the threshold for tooth bud formation.

Investigators later realized that the so-called pure genetic lines of mice they were using were not pure. Others (Deol and Truslove, 1957; Searle, 1954a, 1954b) found that several environmental factors such as diet and maternal environment could alter trait frequencies in offspring generations. Later, Howe and Parsons (1967) looked at the effects of environmental factors on their samples in a different way. They calculated measures of divergence in three inbred strains of mice using the frequencies of 25 nonmetric traits. Before calculating divergence, they separated their groups by several factors such as sex, parity, maternal age, litter size, and age at death. Their results showed that the physiologic and environmental factors had no significant overall effect on the calculated distances when information was combined from all traits, but some traits, when taken individually, were significantly affected, thus suggesting that biological distance analyses using nonmetric traits offer useful genetic information when a large number of traits are employed.²

Two important population studies of human samples appeared in the 1950s. Laughlin and Jorgensen (1956) examined the frequencies of eight cranial nonmetric traits in skeletal samples of Greenland Eskimos, looking for evidence of regional differentiation. Since it is known historically that the native inhabitants of Greenland entered in one area and spread coastally in two opposite directions, the authors hypothesized that a study of divergence should show the greatest population distances between terminal populations of the two lines of migration. Results met expectations; the greatest biological distances were found between the northeast and southeast terminal isolates, suggesting nonmetric traits were very useful for exploring regional differences. A few years later, Brothwell (1959) published a comparison of distances calculated from nonmetric and metric traits for 14 widely dispersed worldwide populations. Although the calculated distances from the two types of traits gave differing results, both patterns made some interpretable sense. He concluded that prospects for the use of nonmetric traits in distance studies were tentatively optimistic, leading him to call for additional work in this area.

In 1967, Berry and Berry published an influential report that stimulated the debate over the advantages of nonmetric traits versus metric ones for studies of humans. The article was important because it provided a concise description of some 30 cranial nonmetric traits, compared widely diverse populations, and claimed that nonmetric traits, unlike metric ones, were unaffected by age, sex, side occurrence and other variables and, hence, better at discriminating genetic relationships. Although these claims were insufficiently supported for human samples (and had been debated for the mouse studies), this article is credited as providing the impetus for subsequent population studies using nonmetric traits.

The 1970s was an important decade for nonmetric trait studies even though there was a decline in the frequency of biological distance studies overall (Buikstra et al., 1990). Important nonmetric trait analyses were still being published in journals and as doctoral dissertations (cf. Buikstra, 1972; Finneghan, 1972, 1978; Ossenberg, 1969; Saunders, 1978) and the debate over their suitability for population comparisons polarized (cf. Corruccini, 1974). But by the end of the decade concerns over the degree to which nongenetic factors could affect the presence of traits in human populations surfaced, and many researchers and

²For a comprehensive review of the mouse studies, see Saunders (1978, 1989) and Wood (1997).
their students turned their focus toward palaeopathology and palaeodemography as interest in the health of past populations grew (Buikstra et al., 1990; Mays, 2000).

Along with debate regarding the genetic nature of nonmetric traits was the discussion of the nature of the skeletal samples used in these analyses, particularly what constitutes “the group,” and whether the group is an equivalent to, or adequate representation of, the population. It was important to verify whether archaeological samples could be treated as biologically related groups. A paper published by Cadien et al. in 1974 noted that skeletal remains recovered from cemeteries are not necessarily contemporaneous. Rather, skeletal populations are samples of lineages, introducing a temporal aspect to the possible environmental influences on nonmetric traits and implying that microevolution can affect the hetero- or homogeneity of a skeletal population that can then, in turn, affect the results when comparing one lineage with another. Temporal sequencing to such a fine resolution is not possible with nondocumented archaeological collections. Although not exactly extinguishing biological distance studies, this paper encouraged researchers to appraise the composition of their population samples more carefully.

The underlying argument for using nonmetric traits to determine familial relationships assumes that related individuals are more likely to possess similar patterns of variants than individuals who are not related (Sjøvold, 1976–1977). This assumption also implies that individuals buried close together are more likely to be related to one another. Lane and Sublett (1972) realized this possibility and took the opportunity to examine the frequency of nonmetric traits among males and females from several Iroquoian cemeteries (Allegheny Seneca cemeteries in Pennsylvania that were relocated after construction of a dam), hoping that they might determine residential and marriage patterns through the use of phenotypic variance analysis. They reasoned that it might be possible to detect clusters of nonmetric traits (i.e., groups of related individuals) if there were closer male–male genetic relationships or closer female–female relationships. For example, if male–male relationships were closer, it could be reasoned that females “married in” to the group. Based on the distribution of 33 cranial nonmetric traits in both sexes from five cemeteries, they calculated distances using the statistic, Smith’s mean measure of divergence (Berry and Berry, 1967) and found greater osteological similarities among females than among males. This finding is consistent with the expectation that female–female (i.e., uxorilocal) inheritance patterns were preserving female relationships over the generations, whereas males moved into the group from outside the immediate area.

A short time later, Spence (1974) examined local relationships within the large precontact urban center of Teotihuacán in Mexico. The subgroup he examined was defined by mortuary patterning within a major house structure and believed to be a well-defined corporate group. These corporate groups were likely based on kin groups (Spence, 1974). Nonmetric traits of the cranial and infracranial skeleton were used to test the hypothesis that within-group similarity implied decreased within-group variability. The sample consisted of 16 males and 16 females dated to one particular phase. Comparisons were made by calculating the ratios of the number of positive matches for traits in each individual by individual comparison (number of times the two individuals share a trait) to the number of potential matches (positive matches plus nonmatches) (Spence, 1974). The results indicated that the men were more closely related to one another than were the women, indicating mainly a male–male or virilocal residence pattern. Avunculocality (residency with mother’s brother) was discounted because this residential pattern was not observed in Mesoamerica ethnohistorically. However, six male and six female skeletons were recovered from one part of the structure that did not adhere to the pattern of postnuptial residence with the husband’s group. The females here exhibited more homogeneity than the males, indicating that some females did
remain in residence after marriage while their partners were from outside the apartment compound. This finding could indicate some degree of flexibility in residence patterns permitted by the society (Spence, 1974). In 1994, Spence reevaluated the skeletal evidence using a different statistical approach. Results confirmed the complicated postmarital residence patterns and migratory nature of Teotihuacan inhabitants (Spence, 1994).

These studies were two of the first attempts at adapting methods of biological distance measures from intergroup to intragroup analyses, particularly to the level of the family or extended family. Later, Sjøvold (1977) outlined a specific statistical procedure for establishing familial patterning based on this spatial assumption, i.e., that related individuals and burial proximity are correlated positively. He suggested using a reference population that differs in time or space from the target sample as a baseline for comparison. If one looks for traits that are usually rare, accumulated genetic effects controlling for those traits will be unexpected in any given population but expected in closely related individuals if the traits appear. However, the occurrence of rare traits is population specific. It is advisable then to calculate the probability that any given individual would exhibit a rare trait as the ratio of the frequency of the trait in the pooled reference and sample populations times overall population size (total “present” and “absent” observations) (Sjøvold, 1976–1977). This method may be used to test individuals suspected of being related, or it may tease out related individuals from a pooled sample. However, it is not possible to determine the internal familial structure based on minor skeletal variants (using pedigree analysis) since no traits are known to follow Mendelian patterns of inheritance.3

Several other researchers later explored the possibility of detecting familial segregation, kin groups, or socially stratified ethnic groups within cemeteries using nonmetric traits alone or in combination with metric variables (Bartel, 1980; Bondioli et al., 1986; Corruccini et al., 1982; Crubézy, 1998; Ortner and Corruccini, 1976; Owsley and Jantz, 1978; Strouhal and Jungwirth, 1979), but the number of non-metric trait studies generally decreased in the 1980s, affected in part by studies that directly investigated genetic control of nonmetric traits using known samples. For example, the skeletons of deceased rhesus macaques from a captive colony on the Caribbean island of Cayo Santiago (Cheverud and Buikstra, 1981a, 1981b; McGrath et al., 1984) provided information on heritability of several cranial traits. Skeletons of animals with known matrilineal relationships were scored for 14 traits, and the methods of quantitative genetics (Falconer, 1965) were used to calculate single trait heritabilities. Estimates ranged from 0 to 1, but half of them were greater than 0.5, indicating a significant amount of additive genetic variation for the traits. Another comparison of the relative heritabilities of nonmetric and metric traits in the same animals found the hyperostotic nonmetric traits to display significantly greater estimates than the metric traits. However, it must be remembered that heritability estimates are population specific. These results can only be claimed for the Cayo Santiago macaques.

In the town of Hallstatt, Austria, a unique opportunity emerged to test the validity of nonmetric traits for examining familial relationships. To create space for new burials, previously interred burials were exhumed and the long bones and skulls were retained, elaborately decorated with the individual’s name, and placed in the cemetery’s charnel house.4 This practice continued for at least 200 years, resulting in a collection of over 700 identified

3 Around this time, Finnegan and McGuire (1979) published a method for assigning individuals to biological groups using nonmetric traits for use in a forensic context. They tested the reliability of statistical tests to classify skeletons empirically in particular groups with the lowest rate of error (Finnegan and McGuire, 1979).

4 Surviving relatives painted the skulls of their ancestors with flowers and other images, including the names and death dates of the deceased.
skulls. Ninety-one pedigrees with two to ten family members were reconstructed through parish registries (Sjøvold, 1984, 1995). Sjøvold then studied the heritabilities of 30 non-metric cranial traits by comparing first-degree relatives. The results showed that almost half of the sample of traits had high heritabilities. Sjøvold reiterated the point made by Cheverud and Buikstra (1981a, 1981b) that heritabilities are population specific. In addition, this approach does not rule out completely the interpretation that relatives could be sharing traits that are produced by a common family environment.

By the end of the 1990s, supporting evidence showed that skeletal and dental nonmetric traits could distinguish biological groups on a regional scale and contribute to discussions of regional migrations (Blom et al., 1998; Christensen, 1998; Johnson and Lovell, 1995; Pardeo, 1991; Prowse and Lovell, 1995). It also became clear that nonmetric traits could be useful in family studies, by showing that several known or suspected family members shared trait expressions (Alt et al., 1997; Spence, 1996). Others used nonmetric trait frequencies to test whether kin groups (matrilineal or patrilineal for example) could be distinguished (Gao and Lee, 1993; Prowse and Lovell, 1996), with varying degrees of success. In 2000, Brothwell stated that the potential for using nonmetric traits was far from realized. The illustration of their feasibility in family studies coupled with the increasing rate of reburial of skeletal collections was an impetus for additional interest in nonmetric trait studies. Ansorge (2001:104) also described a “renaissance in [the] use of morphology” for assessing population relationships in the field of zoology. Others (Stojanowski and Buikstra, 2005) reported an increase in distance analyses (both using metric and nonmetric traits) at this time, and studies of biological relationships within cultural and geographic regions continued (e.g., Cybulski, 2001), as did anecdotal reporting of unusual traits (Usher and Christensen, 2000) and attempts to investigate individual traits using radiography (Bouille, 2001). But there were still critiques of nonmetric studies (cf. Tyrrell, 2000).

The increased population history studies shifted some focus away from group categorization using trait frequencies to more in-depth analyses of variation within and between populations. There was a return to worldwide population studies without the stigma of racism. In one example, Donlon (2000) examined infracranial nonmetric variation in populations from Australia, Africa, East Asia, Europe, and Oceania. Results were compared with population relationships produced using cranial measurements and genetic markers, but the results were not easy to interpret except that there were a small number of nonmetric traits contributing to a separation of the populations based on biological distance. Donlon concluded that the relationship between within-sample variation and between-sample variation is too complex and that these traits would optimally be used with regional or locally related populations rather with than large, unrelated ones.

On the other hand, Hanihara et al. (2003) looked at the characterization of worldwide biological diversity using cranial nonmetric traits. Their sample consisted of over 8000 individuals from different continents around the world. The results supported the expected geographical distinctiveness of populations, but they also showed a gradual but continuous change in trait frequencies across geographic ranges. Several populations were morphologically distinctive, including the Ainu of Japan and Pacific peoples (Oceania). These explorations of variation have provided a theoretical framework for exploring more interesting anthropological questions, within the geographic boundaries of specific regions. A recent example of the incorporation of questions of population history, migration, and admixture is the study by Hallgrímsson et al. (2004). They studied Icelandic populations and considered whether cranial nonmetric traits could show what historical admixture had occurred and the contributions of different parental populations to Iceland through migration and settlement. Their results showed that several different
parental populations likely contributed to the Icelandic population, but the major contributors were the Norwegians (Hallgrímsson et al., 2004).

**ANALYTICAL ISSUES, PHYSIOLOGIC FACTORS, AND NONGENETIC/ENVIRONMENTAL EFFECTS**

The threshold effect articulated by Grüneberg in 1952 is essentially the acknowledgment that it is the integration of environmental influences with the genotype that results in the phenotype. This model is still the basis of nonmetric trait analysis today and has provided the theoretical opportunity to explore possible interrelationships and effects of physiologic factors as well as nongenetic or environmental factors on nonmetric trait expression and biological distance analyses. So how do physiologic factors such as age, sex, and asymmetry, and environmental factors such as activity, socioeconomic status, and pathology influence trait expression in the human skeleton? How do we know which nongenetic or environmental factor is affecting a trait, and is it possible to tell whether there is more than one factor? Is it possible that different factors could produce the same effect? What should be apparent is the interrelated nature of these factors. Biological sex is altered occasionally by chromosomal or genetic anomalies, secondary sex characteristics vary depending on hormones and nutrition, and both of these are influenced by the cultural formation of gender roles. All of these together will influence activity patterns, labor roles, and migration patterns. Aging is the outcome of development; therefore, developmental defects and pathological conditions must be considered. Furthermore, cultural and genetic factors can influence the occurrence, prevalence, and distribution of pathological conditions.

**Sex and Nonmetric Traits**

Nonmetric traits can show a sex bias for both genetic and environmental reasons. Genetically, since human males are the heterogametic sex, carrying one X and one Y chromosome while females have two X chromosomes, there may be effects from sex-linked genes that influence the threshold potential for certain nonmetric traits. No explicit examples have been demonstrated although it has been shown that both extra X and Y chromosomes can influence dentine production and enamel thickness of teeth (Alvesalo, 1997), and this might affect the appearance of accessory dental traits on cusps. During development, the production of gonadal steroid hormones, principally testosterone and estradiol, will contribute to sex differences in the secondary sex characteristics that may influence whether one sex or another is more likely to develop a particular trait. This is reflected mainly in size differences and differences in bone deposition. The outcome of growth and puberty is that males on average are larger and their bones are more robust because they deposit relatively greater amounts of bone during the growth spurt. Therefore, they tend to show higher frequencies of hyperostotic traits, whereas females will have more hypostotic traits.

Biological sex is also culturally interpreted into gender roles that, in situations where the roles are strictly defined, can create different physical environments for males and females. These cultural influences may enhance sex differences or diminish them so that we will not always see expected sex differences in trait frequencies. For example, we expect that males should show more hyperostotic traits and that females will show more hypostatic traits. Thus, there should be a correlation between trait presence and measures of sexual dimorphism or robusticity. However, this has not always been found (Saunders, 1989; Sjøvold, 1977).

Sex differences in trait frequencies within population samples need to be addressed methodologically before a researcher can investigate comparisons between populations. One method is avoidance; significantly sex-associated traits can be discarded from the analysis. A second method is to divide and conquer;
compare males and females separately rather than amalgamating them (this approach may be the ultimate goal if the intention is to look for sex differences in variability within a sample when searching for the effects of kinship patterns). This method essentially doubles the number of statistical tests, which means that interpreting the results can be more complicated. Finnegan (1978) offered the suggestion that, in correcting for sex differences, the ratio of males to females should be equal and dimorphic traits should be removed. Donlon (2000) found that hypostotic traits were associated more frequently with females and hyperostotic traits with males; however, there was no statistically significant association between type of trait and bone size. Contrary to expectations, hypostotic traits were not generally associated with smaller bone size nor were hyperostotic traits associated with larger bone size. It is possible that the particular measurements selected were not those associated with trait occurrence. It is also possible that these traits are sex-linked. The hypostotic traits associated with females may be x-linked, whereas the hyperostotic traits associated with males may be y-linked. Deposition of bone in hyperostotic traits and bone size may not necessarily be under the same genetic control. It should also be emphasized that a trait that is dimorphic in one population is not necessarily dimorphic in another population.

Age and Nonmetric Traits

The form of bone, and therefore the expression of nonmetric traits, cannot be considered without looking at developmental changes in trait appearance. Although bone continues to change and remodel over an individual’s entire lifetime, the major osteological changes are complete by the time an individual becomes a sexually mature adult. Traits that are characterized by lack of fusion such as metopism and the unfused acromion of the scapula are not defined as traits until the age of normal development for that feature has passed. Hyperostotic traits are thought to be more likely to be expressed or stronger in expression in older individuals, but again, this is not universally true (i.e., it is population specific) (Saunders and Popovich, 1978).

The issue facing an observer when considering the effects of age on nonmetric traits can be summarized as follows: Has sufficient time passed in the individual’s development to allow the trait to be expressed? Essentially, is the degree of expression of a trait related to developmental stages—do younger individuals express “weaker” variations of a trait because they have not had sufficient time (i.e., they died young) to develop a stronger form of the trait, or would development have ceased at a particular stage because of the genotype? However, this assumes that each expression is a developmental stage leading to the final, most strongly expressed state. The validity of this assumption must be tested for each trait. A general age trend does exist, however, when looking at categories of traits. Hypostotic traits are generally more common in younger individuals, whereas hyperostotic traits are more common in older individuals (Molto 1983; Ossenberg, 1969; Saunders, 1978; Winder, 1981).

Some traits can be “age stable,” which has been shown by the presence of some traits in fetal and juvenile skeletal remains. For example, Bunning (1964) studied 64 calcanei from 32 fresh skeletons between 6 and 9 months of age and found all forms of squatting facets (Fig. 17.4b). Others have studied juvenile skeletons to see whether traits are present before ossification of the skeleton is complete and by this means have identified age-stable traits (cf. Buikstra, 1976; Molto, 1983; Ossenberg, 1969; Saunders, 1978). Scheuer and Black’s (2000) Developmental Juvenile Osteology provides useful illustrations and explanations of how each bone develops in utero to completion of ossification. Although the developmental descriptions do not focus specifically on nonmetric traits, many traits are included in the anatomical descriptions as examples of variation originating during development (see also Barnes, 1994; Hauser and De Stefano, 1989).
To control for age effects, it is possible, when choosing a skeletal sample, to select only adults (osteological fusion should be complete) and to keep the age distribution of the samples similar. It must also be remembered that a later age of onset does not negate the possibility that a trait is controlled by genetic factors. As with many diseases of complex inheritance, certain genetic influences may not come into play until later stages of development or even adulthood are reached. In the same fashion, it is possible that the formation of some traits in some individuals may be influenced directly by environmental factors affecting bone deposition or resorption at a later stage in life.

**Activity and Nonmetric Traits**

Much of the debate as to the causation of nonmetric traits surrounds the degree to which activity affects morphology. In most cases the infracranial skeletal is subjected to a wider range and stronger degree of activity-induced stress than the cranium, making the infracranial skeleton more susceptible to changes in form caused by activity.

In 1999, Capasso, Kennedy, and Wilczak published the *Atlas of Occupational Markers on Human Remains* as a monograph of the *Journal of Paleontology*. Surveying the literature, they report that a wide range of traits are produced by certain activity patterns. Each trait description provides information on anatomy, stress factors, occupational activities, and a list of bibliographic references. Although the authors do not attribute the causes of all nonmetric traits to specific activities, the purpose of this atlas is to compile evidence for an activity-related component in their etiology. If genetic factors have been demonstrated in previous studies, this is acknowledged in the section on “stress factors.” Sometimes, the confusion over trait etiology may simply be attributable to the vocabulary used. A reference to variations in articular facets implies that abnormal weight bearing or unusual loading produces these features, but this interpretation may be based only on post hoc associations. It is important to explore the developmental basis of trait formation to help clarify these questions. Is the activity the cause of the formation of the trait, or has the activity provided an environment in which the liability for expressing the trait has been pushed over the developmental threshold? (Falconer, 1965, Grünberg, 1952).

An interesting example of this dilemma is illustrated by the extended debate that took place in the mid-twentieth century over whether the septal aperture of the humerus was produced by remodeling activity coming from the effects of the olecranon process of the ulna, its position in the olecranon fossa, and the possibility of hyperextension at the elbow joint (Benfer and McKern, 1966; Glanville, 1967; Hrdlička, 1932; Trotter, 1934). However, when researchers could not find clear relationships between presence of the fossa and various measures of bone size and strength and also recognized that the trait (like other hypostotic traits) was more common in small females, this interpretation was discarded.

In another case, Stirland (2000) studied a large skeletal sample recovered from Henry the VIIIth’s flagship the *Mary Rose*, which had been sunk in Portsmouth harbor in the sixteenth century. She found several examples of unfused acromial epiphyses of the scapula in this all male sample. She hypothesized that this trait, the failure of the epiphysis to fuse, represented the effects of extended use of the longbow. Many men on the ship had been professional archers, serving in Henry’s army. However, the acromion normally fuses between 18 and 20 years (Scheuer and Black, 2000:268), and so only individuals above the maximum age for fusion could be treated as a proper sample for evaluation of the trait. Even if longbow use during adolescence put excessive strain on the scapular acromion, how do we know whether those individuals already at risk for developing the trait were not the ones more likely to express it?

Several nineteenth- and early twentieth-century authors have remarked on the effects of
cranial modification through mechanisms such as cradle boarding on the expression of cranial nonmetric traits (Adis-Castro and Neumann, 1948; Comas, 1942; Dorsey, 1897; Hrdlička, 1935; Montague, 1937; Oetteking, 1930; Sullivan, 1922). Sullivan (1922) stated that cranial modification precludes the use of craniometrics in distance analyses, for obvious reasons. In 1970, Ossenberg compared artificially modified crania from a single Hopewell mound to 1200 unmodified skulls representing 21 Native North American populations (Ossenberg, 1970). Twenty-eight cranial traits were surveyed, and frequencies were calculated. She found a statistically significant increase in sutural ossicles in the regions affected by modification, which was interpreted to mean that modification inhibits the normal growth rate, producing stress, and the stress encourages the formation of ossicles. She also found a tendency for other groups of traits (foramina, canals, grooves) to show either a hypostotic or hyperostotic effect in response to modification; that is, the suite of traits in an affected region showed similar bony responses, either with the reduction (hypostosis) or addition (hyperostosis) of bone. The stress in this situation is cradle boarding, which affects circulation and which, in turn, affects the development of bone through the constriction or redistribution of blood vessels and nerves.

Later work focused largely on extra-sutural ossicles. In one experimental study, Pucciarelli (1974) modified the crania of rats to observe the effects on the production of extra-sutural ossicles. The frequency of ossicles was higher in the group with modified crania, and differences were also noted in the size and shape of the ossicles. El-Najjar and Dawson (1977) examined the occurrence of wormian bones of the lambdoid suture in 200 crania, 120 adult crania from three Southwestern Pueblo Indian skeletal samples, and 80 modern fetal skulls. Two types of cranial modification were analyzed: vertical occipital and lambdoid flattening. By looking at the frequencies and locations of ossicles in the modified skulls and comparing these with the frequencies in the nonmodified adult crania and fetal crania, their results supported the hypothesis that ossicles are affected by stress, but that there must be an underlying genetic component. Later, White (1996) looked at 46 modified and 82 nonmodified crania of adults from Lamanai, a large Maya ceremonial center in Northern Belize. The modification exhibited was anteroposterior, characterized by frontal and occipital flattening, creating bulges in the parietals and a depression along the sagittal suture. Her findings show that sagittal synostosis (premature fusion of the sagittal suture) is highly correlated with cranial modification, and this is interpreted as a response to localized mechanical and tensile forces exerted on the sagittal suture by the cranial modification.

Konigsberg et al. (1993) looked at the effects of different kinds of cranial modification on nonmetric traits, and whether these modifications affect biological distance studies performed on the affected crania. They studied 447 crania from the Hopi, Nootka, Kwakiutl, and one prehistoric Peruvian group. Their results showed that there were localized effects of modification on the frequencies of the traits, but that in most cases, these did not affect the biological distance analyses. They did find that coronal ossicles increased in relative frequency with all types of modification, but other traits, for example the open foramen spinosum and foramen of Huschke (tympanic dehiscence), only changed with one type of modification (Konigsberg et al., 1993). Their research supports the contention that nonmetric traits are partially heritable, and that environmental stress has a moderate proximal effect on these traits; growth is redirected locally. Additionally, traits formed prenatally (e.g., accessory infraorbital foramen, the divided hypoglossal canal, the posterior condylar canal, and the accessory lesser palatine foramen) are not affected by artificial modification since they are formed before the modification can affect their presence or absence.
Pathological Conditions and Nonmetric Traits

Distinguishing Between Pathological Conditions and Nonmetric Traits. An anatomical variant is pathological if normal function is impaired. Many pathological conditions have a genetic component and are therefore useful in nonmetric trait analyses; the pathological nature of a trait does not negate its use in biological distance studies. On the other hand, many nonmetric traits are asymptomatic but are included in palaeopathology texts, either because these variants are considered pathological by the authors or to broaden awareness of nonpathological variation (cf. Auferheide and Rodriguez-Martin, 1998; Mann and Murphy, 1990). Mays (1998) considers nonmetric traits separate from pathological lesions in that the former are asymptomatic; an individual may live his or her entire life without knowing the trait is present. However, radiological studies on living individuals and correlations between symptoms such as pain and the presence of a particular trait are uncommon (cf. Cockshott, 1979; Tilley, 1970).

Clarification in terminology can provide a new analytical perspective in understanding the etiology, ontogeny, and development of nonmetric traits. The issue with terminology goes beyond semantics, particularly because inherent in the definitions of the terms is some implication of etiology. From the perspective of pathology, two influential articles from the Journal of Pediatrics have attempted to classify morphological pathological conditions according to etiology, particularly distinguishing between those conditions resulting from clear genetic abnormalities and those resulting from congenital factors (present at birth but not necessarily genetically controlled) (cf. Smith 1975; Spranger et al., 1982).

In 1994, Barnes published a monograph entitled Developmental Defects of the Axial Skeleton in Palaeopathology in which morphological variations observed in the axial skeleton were described and analyzed. These variants were considered to be developmental field defects. Developmental fields are defined as “the close embryonic interaction of select developing tissues involved in the complex composition of a specific structure or set of closely related structures” (Barnes, 1994:6). She discusses the morphogenesis of the axial skeleton and describes developmental defects observed in skeletal remains from a large prehistoric collection from New Mexico. Differences in frequencies among culturally defined groups were found, although no other statistical measures of biological affinity were performed. Nevertheless, the study is useful in defining and interpreting field defects as well as in illustrating their usefulness in separating local cultural groups. One difficulty, however, is the use of the terms “developmental” and “congenital.” The two terms are not necessarily mutually exclusive, but both imply extrinsic causes (even in utero) even though intrinsic or genetic factors may play a role. Many traits Barnes (1994) discusses do have a heritable component, but it is the palaeopathological aspect that is emphasized rather than the potential for biological relatedness. Recently, Hallgrimsson et al. (2005) used the term, dysmorphology, to articulate morphological representations of the threshold effect. A given phenotype is expressed along a distribution, and the dysmorphology is at the extreme of the distribution.

Effects of Pathological Conditions on Trait Expression. Few studies focus on how pathological conditions specifically affect the presence or expression of nonmetric traits. Typically, it is the pathological condition that is of interest rather than the implications for biological distance comparisons. In addition, there are rarely adequate sample sizes for analyses to be anything more than anecdotal. Although these anecdotal examples are important, it is very difficult to see whether a condition affects a given trait in a consistent manner. It is usually much easier and analytically more sound to remove the affected individual or trait from the study.
The influence of a pathological condition on a nonmetric trait depends on the pathological condition in question and on the particular trait. Localized disorders could simply obliterate an anatomical area and, therefore, the trait as well. For example, severe osteoarthritis of the hip joint may obliterate several femoral neck and hip joint features (Fig. 17.2b). In these cases, the trait would have to be scored as unobservable. When the pathological condition is systemic, however, the results could be quite different rendering many traits visible but perhaps hard to interpret. Systemic disorders affect the entire body, including the skeletal system. These could include metabolic disorders such as rickets and osteomalacia where the composition of the bone is changed. If a condition is clearly genetic or congenital, it will affect morphogenesis from an early age. If the condition is acquired during an individual’s lifetime, it could affect the skeletal system through remodeling. A systemic developmental approach (i.e., looking at soft-tissue structures, functional structures, and total physiology of the area of a given trait) is necessary to understand what is being compared.

There are two studies that have examined the relationship between nonmetric traits and some indicator of quality of life, such as a pathological condition or a more general socioeconomic factor. Bocquet-Appel (1984) compared wheat prices as a proxy for socioeconomic status, or general nutrition/stress to the frequencies of nonmetric traits in cranial remains from a nineteenth to early twentieth century sample from Coimbra, Portugal. He found that the presence of some nonmetric traits (sutural bones and some foraminal traits) did correspond with changing prices of wheat. However, the crania were separated into 20-year age groups so it is possible that the change in nonmetric trait frequencies were the result of secular trends rather than the nutritional consequences of changing wheat prices.

Bergman (1993) used cribra orbitalia and grave equipment as proxies for quality of life in a series of early medieval female crania from Milicz, Poland. The sample was divided into the “haves” and the “have nots,” defined by the presence or absence of cribra orbitalia in each individual. He proposed that the presence of the condition indicated an individual who lived in worse conditions on average. These findings were then correlated with the presence or absence of grave equipment, taken as a separate indicator of standard of living. The premise was that cribra orbitalia provided a biological, organism-dependent source of information on living conditions, whereas the grave equipment was an external source. Cribra orbitalia was found more often in individuals buried without grave equipment, “confirming the opinion that the presence or absence of grave equipment provides information on the discrepancies in living conditions” (Bergman, 1993:68). Bergman (1993) then assessed the relationship among cribra orbitalia, grave goods, and 32 cranial nonmetric traits. Only one trait, ossicle in the lambdoid suture, was more likely to be found in individuals interred without grave goods. No other traits exhibited an association with grave goods. However, this trait, along with 11 others, did exhibit an association with cribra orbitalia, indicating expression of these traits has a relatively high degree of sensitivity to the environment (Bergman, 1993).

Another potential source of skeletal variation is congenital (in utero) in nature. Teratogens are agents that cause malformations in utero. Some interesting work has been done in the field of zoology on the effects of teratogens on nonmetric traits. Russian ecological genetics has focused on the consequences of ecosystem contamination by radiation (cited in Anzorge, 2001). In the Sverdlovsk region of Russia, there is a tract of land called the East Ural Radioactive Track with a recorded radioactive gradient. Twenty-eight nonmetric traits in populations of Clethrionomys rutilus, the red vole, were studied along this gradient. The results showed an increase in morphological diversity as well as a significant epigenetic divergence associated with radiation. There was no relationship with other environmental factors; a change in radioactivity was the only commonality.
CURRENT STATUS OF NONMETRIC TRAIT RESEARCH

The study of nonmetric traits in anthropology has a long and somewhat tumultuous history; late nineteenth and early twentieth century observers focused on the description of anatomical variants, the mid-twentieth century saw a florescence of nonmetric trait research and expansion into studies of population relationships and their theoretical implications. The 1980s saw the decrease in the frequency of studies because of observed low heritabilities for some nonmetric traits in different populations and many complicated results that were difficult to interpret. Finally, a slow increase in research began again in the 1990s, as investigators realized the importance of exploring the use of traits that do show evidence of hereditary control. Nonmetric traits may not be entering a new “boom” cycle, but they are accepted as a useful tool in the analysis of population relationships. The main difference between earlier biological distance studies and more recent ones is the focus on the exploration of intragroup variability rather than on the assumption of homogeneity of populations. This focus provides a new theoretical framework for analyzing admixture, migrations, and population history (cf. Hallgrímsson et al., 2004). The essentialist and racial models that originally inspired analyses are no longer accepted; rather, focusing on phenotypic variance provides a much more detailed explanation of population relationships (Stojanowski and Schillaci, 2006). In general, nonmetric traits are no longer used on a global scale (cf. Donlon, 2000), but they are reserved for interregional, intraregional, and local levels of population interactions.

A recent review article in the Yearbook of Physical Anthropology nicely summarizes different analytical approaches to exploring phenotypic variation in the study of intracemetery biological relationships (Stojanowski and Schillaci, 2006). The authors advocate a move away from the descriptive and methodological analyses most commonly published to an approach evaluating behavioral and evolutionary processes such as social structure, regional migration patterns, microevolutionary processes, genetic admixture, natural selection, and biological and cultural adaptation. They take the refreshing approach that the traditional critiques of biological distance analyses (cf. Tyrrell, 2000) are the very questions that can be explored when using nonmetric trait analyses within an evolutionary and biocultural framework.

This process leads us to the contributions of developmental evolutionary biology. Although most work in this area (evo-devo) does not concern itself with skeletal nonmetric traits, it is leading to an improved understanding of some of the underlying factors that result in some nonmetric traits. Epigenetics, the study of how genes produce their effect on the phenotype of an organism, is a growing field with applications in molecular biology, developmental biology, evolutionary biology, and disease research such as cancer. Epigenetics, as observed in evolutionary developmental studies, focuses on identifying the mechanisms through which the environment affects the expression of the genotype on the phenotype (cf. Hall, 1998; Jablonka and Lamb, 2005). Epigenetics is shifting the focus from isolated anatomical features to systems. To understand the etiology of skeletal nonmetric traits, one must understand the relationships between the skeleton, the surrounding soft-tissue structures, and morphogenetic fields (“regions in the embryo that form modules of developmental precursors that are independent of one another” (Willmore and Hallgrímsson, 2005:211)] through which these systems interact during development.

CONCLUSIONS AND SUGGESTIONS FOR FUTURE RESEARCH

The study of nonmetric traits began as part of the recording of anatomical variants or oddities and progressed to become a useful method of assessing population relationships and histories. The method relies on a suite of traits that are heritable, although environmental influences can affect the expression...
of traits morphologically. It is generally accepted now that nonmetric traits of both the cranial and the infracranial skeleton have an underlying continuous distribution and that the phenotypic expression of these traits is the result of both genotypic and environmental influences. A trait will be expressed if there is a genetic predisposition for it, and if the environment is favorable. Our knowledge of the traits themselves has increased, starting with morphological descriptions. Now, there is a developmental perspective: a need to understand why or how a trait is expressed in a particular way. Both new technological advancements as well as classic, analytical approaches can help, from the use of archives of standard radiographs to more modern imaging techniques such as 3-D imaging (see Hallgrímsson, this volume) that are becoming readily available.

The overall list or suite of traits that may be used in population comparisons has been refined; through time the heritabilities of specific traits have been examined. A vast area of research opportunities is still available for assessing the importance of genetic factors on the formation of traits, an area that may be combined with recent skyrocketing advancements in molecular genetics. It is important to remember that just as a myriad of environmental factors might affect the thresholds for production of traits under different circumstances, so a myriad of different genetic factors may lead to the formation of one trait. What may be a useful explanation for a rare trait observed in the population of one region may not be the same explanation in another region. This process leads to the observation that the geographic scale of exploration has also changed in terms of nonmetric trait analysis. Early on, the scope was continental, particularly in comparing African, South and North American, and European skeletal collections. It is generally accepted that nonmetric trait analysis consists of variation that is most telling at the regional scale, among groups for which there is a possibility of recent relationships.

The future of nonmetric trait analyses lies in two areas: 1) the traits themselves, and 2) population relationships. In terms of the traits themselves, it is important to understand the ontogeny of traits. The use of new imaging techniques and the generation of more information on morphological development, both prenatally and postnatally, will contribute to our broader understanding of skeletal development. At the level of the population, it is important to understand the population-specific behavior of traits. What will be useful in the future is, for example, to correlate nonmetric data with information such as strontium and oxygen isotope data to broaden the scope of population history reconstructions. In this way, we may be able to determine not only the relationships between populations but, also patterns of migration. As new large-scale genome scanning technologies become widely available, correlations between DNA analyses and nonmetric trait analyses would provide a means of examining the relationship between the genotype and the phenotype. As of yet, there are neither any reported analyses linking nonmetric traits with specific genes in humans, nor have there been reports specifically aimed at comparing DNA results of population relationships with those of nonmetric traits from a methodological perspective. This type of analysis would be very informative; however, preservation and cost will likely be prohibitive for some time. The study of skeletal nonmetric traits in biological anthropology is not obsolete. There are still many opportunities for those students who will become the researchers of the future.

REFERENCES


Howe WL, Parsons PA. 1967. Genotype and environment in the determination of minor


Owsley DW, Jantz RL. 1978. Intracemetery morphological variation in Arikara crania from the Sully


Snow CE. 1948. Indian Knoll Skeletons of Site OHZ. Lexington: University of Kentucky Reports in Anthropology.


CHAPTER 18

ADVANCES IN PALEODEMOGRAPHY

GEORGE R. MILNER, JAMES W. WOOD, and JESPER L. BOLDSEN

INTRODUCTION

Among the first questions asked when faced with human skeletons are whether the individual was male or female, and how old he or she was at the time of death. Such concerns are equally important for forensic and archaeological purposes. With regard to the latter, the age and sex of skeletons can tell us about the demographic composition of the group that used a particular burial ground. When combined with contextual information, including artifacts, grave details, and burial location, an individual’s age and sex are central to studies of past societies through mortuary practices. Paleopathology, the study of the diseases of the past, also rests squarely on this information because infections and other debilitating conditions, as we know all too well, are distributed unevenly by age and sex. Its potential significance is the reason the paleodemographic literature has exploded since the 1960s when pioneering studies that laid the groundwork for current research were first undertaken (Acsádi and Nemeskéri, 1970; Angel, 1969; Swedlund and Armelagos, 1969).

Here we pay most attention to the generation and, especially, the interpretation of skeletal age distributions, which have been subjects of intense controversy since the early 1980s. Critiques of paleodemography, countered in turn by equally vehement defenses, have tended to focus on how one goes about estimating the age of skeletons, especially those of adults, the kinds of demographically reliable information that can be derived from age at death distributions, and the issues amenable to analyses of the small and often biased samples typical of archaeological sites. Some authors, with good reason, have thrown up their arms in despair over such issues, especially the methodological problems (Bocquet-Appel and Masset, 1982, 1985; Petersen, 1975). We, however, agree with Konigsberg and Frankenberg (1994; Frankenberg and Konigsberg, 2006) that much can be learned from the age distributions of archaeological skeletons, and that each cycle of criticism and response has, with time, rethinking, and renewed research effort, done much to strengthen the field. The demographic underpinning of paleopathological research—particularly as it pertains to how healthy one ancient population might have been relative to another—has likewise been found wanting, but here too there is room for optimism (Wood et al., 1992b; Wright and Yoder, 2003).

For good reasons, methodological issues have figured prominently in criticisms of paleodemography over the past few decades. Unfortunately, the same level of concern has not been directed toward the theoretical
questions that are the reason we look at skeletons in the first place. The potentially great contributions of paleodemographic information are largely left as vague statements about the benefits or, more usually, the costs of major shifts in diet, population density, settlement size and duration, intergroup relations, and the like. Nobody doubts that the demographic characteristics of past populations provide otherwise unobtainable perspectives on our ancestors’ lives (e.g., Chamberlain, 2006). More attention, however, should be directed toward exactly what can be learned from skeletons that cannot be obtained from other sources, most notably other archaeological materials and written documents. All of them, not just skeletons, have their own problems. It is equally important to understand what cannot be learned from bones, regardless of how interesting a research question might be. We must define culturally and biologically significant theoretical questions that can be tackled with rigorous methods, mortality samples, and groups of skeletons that rarely number more than a few hundred.

For the purposes of this chapter, a narrow definition of paleodemography is adopted as we focus on the demographic characteristics of past populations derived from skeletal samples. Yet there is much more to paleodemography than skeletons (e.g., Hassan, 1981; Paine, 1997). At the very least, it is essential to embed osteological results in a broader theoretical discussion of population processes backed by solid empirical evidence derived largely from archaeological investigations, such as settlement and household numbers indicative of population growth or decline. In short, the study of past populations must include analyses of skeletons, but they are only part of the story.

POINT–COUNTERPOINT

The field of paleodemography in its modern form is only a few decades old. To be sure, bones and teeth have long been used to estimate the sex and age of skeletons from crime scenes and archaeological sites. Yet much of that work would not be considered paleodemography if we mean by that term a rigorous study of the demographic structure of past populations represented by skeletons.

Reports from the early to mid-twentieth century on excavations of cemeteries often include long tables that list the age and sex of burials, along with other information such as body positions and associated artifacts. The accuracy of ages assigned to skeletons was of great concern, as it still is today, and it was found that some parts of the skeleton worked better than others for estimating purposes (Brooks, 1955; McKern and Stewart, 1957). Surprisingly, the information so laboriously collected rarely went beyond simple tabulations. Usually only rather obvious features of the skeletal sample were noted in passing, such as the presence of more adults than juveniles, more men than women, and the like. Perhaps that is fortunate because inferences based on such observations—and here we are not concerned with the accuracy of age estimates, an altogether separate issue—ran afoul of a lack of appreciation for the general demographic characteristics of small-scale societies, even though they were not entirely unknown at that time (Carr-Saunders, 1922; Krzywicki, 1934). For example, in the Chiggerville, Kentucky, skeletal sample, deaths among “infants” (0–6 years) were regarded as high, so much so that they were considered evidence for infanticide in this mid-Holocene hunting-and-gathering population (Skarland, 1939). Yet juveniles of all ages made up 39% of the skeletal sample, so it turns out that there are actually fewer, not more, deaths among young people than one might expect in a preindustrial society (e.g., Howell, 1979; Weiss, 1973).

How data were handled and the means of drawing inferences from them changed abruptly during the late 1960s to mid-1970s. Data began to be presented in the form of life tables derived from deaths (the skeletons), and results such as survivorship were compared with figures taken from model life tables (Ácsádi and Nemeskéri, 1970; Bennett, 1973; Buikstra,
1976; Kobayashi, 1967; Lovejoy et al., 1977; Moore et al., 1975; Nemeskéri, 1971; Swedlund and Armelagos, 1969; Ubelaker, 1974). Data generated from skeletons were presented for the first time in a systematic manner to document the demographic characteristics of ancient populations. This work was part of a broader recognition at that time that there was a real shortage of reliable demographic information on small-scale societies, and this information was essential to improve knowledge about life in the past, as well as the changes that have taken place between ancient and modern times. To help deal with flawed data, model life tables were developed for use with the kinds of populations and the small samples typical of anthropological studies (Weiss, 1973).

Because a life table approach was adopted, skeletal collections were treated as if they came from stable populations. That is not entirely unreasonable because populations usually reestablish their initial age structures after demographic disturbance, such as excessive deaths from famines or epidemics, in several decades provided that fertility and mortality rates remain fixed (Weiss, 1975). From this perspective, there is little reason for concern, but that cannot be said about several other issues stemming from the use of life tables based on cemetery samples.

Perhaps the most basic assumption is that the skeletons came from a single population, although one that might have spanned many generations, that practiced the same way of life throughout the period when the skeletons accumulated. The alternative would be a sample that consisted of deaths from distinct populations that might have practiced different ways of life, hence, had dissimilar fertility and mortality rates. There might well have been continuity among those buried in a particular spot if some coherency in cemetery layout can be demonstrated. Unfortunately, the necessary work on mortuary practices—that is, the cultural context—is not always done before assuming the skeletons represent in some real way a single group of individuals perhaps spanning multiple generations.

For the sake of convenience, and in the near absence of information on population growth, life tables were calculated as if the skeletons came from a stationary population; that is, it was a stable population with an intrinsic growth rate of zero (the growth rate is conventionally designated as $r$). The proportions of a skeletal sample in various age intervals will vary somewhat if the rate of growth is greater or less than zero. More attention should be directed toward estimating change over time in the growth (or decline) of past populations based on independent evidence. That might include the numbers and sizes of houses or sites in archaeologically well-characterized regions, although those data pose their own sets of interpretive difficulties. The problem with assuming stationarity was recognized in this pioneering paleodemographic work as shown by examinations of the effects of population growth on life-table estimates (Asch, 1976; Bennett, 1973; Buikstra, 1976; Moore et al., 1975). It can be readily appreciated that many populations in the past, even if stable over the long run, cannot have been stationary. After all, if they were indeed stationary, we would be wearing skins and living off the land rather than laboring through this sort of essay (this is not a value judgment because you just might prefer to be out fishing rather than slogging through this chapter!).

Researchers also realized that skeletons from archaeological sites were often a biased sample of all deaths that took place in a particular community. Infant under-enumeration, in particular, is a commonly recognized problem. The skeletons of young children might be absent or poorly represented because these individuals, particularly newborns and infants, were not buried alongside everyone else, so they would not be included in cemetery samples even in the best of circumstances. Small bones tend to be poorly preserved relative to those from older people, particularly adults, making them difficult to find. Excavators might not recognize the bones of young children, especially infants, as being human, so they might not be kept for additional study, or they
end up mixed with nonhuman animal bones. Such concerns are well founded because some life-table figures are more severely affected by an under-enumeration of infants than are others (Moore et al., 1975). For example, comparisons of survivorship ($l_x$), which can be found in the literature, might be thrown off if one sample had a deficit of young children because of poor bone preservation, sloppy field methods, or customs that excluded their being buried in the cemetery. An inability to estimate the ages of the elderly was another widely recognized difficulty, one only partly solved through the common use of open-ended terminal age intervals such as 50+ years. Although an honest admission of inaccuracies in estimating ages in the upper end of adulthood, this solution sacrifices potentially useful information on the demographic characteristics of past populations, including overall lifespan.

There was widespread, although not universal, acknowledgment of the uniformitarian principle that underlies all of paleodemography. This topic was covered thoroughly by Howell (1976, 1982) in the early years of the development of the field, partly prompted by inferences drawn from a large skeletal sample studied by Lovejoy et al. (1977). In essence, variation in fertility and mortality rates throughout the human lifespan is constrained in predictable ways. Demographic patterns in prehistory should not deviate wildly from those of modern populations (Chamberlain, 2000; Howell, 1976). Yet we would all agree that demographic processes are not invariant across human populations. If they were, there would be no need to study these processes, including those that characterized ancient populations. The existence of constraints on the range of variation is important because it imposes limits on what can reasonably be expected in demographic reconstructions. That is not to say we cannot learn anything new from skeletal studies—we firmly believe we can, which is why we wrote this essay. But without the uniformitarian principle, it would not make sense to apply models based on the present to investigate the often small, incomplete, and biased samples of skeletons from cemeteries that just happened to survive to modern times.

Paleodemography soon received its share of criticism. Perhaps it attracted such attention because it addressed high-visibility issues such as the demographic characteristics of the kinds of small-scale societies, most notably hunter-gatherers and village agriculturalists, that dominated most of human existence. From this point onward, developments in the field can perhaps best be characterized as point–counterpoint.¹

The first critique, intended as an act of infanticide on an emerging field, was by Petersen (1975) who argued that paleodemography as a whole was based on flawed information, procedures, and samples. Although many early paleodemographic studies, viewed in retrospect, had methodological problems, that hardly seems an inherent limitation of paleodemography. In fact, some problems only became apparent much later, and they were identified almost entirely by paleodemographers rather than by mainstream demographers. There is always some uncertainty about the precise age of a skeleton, in general more so for an adult than for a juvenile. Yet skeletal samples are not the only ones plagued with problems stemming from imprecise age estimates, including widely accepted demographic data sources such as the World Fertility Survey or the Demographic and Health Survey. The fact that paleodemographic samples tend to be small and are produced by a selective, and typically incompletely understood, sampling process was a more telling criticism. The typically small size of samples is largely a result of investigating small communities—even if long-lasting and completely excavated, their cemeteries rarely held large numbers of graves. If we hope to learn something about the life experiences of people in the distant past, we must come to grips with the problems posed by small samples.

¹We fully acknowledge Aldous Huxley’s title that, however appropriate for the current discussion, had nothing whatsoever to do with paleodemography.
samples, including determining how large is large enough for a sample to tell us something meaningful about the demographic structure of past populations (Hoppa and Saunders, 1998). The alternative, to give up on most of human existence, does not seem acceptable.

It was not long before another major criticism with the provocative title “Farewell to Paleodemography” was leveled at the field (Bocquet-Appel and Masset, 1982). It gained more traction than the first and led to a broader recognition of analytical problems, although solutions have been slow in coming and are still not entirely in place. Two principal problems were noted. First, the means of estimating ages were considered too imprecise for a demographic signal to penetrate stochastic noise. Second, age distributions generated from skeletons have a tendency to mimic those of the known-age reference samples used to develop the means of age estimation. That might at first sound odd, but the concept can be illustrated easily through a simple example. Assume the skeleton sample used to generate the original age-estimation method had a badly skewed age distribution, much like the McKern and Stewart (1957) Korean War sample that is the basis of one of the most commonly used pubic symphysis age-estimation methods. As one might expect, the Korean War sample is made up mainly of young men, so any recorded characteristic of the pubic symphysis will be associated with early adulthood. That is true even if in reality the characteristic is mostly found in individuals of middle-to-old age—that is, the trait is usually associated with middle-aged or older people if you had a sample of deaths at all ages from the general population.

The supposed death knell of paleodemography, specifically the difficulty (if not impossibility) of getting demographically useful information from skeletons, was soon answered (Buikstra and Konigsberg, 1985; Greene et al., 1986; Lanphear, 1989; Mensforth, 1990; Piontek and Weber, 1990; Van Gerven and Armelagos, 1983; cf. Bocquet-Appel and Masset, 1985). Although none of these replies was entirely successful, it became apparent that the “farewell” was oversold, although in one form or another, this debate has rattled along to the present. Perhaps the most reasonable view today—the problems have not been wholly resolved, as we shall see—is to remain cautiously optimistic. At the very least, we now have some idea about the way to proceed (e.g., Hoppa and Vaupel, 2002).

The confounding effect of reference sample age distributions, which Bocquet-Appel and Masset (1982, 1985) rightly raised, is indeed a difficulty with traditional age-estimation methods, but it is not a necessary feature of skeletal age estimation (Aykroyd et al., 1997; Bocquet-Appel and Masset, 1996; Chamberlain, 2000; Hoppa and Vaupel, 2002; Konigsberg and Frankenberg, 1992, 1994, 2002; Konigsberg and Herrmann, 2002; Konigsberg et al., 1997; Müller et al., 2002; Wood, 2003). Age estimates derived from skeletons, much like those from living populations lacking hard documentation, will inevitably involve some error. That does not mean the work cannot be done—after all, measurement error is an issue for all sciences. However, two immediate challenges face us: ways must be found to minimize error, and we must learn how to deal with the error that inevitably remains. Progress, in large part spurred on by criticisms of the field, has already been made in addressing the error structure of age estimates (Jackes, 1985; Konigsberg and Frankenberg, 1992; Müller et al., 2002).

At about the same time, it was noted that for cemetery samples population growth or decline—that is, anything other than nonzero growth—distorts the overall age at death distribution. There were two difficulties with how life tables derived from deaths alone (i.e., skeletons) were developed and interpreted.

First, life-table statistics were typically calculated as if the population remained stationary over the entirety of cemetery use, which might cover centuries or even thousands of years. Generally one can do little more than hope that the inevitable stochastic variation in fertility and mortality in small populations eventually equaled out over the long run such that there
was neither growth nor decline in population size over the period of cemetery use. The assumption that ancient cemetery samples approximate stationary populations is based on the observation that growth has indeed been very close to zero for most of human existence (e.g., Alesan et al., 1999). But that fact, incontrovertible as it is, provides little comfort—the difficulty here is one of scale. Growth at or close to zero refers to what took place on continental or global scales, not necessarily at the level of specific communities, the cemeteries we examine. That is, local populations undoubtedly went through periods of increase and decline, even to extinction, related to the vicissitudes of ever-changing natural settings and the efficacy of the social and technological means of dealing with them. Here is one place where knowledge of cultural context is critical for understanding the nature of skeletal samples. For example, in Denmark during much of the Middle Ages and Early Modern period, local peasant populations were tightly regulated by landlords who determined the number of equal-sized farms, hence, the permanent households comprising a community and, in turn, the individuals buried in the parish cemetery (Boldsen, 2000; Porsmose, 1987). But that sort of local population regulation, imposed from above, would be highly unusual considering the broad range of societies and time periods that archaeologists routinely study.

Second, the relative proportions of skeletons in a sequence of age intervals is a result of a more complex process than mortality alone. That is because for an increasing population, one with positive intrinsic growth, there would be an excess of the young relative to the old since not every cohort, each bigger than the last, had moved its way through the entire life-span. The age at death distribution would be correspondingly weighted toward the young. In fact, it was noted that age at death distributions reflect fertility more so than mortality (Johansson and Horowitz, 1986; Sattenspiel and Harpending, 1983). This phenomenon, recognized in demography for a long time (Coale, 1957), has since the mid-1980s become widely appreciated in paleodemography (Buikstra et al., 1986; Corruccini et al., 1989; Konigsberg and Frankenber, 1994; Konigsberg et al., 1989; McCaa, 2002; Milner et al., 1989; Paine, 1989a, 1989b; Paine and Harpending, 1996, 1998). There have even been a few studies that have used this feature of skeletal age at death distributions to estimate crude measures of fertility in prehistory (Buikstra et al., 1986; Konigsberg et al., 1989; McCaa, 2002; Paine, 1989a).

Finally, we move beyond paleodemography to the interface between it and paleopathology. A distinction is drawn here between two separate but related and equally important goals of paleopathology. The first is the identification of skeletal lesions and the disease processes that caused them, as well as their consequences in the absence of effective modern medical treatment (e.g., Aufderheide and Rodrı´ quez-Martı´ n, 1998; Ortner and Putschar, 1985; Steinbock, 1976). For the most part, these objectives can be pursued without regard to paleodemography, although the best studies incorporate age and sex when identifying the diseases that resulted in distinctive bone lesions. That is not true of paleopathology’s second goal: determining how healthy some populations were relative to others. All studies about the biological effects, usually framed as costs, of the transition to agricultural economies and the emergence of organizationally complex societies fall in this second category (e.g., Cohen, 1989; Cohen and Armelagos, 1984a; Steckel and Rose, 2002). Here we are concerned solely with this goal of paleopathology, which is sometimes referred to as paleoepidemiology (Boldsen, 1997).

The demographic conundrum with paleopathology has, for better or worse, come to be known as the “osteological paradox” (Wood et al., 1992b; Wright and Yoder, 2003). The difficulty stems from the fact that mortality samples—in this instance, archaeological skeletons—are not the same as samples of living people. That is because there is heterogeneity in the risk of death and mortality is selective; in other words, the most vulnerable people at each age are the ones most likely to join the ranks of the dead. Even if the cemetery
excavator is fortunate enough to find everybody born into an ancient community, which is a remarkable situation in itself, the sample is still biased toward the people in each age interval who experienced the greatest risk of death, often because they were the sickest. What that means is that the frequency of a skeletal lesion in the cemetery sample is not the same as the prevalence of the disease in the population.

The osteological paradox paper of Wood et al. (1992b) was built in part on Ortner’s (1991) observation that healed bone lesions in archaeological samples were perhaps best regarded as an increased, not decreased, ability to withstand disease. These people had survived a period of ill health only to die later, often for some other reason. The problem in interpretation posed by this situation can be illustrated simply. Say, for example, there were two groups of skeletons: those with a particular kind of bone lesion and those without. The ones with lesions had survived long enough for the bone to respond in a distinctive way, and they perhaps lived for a long time if the bone showed signs of healing. The skeletons without lesions conceivably came from two groups. The first group is made up of people who had never been exposed to the disease (assuming it was an infection of some sort) or were strong enough to fight it off before there was an involvement of bone. The second group comprises those unfortunates who lacked the means of withstanding the disease, perhaps because they were malnourished, and they died quickly before a distinctive bone response developed. The main point here is that paleopathology and paleodemography are linked inextricably when the research question involves the comparison of skeletal lesion frequencies in two or more cemetery samples.

As might be expected, the osteological paradox paper sparked its own round of criticism and response (Byers, 1994; Cohen, 1994, 1997; Goodman, 1993; Jackes, 1993; Milner et al., 2000; Saunders and Hoppa, 1993; Steckel et al., 2002; Wood and Milner, 1994; Wright and Yoder, 2003). One indication of this topic’s importance is the appearance of a separate interchange over the selectivity issue in Britain (Bird, 1996; Waldron, 1994, 1996). Much like earlier critiques of paleodemography, it would be fair to say the implications of the “osteological paradox” have made themselves felt in some quarters but not in others. Some researchers reject them out of hand (Goodman and Martin, 2002), whereas others take them seriously (Wright and Yoder, 2003).

The various critiques of paleodemography over the past 30 years—including a premature “farewell”—have, in the end, succeeded in strengthening the field. No doubt interest in the field remains high, despite vexatious and persistent problems, because most human societies date to a period well before any semblance of documentation, no matter how primitive, that might be used in conventional demographic analyses, including those based on historical records that feature their own problems. It would be a shame simply to write off most of human existence as being forever beyond our reach.

In the pages that follow, we cover several issues that must be dealt with before paleodemography can make strong contributions to population science. First, and perhaps most fundamentally, what theoretical questions can potentially be answered with skeletons from well-documented contexts? To date, the attention of critics has focused principally on methodological issues not on the potential significance of the results. Second, how can we understand and correct for the complex, non-random sampling processes, hence selection biases, that inevitably lie behind archaeological samples? Third, how can unbiased sex and age estimates be obtained, particularly estimates that avoid biases that are a disturbing feature of standard adult age-estimation methods? Fourth, how can we best handle the inevitable age-estimation errors that occur in even the most accurate methods available, including those yet to be developed? Fifth, how can we begin to deal with the nonstationarity problem? Sixth, how can the related problems of heterogeneous risk and selective mortality be handled so that observable pathological
lesions in skeletons can be linked to aggregate-level mortality and morbidity patterns in the original population? These issues are not unrelated to one another, and work on them should proceed simultaneously. The various studies reviewed here represent important steps toward developing such an analytical framework.

THE STORIES SKELETONS CAN TELL

Although paleodemographers have spent much time refining methods, it is best to start with why the work is worth doing in the first place. Without some appreciation of the research questions that can be addressed with old bones, improvements in paleodemographic practices—and there have been some over the years—are little more than techniques in search of problems. Although other researchers are likely to come up with a somewhat different set of reasons for doing paleodemographic studies than the ones we give, the following discussion illustrates the range of cultural and biological information that can be best obtained from the age and sex structure of archaeological cemetery samples.

What should be the first step in the study of a skeletal sample is often relegated to other researchers concerned with mortuary practices or ignored altogether: determining how and why the skeletons ended up in more-or-less close physical juxtaposition where they could be excavated together. Although skeletons from a particular archaeological site are commonly treated as a single “population,” perhaps with considerable time depth, just because burials are found together does not mean the individuals had any necessary connection with one another. The skeletons might have come from a cemetery that contained part or all of one or more groups of people defined socially in any number of ways. A burial ground might contain the residents of one or more communities or perhaps only certain people selected by virtue of their status position or social group affiliation. Sometimes a cemetery sample might have more to say about the circumstances of death, such as battlefield burials that are typically a highly selected subset of all deaths that occurred in a population at a particular point in time (Fiorato et al., 2000; Scott et al., 1998; Thordeman, 1939). In this instance, their sole connection might be the fact their deaths took place at the same time and place, and for the same reason.

Yet all those examples, and the list is by no means exhaustive, share one important characteristic—there is some reason the skeletons are found together. That need not always be the case. Skeletons might simply accumulate at repeatedly occupied sites from the deaths in multiple, separate, and perhaps unrelated groups that just happened to have lived in the same place at different times. Such accumulations of skeletons—as distinct from cemeteries containing individuals who had some biological or social relationship to one another—are not as farfetched as they might seem at first glance. To choose but one example, osteologists and archaeologists alike have treated the large, well-preserved skeletal sample from Indian Knoll in Kentucky as a single hunter-gatherer population for analytical purposes ever since it was first excavated in the mid-twentieth century (Snow, 1948; Webb, 1946). In the absence of a detailed mortuary-practices study that includes the spatial arrangement and depth of burials in a midden that spanned a few thousand years, it seems a stretch to consider the burials of one population. An examination of field records from the temporally equivalent Read site, located along the same river, did not result in the identification of one or more formal cemeteries (Milner and Jefferies, 1998). The burial distribution at Read was instead consistent with graves that held individuals from an unknown number of groups that just happened to settle in a particularly favorable spot occupied intermittently over the course of several millennia.

That brings us to investigations of past mortuary practices that contribute essential information to the archaeological interpretation of the nature of past societies. Such studies, based largely on differences in body treatment,
grave form, cemetery location, and artifact accompaniments, would be hamstrung—in fact, they are not worth doing at all—without reasonably accurate estimates of age and sex. Certainly one of the first objectives of osteological work should be to determine whether collectively the burials deviate from the deaths expected in preindustrial societies. This first step, however, is all too often omitted in archaeological mortuary studies. It is one of the main reasons simple pattern matching of model and skeletal age at death distributions was advocated 20 years ago (Milner et al., 1989; Paine, 1989a, 1989b). Although such comparisons might show that samples are badly biased toward one sex or a certain part of the human lifespan, limiting their use for many biological studies, that does not mean the skeletons are collectively uninformative from a cultural perspective. In fact, the atypical demographic structure may be precisely the clue needed to tell us about what led to a burial area’s formation. As mentioned, the skeletons might represent a special segment of society or people buried together because the circumstances of death were unusual or called forth extraordinary measures, such as bodies being piled unceremoniously in a plague pit. In short, understanding of the social and situational processes that gave rise to a skeletal sample is an essential part of accounting for selectivity bias, even when the research questions are fundamentally biological in nature.

From a biological perspective, there are two main reasons for pursuing paleodemography that are integrally related to one another: how individual health and well-being, especially as they pertain to population structure, vary across time and space; and how those differences are related to population histories, socioeconomic systems, and environmental settings (Wood, 1998). That is, patterning needs to be identified first, and then its correlation with spatial and temporal variation in cultural and natural contexts must be considered in the context of appropriate theoretical models. Research over the past 30 or so years is usually couched in terms of the consequences of environmental, technological (including subsistence practices), and organizational changes that took place in preindustrial societies.

Our current understanding of trends in demography and health from hunter-gatherers, to village agriculturalists, and finally, to the people who lived in civilizations crystallized about 20 years ago (Cassidy, 1980; Cohen, 1989, 1994, 1997; Cohen and Armelagos, 1984b; Lallo et al., 1978). It has received considerable attention, in part because it features an easily comprehended categorization of societies, each with a characteristic health regime, and a straightforward interpretation of much of the skeletal evidence. The general approach focuses attention on questions that are of fundamental importance to our understanding of human existence in the past. In fact, the highly visible work in this area inspired one influential nonpaleodemographer to label the transition to agriculture “the worst mistake in the history of the human race” (Diamond, 1987).

In essence, cultural evolution over great periods of time is characterized, in this view, in terms of two major steps, each resulting in greater morbidity and mortality. First there was the shift to agricultural economies, and it was followed by the emergence of complex societies, with a particular emphasis placed on societies with interconnected urban centers. Throughout this transformation in how people lived, the human condition as measured by morbidity and mortality declined. This view, of course, is not original to modern researchers. For example, it bears some resemblance to Rousseau’s (1950:204, 243–244) writings about humankind’s supposed natural state, which were harnessed to his then-radical political agenda. The “good constitution of savages . . . reflect[s] that they are troubled with hardly any disorders, save wounds and old age;” no doubt this happy state of affairs contributed to their enjoying “free, healthy, honest, and happy lives.” Growing crops, working metal, and developing inequities in property “first civilized men, and ruined humanity.” What is new about the current argument for declines in
health over time is that it is backed by considerable information collected through much hard work by many researchers.

Even though this general scenario is widely accepted—it appears in scholarly writing, textbooks, and the popular media—the categorization of societies as a series of stages may not be the best way to characterize variation in human health, and the interpretation of skeletal data is not as straightforward as it might at first seem. The typological underpinning of the two-step decline in health scenario forces all populations into just a few categories, thereby minimizing variation that must have existed among groups said to conform to a particular stage. To take but one example, many groups that have been characterized as hunters-gatherers or agriculturalists were neither. They were transitional in the sense that they were somewhere between an exclusive reliance on the collection of wild foods and a near-total dependence on the production of agricultural goods. Comparisons based on gross characterizations of ways of life, such as hunter-gatherer or village agriculturalist, conflate profound transformations in diet, sedentism, residential group size, mobility, economic inequality, and interregional contact that should be considered separately since they did not take place in lock-step and could well have had different effects on people’s exposure to pathogens, dietary quality, and susceptibility to food shortages. Furthermore, the commonly accepted scenario does not provide a mechanism that explains why radically different subsistence strategies were adopted, and more importantly persisted, when they had such a detrimental effect—people supposedly became sicker and died earlier each step of the way from hunter-gatherer to urban dweller.

A rather obvious demographic problem should immediately call into question the stepwise, universal decline-in-health scenario (Pennington, 1996). If things got so bad, one is left wondering why humans in aggregate over many thousands of years experienced such a dramatic increase in numbers. After all, growth is commonly considered a measure of a species’ success. Perhaps one might argue that fertility increased despite a decline in the quality of life. That possibility, which smacks of special pleading, faces a problem because we must have an appropriate metric for quality of life in prehistory. If sick people in a mortality sample are to be used for that purpose—and it is the only realistic standard of measurement currently available—then what is observed in skeletons should be readily interpretable. In short, we should know what our yardstick, skeletal lesions in cemetery samples, is measuring. Yet there remains considerable uncertainty over the interpretation of skeletal lesions (e.g., Bird, 1996; Byers, 1994; Cohen, 1994, 1997; Goodman, 1993; Goodman and Martin, 2002; Jackes, 1993; Jackes et al., 1997; Ortner, 1991; Saunders and Hoppa, 1993; Waldrin, 1994, 1996; Wood and Milner, 1994; Wood et al., 1992b; Wright and Yoder, 2003). Unfortunately, one would hardly realize that this lively controversy exists to judge from much of the paleopathological literature in which skeletal lesions continue to be interpreted in a simple and direct manner: Sick skeletons equal an equally sick population.

Of central concern are the roles played by heterogeneous frailty and selective mortality in our understanding of the demography and health of both modern and past populations. We need better understandings of where heterogeneity comes from, how it differs among populations, and whether it changed over time with very different ways of life, as well as the ways

---

2 Indeed, the categorization shares many of the disadvantages of nineteenth-century models of unilineal cultural evolution, albeit with nicer political implications.

3 Here we focus on issues related to selective mortality and its effect on the composition of cemetery samples, but it is important to note they are not the only complications to the study of bone lesions in samples of skeletons from old cemeteries (e.g., Cook and Powell, 2006).
differing distributions of heterogeneity affect population dynamics. We need them because such information has the potential for telling us much about linkages among changes in the natural environment, socioeconomic systems, disease experience, and population processes, which is the main reason paleodemography is of scientific interest (Wood, 1998).

SKELETONS AS A SAMPLE OF DEATHS

We will never be able to examine the skeletons of more than a tiny fraction of all people who were alive at a particular time and place. Archaeological samples, even in the best of circumstances, usually number no more than a few hundred skeletons, although some notable exceptions range into the thousands. Despite what from a demographer’s perspective can only be viewed as risibly small samples, the absolute number of skeletons is not the most troublesome issue that paleodemographers face. After all, extrapolating from small samples is commonplace in the sciences as well as in other spheres of scholarly endeavor. Moreover, sample sizes can be enlarged through additional excavation, although that is often impractical because archaeological work is terribly costly and sites may be destroyed (a frequent occurrence) or access to them denied. The real difficulty is that we only observe skeletons that for one reason or another survived to the current day, happened to be discovered by archaeologists, and were preserved with proper documentation in museum and university collections. Nobody would be so rash as to suggest that every skeleton has an equal chance of being examined by a qualified osteologist. Even in the case of an ideal sample, if such a thing were ever found, it is still a mortality sample so, by definition, it is a highly selected sample of the people who were at one point alive at each age (Wood et al., 1992b). To repeat, a cemetery, no matter how large and well excavated, is not the same as a living community, and it is the latter that is of interest when we wish to make comparisons among different societies in the distant past.

The linkage between a past population and a skeletal collection is long and tortuous. The bones we might examine are the ones that survived a complicated winnowing process that might be summarized by the following sequence: Living → Dead → Buried → Preserved → Found → Saved.

A concern over sampling issues is by no means new, although it has only recently been examined systematically. Thomas Jefferson (1788:105) even recognized it two centuries ago upon excavating an earthen mound in Virginia: “[t]he bones of infants being soft, they probably decay sooner, which might be the cause so few were found here.” Because differential preservation had an effect on the representation of young and old skeletons, the original and excavated samples would differ from one another. He must have had his own experience in mind—high infant mortality in his day—when he noted that too few infants were discovered. Jefferson’s model was implicit and by modern standards intuitive and primitive, but he nonetheless compared observations with expectations and came up with a reasonable explanation for why the two did not match. As in so many things, Jefferson was far ahead of his time. Only in the past several decades have such concerns become an important part of archaeological, including osteological, research.

Although not given as much attention until recently as other parts of the lengthy selection process, the Living → Dead transition calls into question the notion that a skeletal sample can ever be considered a direct reflection of the community where it originated (Saunders and Hoppa, 1993; Waldron, 1994; Wood and Milner, 1994; Wood et al., 1992b; Wright and Yoder, 2003). That is not to say that nothing can be done with bones, only that we must recognize at the outset that skeletal samples are the result of a selection process. Something we can all agree on is that skeletons from
cemeteries are mortality samples. As such they must be biased samples of the living because people are heterogeneous in frailty or relative risk of death. Said another way, not everyone at each age has an equal chance of dying and, hence, of finding their way into the mortality sample that is eventually excavated and studied. Something along the lines of the following is a common response to this problem: “Everyone dies eventually, so ultimately we sample everybody.” Although it is undeniably true that everyone does die, this response misses the point. Just before death, the individuals who are going to be doing the dying are likely to be in poorer condition, on average, than their contemporaries who are not about to die. Thus, people who die will be, once again on average, selected for poor condition at the time (and age) of death. Their healthier contemporaries will indeed die later, perhaps after suffering from some disease, at which point they too will be selected for poor health compared with the still-surviving cohort at that age. Although everyone eventually joins the mortality sample, an individual’s condition at the time of death is not even expected to be representative of his or her own condition in previous years, never mind that of the general population.

It has to be admitted that not all researchers agree that the Living → Dead transition is a significant problem when it comes to studying ancient cemetery samples. Cohen (1994), for example, has claimed that heterogeneous frailty was unlikely to have been important in most preindustrial societies, especially the small-scale “egalitarian” ones that dominated most of human existence and are a major focus of paleodemographic work. Nevertheless, significant heterogeneity and selective mortality have been detected in a broad range of species, both invertebrates and vertebrates—including humans (Boldsen, 2007; Carey, 1997; Carey and Liedo, 1995; Newton, 1998; Thomas, 2003; Usher, 2000; Vaupel et al., 1994). So it is improbable that some humans, by virtue of the type of society they happened to live in, somehow dodged the effects of selective mortality.

To illustrate the general point, we take what can only be regarded as a “worst-case” situation where it might be imagined that selection biases had no effect whatsoever on the people who died. Archaeologists often point to Pompeii—the Roman city destroyed by an eruption of Mt. Vesuvius—as the quintessential example of a rapidly and catastrophically sealed settlement where everything was preserved as it existed on a single day. But even there one cannot merely assume the human remains that have been discovered were truly representative of the city’s inhabitants. To judge from modern disasters, there were any number of reasons, among them physical ability, wealth and social position, place of residence, and a desire to enrich oneself through looting, why some people fled and others did not. However important these remains are for opening a window on life in this city—and there is no denying that they are important—one cannot regard them as a random sample of the people who once resided and worked there.

Some work has been done on whether “catastrophic” age at death distributions can be distinguished from “attritional” ones, using as a test case a mid-fourteenth-century English Black Death plague pit in London (DeWitte, 2006; Gowland and Chamberlain, 2005; Margerison and Knüsel, 2002; Roberts and Grauer, 2001; Waldron, 2001). This site is an excellent test of whether the two kinds of mortality distributions can be distinguished from one another (reality is always messier than any such classification would suggest) and whether the plague deaths differed from the population’s demographic structure. The latter would be indicative of the selective nature of the Living → Dead transition, even during the Black Death, which can only be regarded as a mortality catastrophe for the populations involved. As Waldron (2001) points out, it cannot simply be assumed that a mass disaster even of the horrific dimensions of the Black Death will necessarily produce an age at death distribution that replicates the living population’s demographic characteristics. That can be extended by saying that the set of lesion
frequencies in the cemetery sample need not mirror those in the population that produced the skeletons if prior debilitating conditions had any effect on the chances of succumbing to the Black Death. As it turns out, different studies have produced inconsistent results, reflecting differences in age estimation and analytical methods. DeWitte’s work (2006), which uses the soundest (although not perfect) analytical methods currently available, suggests that mortality from the Black Death was indeed selective, although probably much less so than normal, “attritional” mortality.

Recent studies have shown that mortality risks were heterogeneous even in the small pre-industrial populations known through cemeteries that accumulated over extended periods of time (Boldsen, 1991, 1997, 2007; Milner et al., 1991; Usher et al., 1997). So the skeletons within each age interval were not a random sample of all individuals who were once alive in those same age intervals. We reach a conclusion that should be news to nobody: Cemeteries are full of the weak and sick, precisely because they are the most likely to die at a particular age. That characteristic of mortality samples is separate from other concerns, such as the range of people who might be buried in particular places, the overall number of skeletons, the adequacy of bone preservation, and the skill of excavators. So even if everything else is ideal, which is an unlikely occurrence in itself, selective mortality remains and ways must be developed to deal with it. This issue cannot be ignored just because it complicates interpretations and solutions have proven elusive. After all, dealing with difficult problems is precisely what new research is all about.

The next transition, Dead → Buried, has received much more attention, largely because it is a critical component of the study of mortuary practices and what they can tell us about past societies (Binford, 1971; O’Shea, 1984; Pearson, 1999). Quite obviously, who gets buried and where the grave is located have a big effect on the chances a skeleton survives to modern times and will be found. If people of a certain age or sex were not routinely buried in a particular place, such as a graveyard, this group will be underrepresented when that area is excavated. Unusual circumstances of death can also result in special ways of handling corpses, including burial some distance away from other community members. In many instances, such practices would result in the loss of relatively few skeletons from a cemetery, so few they would be undetectable. What is more significant is the differential treatment of various segments of a community, particularly when that involves body treatments that contribute to poor bone preservation or burial in separate locations. Historical and ethnographic accounts of burial practices indicate that it is not uncommon for newborns and infants to be buried somewhere other than in the village cemetery, so their frequent under-representation in archaeological samples, quite apart from any preservation issue, comes as no surprise. In hierarchically organized societies, people of high rank are often buried in contexts that are spatially separate from cemeteries used by the remainder of the population. Yet another example of how cultural practices influence skeletal lesion frequencies in cemetery samples is the medieval monastery that often served as a hospital for sick and injured people who perhaps traveled long distances to get there (Arentoft, 1999; Boldsen, 2001; Boldsen and Mollerup, 2003, 2006; Møller-Christensen, 1982).

The problem posed by exactly what is represented by a particular cemetery—that is, the relationship between a skeletal sample and the society as a whole—is easy to envision when dealing with societies that consisted of socially  

---

4Goodman and Martin (2002:36), for example, pay little attention to selective mortality because Wood et al. (1992b) “have not yet provided a statistical solution to this problem or any actual testing of [it] with skeletal samples.” It is quite true that this tough nut has yet to be cracked, but progress has been made in dealing with this difficult issue (DeWitte, 2006; Thomas, 2003; Usher, 2000). The results of this work confirm the importance of heterogeneous frailty and selective mortality in paleodemography and paleoepidemiology.
and economically differentiated communities, such as European medieval towns and villages (Petersen, 1997; Petersen et al., 2006). That difficulty, however, is not restricted to these societies, since even in much smaller scale societies it cannot be assumed that each community is the same as the next. That typically implicit assumption is, of course, precisely the one that is made when skeletons from a particular burial ground are considered representative of broader cultural, geographic, or temporal categories.

A specific example brings this particular issue into sharper focus. In the mid-1980s, a 700-year-old cemetery containing many extraordinarily preserved skeletons was encountered during road construction in Illinois (Santure et al., 1990). These people were tribal agriculturalists, and there were no signs that people distinguished by age, sex, or social position had been buried anywhere other than in the completely excavated cemetery. Yet despite the fact that this group came from a kind of society that displays little if any institutionalized political or economic specialization and inequities among settlements—unlike the medieval European example—it cannot simply be assumed that all roughly contemporaneous, geographically proximate, and culturally similar skeletal samples would be much like the one that, by luck, happened to be excavated. That is because this particular community was at the leading edge of a population expansion, a movement probably related to the great loss in life this village experienced in warfare (Milner et al., 1991). At least one third of the adults in the cemetery had been attacked and killed, so if conflict was socially and economically disruptive, which it probably was considering the great losses of life, this community cannot immediately be assumed to be typical of its time, place, or culture. The basic point here is that it would be unwise to regard a single cemetery, no matter how large and well studied it might be, as archetypical of a group of societies, regardless of how they are defined temporally, geographically, culturally, organizationally, or technologically.

The Buried → Preserved part of the transformation from a living population to a museum collection adds more sampling bias. Burial conditions—for example, soil pH in environments similar to those of the eastern United States—influence bone survival, as does the structure of the bones themselves (Bello et al., 2006; Gordon and Buikstra, 1981; Mays, 1992; Waldron, 1987; Walker et al., 1988; Willey et al., 1997). Cancellous bone composed of networks of fine trabeculae does not last as long as dense thick cortical bone. Anyone who has excavated skeletons knows that small bones with thin cortices, such as wrist bones, tend to disappear first. In somewhat acidic soils, it is not at all unusual to find only the diaphyses of large long bones, the major bones of the cranial vault, and teeth, sometimes reduced to enamel caps.

Even when bones survive to modern times, they might not make it through the Preserved → Found and Found → Saved sequence unscathed. Excavators differ greatly in their interest in finding bones and in their experience in recognizing poorly preserved bones or infant skeletons. Skeletal samples are, of course, dictated largely by excavation strategies. A long-standing fascination with impressive artifacts has contributed to an uneven sampling of past populations. Places with the greatest potential for producing many nice artifacts have traditionally been more often targeted for excavation than those with an anticipated low yield of artifacts. So for hierarchically organized societies, tombs belonging to the most important people have attracted more attention than cemeteries for the bulk of the population. Even if found, bones have not always been saved. Obtaining skeletons is often not a principal objective of an excavation, and even today skeletons are often regarded as a time-consuming (hence, expensive) bother. During the nineteenth and early twentieth centuries, bones were discarded in the field, whereas great efforts were made to keep accompanying artifacts for additional study. When bones were retained, it was not uncommon to keep only the most interesting ones,
such as pathological specimens or intact measurable crania.

So it would be naïve to assume that human remains found in excavations are ever a straightforward sampling of the deaths that took place in past communities, or a simple reflection of the conditions in the living populations that produced the skeletons. Sampling biases are always present, and they need to be attended to. Selectivity starts with death in some distant time, the Living → Dead transition, and it continues to when the remains end up on a museum shelf, the Found → Saved transition.

MEASUREMENT CONCERNS

Osteological methods for estimating age and sex from skeletal remains have been summarized by several researchers, and there is no need to replow ground that has been so thoroughly and perceptively covered (e.g., Buikstra and Ubelaker, 1994; Cox, 2000; Iscan and Loth, 1989; Jackes, 1992, 2000; Kemkes-Grottenthaler, 2002; St. Hoyme and Iscan, 1989). Here we emphasize the process of estimation and the errors associated with it. Accuracy, of course, is of great concern, but so too is the confidence we can place on assessments of age and sex.

Sex Estimation

The assignment of sex to a skeleton is often called sex determination, as if it could be done without error—we suggest that sex estimation, although it sounds cumbersome, better reflects what is actually done. Estimates of sex from bone size and shape are usually limited to adult skeletons because clearly distinct morphological features first appear late in the second decade of life. Today the pelvis is more heavily emphasized for this purpose than the cranium, mandible, or other bones. That was not always true; for example, through the mid-twentieth century, osteologists in the United States tended to place greater attention on the skull. The size and shape distributions of skeletal features used to differentiate males and females overlap considerably, although some bony characteristics are more reliable than others when it comes to ascertaining sex. Furthermore, the degree of sexual dimorphism in the skeleton varies among human populations, so standards developed for one part of the world are not necessarily applicable to another. Skeletal size and shape for sex estimation can be characterized subjectively by eye—and here experience counts—as well as quantitatively by multivariate statistical techniques, with about the same accuracy (St. Hoyme and Iscan, 1989).

It is also possible to extract and amplify ancient DNA from bones, which can be used to ascertain the sex of an individual regardless of age (Stone et al., 1996). Yet despite rapid and impressive advances in this field, practical limitations remain to the routine use of DNA for estimating the sex of large numbers of archaeological skeletons. Not all prehistoric or early historic skeletons yield DNA, contamination is a particularly vexatious problem, and the exacting procedures take time and money.

Several cranial characteristics are indicative of an individual’s sex, most notably the supraorbital ridges, mastoid processes, and superior nuchal line. When one or more of these features are robust, the cranium is generally considered to have come from a male. These parts of the cranium are less likely to be thought of as distinctly female, except when the individual was quite gracile. Furthermore, in older females, these cranial characteristics, often the supraorbital ridges, can approximate those of males, making it more likely to misclassify them than younger adult females (Meindl and Russell, 1998; Meindl et al., 1985a). The pelvis is different insofar as several characteristics can be considered distinctly male or female. Over the past few decades, when establishing the sex of a skeleton osteologists have tended to rely on both the cranium and the pelvis, with an emphasis on the latter, as well as on other parts of the skeleton such as the mandible. Earlier osteologists, however, commonly focused on cranial morphology, sometimes to the near exclusion
of the pelvis. An emphasis on the skull probably contributed to an overabundance of males in earlier osteological reports, as noted by Weiss (1973). In fact, several studies of skeletons first examined several decades ago have found that discrepancies in sex estimates were more likely to be “females” classified as “males” in the original work than the reverse (Milner and Jefferies, 1987; Powell, 1988; Ruff, 1981). This issue is worth examining in more detail, especially in terms of the degree to which sex estimates in early to mid-twentieth-century reports can be relied on. Continually questioning assumptions and results is a much better way to proceed than simply dismissing uncomfortable findings, as has been done in this instance (St. Hoyme and Iscan, 1989). This issue has some bearing on our interpretations of life in the distant past because the general bias toward males is still sometimes interpreted as behaviorally or biologically significant, such as when it might be attributed to disproportionate juvenile female mortality, including selective infanticide (Meindl and Russell, 1998).

Estimating sex from bones is a statistical problem similar to other forms of estimation (Giles and Elliot, 1963; Holman and Bennett, 1991; Konigsberg and Hens, 1998; Meindl et al., 1985a; Wood et al., 2002). Thus far, two statistical approaches have predominated: discriminant function analysis (e.g., Giles, 1964; Robling and Ubelaker, 1997) and finite mixture analysis (e.g., Dong, 1997; Pearson et al., 1992). As pointed out by Konigsberg and Hens (1998) and by Wood et al. (2002), sex estimation is perhaps best viewed as a Bayesian problem and, thus, involves inverse probability. Bayes’s theorem gives us the means of estimating the unknown probability of a trait that cannot be observed directly, such as the sex or age of a skeleton, given that it has some observable characteristic (e.g., greater sciatic notch angle or degree of cranial suture closure). The joint distribution of both traits in a reference sample is needed, and it can be derived from one of the anatomical or forensic known-sex and age skeletal collections housed in various universities and museums. The so-called posterior probability is the conditional probability of having the unobservable trait if the observable one is present. The inverse probability issue is important because it addresses a concern outlined in the “farewell to paleodemography” article that sparked so much controversy (Bocquet-Appel and Masset, 1982). If the proper inversion is not done, the sex composition of the sample of interest (skeletons from a cemetery) will be biased toward that of the modern reference sample used for the original osteological standards.

### Age Estimation

For paleodemographic work, estimating the ages of adults is a much bigger problem than doing the same for juveniles. Juvenile ages can usually be estimated with acceptable error, considering the fact that samples are small and suffer from many biases, as discussed above (see Saunders, this volume). Using tooth crown and root development, tooth eruption, and epiphyseal closure, it is generally possible to estimate the ages of juveniles to within several years. For the first few years of life, it is routine to do so to within a single year.6

---

6Brothwell (1971) also noticed this tendency at about the same time, although unlike Weiss (1973) he chose to emphasize the fact that most samples did not depart too far from roughly equal numbers of males and females. Nevertheless, in Brothwell’s (1971) 27 samples, there was a decided tendency toward male over-representation since there were more of them than females in two thirds of the skeletal collections, and the greatest departures from an equal representation of the two sexes were in those studies where males were said to predominate. For a few samples, plausible explanations based on the nature of the sites were provided for unequal sex ratios, underscoring the importance of cultural context when studying human remains.

6Wood et al. (2002) have advocated estimating the ages of adult and juvenile skeletons simultaneously, arguing that age and sex should be estimated jointly at all ages using Bayesian methods. The results of this approach, if it could be implemented, would presumably be juvenile skeletons with small standard errors for age and large ones for sex, and vice-versa for adult skeletons. But at least the overall error would be minimized, and just as importantly, a clear picture of the total error structure would be obtained.
Adults, however, have proved to be a persistent problem, and the problem worsens with advancing age. A common solution to the difficulty of estimating the ages of old people is to put them in a single open-ended terminal interval, such as 50+ years. Although this practice is an admirable acknowledgment of problems with estimating the ages of the elderly, it prevents us from saying much about senescent patterns of death.

Ages can be reported either as point estimates or intervals spanning several years or more. The open-ended terminal intervals that are often used (e.g., 50+ years) obviously cover much of the possible human lifespan, although in preindustrial societies, old-age attrition was certainly higher than in modern times. Multiple-year intervals are generally preferred over single-year figures since they take into account imprecision in the estimation process. A common practice is to put a skeleton, based on its particular characteristics, into a multiyear interval defined by the original work with a known-age reference sample. The timing of changes in key age-related skeletal features is thereby treated as the same in the reference and archaeological samples, and the interval widths are considered sufficient to accommodate uncertainty in the actual age. It is difficult to test the first simplifying assumption because few well-documented known-age samples, including individual skeletons, date to a time before the twentieth century. The degree to which age-sensitive skeletal characteristics vary in terms of lifestyle, specifically ancient versus modern work load resulting in wear-and-tear on the skeleton, is largely unknown. The second concern has to do with the interval widths used in several commonly used age-estimation methods—including Todd (1920, 1921) and Meindl et al. (1985b) for the pubic symphysis, and Lovejoy et al. (1985b) for the iliac auricular surface—since the chance that a skeleton assigned to a category actually belongs there cannot be estimated from information in the original publications.

No osteologist today would believe that an exact age can be assigned to an adult skeleton. Nor would it be thought that all skeletons that look like they might be about a certain age could be assigned with equal confidence to the same age interval; for example, by assigning all skeletons that look as if they were in their thirties to the 30–40-year category. Each skeleton has its own degree of error depending on the suite of traits that happen to be present. Practical experience shows this to be true because it is not unusual for osteologists to note that one part of the skeleton, such as the pubic symphysis, does not agree with what might be found elsewhere, perhaps the cranial sutures.

What we want to estimate is the probability that the person died at a certain age given that it has one or more skeletal characteristics (Konigsberg and Frankenberg, 1992, 1994, 2002). To do that we need an appropriate reference sample, a skeletal collection where age at death is reported and age-related markers can be scored. Two commonly used reference samples are the Hamann-Todd and Terry collections assembled from medical autopsies (Hunt and Albanese, 2005; St. Hoyme and Iscan, 1989; Usher, 2002). These collections are highly selected samples in terms of the socio-economic position of the individuals they contain. The remains of Korean War soldiers were used to develop another frequently used age estimation technique based on the pubic symphysis (McKern and Stewart, 1957). This age distribution, naturally skewed toward young men, results in the reference-sample age biases mentioned previously (Bocquet-Appel and Masset, 1982; Konigsberg and Frankenberg, 1992, 1994, 2002). The problem can make adult mortality rates seem to be extraordinarily high and to accelerate much more quickly than in historically well-documented populations.

Technical explanations of how these biases occur have appeared elsewhere, and interested readers are referred to those sources (Aykroyd et al., 1997, 1999; Hoppa and Vaupel, 2002; Konigsberg and Frankenberg 1992, 1994, 2002; Konigsberg et al., 1997; Müller et al., 2002). For this brief discussion, it is perhaps best to start with a caricature of the logic
behind traditional analyses, although no one actually estimates ages in the naive way described here. One starts with all individuals in the reference sample (the original skeletons used in developing the method) that display a particular trait. The trait, in this instance, can be either a particular anatomical feature, such as one of a sequence of bony changes subsumed under a single “component” in McKern and Stewart’s (1957) system, or a group of associated features much like those subsumed under an iliac auricular stage described by Lovejoy et al. (1985b). It follows that estimates of the age distribution of the trait in question is determined partly by the age composition of the reference sample. It is important to emphasize that the reference sample age distribution will not be duplicated, only that estimates biased in the direction of the reference sample will be produced (Konigsberg and Frankenberg, 1992).

Suppose, for example, that there are many 20- to 30-year-olds in the reference sample but few people 70 years or over. Such a sample is not far fetched because it is much like the Korean War skeletons in McKern and Stewart’s (1957) influential pubic symphysis age estimation procedure. If the sample was indeed as heavily weighted toward the young end of adulthood as it is in this example, then the trait, even if actually more typical of living old people than those in their twenties and thirties, would appear characteristic of the latter because it is simply more common among them in this particular sample. Because of sampling, the trait might not even appear among old people in the reference sample if there were only a handful of them. Our prehistoric skeleton, therefore, would seem more like a 30-year-old than a 70-year-old based on this trait, simply because of the biasing effect of the original reference sample’s age structure. Putting all skeletons from a cemetery together, one would arrive at an erroneous age distribution in which mortality seemed to be quite high early in adulthood and to skyrocket from that point onward. Note that the problem is not in the application of the method, including the correct classification of traits measured in terms of low inter- and intraobserver error. It is, instead, a problem with the method itself, specifically the characteristics of the skeletal sample originally used in its development.

These difficulties do not have to signal an end to paleodemography. In fact, several avenues of research are being actively pursued that promise to put the field on a more solid footing. Perhaps the first step is to obtain unbiased estimates of a trait’s distribution in a way that is not affected by the reference sample’s age distribution. One general approach to this problem is to use ordinary least-squares regression to relate a

**BOX 18.1. WHO WAS IN PAKAL’S TOMB?**

In this chapter, we focus on gaining a better understanding of populations, albeit archaeological ones that accumulated over extended periods that often spanned many generations. Osteologists, however, are not always interested in large groups of individuals, since single skeletons, or only a few of them, are often the focus of study. That is common in forensic applications—missing people, homicides, and the like—but this also happens in archaeological investigations. One such example is the mystery of who was buried in an elaborate tomb at the Maya site of Palenque in Mexico (Tiesler and Cucina, 2004, 2006).

The controversy over who was buried in a tomb attributed to Pakal centered not on the individual’s sex but on the age at death (Buikstra et al., 2004, 2006). There was no question that this skeleton was likely to have been from a man. His age, however, presented a problem. When the age was first estimated half a century ago, the skeleton was said to have been from a man who had died somewhere around his fifth decade of life. Epigraphic evidence on Pakal’s reign, however, indicated he was about 80 years old when he died. There are three possible
reasons for such a wide discrepancy in ages: The inscriptions are wrong or misinterpreted, someone else was buried in the tomb, or the original age estimate is incorrect. This problem seemed like a natural one for transition analysis because traditional age estimation methods with open-ended terminal intervals would not be of much help, except to say that the skeleton was indeed old if it conformed to the characteristics of an old-age category (e.g., 50+ years).

Fortunately, the pubic symphyses were, almost miraculously, preserved, and they could be cleaned sufficiently for observation. The transition analysis estimate, based solely on the left and right pubic bones, indicated the man was over 79 years old (lower 95% confidence limit) when he died, using a uniform prior distribution. Although a conservative approach, such a flat distribution is unrealistic since it assumes there was an equal likelihood of a skeleton being found from an individual, say, in his 60s as in his 90s. In the absence of information on the actual age distribution of this Maya population, another preindustrial prior distribution was used, one based on seventeenth-century Danish parish records. Keep in mind that an exact specification of age was not really of interest, only one that was somewhere in the ballpark. The question was simply whether this person happened to have died in middle age or did so as a truly old man (at least by the standards of his time). The revised age estimates ranged from 68 to 91 years (95% confidence interval), with a maximum likelihood of 81 years. Although the point estimate is almost too good to be true, it is the age range that is of most interest.

There can be little doubt that the man in Pakal’s tomb was as old as the epigraphic evidence would indicate. So perhaps this was really Pakal after all.
Frankenberg, 2002; Schmitt et al., 2002). In particular, the reference-sample problem is not eliminated simply by using multiple indicators since one cannot assume that biases from different reference samples simply disappear because they cancel each other out.

When relying on multiple indicators of age, the possibility always exists that separate traits are correlated with each other, so the information they provide is not independent. As a first approximation, we might simply treat the correlation among traits as purely attributable to age, so that if conditioned on age, the traits would be independent (Boldsen et al., 2002; for similar assumptions, see Lucy and Pollard, 1995; Roche et al., 1988). Although the assumption of conditional independence may work for some traits, it is unlikely it holds true for all of them. When traits are correlated at each age, the standard errors of age estimates are biased downward; that is, the age ranges assigned to skeletons are too narrow (Boldsen et al., 2002). This problem has not escaped the attention of researchers, although it is not yet clear whether much more computationally demanding approaches produce markedly different age estimates from those where conditional independence is assumed (Holman et al., 2002; Konigsberg and Frankenberg, 1992).

Finally, we come to the issue of what sort of skeletal markers would be best for age estimation purposes. Two fundamental criteria exist: what Müller et al. (2002) call invariance, and what might be called (if no one objects to some hideous jargon) age monotonicity. Invariance means that the tempo of development of each marker is sufficiently constant across human populations to justify the use of some modern reference sample. Osteologists have for decades sought traits that are relatively age invariant in this sense of the term. Age monotonicity, in contrast, means that the sequence of change in marker state is sufficiently unidirectional so that one state can be classified reliably as younger or older than another. Currently some debate exists about the best statistical approach for capturing age monotonicity. Traditionally, osteologists have relied on parts of the skeleton that show morphological changes that can be classified into a series of distinct, ordered stages, from young to old. That is true of age estimation methods that rely on complexes of traits as well as on individual characteristics of bony structures. Using the pubic symphysis as an example, the studies of Todd (1920, 1921), Meindl et al. (1985b), and Brooks and Suchey (1990) are examples of the former, whereas in McKern and Stewart’s (1957) work, the symphyseal face is broken down into three “components,” each consisting of several stages. It is perhaps time to rethink the overall approach, relying instead on traits with two stages (a single transition), as long as there are many of them and they can be combined in some statistically rigorous manner. There are two reasons for suggesting—and at this point it is only a suggestion—that we would do well to shift our focus to binary traits, each necessarily having only low information content. First, such skeletal characteristics make it computationally easier to put into effect some of the promising mathematical methods that have been proposed to deal with the problems of age estimation (Holman et al., 2002; Konigsberg and Herrmann, 2002). Second, these traits—lots of them in combination—are highly informative about age as they form the basis of an experienced osteologist’s subjective appraisal of the age of a skeleton. There is potential here because such appraisals can be as good as, or better than, traditional methods as shown in blind tests (Boldsen et al., 2002).

ANALYTICAL CONCERNS

Mortality Analysis and the Nonstationarity Problem

Before the 1980s, it was assumed that age at death distributions generated from cemetery samples were essentially a product of mortality. Age-specific mortality rates could then be estimated by aggregating observed skeletal ages at
death into intervals that were often five or ten years wide (the width is irrelevant to the general point) and then treating them as if they were equivalent to the \( \mu_{d_c} \) column in a life table (the stationary age at death distribution). Simple modifications of standard life-table analysis could then be used to back-estimate age-specific mortality rates (Acsádi and Nemeskéri, 1970).

There are problems with a life-table approach to paleodemography. Computing a life table involves estimating one parameter (either \( \mu_{d_c} \) or a central mortality rate) for every age interval, often for each sex separately. Few paleodemographic samples are large enough to support such a data-hungry method. In addition, the use of fixed-age intervals in the life table implies that the ages of all skeletons are known with the same margin of error, including those of fragmentary skeletons that have only a few, perhaps unreliable, indicators of age. Such intervals, fixed in advance without regard to the inherent uncertainty in age estimates for individual skeletons, inevitably misrepresent the error structure actually involved in paleodemographic analysis.

Perhaps the most serious shortcoming of classic paleodemographic life-table analysis is that the original living population, the one from which the skeletons were derived, is treated as if it was stationary. That is, over the duration of cemetery use the ancient population was closed to migration, age-specific fertility and mortality rates were unchanging, and the intrinsic rate of increase equaled zero (Lotka, 1922). Two separate, but related, difficulties exist with that approach. First, assuming a growth rate of zero means that one of the most important things we could hope to learn about a population remains uninvestigated. Second, by making that assumption, one runs the risk of mistaking a difference in fertility and population growth for one in mortality (Horowitz et al., 1988; Johansson and Horowitz, 1986; Sattenspiel and Harpending, 1983; see also Asch, 1976; Wood et al., 2002:Appendix 7.1). That can even occur when the two populations have identical age-specific mortality rates. If mortality was the same in those two populations but fertility was not, then there would have been different growth rates, age at death distributions, and mean ages of death (the last occasionally used as a crude, but flawed, characterization of mortality in ancient times).

To understand the difficulty posed by nonstationarity, one must first delve briefly into the determinants of age at death distributions. Here it is useful to assume the population was stable, although that need not mean it was stationary. As it turns out, most populations closely approximate a stable age distribution at any particular point in time even when fertility and mortality rates change and migration takes place (Bourgeois-Pichat, 1971; Coale, 1972:117–161; Keyfitz, 1977:89–92; Parlett, 1970). This characteristic of age distributions, which is referred to as weak ergodicity (Lopez, 1961:66–68), ensures that stable models usually fit reasonably well, except when populations have experienced major disturbances in the recent past. Although that is always possible in ancient samples, in most instances it would seem appropriate to treat skeletal collections as being from stable populations as long as the cemetery was large, had accumulated over lengthy periods of time, and did not otherwise show signs of containing only a select fraction of the original population.

In addition to being a function of age-specific mortality, skeletal age at death distributions are influenced by the number of individuals at risk of death at each age and that, in turn, is affected by population growth. So the number of deaths at a given age is the product of the age-specific hazard of death at that age and the number of individuals who had reached that age. That is why one might have few skeletons of very old people, such as those in their nineties, even though the risk of death would be quite high compared with what is found at much younger ages. In a stationary population, the fraction of a population that is a particular age is proportional to the probability of surviving from birth to that same age, so the age at death distribution is a reflection of mortality alone. That is not true of a
stable population where the growth rate is something other than zero. Here the number of newborns entering the population each year changes over time through population growth, distorting the age distribution expected under conditions of stationarity. Where there is positive growth, more individuals are born in a particular year than the one before, and so on. So the number of people dying at each age is a function of both the hazard of death and the growth rate.

Suppose we have a cemetery used for an extended period, say a century, by a community where age-specific fertility and mortality rates remained fixed. Assume also that everyone who died in the village was, without exception, buried in the graveyard, all bones were preserved, each individual was excavated, and the age of each skeleton was estimated without error. Under conditions of positive growth, there would be more young people relative to old ones in the cemetery sample than if the population had remained stationary over the same period. That is because each succeeding cohort was larger than the last, and the people of each age buried in the cemetery is a function of the chance of dying and the number of individuals that had attained that age. To take but one age category, say 5-year-olds, there were simply more of them in the population when the cemetery ceased being used than there had been earlier in time. To illustrate this point, consider the last year of cemetery use when there were more deaths of 5-year-olds, in absolute numbers, than took place a half-century earlier when the people who were now 55 were themselves children.

Figure 18.1 depicts the effect of population growth on the age at death distribution. In the upper panel, the solid line is the age at death distribution in a stationary population \( r = 0 \) under a mortality regime specified by a mixed Makeham model (O’Connor et al., 1997; Wood et al., 2002). Different growth rates produce different age at death distributions, shown by broken lines. Two populations are declining \( r = -0.01 \) or \(-0.02\), whereas the other two are increasing \( r = 0.01 \) or \(0.02\). Positive values of \( r \) increase the deaths at early ages relative to those at later ages, as described above. That would, of course, reduce the mean age at death as calculated from the skeletons that accumulated in a cemetery. The lower figure is purely a result of larger successive birth cohorts because the population is growing. Population decline has the opposite effect: The numbers of the young are depressed relative to older people. If one were to assume all five populations were stationary, as is commonly done in paleodemographic studies in the absence of information about the population’s growth rate, the age-specific survival curves shown in the lower panel of Fig. 18.1 would indicate that survivorship was different despite the fact mortality rates were identical. Chamberlain (2006) points out that the problem is easy to fix if \( r \) is known, but it almost never is in paleodemography (see also Frankenberg and Konigsberg, 2006; Moore et al., 1975).

In the absence of migration, only some difference in fertility can account for two stable populations having identical mortality rates but different growth rates. Thus, fertility as well as mortality can contribute to different age at death distributions in cemetery samples. In fact, age at death distributions derived from skeletons are more sensitive to small changes in fertility than to equivalent ones in mortality, as pointed out by Sattenspiel and Harpending (1983). This property has led to age at death distributions being used to estimate fertility. Yet the most that can be estimated about fertility from skeletons is the crude birth rate. Age-specific fertility rates are not identifiable from data derived from skeletons alone. Nevertheless, we may be able to obtain unbiased estimates of population growth, which would be of great importance for the study of ancient societies. Estimates of \( r \) from skeletal analyses would complement those from other sources, such as settlement numbers and sizes. Having multiple independent lines of evidence in such instances is highly desirable as all of these data are subject to a variety of problems.

Several corrections for population growth have been made within the context of life-table
analysis. Half a century ago, Carrier (1958) developed an approach to incorporate growth rates into a life table, making it attractive to paleodemography. Through such a procedure, multiple mortality distributions can be computed from a reasonable range of growth rates (Bennett, 1973; Jackes, 1986; Moore et al., 1975; O’Connor, 1995). This approach prevents us from concluding that two populations had different mortality rates if, in fact, these rates were identical, but one population was growing, whereas the other was not. Unfortunately, it does not provide a means of estimating the intrinsic rate of increase and that, as mentioned, is a figure of great significance in studies of ancient societies. Another

Figure 18.1 The effects of population increase or decrease on estimates of the age at death distribution and age-specific survival (the probability of surviving from birth to each subsequent age) are shown. The upper graph gives the distribution of ages at death in five stable populations with identical age-specific mortality rates but different growth rates ($r$). Age-specific mortality rates are generated by a mixed Makeham model (O’Connor et al., 1997; Wood et al., 2002) with parameter values $a_1 = 0.5$, $a_2 = 0.0005$, $a_3 = 0.005$, $b_3 = 0.05$, and $r_0 = 0.1$. In the lower graph, the probability of survival from birth to each subsequent age is estimated from the same age at death distributions under the erroneous assumption that all five populations are not only stable but stationary ($r = 0$). Positive growth makes it seem as if survival is lower at each age, whereas negative growth has the opposite effect.
way to tackle the problem, which is quite different from Carrier’s, was advocated by Paine (1989a), and it builds on a manual pattern-matching approach initiated by Henry Harpending (Milner et al., 1989). Paine (1989a) starts with a mortality regime consistent with regression coefficients used to generate the Coale and Demeny (1983) “West” Level 7 mortality function. Maximum likelihood methods are employed to estimate mortality rates and the single best estimate of the growth rate conditional on those mortality rates. Although it is an improvement on the earlier life-table approach, it still requires a priori assumptions about the shape of the mortality function. It does not allow unrestricted estimation of both the age pattern of mortality and the population growth rate simultaneously. There is also a practical difficulty: The likelihood surfaces produced are generally relatively flat, preventing the precise estimation of population characteristics, including the growth rate.

These methods involve either directly estimating life tables from skeletons or fitting model life tables to skeletal age at death distributions. But the statistical properties of life-table estimates are well characterized only when calculated from the ratio of observed deaths at a given age during a precisely defined reference period to the number of person-years of exposure within the corresponding age interval (Smith, 1992:108–117). These data are unavailable to the paleodemographer. The other approach, starting with model life tables and fitting them to skeletal age distributions, does not entirely solve the problems paleodemographers face because they are selected by ad hoc methods, and every set of model life tables—including those of Coale and Demeny (1983), the United Nations (1983), and Weiss (1973)—has limits on the range of mortality patterns it can fit.

**Parametric Mortality Models and Maximum Likelihood Estimation**

Parametric models of age-specific mortality (along with the growth rate) estimated by maximum likelihood methods are an alternative to life-table analyses. They provide greater freedom to study the age pattern of mortality by making fewer initial assumptions about the details of the mortality curve. To be useful, these models must be sufficiently flexible to approximate all human mortality experience because the ancient population we are trying to reconstruct must fall within the range covered by the parametric model.

Siler’s (1979, 1983) five-parameter competing hazards model, investigated thoroughly by Gage (1988, 1989, 1990, 1991, 1994, 2005), has found some application in paleodemography. In this model, an individual’s risk of death at each age is determined by three sets of competing causes: juvenile mortality, senescent mortality, and mortality independent of age (“accidental” causes). The juvenile component is specified as a negative Gompertz function and the age-independent and senescent components as a Gompertz-Makeham model (Siler, 1979). Their joint effect varies with age.

The applicability to humans of the Siler model has been examined by Siler (1983), Gage (1990), and Gage and Dyke (1986), and it has been shown to yield important insights into the biology of death (e.g., Gage, 1989, 1991, 1994, 2005; Gage and O’Connor, 1994). It is more flexible than the four-parameter Brass logit model that demographers frequently use (Gage and Dyke, 1986), and it covers a wider range of mortality patterns than the Coale–Demeny model life tables (Gage, 1990). More to the point of this chapter, the model has been shown to span a wide range of mortality patterns reported in the paleodemographic literature (Eshed et al., 2004; Gage, 2005; Nagaoka et al., 2006; O’Connor, 1995). That is extraordinary because at least some, and perhaps most, of these age at death distributions are in part the product of selective burial practices, differential skeletal preservation by age (especially with regard to the very young), and biased age estimates (particularly for adults). The only features of human mortality that cannot be captured routinely are an adolescent mortality hump (Mode and
Busby, 1982) and the apparent deceleration in mortality sometimes observed in advanced ages (Manton et al., 1986). Although more complicated versions of the Siler model can accommodate these deficiencies (Gage, 1989; Wood et al., 1992), it is not clear whether the adolescent mortality hump and the senescent deceleration in mortality are of sufficient concern to warrant a more complicated specification. When all is said and done, additional complexity is pointless when balanced against the small size and biased nature of most cemetery samples, the fact that neither feature is universally found in human populations, and the old-age deceleration occurs in a part of the lifespan that would be represented poorly in the distant past, especially in samples typically consisting of no more than a few hundred individuals. The Siler model would seem to strike the right balance between its potential for telling us something about mortality experience in the past and its robustness in terms of preventing us from being distracted by minor squiggles in the age-specific curve that are of dubious biological significance. Thus, there is at least one flexible model for paleodemographic purposes, although other potentially useful model specifications are discussed by Holman et al. (1997), O'Connor et al. (1997), and Wood et al. (2002).

An additional advantage of parametric models, like the Siler model, is that they provide a means of estimating mortality throughout the entire human lifespan, even where precise age estimates are impossible. Once model parameters are estimated, age-specific mortality for any age of interest can be generated. This possibility is of particular interest because standard age estimation methods typically feature open-ended terminal categories, such as 50+ years, the use of which means that patterns of mortality in old age are effectively inaccessible.

The Siler model and other parametric mortality models can be fit to age at death distributions through maximum likelihood estimation, which has already found a place in the paleodemographic literature (Boldsen, 1984, 1988; Herrmann and Konigsberg, 2002; Holman et al., 1997; Konigsberg and Frankenberg, 2002; Konigsberg et al., 1997; O'Connor, 1995; Paine, 1989a, 1989b). As discussed elsewhere (Wood et al., 1992, 2002), maximum likelihood methods provide a powerful means of fitting parametric mortality models. Furthermore, they allow for estimation of the full distribution of errors that will always be part of skeletal age estimates, and they potentially correct for the distorting effects of nonstationarity. In principle, all we need to yield maximum likelihood estimates of the parameters of interest is an observed distribution of age-informative skeletal markers in the archaeological sample and the distribution of the same markers in a known-age reference sample.

This approach is conceptually fairly straightforward, and special-purpose software is now available to implement it (Holman, 2000). A parametric model is chosen that describes the probability of obtaining each suite of age markers in a skeletal sample, conditional on the prior distribution provided by our parametric model (to be estimated). The product of the individual probabilities is referred to as the likelihood function. New parameter estimates are calculated repeatedly until the parameter values that maximize the overall likelihood function are obtained. Those parameter values (maximum likelihood estimates) have the lowest possible standard error for each parameter estimate, which makes them useful for small paleodemographic samples. Holman et al. (1997) have developed likelihood functions for use with age at death data that can be maximized numerically such that they yield simultaneous estimates of the age-specific mortality function and the population growth rate. Simulation studies verify that the parameters of interest, including growth rate, can be recovered without bias (Holman et al., 1998).

That leaves the question of how best to deal with skeletal ages that are always inexact in the sense that we can never be sure that an individual at the time of death was precisely any given age, which is why osteologists typically use intervals with widths varying from months...
for infants and years or even decades for adults. One way of doing so is to calculate the likelihood of a skeleton appearing in the mortality sample as if it were a particular age, weighted by the probability that it actually is that age based on the observed traits, and integrated over all possible ages. One can think of this procedure as smearing out each skeleton across the age at death distribution, thereby directly incorporating into the likelihood function the range of errors inherent in age estimates. Although this approach is feasible, it will require information on trait and age distributions to be collected that takes into account issues raised earlier about biases introduced by reference samples in standard methods.

**Heterogeneous Frailty and Selective Mortality**

Individual-level heterogeneity in the risk of death is of concern in paleodemography and paleopathology because it implies that all individuals do not have an equal chance of entering the skeletal sample at each age. Eventually all people might be buried in a community’s cemetery, but they do not face the same risk of death at each age as other members of their cohort. Such heterogeneity stems from many sources separated into those that, for sake of convenience, can be regarded as principally cultural or biological in nature. Mortality is influenced through differences in one’s social and economic standing, the work environment, nutrition, growth and development, exposure to infectious agents, and the like. Constitutional differences also affect mortality, such as a familial predisposition to certain cancers or cardiovascular disease, and prior experience with diseases that compromise the immune system leading to greater susceptibility to subsequent infection.

Most demographic, epidemiological, and statistical work on heterogeneity, including the “osteological paradox” (Wood et al., 1992b), emphasizes the interpretive difficulties it can cause. If heterogeneity is “hidden”—that is, not captured by measured variables—it can confound and bias the results of aggregate-level mortality analyses (Manton et al., 1992; Vaupel and Yashin, 1985a, 1985b; Wood et al., 1992). In other words, the apparent behavior of the aggregate, in this instance a human population, need not match that of all the individual components that make up the whole.

Once again, a deliberately simplified example illustrates the problem. It is best to start with a term, frailty, that refers to an individual’s age-standardized relative risk of death (Vaupel et al., 1979). Next imagine a cohort of newborns, each of whom experiences a constant risk of death; that is, it is set at birth and does not change thereafter. We also assume that the relative risk of death varies among newborns so that frailty is heterogeneous.

So we now have a situation in which an individual’s risk of death does not change throughout life, and it is dissimilar to what others, whose risks are also fixed, experience. That does not mean, however, that the aggregate frailty distribution remains the same as the cohort ages. As time goes on, mortality progressively removes the unfortunate children of higher frailty, so the mean risk of death for survivors changes in each successive age interval. Once again, mortality is selective, and because of that, the frailty distribution shifts with age as individuals at greatest risk are preferentially removed from the population (Fig. 18.2). In this example where individual frailty remains unchanged, the mean frailty of survivors declines with age, and so does the aggregate death rate. Thus, changes in aggregate-level death rates, as estimated using conventional life-table methods, are not necessarily indicative of the risks experienced by any single individual who is part of the population of interest.

In this simple example, heterogeneity can only lead to decreasing aggregate mortality as the cohort gets older because the risk of death for every individual is unrealistically assumed to remain fixed throughout life. In human populations, however, the aggregate-level risk of dying increases in old age, an indication that individual risks of dying, at least some large fraction of them, also increase with advancing
Dealing with individual as opposed to aggregate-level risks of dying will be by no means easy with archaeological samples, for all the reasons discussed above. Nevertheless, a recent analysis of medieval Danish skeletons found there was an individual-level decrease in the risk of dying from age 1 to 3. From that point onward to adulthood, however, continuous variation for an age-constant individual risk of dying could account for the aggregate-level decrease in the risk of dying that was observed (Boldsen, 2007).

Although heterogeneous frailty and selective mortality affect all areas of population biology, there is an additional twist in the tale for

---

7This is not to say that heterogeneous risks and selective mortality do not occur at later ages; see the issue of old-age deceleration in mortality in Horiuchi and Wilmoth (1998).
paleodemography and paleopathology. The difficulty is shown in Fig. 18.2 where the smooth curve tells us about the health of the living population, whereas all the osteologist has to work with are the black dots, the people who died at a particular age. Quite obviously, any skeletal traits that may reveal something about an individual's health are observed only in the dead, which are not representative of the living at each age because the mortality sample is highly selected with respect to the overall population's frailty distribution.

The principal point here is that heterogeneous frailty and selective mortality must be added to the paleopathologist's long-standing concern over the low sensitivity and specificity of bony lesions as markers of particular diseases. By that we are referring to the fact that not everyone who has a disease also develops a characteristic bony lesion, and not all observable skeletal lesions are equally indicative of one particular disease. These two issues are widely recognized by osteologists as creating difficulties when trying to characterize the disease experience of past populations.8

Several researchers have been concerned with the effects of heterogeneous frailty and selective mortality in cemetery samples (e.g., Cook, 1981; Cook and Buikstra, 1979; Guagliardo, 1982; Ortner, 1991; Palkovich, 1985; Saunders and Hoppa, 1993; Wood et al., 1992b). Despite this work, it is commonly assumed, implicitly if not explicitly, that skeletal or dental indicators of poor health in mortality samples are a direct reflection of the prevalence of particular conditions in once-living populations. Cohen (1994) has gone as far as arguing that this straightforward relationship is indeed fully warranted because the conditions that caused pathological skeletal or dental features commonly observed in archaeological samples had little, if any, effect on the risk of dying. Waldron (1996), although acknowledging the effects of selective mortality for diseases that increase the risk of death, also took the position that for many diseases the skeletal lesion frequencies are equivalent to their prevalence in living populations. Something of the reverse of this argument would be to believe that skeletal lesions among modern (or pre-antibiotic era) patients suffering from particular diseases would occur in the same proportions in cemetery samples, minus the bony lesions that are not sufficiently distinctive to be attributed to a specific cause. Following this logic, if the archaeological sample would seem to have too many skeletal lesions indicative of a particular disease, then the identification of the disease either must be wrong or more than one disease was necessarily present.

It is undeniably true that if a certain condition had absolutely no direct or indirect effect on the risk of death, signs of that condition in a mortality sample should be proportional to their prevalence in the living population. But then there would be no point in treating those particular skeletal or dental traits as a means of assessing the health of prehistoric people, as viewed through the diseases they experienced—this sort of work is typically phrased as the biological costs of pursuing particular ways of life. In other words, there must have been some associated discomfort or disability since the existence of a disease indicated by a bony lesion implies a departure from some ideal state (pure health, if there ever was such a thing). That, in turn, translates however weakly and indirectly into the risk of dying, bringing us right back to where we started: a concern over selective mortality.

It cannot be emphasized enough that we are not referring only to those conditions that directly lead to death but to any that are associated with an increased risk of death for whatever reason. For example, short stature, unlikely to kill you in and of itself, may well be correlated with socioeconomic status, which may indeed

---

8 Recent work promises to do much to solve the problem of underestimation resulting from reliance on classic lesions—those arguably attributable to a particular disease and only that disease—through a new approach that simultaneously estimates specificity and sensitivity (Boldsen, 2001, 2005b; Boldsen and Mollerup, 2006).
influence the relative risk of death. An archaeological example mentioned earlier—the 700-year-old community in Illinois that suffered so terribly at the hands of their enemies—illustrates this point (Milner et al., 1991). The skeletons of people who were killed often showed signs of debilitating conditions, including active infections, partially healed bone fractures, and dislocated limb joints. Although these unfortunates did not die from these illnesses and injuries, their ability to fight or flee when threatened was certainly impaired. An incompletely healed rib fracture or dislocated hip did not by itself cause an individual’s death, but it could certainly increase that person’s vulnerability when attacked. So for any condition resulting in diminished physical capacity, even the victims of violence in a cemetery sample cannot be assumed to be representative of all people once alive in a particular community.

The relationship between a pathological lesion and the individual’s frailty is by no means straightforward. Osteologists often differentiate bony lesions that were active at the time of death from those that had healed. Yet this distinction between coarse woven bone and smooth remodeled bone highlights a deeper question that, to our knowledge, was first raised by Ortner (1991): Is it good or bad to have a skeletal or dental indicator of earlier ill health? Such conditions indicate survival after disease, nutritional deprivation, or trauma long enough to undergo a recognizable alteration of the normal structure of bones or teeth. Take, for example, skeletal signs of an infectious disease. The mere fact that some people had a skeletal lesion indicates they were healthier than others who succumbed before one had time to form. But undoubtedly some community members shrugged off the infection before it affected their skeletons in any recognizable way. Of course, there would always be those who might never have been exposed to that particular pathogen, so they too would lack skeletal signs of infection. The same could be said for people who never experienced nutritional hardship or rarely engaged in activities that were likely to lead to particular kinds of bone fractures. Thus, it is by no means a simple task to sort out what the presence or absence of any skeletal or dental indicator of ill health might mean, regardless of whether it is directly related to a specific disease or the result of a number of causes, such as stunted growth.

The first step, as is commonly done, is to look at the age distribution of individuals with particular skeletal or dental markers of poor health. That is because lesion frequencies in two samples can be similar when individuals of all ages are considered in aggregate (or lumped into broad categories such as juvenile versus adult), but the patterning across age groups may be much different. For example, it has been shown that the age-specific risk of developing dental decay in permanent molars was not the same for men and women in a medieval Danish village (Boldsen, 1997). Yet the proportions of men and women with decayed teeth were similar when they were treated simply as adults. So when the age structure of the skeletal sample was ignored, the difference between the sexes in their experience with dental decay disappeared. Thus, mortality effects are demonstrable even with conditions that are not normally thought of as life threatening. Another example, also from medieval Denmark, underscores the association between bad teeth and mortality, since advanced dental attrition was associated with an increased risk of death (Boldsen, 1991, 2005a). Occlusal attrition, in this instance, serves as a marker of whatever social or environmental conditions acted to shorten the lifespan. Although people presumably did not die because they had badly worn teeth, they were nonetheless winnowed out of the population at a higher rate than their fellow villagers, although we do not know from their teeth alone why that might have taken place. The result is that they were disproportionately represented among the young adults in the cemetery sample. In short, skeletal lesions, when distributed across the
lifespan in some patterned way, can provide information about heterogeneity and selectivity. Other such examples can be found in the osteological literature, such as higher stature in adults being related to longer life in medieval to early modern Europe (Kemkes-Grottenthaler, 2005).

Several steps need to be taken. First, measurable characteristics of skeletons must be identified that are informative about an individual’s frailty. Saunders and Hoppa (1993), in a thoughtful review of the role played by selective mortality in the formation of cemetery samples, concluded that other difficulties that plague archaeological skeletal samples, such as differential preservation and biased age estimation, are likely to be more worrisome. That will undoubtedly often be the case, but the additional effects of selective mortality do not simply disappear. Second, a robust theoretical framework needs to be developed about the distribution of frailty and how it might change with environmental, economic, and social conditions. Wood (1998) has started to examine how frailty distributions vary with economic and demographic change, and Weiss (1990) has drawn on models from population genetics to help explain variation in frailty.

The possibility that heterogeneity and selectivity can be identified in archaeological samples with skeletal traits has also begun to be explored. In one study, the proportional hazards model was generalized to allow for both a main effect of the presence of skeletal indicators of health on the age at death, and an interaction between the bony traits and age at death (Usher et al., 1997). The interaction term captures age-specific changes in the frailty distribution attributable to selective mortality. After being validated on simulated data, the model was applied to two archaeological collections, one being a medieval Danish cemetery and the other a 700-year-old pre-Columbian site in Illinois. Even with comparatively small samples similar to those from many archaeological sites, there was clear evidence of heterogeneity and selectivity in femur length, with more equivocal results using porotic hyperostosis and cribra orbitalia. More recently, Thomas (2003) has detected significant heterogeneity and selectivity in enamel microdefects in teeth from the same medieval Danish sample.

What is needed is a comprehensive analytical framework to handle all aspects of the heterogeneity and selectivity problem simultaneously, and a step in that direction has been taken by Usher (2000) who developed a modeling framework based on stochastic renewal theory and multistate survival processes to study heterogeneity and selectivity. With respect to a particular bony lesion, a skeleton can be classified as “well” at the time of death (the lesion was absent), “ill” (the lesion was active at death), or “healed” (the lesion had healed). Living newborns enter the “well” state through a process adjusted for population growth, and they are assigned an initial frailty value drawn at random from a gamma distribution. The hazard rate for each transition (e.g., “ill” to “healed” or “dead”) is modeled as conditional on the individual’s frailty, which allows the frailty distribution to change with age and state. “Dead” individuals, the only ones osteologists observe, are a mixture of those from the “well,” “ill,” and “healed” states. The age at death distribution and age-specific frequencies of active and healed skeletal lesions provide the information from which the model can be estimated by maximum likelihood methods (for examples of multistate model estimation, see Manton and Stallard, 1988; Wood et al., 1994). Parameter estimates permit inferences about the prevalence of lesions in the living population, the effect of the conditions that produced the lesions on the risk of death and the age-specific frailty distribution. More recently, DeWitte (2006) has used Usher’s model to examine selective mortality associated with the medieval Black Death.

Something like this model should move us from an exclusive concern about the possible confounding effects of heterogeneity to an understanding of its central role in population dynamics. Such models require us to use both paleodemographic and paleopathological
data—that is, ages at death and age-specific lesion frequencies—in the same analysis. That is why paleodemography has such a major role to play in the future of paleopathology, especially when seeking to determine how healthy one population (cemetery sample) was relative to another.

CONCLUSION

Several troublesome issues and ways to address them have been identified in this overview. These issues include age at death estimation and population nonstationarity in paleodemographic analyses as well as heterogeneous frailty and selective mortality at the intersection of paleodemography and paleopathology (paleoepidemiology).

This chapter is intended to be an optimistic appraisal of the field. The theoretical issues—what we would like to know about past societies and how we can go about doing so with skeletal samples—are significant. Difficult methodological problems have been identified over the past few decades, and some of them have been solved, whereas others are the subject of active study. Because of those difficulties, the field has been repeatedly pronounced dead. But however appropriate that description might be of the subjects we study, it is at best premature when applied to the field as a whole. The methodological problems that remain are not easy to solve, but progress is being made. Paleodemography will not be killed by them, but it will be if they are simply ignored or declared unimportant at the outset. That is why it is essential to continue probing them with new analytical approaches, such as those identified here. Paleodemography—and by that we mean a better understanding of population characteristics and health throughout the greater part of human existence—will be forever impoverished if we do not do so. After all, for the most part, the kinds of information provided by rigorously conducted paleodemographic analyses cannot be obtained from any other source.

ACKNOWLEDGMENTS

We thank the editors for inviting us to revise and update an earlier chapter in a volume they also edited (Milner et al., 2000). The two versions are similar, and sometimes identical, although in this chapter certain points are discussed more fully and mathematical detail has been removed throughout to make it more accessible to the general reader. Those who are interested can refer to our earlier chapter and a review by one of us (Wood, 2003) for more technically demanding introductions to the field and where it is going. Several people, including Rebecca Ferrell Thomas, Darryl Holman, Lyle Konigsberg, Kathy O’Connor, and Bethany Usher, kindly made much-appreciated comments on drafts of the original chapter, so by extension they contributed to this version as well. Figure 18.1, reproduced from the original, was produced by Darryl Holman.

REFERENCES

Aykroyd RG, Lucy D, Pollard AM, Roberts CA. 1999. Nasty, brutish, but not necessarily short: A reconsideration of the statistical methods used...
to calculate age at death from adult human skeletal and dental age indicators. Am Antiq 64: 55–70.


REFERENCES


Snow CE. 1948. Indian Knoll Skeletons of Site Oh 2, Ohio County, Kentucky. Reports in Anthropology 4(3, 2). Lexington: University of Kentucky.


INTRODUCTION

It has been a quarter of a century since Bocquet-Appel and Masset (1982) raised several questions concerning the recovery of mortality, fertility, and age structure from skeleton-based demographic studies. These questions have resulted in constructive debates and reexaminations of models and assumptions. Although there have been advances in method and theory in paleodemography (Bocquet-Appel and Demars, 2000; Hoppa and Saunders, 1998; Hoppa and Vaupel, 2002; Jackes, 1992, 2000; Konigsberg and Frankenbury, 1992; Meindl and Lovejoy, 1989; Meindl and Russell, 1998; Paine, 1997; Sattenspiel and Harpending, 1983; Wood et al., 1992), some debate still exists over whether useful demographic inferences can be drawn from skeletons. Moreover, demographic studies of modern, so-called “traditional,” populations have their own unique challenges (Leslie and Gage, 1989) and biases (Campbell and Wood, 1988). We maintain that archaeological populations offer special and important views of the biology, cultural evolution, and vital structure of our species that are not available from other fields of anthropology.

The questions raised in 1982 relate directly to three major biases in the way we have conducted the demographic analysis of prehistoric societies and to the structure of this chapter. First and foremost, paleodemography has always been a sampling process that may involve several potential sources of error. The magnitude of this potential bias is difficult to quantify, and of course, depends on the particular site under investigation. Aboriginal inhumation practices, soil and other conditions of interment, subsequent disturbances of the burials, and the care and skill of excavators may all affect how accurately a cemetery represents the population that produced it.

It bears repeating that if “an unbiased, representative sample . . . cannot be assumed, further demographic analysis is not likely to be productive” (Weiss, 1973:58). It is not the purpose of this chapter to present and evaluate paleodemographic methods for sites with scattered or poorly preserved remains. Nor is it
possible to estimate the levels of bias in preservation by age and sex and to attempt corrections, when it is clear that some differential loss of skeletons has occurred. However, our experience with skeletal sites in the eastern United States leads us to conclude that the effects of differential preservation on the field of paleodemography (Jackes, 1992; Masset, 1973; Walker et al., 1988) are overstated. This is not to say that all sites are demographically useful. In fact, most are not; however, the potential for this kind of bias is obvious from simple inspection of the burial pits, the frequency of disarticulation, postmortem bone damage, and general conditions of the site. Bone and tooth survival is a function of root and rodent displacement, soil drainage, and especially pH (Gordon and Buikstra, 1981). Unfavorable conditions cause the destruction or movement of hard-tissue remains with such rapidity that sites of moderate antiquity fall essentially into two groups—those with demographically useful skeletons, and those without. An intermediate site in which the human remains are actually in the process of disappearing would have such fragile contents that even the adult sample would be in poor condition.

Both Sundick (1978) and Saunders (1992 and this volume) have argued that in such sites the hard evidence required to recognize and determine the age of a subadult (i.e., primarily teeth) is preserved at least as well as the bones of a robust adult skeleton. Nevertheless some added provision may be needed in the wake of moderate mechanical disturbances to count accurately the numbers of infants and very young children. In nearly all instances the archaeological data themselves should demonstrate clearly which sites are useful to paleodemographers.

Certainly, burial customs associated with child homicide, especially infanticide, have skewed the findings from many cemeteries (Saunders, 1992). Conditions that predispose cultures toward (typically female) newborn and child homicide include very short interbirth intervals, high costs of transporting infants over large distances, customs of killing children who are orphaned or malformed, efforts to limit family sizes, and greater values placed on males as economic providers or as combatants (Daly and Wilson, 1988; Divale and Harris, 1976; Hrdy, 1995). Infanticide rates seem to be associated moderately positively with fertility levels.

Ethnographic evidence indicates that the old and infirm may be killed occasionally, but their remains rarely are treated differently from those of other adult decedents. Graves containing crippled or injured elderly individuals imply that efforts were made to place all of that society’s members in a common burial ground. A few, especially young males, may have died some distance away from the habitation/cemetery. The presence of bundle burials (i.e., only crania and assorted long bones) testifies to a prehistoric sense of place, to a general effort to return the kinsmen’s essential remains to common burial, and to an important piece of evidence for that cemetery’s value to demographic study.

A second central problem in paleodemography involves potential biases in the direct estimation of adult skeletal age. Whenever traditional bony sites play even a moderate role in skeletal aging, a tendency exists to underestimate the ages of the cemetery’s oldest burials. One result of this well-known regression problem is usually an overestimate of the proportion of adults between ages 30 and 40 years.

A large body of literature is devoted to skeletal age and to model-based corrections of age distributions during the past decade (in English, and too numerous to list individually), but many are contained in a few edited texts (Hoppa and Vaupel, 2002, this volume; Katzenberg and Saunders, 2000; Paine, 1997; Saunders and Herring, 1995). Additional general and/or statistical critiques have also appeared recently (Aykroyd et al., 1999; Bocquet-Appel and Masset, 1996; Koningsberg and Frankenberg, 2002; Lucy et al., 1996; Meindl and Russell, 1998; Paine and Harpending, 1996). New modifications, recalibrations, and assessments of our auricular
surface method (Lovejoy et al., 1985b) have continued to emerge well into the new millennium (Buckberry and Chamberlain, 2002; Igarashi et al., 2005; Mulhern and Jones, 2005). However, numerical scoring procedures for the auricular surface of the ilium as well as other bony sites remain problematic. We sympathize with the need for repeatable objective criteria for recovering cemetery age distributions (Jackes, 2000); however, it is not yet clear that the numerical scoring methods have offered any improvement in the accuracy of age estimation so far (e.g., see comment, Storey, 2005). It is not even clear whether they have resulted in estimates (Herrmann and Konigsberg, 2002) that differ much from previous ones (e.g., Kelley, n.d.), although some do represent improvements in the appropriate direction, i.e., increased survivorship and reduction of the “mid-age bulge.” In any case, despite calls for new directions in paleodemographic research (Hoppa and Vaupel, 2002), a much greater problem in paleodemography remains than the adult aging bias as highlighted in the “Rostock Manifesto.”

Aside from the representativeness of a site, the third major issue in paleodemography is that for any given prehistoric cemetery, each of a continuum of stable populations could have filled it with exactly the same age proportions. Whenever the typical assumption of stationarity or near stationarity (i.e., the intrinsic rate of growth is zero or very small) is imposed on the paleodemographer’s model of an extinct population, which in fact had been growing during the occupation of the site, an inevitable and serious bias occurs. The result of underestimating prehistoric growth is that measures of health and longevity, measures of fertility, and the proportion of young persons in the series of once-living cohorts are biased downward, sometimes by very great margins. The implications are important. For example, what does a very young cemetery age distribution from the period of early adoption of agriculture imply? A moderate increase in fertility signaling the success of a cultigen-based economy and a healthy population growth rate is certainly one possibility. But another is exactly the opposite: the nutritional limitations of undiversified agricultural products, a considerable increase in childhood mortality, and a reduced population growth or even decline. By itself, the cemetery could be interpreted in either fashion or, more appropriately, in a whole range of fashions.

We suspect that this bias is far greater than most instances of the cemetery-aging bias addressed by recent workers (Hoppa and Vaupel, 2002; Jackes, 1993). We would also make the case that recognizing this bias will go a long way toward explaining decades of unreasonably low estimates of life expectancies in the field of paleodemography. In fact, new views about the salient demographic characteristics of the human primate, its Pliocene ancestors, and its hominoid relatives suggest that both the levels and variances of human fertility rates in the prehistoric record have been underestimated substantially.

This chapter offers an analysis of the Libben site of Ottawa County, Ohio, and it represents an attempt to address all three biases. In addition, we emphasize that paleodemography requires perhaps the most interdisciplinary efforts of all the fields of anthropology. 1) The archaeology of the Libben site addresses the first bias (above) and includes a brief description of the western Lake Erie Basin material culture, the nutritional resources that were exploited by the population in this woodland setting 1000 years ago, the pattern of settlement spanning some ten generations, the excellent condition of the burials, and the extreme care exercised in recovering all aspects of this site. 2) Traditional methods of skeletal biology were used in estimating the ages of infant, children, and adolescent burials, for analyzing the remains of hard-tissue lesions and for determining the sexes of the adults. However, to address the second bias we depart from traditional skeletal biology in one important way: The older adults in the Libben cemetery were aged using the auricular surface of the ilium, with very little reliance on other commonly used osteological sites (e.g., cranial suture closure, pubic symphyseal metamorphosis, or fourth sternal
Ethnographic anthropology has provided surveys of the total fertility of women of non-contracepting populations from Africa, Asia, and especially South America. To address the third bias, we have selected the fertility estimates of those Neotropical populations properly termed “colonizers,” and we have used these in conjunction with the cemetery age distribution to complete the paleodemographic reconstruction (age structure, mortality, and growth) of the people of Libben.

THE LIBBEN SITE, OTTAWA COUNTY, OHIO

The Libben site is located on the north bank of the Portage River in Ottawa County, northern Ohio, and it represents one of the largest, best censused, and homogeneous skeletal populations ever recovered in North America (Lovejoy et al., 1977). Radiocarbon dates indicate an occupation from the eighth to the eleventh centuries A.D. A thousand years ago this site was situated within an extensive elm-ash swamp forest, covering freshwater marshes and streams with very abundant fish and mammal resources. As in many riverine Late Woodland sites, cultivated maize and small amounts of shellfish were present, but the people at Libben seem to have relied very heavily on a trap-and-weir economy. They captured small game and deer from the marshes and prodigious quantities of fish from the Portage River, streams, and Lake Erie. Refuse pits contained migratory waterfowl and mammals of various sizes. These pits included white-tailed deer and muskrat. Quantitative analysis of pit contents revealed proportions of meat resources by weight: For every kilogram of dressed poultry, there were 4 kgs of venison and muskrat meat, and more than 20 kgs of edible fish.

The primary goals of the Libben project remain the estimation of the local late Woodland population structure as well as the elucidation of human demographic evolution. To these ends, this site provides a unique opportunity to estimate the mortality and age composition of a prehistoric, precontact population. Excavation of the cemetery/living site proceeded with the skeletal sample as its primary objective. Each skeleton was fully exposed and then drawn, mapped, and photographed by a senior excavator. Preliminary age, sex, and pathology were noted before each burial was removed for additional analysis. The postcrania of subadults were monitored carefully for all possible secondary centers of ossification. Gall, kidney, and bladder stones were collected. Great care was taken to recover every element of all juvenile dentitions. The bulkheads separating excavation squares were removed systematically. These procedures yielded several embryonic and fetal skeletons, the earliest being from the first trimester. More systematic examinations of these skeletal remains have been carried out extensively over several years of postexcavation.

Libben was a permanent fishing village with clear evidence of long-term, year-round habitation. It was situated on a sandy knoll and completely surrounded by river and marsh. A partial enclosure was built on the marsh side that may have been replaced or repaired as many as three times during the site’s occupation. The largest dimension of this two-acre site is east-to-west, with the greatest density of burials near its center. About 1300 articulated burials were recovered from Libben. Almost all burials were extended, there being only very small numbers of bundles and cremations. Most burials were oriented north-to-south, but no relationships to age, sex, or artifacts have been uncovered (Schick, n.d.). Burials were found throughout the habitation site, and unlike many other Ohio Woodland sites—or modern Christian, Islamic, and Jewish cultures for that matter—no segregation of living and burial spaces occurred. At Libben, there was continuity from life to death.

Libben’s continuous habitation for a quarter of a millennium has resulted in one of the richest recovered assemblages in North American prehistory. It includes extensive
domestic artifacts, ceramics, and faunal remains (Lovejoy et al., 1977). Pottery was primarily coiled and grit tempered, relatively thick, breakable, utilitarian, and only minimally decorated. No statistical association between type and/or characteristic of artifacts with location or depth within the site has been demonstrated. Clearly, the site is essentially homogeneous, with a stable occupation that seems to have spanned roughly ten generations, which is a time period that represents about half of the Late Woodland phase of northern Ohio.

Archaeological and systematic analysis of the entire population has revealed almost no evidence of violent conflict between other villages or within the population itself (Lovejoy and Heiple, 1981). The site precedes the first ethnographic accounts of Algonquian-speakers (who staunchly defended tracts of hunting land) by some 700 years. Although long bone fractures were numerous, almost all were caused by accident, and a high association exists with age, most fractures occurring either during adolescence or old age. Only two cases of blunt trauma to the cranium were observed in the entire population. Projectiles imbedded in skeletons were rare and no “battered individuals,” as are sometimes recovered from other sites, were recovered. Detailed analysis of fracture patterns showed that most were associated with general activity and probably were not defensive injuries; for example, Colle’s fractures (compression of distal radius) were frequent; however, only two “parry” fractures (ulnar midshaft) exist in the entire cemetery. No evidence was found of any kind of child “abuse” as small fractures were virtually never observed in infants or small children. Nevertheless, as in the cases of the first sedentary populations from the western Kentucky Late Archaic (Meindl et al., 2001; Mensforth, 2001), a depression of adult male survivorship exists compared with that of females.

Additional instances at Libben exist for which causes of death have been diagnosed readily. These instances included various neoplasms and several obstetric outcomes, including a shoulder dystocia, a transverse lie, and a tubal pregnancy that were likely sources of fatal hemorrhage. A few graves contained multiple and obviously contemporary burials, suggesting brief episodes of infectious disease. Several Late Archaic and Woodland archaeological sites have also revealed evidence of sedentism in the eastern woodlands (Meindl et al., 2001), and like these, Libben manifests high frequencies of incipient periostitis, sometimes having clearly advanced to systemic osteomyelitis. Patterns of porotic hyperostosis suggest an auto-induced iron deficiency anemia as a probable defense mechanism against common infections (Mensforth et al., 1978).

A few cases of true chronic immobility recovered, together with systematic analysis of fracture healing indicate a lack of any significant migratory activity on the part of the population. In recent years it has become popular to cite various long bone injuries and to suggest that these indicate “caring” on the part of other society members, and that such individuals could not have survived without a strong, typically human, social network. However, similar injuries have been reported for many nonhuman primates, which obviously lack such a network (see Bramblett, 1967; Buikstra, 1975; Schultz, 1939). In addition, indirect evidence exists of essentially the opposite, viz., that attention and immobility were sufficiently prized as “patient care” attributes that some injuries unlike those observed in other primates are recovered. In this regard, one adult Libben burial of advanced age exhibited profound lower limb atrophy typical of spastic paraplegia, which testifies to a complete absence of any seasonal mobility. Moreover, the pattern of fracture healing in the population was remarkable, with very high levels of proper alignment and only minimal loss of long-term mobility. In contrast to our surveys of late Archaic skeletons of the eastern Woodlands, the people of Libben knew very well how to “set” long bone fractures. Clearly, enlightened, true instances of “social care” existed, and many examples of lengthy immobilization were required for proper fracture healing (Lovejoy and Heiple, 1981). The
accumulated skeletal evidence for absence of seasonal mobility is in full accord with the analysis of vegetal and mammal remains, the rich abundance of local resources, the partial enclosure, and the very presence and apparent completeness of the cemetery itself.

Libben was an egalitarian society, and death and burial were important rites of passage for all of its members. Only two kinds of nonarticulated burials were included in the demographic sample. First a few bundle burials were composed only of skulls and major long bones, suggesting a general effort to return all kinsmen back to the home burial ground (see the earlier discussion). A second exception to the “articulation rule” for demographic inclusion was the recovery of patently fetal or newborn long bones, which might have been destroyed partially or scattered over a small area by the action of plant roots, rodents, or other agents (e.g., the Libben farm contained a peach orchard for many years). Such tiny burials were enumerated in the census, but care was taken not to double-count them (i.e., each along with any other associable long bones of the same age found within a 2-m radius became a single addition to the demographic cohort).

THE OSTEOLOGY OF AGING

Subadults were aged primarily by analysis of their dentitions, each of which was scored for both crown and root development—eruption data were employed only rarely. Epiphyseal closure, long bone lengths and breadths, and basicranial dimensions were used in those rare cases in which dentitions were not recovered. All burials with unfused root apices and/or partially calcified crowns were seriated systematically within the entire population of subadults (Lovejoy et al., 1985a). Long bone information may have been used to confirm or adjust their placements within this series. Nearly 18% of the Libben burials are infants, and a large proportion of these are neonates.

Standards for adult age based on the auricular surface of the pelvis were developed in the laboratories at Kent State University and the Cleveland Museum of Natural History (Todd–Hamann Collection) (Lovejoy et al., 1985b). Unlike other bony sites, the auricular surface exhibits developmental-like changes well into the fifth decade of life, before it finally becomes dominated by obvious degeneration (Meindl and Lovejoy, 1989; Meindl and Russell, 1998; Schwartz, 1995). Moreover, there are very similar, protracted sequences of change in the auricular surfaces of adult African apes with age, unlike fourth rib articulations, ectorcranial suture closure, or pubic symphyses (Bedford et al., 1993; Lovejoy et al., 1995, 1997). Such analogy suggests that predegenerative age metamorphosis of the auricular surface is uniform across hominoids and therefore independent of biomechanical or any other environmental factors within the human species.

Buckberry and Chamberlain (2002) have devised a quantitative scoring system using some of our components on the presumption that such objective techniques make it easier and more replicable for paleodemographic analyses. Unfortunately, their series of Bayesian probabilities seems to be unsmoothed, probably because of limited sample sizes (e.g., a zero probability of an individual being 55–64 years in their stage 7, yet about 25% probability for both adjacent decennial categories, 45–54 and 65+; see Storey, 2005). Limitations exist to the use of the auricular surface for assessing extreme age, and we never intended to quantitatively score or overly interpret instances of advanced degeneration, in which environmental factors may indeed play a role.

The original multifactorial age distribution of the Libben adult burials (Lovejoy et al., 1977) has now been recalibrated by means of the auricular surface using our original standards (Lovejoy et al., 1985b). First, all adult auricular surfaces were seriated within sexes (Lovejoy et al., 1985a). Next, in both the male and the female adult sequences, absolute ages were assigned. Some of the new ages, especially the very oldest, were increased over previous estimates by more than five years. Degenerated surfaces at the ends of both sequences were...
not very numerous and were given our maximum estimates for extreme age in Libben adults. These surfaces in fact received age estimates commensurate with reasonable terminations for Type I mortality and survivorship functions. By contrast pubic symphyseal methods, which derive nearly all of their correlation with age from one developmental event, played no role in assessing the ages of the oldest Libben decedents. The closure of the “ventral rampart,” an epiphysis that is uniquely delayed in humans, has no real age-predictive power much after full fusion (Meindl and Lovejoy, 1989). Nevertheless, methods based partly or largely on the pubic symphysis continue to be used to attempt estimation of entire adult cemetery age pyramids (Boldsen et al., 2002; Herrmann and Konigsberg, 2002). Some pubic data were included in the Libben analysis but only for alignment of the ages of the younger adults, which often amounted to a few years downward (Meindl et al., 1985). With the discarding of other skeletal age indicators, the average age of the Libben adults was increased by about four years; the average age of the whole cemetery by two. This increase in the average skeletal age and its variance represent an improvement over the previous estimates.

TOTAL FERTILITY AND INTRINSIC GROWTH

In Table 19.1, the no-growth paleodemographic model for Libben is compared with South African !Kung, Paraguayan Ache, and Venezuelan Yanomamo mortality and fertility structures (Hill and Hurtado, 1996; Howell, 1979; Weiss, 1975). If Libben had been indeed stable throughout its occupation and reflective of a nongrowing population (i.e., demographically stationary), then the conclusion must be that it had low total fertility like the !Kung as well as poor adult life expectancy like the Yanomamo. Such levels are artificially extreme, and other reasons exist why such a demographic profile is unlikely to be accurate.

A fundamental problem in paleodemography exists that greatly surpasses osteological aging bias. It involves a primary demographic metric, which generally is not known. That is, for any prehistoric population, how is its annual intrinsic rate of increase ($r$) to be estimated? Without some notion of the Malthusian parameter, the paleodemographer has one too many unknowns or one too few equations. Modern colonizing populations such as the horticulturalist Yanomamo and the forager Ache experienced sizable but variable rates of intrinsic increase during the previous century. Twenty-nine Yanomamo villages may have produced annual growth rates between 0.5% and 2%, although this very general estimate is “based on consideration of the total tribe” (Neel and Weiss, 1975:34). Greater local variation may have occurred, but the Yanomamo data are inadequate for consideration of this question. Ache intrinsic growth, determined from census data, ranged between 1.5% and 3.5% annually, and they were especially high after 1973 because of sedentary confinement to the missions (Hill and Hurtado, 1996:415). This process apparently greatly increased life expectancies, but left the fertility function virtually unchanged. Although the

<table>
<thead>
<tr>
<th>Population</th>
<th>Infanticide?</th>
<th>Longevity ($e_{15}$)</th>
<th>Fertility (TFR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Libben (Prehistoric Ohio, no growth)</td>
<td>Unknown</td>
<td>24.5</td>
<td>5.2</td>
</tr>
<tr>
<td>Dobe !Kung</td>
<td>Very low</td>
<td>54.1</td>
<td>4.9</td>
</tr>
<tr>
<td>Forest Ache (Paraguay)</td>
<td>High</td>
<td>43.3</td>
<td>8.1</td>
</tr>
<tr>
<td>Yanomamo (Venezuela)</td>
<td>High</td>
<td>27.5</td>
<td>8.0</td>
</tr>
</tbody>
</table>

Hutterites, despite their late age at marriage, grew at the astonishing rate of about 4% for the past century and a quarter, nearly all in the sample of 70 traditional (noncontracepting) populations surveyed by Cambell and Wood (1988) do not approach the Ache extreme. Yet what does this tell us? Except for the Mormons, some eighteenth-century Canadian pioneers (and of course the Yanomamo and the Hutterites), the other 66 populations reveal low growth and “remarkably heterogeneous” fertility rates (Campbell and Wood, 1988:43). In fact, so much variation exists that the search for any systematic fertility differences among foragers, horticulturalists, and agriculturalists was unsuccessful.

Campbell and Wood appropriately focus on the total fertility rate (TFR, corrected for its discrepancy with marital fertility). The TFR is an ideal measure of completed fertility performance, representing the average number of live-births to women who survive to the end of their reproductive span. Bentley et al. in 1993 joined the effort to partition the massive variance of the TFR among subsistence categories. They isolated what they determined were unreliable primary data, transitional populations, and considerable pseudo-replication in the Campbell/Wood sample; however, “most of the corrections in fact [made] little difference” in the explanatory power of subsistence models (Bentley et al., 1993:783).

Most important, the original and corrected samples of those studies produced mean fertility values and intrinsic growth rates that were considerably lower than those of the few extant expansionist examples (in which the real explanation for the variance in TFR probably lies). Given what we now know about the evolution of the human species and its colonizing characteristics compared with other hominoid primates, the customary paleodemographic presumption of no growth (i.e., stationary demographic conditions) is not realistic in most instances, and certainly not with the resource abundance characterizing the Libben area a millennium ago. And yet, even moderate variation in the hypothetical intrinsic growth rate of any extinct population results in great ranges of predicted living age structures and mortality profiles.

A cemetery-based stable demographic model requires that each age-class of skeletons be discounted as a function of the intrinsic growth rate, so that the estimation of prehistoric mortality rates takes proper account of the age structure of the once-living population. (See Weiss, 1973:72; or Moore et al., 1975, for exponential corrections to \( n \) to account for any nonzero value of \( r \). Calculation of a new growth-corrected life table follows directly from this.) Stable growth correction must be applied to the estimation of the age-specific maternal fertility function as well.

As no predictive relationship exists between the pre-auricular groove or any other element of the adult female skeleton to past parity (Spring et al., 1989), we use Weiss’s (1973, ch. 6) archetypal fertility pattern averaged from 79 nonindustrial populations and “then adjust[ed] the level by a constant factor” (Keyfitz and Murphy, 1965:152). A variety of mortality, fertility, and age structure measures can be calculated from any such generated life table and fitted fertility function (Weiss, 1973). Using these algebraic procedures while manipulating the value of \( r \) generates a continuum of complete demographic solutions for any given cemetery (Fig. 19.1). The next question is: which of these solutions best describes the average prehistoric demographic profile?

Sattenspiel and Harpending (1983) examined variations in life expectancy and crude birthrate within several of the West series of the Princeton models of mortality and age structures (Coale and Demeny, 1966) and offered an interesting observation: The cemetery age distribution of a stable population determines the crude birthrate but makes no prediction of life expectancy. Mean age at death of the cemetery predicts the reciprocal of the crude birthrate nearly exactly for all humanly possible values of the intrinsic growth rate (i.e., \(-1\% < r < +4.0\%\)). In other words, a given cemetery fixes the crude birth rate, but both mortality
and growth are free to vary (Paine and Harpending, 1998; Sattenspiel and Harpending, 1983). Since the cemetery fixes crude fertility, how can paleodemographers choose a specific mortality level on a continuum of solutions? Bennett’s (1973) analysis of Point of Pines, Asch’s (1976) approach to middle woodland groups in the Lower Illinois Valley, Muller’s (1997) models for late prehistoric populations in the Eastern Woodlands, and Hernandez’s (2005) analysis of Classic Maya cemeteries all argue for solutions based on a hypothesized growth rate. Milner et al. (2000) introduce the problem, explore models of constant mortality and differing fertility, and recognize the extreme variations resulting from this variable, but they foresee no empirical solution other than recovery of the crude birth rate. In any event they seem unwilling to depart very far from the traditional stationary models of paleodemography. Of course, archaeological support for a single value of Malthusian $r$ is very difficult to find (Horowitz et al., 1988).

Life expectancy at birth is a function only of mortality. By contrast the crude birth rate depends on age-specific fertilities, the once-living age structure of the population, and even the mortality function. A different fertility measure—one stripped of any influence of maternal mortality and age structure—would address the problem. The TFR has been emphasized not only in proximate-determinants models in anthropological populations (Bentley et al., 1993; Campbell and Wood, 1988; Wood, 1990) but also in theoretical approaches to paleodemography (Harpending, 1997; Keckler, 1997; Meindl et al., 2001). Also calculated as twice the sum of the one-sex age-specific fertility rates, or gross reproductive rate, (GRR) (Weiss, 1973:39), the TFR is independent of the population’s age structure or maternal mortality.

Whatever the average growth rate of any prehistoric population may have been, it was filling its cemetery on the basis of both its death rates and its once-living age structure. That is, many different stable population profiles exist, from poor longevity/medium fertility/low growth or decline (Fig. 19.1, lower left) to good survivorship/high fertility/high growth (upper right), and every gradation in between, that could have filled the single Libben cemetery with the same age-class proportions. In fact, a continuum of solutions to Libben or to any single prehistoric cemetery exists. The determination of the numerical values for a complete demographic profile requires stable demographic theory and at least one other datum. The choice of a total fertility rate from modern populations coupled with the cemetery age-sex distribution of an extinct population will produce a full demographic solution, including the average stable growth rate, the once-living age structure, and the mortality profile.
The Campbell and Wood (1988) survey of 70 noncontracepting populations (CW sample), which is often cited for its implications about natural fertility and population transitions in subsistence, has an average TFR of 6.1 children. However, the variance of this empirical distribution is surprisingly large and may explain the choice to substitute “traditional” or “anthropological” for the old term, “natural” fertility (Wood, 1990). In this survey, the lowest values were poorly reported to them, repeatedly sampled (Bentley et al., 1993), or strongly influenced by endemic venereal diseases (Caldwell and Caldwell, 1983); the highest had a history of unchecked population expansion (Campbell and Wood, 1988), or modern diets and health combined with religious mandates for large families. From the interior 35 cases of the CW sample, we have applied three modern fertility performances to the Libben cemetery—the end of the first quartile (TFR for the eighteenth population), the mean, and the beginning of the last quartile (fifty third) (Table 19.2).

Infanticide is a widespread cultural practice associated with high fertility and studied as well as possible in several Neotropical foraging and horticultural populations. It can be individual or community-based, and both the intensity of and reason for its practice vary as much as the cultures of South America. Hill and Hurtado (1996) estimate that perhaps 5% of all children of the forest Ache were victims of homicide, with more than a few cases occurring at a considerably older age than infancy. Perez and Salzano’s study (1978) of five Ayoreo populations revealed high levels of infanticide as well as a weak positive correlation of this with total fertility. The very high levels of Yanomamo infanticide showed a strong female component (Neel, 1978). Perhaps it is difficult to care for closely spaced children. More probably, girls are perceived as ineffective combatants or (incorrectly) as less efficient food gatherers/producers. Therefore, increasing amounts of especially female infanticide, “unrecovered” by archaeologists, were added to the composite (multigenerational) cohort depending on hypothesized fertility level (Table 19.2).

Each fertility level produces a net reproductive rate (in daughters per woman), and the three crude rates (per person per year), of birth, death, and intrinsic increase (Table 19.3). Note that even the truncated 50% range in total fertility produces great variation in the net reproductive rate, which represents the effective number of daughters born to women, taking into account their own mortality rates to and through the reproductive years. Note also that the crude birth rate changes slightly in our model because of our manipulation of the degree of unrecovered infanticide (Sattenspiel and Harpending, 1983).

Not only does fertility predict annual growth rate, it also predicts the once-living age structure and the complete mortality profile as well. For sedentary populations in rich environments with a paucity of human competitors, the higher fertility model is favored—in fact, it might not be high enough—however, the two lower fertility models are presented to demonstrate the impact of a different TFR on all other aspects of a cemetery-based demography. For instance, very different mortality structures follow from variations in assumed fertilities (Table 19.4). If the prehistoric TFR was 7, then life expectancy at birth at Libben was

<table>
<thead>
<tr>
<th>Libben Site Fertility Model</th>
<th>TFR</th>
<th>“Unrecovered” Infanticide</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>7.0 children</td>
<td>7% Female, 2% Male</td>
</tr>
<tr>
<td>Mean</td>
<td>6.1 children</td>
<td>5% Female, 2% Male</td>
</tr>
<tr>
<td>Low</td>
<td>5.5 children</td>
<td>4% Female, 2% Male</td>
</tr>
</tbody>
</table>
much better than seventeenth-century London even during the nonplague years (Hacking, 1975) but not as good as the time-specific figures for some western cities during the 1918–1919 influenza pandemic. Regardless of which TFR value is used to generate a model, the Libben sex-specific mortality profiles contrast with that of the Paraguayan Ache: 1) Libben life expectancies are lower, 2) the ratios of infant mortality to young adult mortality are less for Libben (even when increased by hypothetical levels of infanticides), and 3) the survival of postreproductive adults is depressed.

Before 1960 !Kung fertility was actually less than Howell’s very low estimates for this society in the 1970s. These levels were at least partly the result of infectious diseases. Syphilis and other venereal diseases had been endemic in much of sub-Saharan Africa since their introduction by Moslem traders generations ago. In fact, a large area of west-central Africa north of the lands peopled by the !Kung had been termed the “sterile crescent” by demographers until the infertility epidemic was cured by penicillin (Caldwell and Caldwell, 1983). The presence of venereal diseases in the !Kung has long been known (Howell, 1979), but the impact of infectious sterility on the reproductive histories of !Kung women was only fully appreciated later (Harpending, 1994). Syphilis, gonorrhea, and other sexually transmitted diseases affected women’s age patterns of fertility as well, particularly the parity-progression ratios of the !Kung, which mirrored those of populations with no fertility control (Howell, 1979), albeit ones with greatly reduced TFRs. Until DeBeers pressured for the relocation of the !Kung into several refugee camps at New Xade in 2002 (Price, 2005), they had exploited the desert scrub environments of Namibia and Botswana, and they deservedly enjoyed a special place in anthropological studies. However, it seems that they were never a good model for Stone-Age demography. Fertility is the real determinant of age structure, and !Kung fertility has long been artificially low. The !Kung census of 1975 reveals very few dependents per producer (Fig. 19.2). If Libben’s tenth-century TFR was 7, then its living population was similar to the high-growth twentieth-century Yanomamo and unlike the !Kung.

### TABLE 19.3 Three Hypothetical Fertility Models for Libben: Net Reproductive Rate (Daughters per Woman), Crude Demographic Rates, Birth, Death, Increase (per Person per Year)

<table>
<thead>
<tr>
<th>Libben Site Fertility Model</th>
<th>Crude Rates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$b - d = r$</td>
</tr>
<tr>
<td>High (7.0)</td>
<td>1.98</td>
</tr>
<tr>
<td>Mean (6.1)</td>
<td>1.51</td>
</tr>
<tr>
<td>Low (5.5)</td>
<td>1.15</td>
</tr>
</tbody>
</table>

### TABLE 19.4 Three Hypothetical Fertility Models for Libben: Life Expectancies at Birth and at Age 15 years, Adolescent Survivorship, by Sex

<table>
<thead>
<tr>
<th>Fertility Model</th>
<th>Mortality $\bar{e}_0$</th>
<th>$\bar{e}_{15}$ (fem)</th>
<th>$\bar{e}_{15}$ (male)</th>
<th>Survivorship $l_{15}$ (fem)</th>
</tr>
</thead>
<tbody>
<tr>
<td>High (7.0)</td>
<td>31</td>
<td>29</td>
<td>25</td>
<td>.69</td>
</tr>
<tr>
<td>Mean (6.1)</td>
<td>27</td>
<td>27</td>
<td>24</td>
<td>.63</td>
</tr>
<tr>
<td>Low (5.5)</td>
<td>23</td>
<td>25</td>
<td>23</td>
<td>.55</td>
</tr>
</tbody>
</table>
The Human Species in Demographic Context

There is no reason why women at Libben would have averaged fewer than seven children at the end of their reproductive period, should they have survived it. Indeed, this value is conservative compared with those of some of the best-studied modern foragers and horticulturalists. Nevertheless, even with a TFR value as low as seven children, Libben’s average demographic profile is remarkable: Prehistoric life expectancy is predicted to have been about 30 years, the living population was very young and characterized by a high dependency ratio, and the annual growth rate was substantial. Such growth in turn predicts intrinsic doubling-times averaging less than 30 years, with occasional outmigration of related, extended families, leaving behind a core village of fewer than 150 people who continued to use the cemetery. This estimate of the limit to the size of the village plus the cemetery age distribution and a modern TFR combine to predict other aspects of the structure and dynamics of this Late Woodland society as well (Table 19.5). These societal implications will be the subject of a future demographic report.

As paleodemographers we have often failed as evolutionary biologists. One of the most salient issues in evolution, if not all of science, is human origins. The primary basis of the emergence of clearly identifiable human ancestors more than six million years ago was not an elaboration of material culture. Nor was it a consequence of brain evolution. Rather, it was a demographic revolution. Modern human fertility rates are biologically greater than those of even well-fed, captive chimpanzees (Table 19.6). In fact, free-ranging chimpanzee females have interbirth intervals that exceed four years (Sugiyama, 2004); orangutans exceed six years (Gilders, 2000). The recovery of new fossil materials from Kenya and Ethiopia suggests that hominids were habitually bipedal at least by the early Pliocene.

Table 19.5  Three Hypothetical Fertility Models for Libben: Age Structures

<table>
<thead>
<tr>
<th>Libben Site Fertility Model</th>
<th>Mean Age (years)</th>
<th>&lt;15 years %</th>
<th>Dependency Ratio</th>
<th>Sibship Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>High (7.0)</td>
<td>18</td>
<td>51</td>
<td>1.17</td>
<td>5.8</td>
</tr>
<tr>
<td>Mean (6.1)</td>
<td>19</td>
<td>48</td>
<td>1.04</td>
<td>4.8</td>
</tr>
<tr>
<td>Low (5.5)</td>
<td>20</td>
<td>45</td>
<td>.95</td>
<td>4.2</td>
</tr>
</tbody>
</table>
(Haile-Selassie, 2001; Ohman et al., 2005) and must have thereby already adopted novel means of environmental interaction and exploitation (White et al., 2006). It is therefore probable that such adjustments of their adaptive profile included substantial alterations of their fundamental demographic characters (Lovejoy, 1981, 1993). Indeed, it is likely that they had already become biologically the most \( r \)-selected of any hominoid primate yet to appear in Africa. By the same point in time, even their closest living relatives, the ancestors of chimpanzees and bonobos, were declining and had become largely relict populations in Late Miocene refugia, unable to compete successfully with the relatively recent radiation of cercopithecoid monkeys, another demographically successful group of primates.

It is the evolutionary nature of human populations to grow. For mammals, a large variance in intrinsic growth rate may be manifested in several ways, but potentially high net fertility is the hallmark of a colonizing species, and certainly applies not only to \textit{Homo sapiens} but also very likely to their Plio-Pleistocene ancestors. This unique suite of demographic characteristics has resulted in both Pleistocene success and, ultimately, modern population crisis.

For too long, paleodemographers have overly interpreted the osteology of senescence. We have limited our perception of the expansive demographic evolution of a secondarily \( r \)-selected higher primate by the application of restrictive reference sets of life tables and functional graduations. Most important we continue to accept stationary models and the mortality and age profiles that they generate without question. It is time to turn attention, instead, to the fertility performances of carefully chosen anthropological populations of the mid twentieth century and to incorporate these into our models of prehistory. Only when this research is done will stationary demographic conditions come to be regarded as the exception, not the rule, in human history, prehistory, and evolution.

**ACKNOWLEDGMENTS**

The authors thank Melanie McCollum, Andrea Cucina, and Maria Serrat for their valuable assistance in the preparation of this chapter.

**REFERENCES**


Bennett KA. 1973. On the estimation of some demographic characteristics on a prehistoric population

<table>
<thead>
<tr>
<th>Population</th>
<th>( \hat{e}_0 )</th>
<th>TFR</th>
<th>Annual Growth?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Common Chimpanzee</td>
<td>29</td>
<td>6.2</td>
<td>none</td>
</tr>
<tr>
<td>Ache (Paraguay)</td>
<td>43</td>
<td>8.1</td>
<td>+.025</td>
</tr>
<tr>
<td>Libben (Ohio, 900 A.D)</td>
<td>31</td>
<td>7.0</td>
<td>+.025</td>
</tr>
<tr>
<td>Costa Rica (1960)</td>
<td>65</td>
<td>6.4</td>
<td>+.025</td>
</tr>
<tr>
<td>Sweden (1960)</td>
<td>76</td>
<td>2.4</td>
<td>+.005</td>
</tr>
</tbody>
</table>


Kelley MA. n.d. Pathology inventory for Indian Knoll, Oh2. Lexungton: William S. Webb Museum of Anthropology, the University of Kentucky.


Aboriginal populations, trauma analysis, 377
Abrasion, dental, 305, 309–310, 453
Abscess, dental, 322–323
Accessory ridges, 280
Acetabular fracture, 358
Acetic acid, 454
Acid washing, 453–454
Acromioclavicular joint, 357, 366
Acromioclavicular separation, 357
Adaptation, 9, 26, 183–184, 389, 553
Admixture analysis, 287, 292, 497, 546
Adobe Photoshop, 254
Adult modeling, 151
Afterlife, 5–6
African-American population, 157–158
Age assessment
characteristics of, 49–50
subadult, using dental development, 237–258
Age at death, 49–51, 134, 566, 572, 581–583, 586, 590, 608–609
Age estimation
human cortical bone, 152–182
methodologies, 565, 567, 576–580, 590
significance of, 125–133, 138–139
Agglomerative clustering, 496
Aging process
bone loss, 399–400, 403
bone mass declines, 196–197
nonmetric trait variation, 547–549
osteology of, 606–607
Agricultural populations, 394
agricultural economies, 566
agricultural lifestyle, 396–397
agriculturists, 570
AIDS, 428
Alfred Packer site, 102
Allometry, 220–222, 233, 497
Aluminum, bone composition, 453
Alveolar bone, 97
Alveolar process
bone loss, 317, 319–324
fractures and, 354
Amboseli Reserve (Kenya), 80
Ameloblasts, 239, 245–246
Amelogenesis imperfecta, 303
American Academy of Forensic Sciences, 43, 46, 58
American Anthropological Association, 59
American Board of Forensic Anthropology, Inc., 43, 58–59
American Indians, dental morphology, 273
American Journal of Physical Anthropology, 42
American Museum of Natural History, 11
Amino acids, 417–419, 468, 475
Amplification, DNA analysis, 467
Ana Manuku Cave (Cook Islands), 100
Analysis of variance (ANOVA), 218, 223, 225
Anasazi Indians, 30–31, 72, 76
Anatomical Preservation Index, 79

Copyright © 2008 John Wiley & Sons, Inc.
Anatomy, dental, 266–267
Ancestor-descendant relationship, 26–28
Ancestry, 52–53, 494
Ancient DNA (aDNA)
   analysis of. See DNA analysis characterized, 125
Ancient Egypt, 10, 161, 291
Ancient Greece, 4–7
Anemia, 474
Angel, J. Lawrence, 44–45
Animal scavenging, 71
Animal studies
   chewing damage, 96
   bone maintenance, 394
   bone modifications, 87
   DNA analysis, 476–477
   postmortem alterations, 55
Ankle
   dislocation, 367
   fractures, 358–360
   ossification of, 344
Anthropological Research Facility (University of Tennessee), 55
Anthropometry, 487
Anthroposcopic traits, 533
Apatite, 444, 452
Approximal attrition, scoring schemes, 326–327
Archaeological assemblages
   archaeological record, 73–75
   human bone
      decomposition of, 78–86
      modification, 86–104
      preservation, 78–86
   human remains, archaeological recovery of, 76–78
   museum collections, study of, 104–105
   overview of, 71–72
   mortuary programs, 73–75
   taphonomy, 72–73
Archaeological populations, bone loss research, 390–404
Archaeological records
   cannibalism in, 98–100
   components of, 73–75
Arikara Indians, 157
Aristotle, 7
Arizona State University Dental Anthropology System (ASUDAS), 271–272, 275, 291
Army Medical Museum, 11
Arthritis, 372
Articulation rule, 606
Artifacts, 26, 80, 445, 574
Ashing, 420, 453
Asia, East/Northeast
   cranial series, 505–506, 511, 516, 525–526
   jackknifed classification, 508
   nonmetric trait variation, 546
   smallest Mahalanobis’ distance, 521–524
   stepwise discriminant function analysis, 514–515
Asian population, dental morphology, 287, 289
Asymmetry
   biomechanical analysis, 198–199
   dental morphology, 284–285
   directional, 284
   morphometric analysis, 225, 233
Atavism, 534
Atlas of Occupational Markers on Human Remains
   (Capasso/Kennedy/Wilczak), 549
Atmospheric Nitrogen (Air), 423
Atomic absorption (AA) spectroscopy, 455–456
Atomic emission spectroscopy (AES), 455–456
Atra Kadisha, 29
Australia
   cranial series, 503, 510, 524
   smallest Mahalanobis’ distance, 517, 519–520, 524
   stepwise discriminant function analysis, 513
Australian Aborigines, 4. See also Aboriginal population
Autopsy, 11
Avulsion fracture, 346
Avunculocality, 544
Aztec Indians, 91
Barium, bone composition, 444, 446, 448–450, 452–454, 456
Basic multicellular unit (BMU), 152, 154, 161
Basilar fractures, 352
Bayes’ theorem, 576
Bayesian probability, 606
Bedrock, 81
Bending, biomechanical analysis, 185–186, 194–195
Berry, A. C., 273
Bertillon anthropometric trait system, 42
β-globin, 474–475
Bias, impact of, 78, 118, 120, 163, 567, 575–576, 580, 586, 602
Bioarchaeological ethics
   beliefs about the dead, 4–5
   conflict
      resolution strategies, 29–31
   sources of, 26–29
   descendant rights, 21–24
   development of, 3–4
human dignity, 20–21
human remains
dissection of, 5–6, 10, 12
research, history of, 5–9
skeletal, value of, 3, 13–18
preservation ethic, 24–26
skeletal biologists, ethical responsibilities of, 18–26
skeletal collections sources, 9–13
Biodistance analysis, 80, 489, 497
Bioinformatics, 233
Biological apatite, stable isotope analysis, 419
Biological continuity, 24
Biological distance
defined, 538
research studies, 487–488, 543–544, 553
see also Biodistance analysis
Biological diversity, 7
Biological integrity, 453
Biomechanical analyses of human skeletons
bone functional adaptation, 183–184
future research directions, 200–201
individual variations
bilateral asymmetry, 198–199
ontogeny, 195–198
long bone
diaphyseal structure analysis methodologies, 184–189
robusticity, evolutionary trends, 189–190
population variations
sexual dimorphism, 192–195
subsistence strategy, 190, 193
terrain, 191–192
Wolff’s Law, 183–184
Biomechanics theory, 183
Biometricians, 541–542
Biomolecular archaeology, 81
Biopurification, 445, 448–449, 454
Bipartite patella, 358
Black Death plague, 572, 590
Blind tests, 580
Bone
absorptiometry, 188
burned, 57
chemistry of. See Bone chemistry
cross-sectional geometric properties, 185
density
element survival, 82–83
influential factors, 397–398
paleodemography and, 83–85
fragmentation, 79, 102–103, 352, 362
human. See Human bone
loss of. See Bone loss
mineralization, 402
modification stages, 72
pathogenic-specific, 26
postmortem interval research, 56
quality of, 392
remodeling process, 151–152, 157–158, 161, 165
resorption, 389
rigidity, 185–186
stable isotope analysis, 430
see also specific types of bones
Bone chemistry, trace element analysis of mineral
fraction of bone
categorized, 443–444
current perspective, 456–457
elements in, 444–451
dental enamel, 443, 454–457
dentine, 443
instrumentation, 455–456
post-mortem alteration, 451–454
Bone loss
biomechanical analysis, 184
bone quality, 392
characterized, 388–391
influential factors, 393–403
lifestyle factors, 396–397
patterns of, impact of, 403–404
reproductive factors, 398–399, 403
research methodologies, 388
Bone mineral density (BMD), 119, 390–391, 398, 402
Bone Representation Index, 79
Bone structural units (BSUs), 151, 161
Bone-seeking elements, 444, 449
Bootstrap analyses, 492
Boundary Mound, Middle Woodland site, 90
Bovine serum albumin (BSA), 465
Boxgrove, 164
Boxplots, 495
Boyd, W. C., 290
Brachycranic, 489
British Anglo-Saxon population, 316
INDEX

British Association for Biological Anthropology and Osteoarchaeology, 79
British Museum, 15
Broca, Paul, 41
Bronze Age, 90, 194
Brushite, 444
Buccal cusp, 280
Buccal plate, 329–330
Buckling fracture, 345
Buddhism, 15
Bundle burials, 606
Burial
  customs, 602
  location selection, 75
  practices, selective, 584
  rituals, 5, 28
  sites, flooding of, 80
  types of practices, 573
Burial Terminology, 74
Buried → Preserved transition, 574
Burst fracture, 346
Bushman canine, 276

Cadmium, as toxin, 455
Calcaneal fractures, 358
Calcification, 126
Calcium carbonate, 444
dietary, 393–396, 445
Calcus, dental, 312–313
Canadian Society for Forensic Science (CSFS), 43
Canalization, 233
Cancer, 586
Canines
  development of, 267–268, 301, 304, 310, 314
  mandibular, 279
  maxillary, 275–276
Cannibalism, 71, 91, 93, 98–100
Canonical coefficients, 493, 510
Canonical plots, 495
Canonical variates analysis (CVA), 216–217, 233
Canyon Diablo meteorite troilite (VCDT), 423
Carabelli’s cusp, 289
  Carabelli’s plaque, 271
  Carabelli’s trait, 269–270, 273, 277, 284, 288, 290
Carbohydrates, 313
Carbon
  isotopes, 422
  primary and reference standards, 423
Carbonate
  bone composition, 424, 450, 453
  stable isotope analysis, 417, 419
Cardiovascular disease, 586
Caries (see also Cavities, dental)
  dental, 313–319, 325
dentine, 327
  scoring schemes, 326–329
Carnivores, 86–87, 445–446, 451, 453
Carp, 427
Casts, dental morphology, 288, 291
Canibalism, 428
Catacombs, 74
Cavities, dental, 331–332, 450
Cement dentine junction (CDJ), 241
Cement-enamel junction (CEJ), 254, 256, 314, 317
Cemeteries
  aging bias, 603
  characterized, 75
  historic, 12–13, 84–85, 139
Centroids, 495
Cervical cancer, 474
C4 plants, 423–424, 426, 433
14C, 415
Chamorro culture, Latte Phase, 77–78
Cheek cells, 312
Chemical weathering, 80–81
Chi-square distribution, 495
Chicago Field Museum, 11
Child abuse, 375
Children
  dental morphology, 290
  diet and nutrition, 454
  fractures in, 345, 350–351, 368, 375
  neonates. See Infants
  skeletal assemblages, 84–85
  skeletons of. See Juvenile skeletons
  skull fractures, 350–351
  traumatic injuries, 361–362, 377
China
  cranial series, 506, 511
  customs for funeral arrangements, 22
  jackknifed classification, 508
  smallest Mahalanobis’ distance, 521–523
  stepwise discriminant function analysis, 515
Chipped teeth, 306–309
Chlorine, bone composition, 450
Cholesterol, 421
Christianity, 6–7, 9–10, 27
Chronic pulpitis, 322
Chronic renal disease, 161
Chumash Indians, 30
Circadian rhythm, 242
Clavicle
  age-estimation methods, 182
  fracture of, 345, 357, 365
Clay-shoveller’s fracture, 355
Clays, 43
Cleft lip/palate, 270
Clethrionomys rutilus, 552
Cloning, 465, 467
Cluster analysis, 496–497
Cluster identification, 495
Clustering algorithms, 496–497
Clustering methods, 492
Cobb, William Montague, 12
Coccyx, fracture of, 354–355
Coimbra Identified Skeletal Collection, 4, 51–52
Cole, Tim, III, 219, 235
Collagen
  bone composition, 451
  characterized, 80–81, 84, 392
  in fossil bones, 444
  stable isotope analysis, 416–419, 424–425, 427, 434
Collection methods, 77
Colles, Abraham, 362
Colles’ fracture, 362–363, 366, 399–400, 605
Columbian Museum of Chicago, 11
Comminuted fracture, 346, 360
Competing hazards model, 584
Completeness Index, 79
Complex societies, 566
Compound-specific analyses, 418–419
Compression fractures, 368
Computed tomography (CT) studies, 187–188, 343–344
Computer software programs
  biomechical analysis of skeletons, 188
  dental development, 254
  forensic anthropology, 54
  metric analysis, 497
  morphometric analysis, 218–219, 222, 224–225, 233–235
  scoring schemes in dental pathology, 325
  statistical packages, 489, 494, 499
Condylar fractures, 361
Conflict
  resolution strategies, 29–31
  sources of, 26–29
Connective tissue, 74
Conservation strategies, 76, 105–106
Contact area caries, 332
Contamination
  chemical, 454
  of ecosystem, 552
Continuous variation, 488
Coordinate points, 488
Correlation coefficients, 163, 498
Correlation matrices, 222
Correlations
  morphometric analysis, 234–235
  tooth morphology, 269
Cortical bone, 319, 323
Counter-winging, 273
Covariance, 222–223, 490
Covariance matrices, 222, 492, 495
Cranial series, metric analysis, 499–507, 511, 516–517
CRANID2/CRANID3, 494
Cranio logical studies, 78
Craniometer, 105
Craniometry
  characterized, 487
  landmarks, 488
  research, 7–8
  variation, 488
Cranium
  fractures, 95, 350
  modification patterns, 4, 91–94
  morphology, 489
  nonmetric traits, 549–550
  traumatic injury, 86, 605
Cre-lox mice, 208
Cremation, 73, 86
Cribra orbitalia, 552
Cro-Magnons, DNA analysis, 468
Cross striations, dental development, 242, 245, 247, 252–254
cross-sectional moments of inertia, 185
Cross-validation procedure, 508
Crow Creek site (South Dakota), 82–83, 88, 96
Crown
  dental anatomy, 239
  dental caries, 315
  dental development defects, 301, 303–307
  dental morphology, 273–283, 288–289
  formation of, 245, 250
Crush fractures, 346, 355
Crystallinity, 453
C-scores, 499
$\delta^{13}$C, 424–427, 429–431, 433–434
C$_3$ plants, 423–424
Cultural affiliation, 24
Cultural background, 22–23
Cultural bias, 14
Cultural continuity, 23–24
Cultural diversity, 3, 7, 18
Cultural values, 14, 27
Cuspal striae, 244
Cusp number, dental morphology, 277, 279–280, 282, 289–291
Cutmark
distribution analysis, 87, 91, 96–97
morphology, 89
trauma analysis, 375–376
Cutting machine, 58
Cyst, periodontal, 322–323
Dahlberg, Albert A., 266, 268, 271–272, 288
Dahlberg, Thelma, 288
Darwinian theories, natural selection, 8
Data
archiving, 73
gathering strategies, 210
missing, 499
standardization, 209
Data banks, 45–47
Databases, Online Mendelian Inheritance in Man (OMIM), 209
_Daubert v. Merrell Dow Pharmaceuticals_, 56
Dead people, cultural beliefs, 4–5
Dead → Buried transition, 573
Dean, M. C., 241
Death
age at, 49–51
time since, 55–56
“Death-care” industry, 5
Deciduous teeth, 289–290, 455
Declaration of Human Rights of the United Nations, 19
Decline-in-health scenario, 569–570
Decomposition, 43–44, 51, 74–75, 454
Decussation, 256
Deflecting wrinkle, 282–283
Defleshing, 4, 91
Degrees of freedom, 218, 495
Dehiscence, 330
Demineralization, 81, 420
Demirjian dental maturity standard, 128
Demographic and Health Survey, 564
Demographic evolution, 604
Demographic revolution, 612
Demographics, significance of, 26, 49, 51, 612–613
Dendograms, 496, 523
Densitometry, 188
Dental anatomy, 266–267
Dental decay, risk for, 589
Dental development
growth layering, 302
histological techniques, 257–258
influential factors, 302
phases of, 301–302
subadult age assessment
dental anatomy, 239–241
characterized, 237–238, 256–258
enamel microstructures, classification of, 247–248
estimates, histologically established, 251–256
future research directions, 256–258
growth markers, microstructural, 241
histological exam, preparation for, 248–250
historical perspective, 238
long period markers, 243–246
noninvasive estimates using perikymata, 250–251
short period markers, 241–243
striae of Retzius, 243, 246–247
timing of developmental events, 250
Dental disease, 86. _See also_ Caries; Periodontal disease
Dental enamel. _See also_ Enamel
Dental formation stages, 127, 129–131
Dental maturity, implications of, 126
Dental morphology
characterized, 265–267
dental anatomy, 266–267
historical perspective, 266
human dentition, 267–271
population studies, 286–292
research methodologies
asymmetry, 284–285
crown and root traits, 273–278
counting methods, 284–285
errors, types of, 285–286
observations, 283–284
trait classifications, 271–273
Dental pathology
characterized, 301
dental development defects, in enamel of tooth
crown, 301–306
impact in archaeology, 324–325
influential factors, 325
plaque-related diseases, 301, 311–325
tooth wear, 306–311
Dental work
juvenile skeletons, 124
in positive identification, 56
Dentine
- bone chemistry and, 443, 454
- caries, 327
- composition of, 239
- formation of, 239
- stable isotope analysis, 430

Dentition, human
- dental genetics, 270–271
- deciduous, 289–290
- interactions among size, morphology, and tooth number, 268–270
- morphogenetic fields, 268
- permanent, 289–290
- terminology of, 267
- traits, types of, 268

Depression fractures, 346, 351

Dermatoglyphic studies, 42

Descendant rights, 21–24

Developmental analysis
- dental development and subadult age assessment, 237–258
- juvenile skeletons and growth-related studies, 117–139

Developmental biology, 538, 553

Developmental Defects of Enamel (DDE), 303, 305–306

Developmental Juvenile Osteology, 548

Diabetes, 161

Diagenesis studies, 81, 102, 161, 241, 446, 450–454

Diaphysis, significance of, 345

Diet
- bone maintenance and, 393–396
- childhood, 454
- deficiencies in, 303
- DNA analysis and, 477
- marine-based, 192, 425, 431, 448
- prehistoric, 444, 447, 450
- terrestrial-based, 192, 415, 425, 431, 448

Differential preservation, 84, 118

Diffusion, 452

Digitizers, 212

Digitizing pads, 488

Dimorphism, 138

Disarticulation, 89

Discriminant coefficients, 493

Discriminant function analysis, 490, 493–494

Diseases, genetic, 474–475

Dislocations
- lower limbs, 367
- spinal, 368
- types of, 345, 361, 366, 589
- upper limbs, 366

Dismemberment, 89

Disparity Box, 224, 233

Dispersion, multivariate analysis, 509

Displacement, 373

DISPOP, 494

Dissection, 5–6, 10, 12

Distal accessory ridge, 275–276, 279

Distance analyses, 493. See also Biological distance

Distance studies, morphological data, 497

Distribution-free inferential techniques, 492

Diversity, mtDNA analysis, 471–472

DMF score, 316

DNA
- amplification, 138
- analysis process. See DNA analysis evidence, 51
- extraction, forensic, 57
- preservation, 312, 464–465, 467–468, 471
- research, 51
- sequencing, 472, 476

DNA analysis
- applications, generally, 461
- authenticity criteria, 461, 466–467
- case examples, 468–477
- characterized, 25–26, 51, 60, 138
- future research directions, 477–478
- methodologies, 464–468
- nonmetric trait variation, 554
- sample size, 468–469
- sex identification, 472–473

Documentation
- excavation notes, 104
- field notes, 104
- museum registration records, 104
- quality of, 78
- significance of, 76

Dolichocranic, 489

Domestic violence, 376–377

Dorsey, George A., 42

Double diffusion tests, 48

Double-shoveling, 275

Double-zonal osteons, 154

3DPCA, 233

Dry ashing, 420

Dryopithecus, 282

Dwarfism, 474

Dwight, Thomas, 41–42

Dying, risk of, 586–587

Dysmorphology, 551

Early Bronze Age, 85

East Ural Radioactive Track (Russia), 552
Ecological research, 421, 431
EDMA, 216, 218–220, 222
Egyptian Book of the Dead, 4
Eigenvalues
  morphometric analysis, 232–233
  multivariate analysis, 509
  variance of, 235
Eigenvector, morphometric analysis, 233
Elbow, dislocation of, 366
Elder abuse, 376
Elderly, age estimation, 158–160
Electrophoresis, 466
Element survival, 82–83
Enamel
  bone chemistry and, 443, 454–457
  cuspal formation, 255–256
  dental defects, 301–306, 309
  development of, 245
  extensions, 278–279
  hypoplasia, 238, 248
  microdefects, 590
  microstructure classifications, 247–248
  mineralization of, 241
  postmortem interval research, 56
  stable isotope analysis, 419, 430
Enamel–dentine junction (EDJ), 239, 241, 243–244, 246, 252, 254, 256
Enlightenment, 7, 10
Entoconid, 281–282
Environmental conditions, impact of, 554
Epicondylar fractures, 360–361
Epidemic disease, 137
Epigenetics, 535, 553
Epipyseal–metaphyseal fractures, 375
Epiphysial union, 131
Erasistratus, 5
Errors
  casting, 289
  metric analyses, 498–499
  observation, 283–286
  paleodemographic research, 578
Eskimo population
  age estimation, 157
  nonmetric trait variation, 543
  odontomes, 270, 276–277
Estrogen levels, implications of, 84
Etching, 249–250
Ethnic identification, 287
Ethnic identity, 53
Ethnocentrism, 3
Ethnographic anthropology, 604
Ethnographic evidence, 602
Euclidean distance, 494–495
Euro-American population, 23, 157–158
Eurowestern values, 27
Evidence
  of foul play, 57–58
  recovery strategies, 47–58
Evolution, 209
Evolutionary biology, 538, 553, 612
Evolutionary genetics, 541–542
Evolutionary synthesis, 538
Evolutionary trends, 292
Excavation tools, 76
External enamel defect (EED), 248
Extinct animals, 445
Face
  fractures of, 352–354, 365
  reproduction of, 54
Factor analysis, 490, 492–493, 495
Familial relationships
  DNA analysis, 473–474
  nonmetric trait variation, 546
Fatigue fracture, 346
Fatty acids, 55
Federal Bureau of Investigation (FBI), 42, 45
Fédération Dentaire Internationale DDE Index, 305–306
Femoral fractures, 363–364, 367, 399
Femur
  age-estimation methods, 176–179
  biomechanical analysis, 192, 194–195
  bone mineral density, 398
  characterized, 86, 90, 132, 136, 156, 158, 163–164, 590
  fracture of. See Femoral fractures
  nonmetric traits, 536
Fenestration, 330
Fertility rates, 570, 581–582, 603, 607–611
Fetus/fetal
  age, 132
  skeletons, 548
Fibula
  age-estimation methods, 176
  fractures, 364
Field archaeology, 76
Field notes, 104
Finite element method (FEM), 200
Fisher z-transformation, 222
Fluctuating asymmetry (FA), 225, 233
Fluorapatite, bone composition, 450–451
Fluoridation, 319
Fluoride, bone composition, 451, 453
Fluorine, 318, 450
Follicle-stimulating hormone (FSH) levels, 122
Food chain, 445–446
Food shortages, 570
Foot, fractures of, 358–360, 367
Footwear, 51
Foramen magnum, 93
FORDISC 1.0, 46
FORDISC 2.0, 52
FORDISC 3.0, 46–47, 60, 494
Forensic anthropology
  age at death, 49–51
  ancestry, 52–53
  case study, 59–61
  evidence
    of foul play, 57–58
    recovery of, 47–48
  facial reproduction, 54
  forensic data bank, 45–47
  future prospects, 58–59
  historical development, 41–44
  living stature, 53–54
  molecular approaches, 57
  photographic superimposition, 54–55
  positive identification, 56–57
  remains, human vs. nonhuman, 48–49
  sex, 51–52
  skeletal biology, relationship with, 44
  theoretical approaches, 44–45
  time since death, 55–56
  training in, 59
  Forensic Data Bank, 53
  Forensic investigations, 31
  Forensic odontologist, 60
  Forensic taphonomy, 72
  Formaldehyde, 465
  Fossil(s)
    bone composition, 444
    discriminant function analysis, 494
    fluorine in, 451
    hominids, 292
    sites, 87
  Foul play, 57–58
  Found → Saved transition, 574
  Fourier transformations, 495
  Fractionation, 416
  Fractures
    dental, 310–311, 323
    description of, 95–97, 347–348
    facial, 352–354
    flat and irregular bones, 354–360
    fragility-related, 387, 399–401
    healing process
      complications of, 371–373
      duration of, 368–371
      long bone, 360–364
      perimortem, 90, 95–97
      skull, 350–352
      summary of, 365–368
      types of, 346–350
    see also specific types of fractures
  Fragility fractures, 387, 399–401
  Fragmentary osteons, 158
  Fragmentation patterns, 71
  Frailty, 570, 572, 586–590
  F-ratio, 494
  Freshwater fish, 426–427
  Frontal bone, fracture of, 353
  Funeral(s)
    arrangements, 22
    characterized, 74
    rites, 13
    rituals, 4, 74
  F-values, metric analysis, 499, 508
  Galera, Virginia, 50
  Galton, Francis, Sir, 207–208
  Gamma radiation, 465
  Garden of Eden, 7
  Garn, Stanley M., 266
  Gas chromatography (GC), 421
  Gaucher’s disease, 474
  GC-C-IRMS (gas chromatography combustion isotope ratio mass spectrometry), 421–422
  GC-IRMS, 432
  Gender differences
    age estimation, 155
    bone loss, 398–399, 403
    bone mineral density, 83
    bone strength, 197–198
    nonmetric trait variation, 547–548
    paleodemographic research, 576
    traumatic injury, 377–378
  Gene flow, 497
  Generalized distance, 490
  Generalized shape image (GSI), 226, 233
  Genetic admixture, 553
  Genetic drift, 497
  Genetic relationships, 22–23
  Genetics
    dental, 270–271
    evolutionary, 542
  Genocide, 15, 23, 31
  Genome sequences, 476–477
Genomics, 233
Genotypes, 285, 548
Geochemistry applications, 421–422
Gleno-humeral joint, 366
Gnarled enamel, 256
Gompertz-Makeham model, 584
Goodall’s F-test, 218
Granuloma, 322–323
Grasshopper Pueblo (Arizona), skeletal assemblage, 76–77, 84, 431
Grave goods, 552
Grave-robbing, 6
Graves, bedrock-floored, 81
Great Basin studies, biomechanical analysis of skeleton, 184
Great Chain of Being, 7
Greenberg, Joseph, 287
Greenstick fracture, 345–346
Gregory, William King, 266, 281
Groove pattern, dental morphology, 281–282, 289–290
Gross occlusal caries, 332
Gross reproductive rate (GRR), 609
Groundwater, composition of, 81, 450–451
Grouping analysis, 491
Guidelines to the Standards for Recording Human Remains, 72, 79
Gunshot trauma, 58, 60–61
Gustatory cannibalism, 98–99
Hafnium, bone contamination, 452
Hamann-Todd human remains collection, 46–47
Hand fractures, 355, 358–360
Hanihara, Kazuo, 266, 289, 291
Hanihara, T., 291
Hard tissues
chemical and genetic analysis
DNA analysis of archaeological remains, 461–478
stable isotope analysis, 413–434
trace element analysis, bone chemistry and, 443–457
decomposition, 72–73
Harpending, Henry, 584
Harvard Peabody Museum of Archaeology and Ethnology, 11
Haversian system, 153–154, 158
Hawaiian population, 31
Hayek, Lee-Ann, 50
Hematoma ossification, 372
Hemoglobin, 474–475
Herbivores, bone composition, 445–446, 449, 453, 456
Heritability, 265, 488, 497, 545–546, 554
Herodotus, 6–7
Herophilus, 5
Heterogeneity, 586, 590
Heterozygosity, 270, 474
Hierarchical clustering, 496
High-resolution microcomputed tomography (micro-CT), 166
High-throughput analysis, 225–227
Hip
dislocation, 367
fractures, 395, 399–400
Hispanic population, 23
Holocene Period, 189
Holography, 488
Homer, 7
Hominid evolution, 238
Hominids, dental morphology, 250–251, 292, 456, 603
Hominin, 98
“Homo monstrosus”, 7
Homo sapiens, 257, 393, 613
Homozgyosity, 270, 474
Honokohua site (Hawaii), 84
Hooton, Ernest, 8
Hopi Indians studies, 30, 270
Hornbostel, Hans, 77–78
Howells, W. W., 52–53
Howship’s lacunae, 152
Hrdlička, Aleš, 8–9, 42, 44, 266
Human adaptation, 9
Human-animal relationships, 99
Human bone
cortical. See Human cortical bone decomposition of, 78–86
density of. See Bone density fragments. See Bone fragments modification of. See Human bone modifications
deposition, 78–86, 100, 102–103
shape of, 81–82
size of, 81–82
toolmarks on, 89
Human bone modifications
animal agents in, 86–89
skeletal modifications, mortuary rituals, 93–94
toolmarks
cranial modification patterns, 4, 91–93
postcranial modification patterns, 89–91
thermal alteration, 97–98
trauma, perimortem vs. postmortem, 58, 95–97
<table>
<thead>
<tr>
<th>Human cortical bone, histomorphometric age estimation methods</th>
<th>Human dignity, 20–21</th>
</tr>
</thead>
</table>
INDEX

MACE, 222, 234
Mace3D, 222
Magnesium, bone composition, 450, 457
Magnetic resonance-based tomography, 215
Magnetic resonance imaging, 54
Mahalanobis’ generalized distance (D²), 494–496, 498, 516–524
Maillard reaction, 465
Maize, 415, 424
Malunion, of fracture, 373
Mana, 31
Mandible
age-estimation methods, 175
characterized, 87, 94, 101, 104
dental pathology, 308–309
fracture of, 353–354
Piltdown, 451
Manganese, bone composition, 452–453
Manouvrier, Leonce, 41
Mantel’s test, 222, 234
Marginal ridge, 279
Marianas Archipelago, archaeological assemblages, 77
Marine-based diets, 192, 425, 431, 448
Marquez, Eladio, 222, 234
Mass graves, 93
Mass spectrometer, description of, 420–421
Mass spectrometry, 414, 420–422
Matching asymmetry, 225
Matrix correlation, morphometric analysis, 234
Mausoleums, 73
Maximum likelihood estimation, 584–586
Medico-legal examiners, 51
Medico-legal studies, 41–43, 56, 59, 97
Medieval Europe, 574, 589
Meiosis, 463
Melanesia
cranial series, 501–502, 510, 524
mortuary ritual, 9, 93
multivariate discriminant analysis, 509
smallest Mahalanobis’ distance, 517–519, 523
stepwise discriminant function analysis, 512–513
Melanin, 157
Mendelian geneticists, 541–542
Mendelian inheritance, 207
Mendelian traits, 270
Mercury intrusion porosimetry, 81
Mesial drift, 319
Mesial marginal accessory tubercles, 278
Mesolithic (MES) Period, 194, 425
Metabolic disorders, 161, 552
Metacarpal
age-estimation methods, 175
bone mineral density, 398
fracture, 358–359, 366
Metaconid, 281
Metapathesis, significance of, 345
Metatarsal fracture, 358–360
Methylmethacrylate, 318
Metopism, 548
Metric analysis of skeletal remains applications
cranial series, 499–507
data preparation, 498–499
multivariate analysis results, 499–500, 507–511, 516–517, 523–525
benefits of, 525–526
biological distance studies, 487–488
cluster analysis/clustering algorithms, 496–497
errors, 498–499
justification of metric data, 488–489
landmark measurements, 488
measurement in anthropology, 487
missing data, 499
multivariate data, assumptions of, 491–492
multivariate statistical procedures
classic, 492–496
types of, 490–491
research design, 497–498
size-based component, removal of, 499
univariate methods, 489–490
variable selection, 498
Mexican American population, 23
Microenvironments, 465
Microevolutionary processes, 553
Micronesia
cranial series, 504, 524
jackknifed classification, 508
smallest Mahalanobis’ distance, 520
stepwise discriminant function analysis, 513
Microsatellites, 463
Microscopy, 48, 58, 81, 104–105, 254, 304, 306, 318, 343, 395
Microsoft Excel, 325
Middle Ages, 7
Migration studies, 430–431, 497, 553
Miles, A. E. W., 133
Mineralization, 241, 312
Minimum number of elements (MNE), 79–80, 82
Minimum number of individuals (MNI), 79–80, 82
Minor skeletal variants, 533–534
Misclassification, 575
Missing data, 499
Mitochondrial DNA (mtDNA), 60, 287, 461–464, 467–471, 474, 477
Moche iconography, 97
Modeling
age estimation methods, 149–157
drifts, 150, 161
Moire fringe patterns, 488
Molars
caries, 333
decay in, 589
dental caries, 314
dental morphology, 288–290
dental pathology, 326–327
development of, 267–268, 301, 310
mandibular, 281–283
maxillary, 277–279
Molecular analysis, 56–57, 471, 476
Molecular biology, development of, 208
Monogenesis, 7
Mononuclear cells, 152
Monotonicity, in paleodemographic research, 580
Monteggia, Giovanni, 362
Monteggia fracture-dislocation, 362
Moon’s molar, 306
Moorees, C. F. A., 266
Moral authority, 27
Morbidity, 569
Morphological analysis
biomechanical analyses, 183–201
dental morphology, 265–292
human cortical bone, histomorphometry, 149–166, 172–182
morphometric and biological archaeology in postgenomic age, 207–227, 232–235
Morphometric analyses
allometry, 220–222, 232
characterized, 207–210, 227
covariance structure, 222–223
data types, 210–211
differences, localizing, 219–220
eigenvalues, 232–233, 235
fluctuating asymmetry, 225, 233
high-throughput analysis, 225–227
landmarks, 210–215, 234
magnitude, 222–223
shapes
comparison of, 217–219
implications of, 216–217
variance, comparison of, 223–225
size, 215–216
terminology and acronyms, 233–235
traditional, 210–211
Mortality
age-specific, 584–585
bias, 136
juvenile, 584
parametric models, 584–586
rates, 76, 569, 577, 580–582, 613
selective, 570, 572–573, 586–590
Morton, Samuel George, 8
Mortuary
processing, 89
program, 85, 104
recording features, 74–75
rituals, 4, 72, 86, 89, 93–94
Most Likely Number of Individuals (MLNI), 79–80
Mouse Phenome Project, 208
Mouse studies
morphometrics, 208–209, 226–227
nonmetric trait variation, 542–543
mRNA, morphometric studies, 215
MSA (mean score of affected), 284
MSCT, 343
Multiculturalism, 3
Multidimensional scaling techniques, 492, 495
Multiple discriminant function analysis, 493, 495–496
Multivariate analysis
characterized, 499–500, 507–511, 516–517, 523–525
stepwise discriminant function analysis
canonical coefficients, 510, 516
characterized, 499–500, 507
jackknifed classification, 508–515
Multivariate analysis of variance (MANOVA), 218
Multivariate data, assumptions of, 491–492
Multivariate statistical procedures
classic, 492–496
types of, 489–491
Mummies
age estimation, 157
characterized, 15
DNA analysis, 461
Egyptian, 10
genetic research, 22
juvenile skeletons, 124
tuberculosis DNA research, 476
Museum collections
historical perspectives, 10–11
study of, 104–105
value of, 12
Myositis ossificans progressiva, 344
Myositis ossificans traumatica (MOT), 344
Nariokotome juvenile skeleton, 196–198
National Bureau of Standards (NBS), 422
National Institute of Mental Health (NIMH), 254
Native American Graves Protection and Repatriation Act (NAGPRA), 24, 29, 525
Native American population, see specific Indian groups
ancestors, 23
categorized, 3, 9, 11, 14, 16–18, 21, 316
cranial traits, 549–550
DNA analysis, 469
treatment of human remains, 23–24
tuberculosis in, 475
Natural historians, 7
Natural history museums, 11
Natural selection, 8, 553
atual theology, 7
Navajo Indians, 28, 30
Neanderthals, 29, 72, 199, 461, 468–469
Neighbor-joining (NJ) algorithm, 496–497
Neolithic (NEO) Period, 194, 425
Neotropical populations, 604
Nerve damage, 364
Nerve injuries, 372
Neuromesis, 372
Neutron activation analyses (NAA), 456
New Age movement, 23
New Guinea
cranial series, 502–503, 510
jackknifed classification, 508
smallest Mahalanobis’ distance, 517, 519–520, 523
stepwise discriminant function analysis, 513
New Guinea Research Program (NGRP), 102–103
NIH Image analysis, 254
$^{15}$N, 428–429, 433
Nitrogen
primary and reference standards, 423
isotopes
diet and, 425–426
protein stress and, 428, 431
water stress and, 427–428
Nonhuman remains, 48–49
Nonmetric trait variation in skeletons
activity, 548–550
age, 548–549
categorized, 533–538
classification systems, 536, 539–541
current research, 553
future research directions, 553–554
historical review, 538, 541–547
nonmetric defined, 534
pathological conditions, 551–552
physiologic factors, 546
sex, 547–548
types of, 538
Nonstationarity problem, 580–584
Nonunion, of fracture, 372–373
Normal distribution, 492, 495
Norris Farms site (Illinois), 88, 467–468, 470–472
NTSYS-pc, 497
Nuclear DNA, 461, 463, 474
Number of identifiable specimens (NISP), 79, 102
Nutrition, bone maintenance and, 393–396.
See also Diet
Nutritional deficiencies, 347
Object asymmetry, 225
Oblique fracture, 346, 368
Observability, 265
Occipital bone, age-estimation methods, 175
Occlusal attrition, scoring schemes, 327, 332
Occlusal caries, 332
Octocalcium phosphate, 444
Odontoblasts, dental anatomy, 239, 241
Odontochronology, 126, 256
Odontomes, dental morphology, 276–277, 280, 288
Ohlone people, 28
Old age, 586–587
Oligonucleotides, DNA analysis, 476
Omnivores, 453
Omo Basin studies, bone composition, 453
Online Mendelian Inheritance in Man (OMIM), 209
Ontogeny, 195–198
Open fracture, 364
Optical microscopy, 81
Orbital plates, 352
Orfila, Matthieu-Joseph Bonaventure, 41
Organic matter, 452–453, 465
Organization of American States (OAS), 19
Ossification, 344, 372, 536–537, 604
Osteoarthritis, 358–359
Osteochondritis dissecans, 372
Osteodystrophy, 161
Osteogenesis imperfecta (OGIM), 160–161
Osteological paradox, 566–567
Osteologist, functions of, 575, 577, 585, 588–589
Osteology, 72, 613
Osteomalacia, 552
Osteometry, 79, 487
Osteomyelitis, 371, 605
Osteon population density (OPD), 153, 156, 159–160, 166
Osteopenia, 390
Osteoporosis
age-related, 390–391, 404
biomechanical analysis, 190
categorized, 388–389
etiology, 84, 347, 356, 364, 377, 396
research methodologies, 388
risk factors, 391, 393
secondary causes of, 390
types of, 390–391
$\delta^{18}O$, 430
Oxygen
isotopes, 417, 429–430
primary and reference standards, 423
Paget’s disease, 161
Paiute-Shoshoi population, 469
Pakal’s tomb, 578–579
Palaeohistologist, functions of, 166
Paleopathology, 543–544
Paleodemography
advances in. See Paleodemography advances
biases, 602
categorized, 83–85, 543–544, 601–602
Paleodemography advances
frailty, heterogeneous, 586–590
measurements
age estimation, 565, 576–580
sex estimation, 575–576
mortality analysis, 580–584
overview of, 561–562
parametric mortality models, 584–586
point-counterpoint, 562–568
research design, 568–571
selective mortality, 586–590
skeletons, as sample of deaths, 571–575
Paleodiet, 415, 423–425
Paleoepidemiological research, 80, 348–350
Paleopathology, 44, 85–86, 378, 394, 588
PAM, 499
Pandemics, 475, 611
Papilloma virus, 474–475
Papua New Guinea, 93
Paracorne, 278
Parametric mortality models, 585
Paranthropus robustus, 419
Parathyroid hormone (PTH), 396
Parry fractures, 362, 605
Partial correlation, 234
Passive infanticide, 118
Patellar fractures, 358, 367
Pathological fracture, 346
Paul H. Rosendahl, PhD Inc., 77
PCR. See Polymerase chain reaction (PCR)
PCR amplification, 312, 465, 471–472
Peabody Museum of Archaeology and
Ethnology, 11
Pearson, Karl, 41
Pearson’s Coefficient of Racial Likeness (CRL) statistic, 490, 495
Pecos Pueblo studies, 195–196
Pedersen, P. O., 266
PeeDee Belemnite (VPDB), 423
Pelvic fractures, 357–358, 366
Penetrating fracture, 346, 352
Penetrating injuries, 341
Penrose’s Size and Shape statistic, 490, 495
Periapical
bone loss, 317, 330–331
periodontitis, 322
Perimortem
bone modifications, 80
trauma, 58, 75, 95–97
Periodontal disease, 319, 321, 325
Periostitis, 605
Peripheral quantitative CT (pQCT), 188
Permanent teeth, 289–290
Peterson fracture classification system, 362
PCAgen, 217, 234
pH, 464
Phalanges, fracture of, 358–359, 366–376, 376
Phenotypes, nonmetric trait variation, 553
Phoebe Heart Museum, 12
Phosphate, bone composition, 450
Phosphorus, bone composition, 396, 444, 450, 452–454
Photographic superimposition, 42, 54–55
Photomicrographs, 254
Phototropism, 423
Phrenology, 8
Phylogeny, 496
Physical activity, 396–398
Physical anthropology, 44
Physical weathering, 80–81
Piltdown hoax, 451
Pima Indians, dental morphology, 288
Pinocchio effect, 219
Plague’s, 475, 572
Plains Indians, 31
Plant-based diets, 415
Plaque-related diseases
  bone loss from alveolar process, 319–324
dental calculus, 312–313
dental caries, 313–319, 332–333
  overview of, 311–312
Pleistocene era, 393
PLS3PAC, 223, 234
Pneumothorax, 356–357
Point mutations, 476
Polygenesis, 7–8
Polymerase chain reaction (PCR), 465–467,
  473, 476
Polymorphisms, genetic, 467
Polynesian Islands
  cranial series, 501, 510, 524–525
  smallest Mahalanobis’ distance, 517–518, 523
  stepwise discriminant function analysis, 512
Pompeii, 572
POPKIN, 494
Population studies
  dental morphology
    casts, 288, 291
    deciduous dentition, 289–290
    permanent dentition, 289–290
    problem definition, 286–287
    sampling, 288
    skulls, 288–289
    traits, types of, 290–291
  growth rates, 582–583
  paleodemography, 561–591, 601–613
  regulation, DNA analysis, 555
  replacement, 22, 469–470
POCON, 494
Positional plagiocephaly, 344
Postcranial modification patterns, 89–91
Postgenomic age, morphometrics and biological
  anthropology, 207–227, 232–235
Postmortem
  alterations, 55–56, 72
  bone modifications, 80
  damage, 85
  trauma, 58, 95–97
Potassium, bone composition, 453
Pott’s disease, 347, 356, 475
Pre-Columbian studies, genetic diseases, 474–475
Pregnancy, 84, 398–399
Prehistoric health and disease
  bone loss and osteoporosis, 387–404
  dental pathology, 301–333
  skeletal trauma, 341–378
  Premolars
    caries, 314–315, 333
    dental morphology, 288–289
    dental pathology, 326
    development of, 267–268
    mandibular, 279–280
    maxillary, 276–277
  Premortem alterations, 72
  Preservability, 265
  Preservation ethic, 24–26
  Preservation of bone
    in children, 119–120
    extrinsic factors, 80–81
    intrinsic factors, 81–85
  Preserved → Found transition, 574
  Primary burial, 100
  Primate studies
    dental morphology, 279, 290
    morphometrics, 208–209
    paleodemographics, 605, 608, 612–613
  Principal components analysis (PCA), 216–217,
    221–222, 223, 492–493, 495
  Principal coordinate analysis, 495
  Prior distribution, 579, 585
  Procrustes superimposition, 215–216, 219–220,
    222–223, 225, 234
  Professional associations, 19
  Programming languages, R, 234
  Projection radiography, 342
  Protein
    dietary, 313, 395–396, 417
    morphometric studies, 215
  Proteinase K, 464
  Protein radio immunoassay (pRIA), 49, 60
  Protestant Reformation, 4–5
  Protocone, 279
  Protoconid, 281
  Protostylid, 282
  Pseudofractures, 374
  Pseudopathology, 374
  Pubic symphysis, 333, 565, 577–578, 580, 607
  Pueblo Indians, 157
  Pyloric stenosis, 270
  Q-mode analysis, 495–496
  Qualitative Bone Index, 79
  Quality of life, 387, 552, 570
  Quantification, DNA analysis, 467
  Quantitative computed tomography (QCT), 395
  Quantitative methodologies
    metric analysis of skeletal remains, 487–522
    nonmetric trait variation, 533–554
Quarangni Cave (Fiji), 102
Quartz, 452
R, statistical programming, 222, 234
Racial differences
  age estimation, 157
  bone mineral density, 83
Racial typology, 487
Racism, 17, 23, 31, 546
Radioactivity, impact of, 552
Radio carbon
  analysis, 56
dating, 104, 414, 605
Radiographic studies, 56–57, 60, 123, 126,
  128, 138, 188, 196–197, 342–343,
  369–370, 546
Radiolucency, periapical, 331–332
Radius fracture, 362–363, 366
Rat studies, nonmetric trait variation, 550
Ray, John, 7
Reburials, 82
Reference sample, paleodemographic research,
  577–578, 580, 585
Regional acceleratory phenomenon (RAP),
  161, 165
Regression, 163
Reincarnation, 5
Relatedness, 27
Reliability studies, 498
Religious beliefs, 3–5, 23
Remodeling, of bones, 389
Renaissance, 6
Repatriation, 71, 106
Replication studies, 498
Reproducibility, DNA analysis, 467
Resampling, morphometric analysis, 222–224, 232
Residence studies, 430–431
Restriction fragment length polymorphisms
  (RFLPs), 465, 476
Resurrectionists, 6
Rhizobium, 425
Rholf, Jim, 235
Ribosomal RNA, 462
Ribs
  age-estimation methods, 181–182
  fractures, 356–357, 365, 375, 589
Rickets, 377, 552
Rigidity of bone, biomechanical analysis, 185–186
Ritual cannibalism, 98–99
R-mode analysis, 495–496
Rollet, Etienne, 41
Root(s)
  characterized, 267
dental anatomy, 239
dental morphology, 273–283, 289
development of, 250
number, 277, 279, 283
  surface caries, 329, 332–333
  Rotation, 373
Ryukyu Islands, 505, 515, 521, 524
Sacral fractures, 354–356
SAGE3D, 234
Sagittal synostosis, 550
Sample size
  metric analysis, 495
  multivariate analysis, 492
Sampling
  age estimation, 162–163, 166
  paleodemographic studies, 567, 575, 601
  significance of, 118–121
San Cristobal Pueblo (New Mexico), 104–105
San Diego Museum of Man, 12
SAS, 494
Scalping, 91–92, 97, 377
Scanning electron microscopy (SEM), 242, 244,
  251, 343
Scanning electron microscopy/energy dispersive
  spectroscopy (SEM/EDS), 49, 60
Scaphoid fractures, 358, 366
Scapular acromion, 549
Scapular fractures, 357, 365
Scatterplots, 216
Scavenging behavior, 88
Sciatic nerve, 364
Sciulli, Paul, 289
Sclerosis, 331–332
Scratching, dental, 310
Seals, 427
Secondary burials, 77–78, 89, 100
Secondary dentine, 332
Secondary fracture, 346
Second moments of area (SMA), 185–186
Self-correction process, 25
Seminars, 30
Seneca Indians, nonmetric trait studies, 544
Senescence, 613
Sex
  determination, 51–52, 125, 575
  estimation, 122–125, 138, 575–576
  identification, molecular, 472–473
Sexual dimorphism, 122–125, 138, 192–195
Shape
  in morphometric analysis, 216–219, 223–225
  in sex estimation, 575
  in trauma analysis, 344–345
Shapiro, H. L., 42
Sheets, David, 218, 223, 233
Short tandem repeat (STR), 461, 463, 474
Shoveller’s fracture, 368
Sickle cell anemia, 356, 474–475
Siler model, 584–585
Silica/guanidine isothiocyanate, DNA analysis, 464
Silen’s method, 420
Silumiut (Canada), 81
Simple3D, 234
Simulation studies, 585
Single nucleotide polymorphisms (SNPs), 463
Sinus, periapical bone loss, 331
Size
  in morphometric analysis, 215–216
  in sex estimation, 575
  tooth morphology, 268–270
Skeletal assemblage, historical stages, 72–73
Skeletal biologists, ethical responsibilities of, 18–26
Skeletal biology
  characterized, 603
  forensic anthropology distinguished from, 44
Skeletal collections, legal perspectives, 28–29
Skeletal maturity, 150–151
Skeletal trauma analysis
  causative interpretation, 374–378
  dislocation, 345
  fractures, 346–374
  methodologies, 342–344
  paleoepidemiology, 348–350
  pseudopathology, 374
  purpose of, 341
  shape, 344–345
  soft tissue, ossification of, 344
  trauma
    defined, 341
    types of, 341–342
    violent injuries, 376–377
Skeletonization, 72, 74, 93
Skull
  collections, curated, 93–94
  dental morphology, 288–289
  fractures, 350–352, 365, 372
  racks, 91–92
Smith scoring system, 327
Smithsonian Institution, 11
Smithsonian’s National Museum of Natural History, 46, 50, 52
Smooth-surface caries, 332
Snow, Charles E., 42
Social classification, 53
Social organization, reconstruction using DNA, 470–472
Social structure, 553
Society for American Archaeology Statement Concerning the Treatment of Human Remains, 25
Sociocultural environment, significance of, 377–378
Socioeconomic status, impact of, 588–589
Soft tissue
  decomposition of, 72, 74
  development of, 552
  DNA analysis, 464–465
  in forensic anthropology, 48, 54
  ossification of, 344, 536
  plaque-related disease, 321
  preservation of, 55
  rib fractures and, 356
Soil
  microbiology, 81
  oxides, 453
  pH, 119
  screening, 77
Southeast Asia Mainland
  cranial series, 504–505, 510–511, 524
  smallest Mahalanobis’ distance, 521, 523–524
  stepwise discriminant function analysis, 514
Southeast Asian Islands
  cranial series, 504, 510–511, 524
  jackknifed classification, 508
  smallest Mahalanobis’ distance, 517, 520, 523–524
  stepwise discriminant function analysis, 514
Southwest Indians, dental morphology, 271, 273
Spacer-oligonucleotide typing, 476
Spina bifida, 270
Spinal cord, 368
Spinal fractures, 368
Spinal TB, 475
Spiral fracture, 95, 346, 368
Spitalfields studies, 130, 138, 164, 397
Splintering of bone, 352
Split Rock Creek, Late Woodland Site, 91–92
Spoligotyping, 476
Spondylolysis, 355
Spreadsheet applications, 255, 325
SSPS, 494
Stable isotope analysis
  benefits of, 431–433
  developments of, 414–415
Stable isotope analysis (Continued)
future directions for, 432
historical perspective, 415–416
isolation methodologies, 417–420
mass spectrometry, 414, 420–422
in skeletal biology
applications, 413–414, 422–423
infant feeding/weaning studies, 428–430
freshwater resources, 426–427
paleodiet, 415, 423–425
residence and migration studies, 430–431
of tissues, 416–417
variations, 415–416
Standard Mean Ocean Water (VSMOW), 423
Standard of living, 552
Standards for Data Collection From Human Skeletal
Remains, 72, 79
Staphylococcus aureus, 371
Statistical testing, 50
Stepwise discriminant function analysis, 499–500, 507–511
Stepwise regression analysis, 499
Sterile crescent, 611
Sterno-clavicular joint, 366
Sternum, fracture of, 357, 365
Stewart, T. Dale, 42–45
Stress
activity-induced, 548–549
fracture, 346, 355
levels, 133
Striae of Retzius
characterized, 304–305
tooth development
functions of, 243–244, 247, 250, 252–254
irregular, 246–247
Strontium
bone composition, 444–453, 455–457
dietary studies, 454
isotopes, 430–431
stable isotope analysis, 419–420
Subadults, age estimation, 158–160
Subgingival calculus, 312, 317, 329–330
Subsistence strategies, 190, 193
Suchey-Brooks pubic symphysis stages, 50, 333
Sulfur, primary and reference standards, 423
Supracondylar fractures, 360, 364
Supragingival calculus, 312, 326
Surface contouring, 488
Survivorship, 603, 607
Sutton Hoo site, 76
Swartrkans site (South Africa), 87
Symmetry, tooth morphology, 269
Syphilis, congenital, 306
Systemic diseases, 347
Systemic disorders, 552
Systems biology, 208
Tanning agents, 465
Taphonomics, 51
Taphonomy
as assemblage history, 72–73
characterized, 55, 71–72, 100–102
defined, 71
forensic, 72
paleopathology and, 85–86
quantitative, 102–104
systematic analysis, 104
variables in data collection protocol, 101
Taq polymerase, 466
Tasmania
cranial series, 503, 524
jackknifed classification, 508
smallest Mahalanobis’ distance, 517, 520, 524
stepwise discriminant function analysis, 513
Technological advances, types of, 47–48
Teeth
dental development. See Dental development
fracture of, 354, 365
see also specific types of teeth
Teotihuacán population, nonmetric trait studies,
544–545
Teratogens, 552
Terrain, significance of, 191–192
Terry, Robert J., 46–47
Terry human remains collection, 46, 50–53
Thalassemia, 474
Theories
bioarchaeological ethics, 3–31
forensic anthropology, 41–61
taphonomy, 71–106
Thermal alteration, 71, 97–98, 102
Thermal ionization mass spectrometry
(TIMS), 421
Thermonuclear devices, 56
Thin-plate spline, 235
Tibia
age-estimation methods, 179–180
biomechanical analysis, 194–195
characterized, 53, 82–83, 85, 96, 163–164
fractures of, 364, 367
Titanium, bone contamination, 452
Tomes, C. S., 281
Tomes’s root, 280–281
Tool marks, implications of, 71, 96. See also Cutmarks

Tooth
  development, 138
  emergence, 126
  enamel. See Enamel
  number, 268–270
  see also specific types of teeth

Topinard, Paul, 41

Torus fracture, 345–346

Total fertility rate (TFR), 607–611

Toxins, bone-seeking, 455

tpsRegr, 218, 234

Trace correlation, 235

Trace element analysis, 104, 416–417

Transgenic mice, 208

Transition analysis, 579

Transmission electron microscopy (TEM), 248

Transverse fracture, 346, 351, 360, 368

Trauma
  blunt force, 351, 605
  identification of, 57–58
  indirect, 346, 352–353, 357
  interpretation, 42, 58
  severe, 363

Traumatic injury, 89, 91, 97

Tree structures, 496

Trehphination, 86

Trophic level, nitrogen isotopes, 425–426

Trophy skulls, 9

Trotter, Mildred, 42, 46, 53

T-tests, 223

Tuberculosis, 347, 356, 475–476

Tuberculum dentale, 269, 272, 275, 279, 288
  intermediate, 282
  sextum, 282

Turner, Christy G., II, 271–272, 287

Twin studies, 288

Type II osteons, 153–154, 158

Tyrolean “Iceman”, 186–188, 194–195, 477

Tyson, Edward, 7

Ulna
  age-estimation, methods, 176
  fracture of, 362, 366
  olecranon process, 549

Ultrasound studies, 54

Ultraviolet light (UL), 242, 246, 464, 466

Ultraviolet radiation, 80

Uniform Anatomical Gifts Act, 21–22

United Nations (UN), 19

Univariate statistics, 489–490

UPGMA, 496–497, 524

Upper Paleolithic Period, 4, 75, 194, 199

Uranium, 452, 454, 456

Uto-Aztecan premolars, 276

Variance-covariance matrix, 216

Value systems, 3, 13

Variability, significance of, 235, 265

Variance
  between-group, 495
  of eigenvalues, 235
  significance of, 498

Variance-covariance matrices, 222
  within-group, 495

Vascular damage, 371–372

Velim Skalka site (Czech Republic), 90

Venereal diseases, 610–611

Vertebrae
  fractures of, 354–356, 399
  injuries to, 377

Very old population, 581

Video-eulogies, 73

Vienna Standard Mean Ocean Water (VSMOW), 423

Violence, interpersonal, 376

Vitamin D, 83–84, 396

Vitamin deficiencies, 373, 396

Volumetric data, morphometric studies, 215

von Ebner’s lines, 242

Voxels, 233

War crimes, 59

Warfare, 574

Warping, postmortem, 86

Washburn, Sherwood, 542

Weak ergodicity, 581

Weathering
  patterns, 71, 74, 103–104
  physical and chemical, 80–81
  thermal alteration vs., 102

Western Reserve University, 12

West Level 7 mortality function, 584

Wharram Percy population, bone loss studies, 398–399

Whitlockite, 444

Wilder, Harris Hawthorne, 42

Wilson bands, 238, 246–248, 255

WinEDMA, 219, 235

Wireframe, 235
Witch execution, 93
Wolff’s Law, 183–184
Women’s health
   bone maintenance, 395
   breastfeeding, 428–430
   lactation, 398–399
   mammary glands, 454
   pregnancy, 84, 398–399
Workshops, 30
World Fertility Survey, 564
World Health Organization (WHO), 390
Worldview, conflict in, 15–16, 19
World War II, 77
Wrist fracture, 358–360, 399
X chromosomes, 125, 312, 463, 472–473, 547
X-ray absorptiometry (DEXA), 395, 401–402
X-ray diffraction studies, 451–452, 455
X-rays, 126, 128, 215, 257, 308, 318, 323, 342
Y chromosomes, 125, 312, 461, 463–464, 470, 472–473, 547
Yearbook of Physical Anthropology, 553
Young, Nathan, 222
Young Mortuary Complex burials, 93
Zegura, Steven, 287
Zinc, bone chemistry, 444, 449–450, 456–457
Zirconium, bone composition, 450, 452
Zooarchaeology, 71–72, 87, 98–99
Zoology, nonmetric trait variation, 546, 552
Z-scores, 191, 223, 499
Zubov, A. A., 265
Zuni Indians, 28