To the Graduate Council:

I am submitting herewith a dissertation written by Natalie Renee Shirley entitled "Age and Sex Estimation from the Human Clavicle: An Investigation of Traditional and Novel Methods." I have examined the final electronic copy of this dissertation form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Anthropology.

Richard Jantz, Major Professor

We have read this dissertation
and recommend its acceptance:
Andrew Kramer
Steven A. Symes
Mohamed Mahfouz

Accepted for the Council:
Carolyn R. Hodges
Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

AGE AND SEX ESTIMATION FROM THE HUMAN CLAVICLE: AN INVESTIGATION OF TRADITIONAL AND NOVEL METHODS

A Dissertation
Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Natalie Renee Shirley
May 2009

I dedicate this dissertation to Linda Finley, because without her none of this would have ever happened. "Nino" has been more like a mother than a cousin to me since the day I was born, and her unique outlook on life continues to amaze and inspire me to this day. She has been my mentor, my teacher, and my friend. She taught me the value of music, education, family, and faith. Throughout my childhood, she took me on fascinating and educational vacations, helped me with my homework, made me silver dollar pancakes, took me shopping on my birthday, and even gave me baths in her kitchen sink as a baby. I learned from Nino that when I feel frustrated and ready to give up, I should just take a nap—it will all be better when I wake up. Her genuine love of life, learning, travel, and people is just a small part of what makes Nino special. She does not judge, she loves unconditionally, and she is always willing to lend a patient ear. I can honestly say that if everyone were like Nino the world would be a better place. This dissertation is just a small symbol of the passion, dedication, hard work, and perseverance that Nino has instilled in me. I dedicate it to her in remembrance of her mother and father, Delia and Russell Finley (DeeDee and Uncle Russell).

Acknowledgements

The following list of people and organizations is in no particular order, as I extend my sincerest gratitude to each of them equally. First, I would like to express thanks to my committee, Dr. Richard Jantz, Dr. Andrew Kramer, Dr. Steven Symes, and Dr. Mohamed Mahfouz. Thanks also to the National Institute of Justice (Grant Number 2007-DN-BX-0004), the William M. Bass Endowment, and the Kneberg-Lewis Anthropology Graduate Fellowship for funding this research project. I especially appreciate Dr. Kramer's efforts in making the NIJ grant a feasible funding option. Additionally, thanks to the Graduate Student Senate, the Anthropology Department, and the American Academy of Forensic Sciences for funding travel to disseminate preliminary results of this research.

I would also like to thank Dr. Lyle Konigsberg for his help with the transition analysis and hazard models and Bridget Algee-Hewitt for her patient advice with "R". I'm grateful to have friends and colleagues who willingly take time out of their hectic schedules to score clavicles—thanks Dr. Megan Moore, Angela Dautartas, and Kathleen Alsup. And I'd be remiss if I didn't thank Dr. Katie King and Rebecca Wilson for their help with the McCormick Collection, as well as my office mates Christine Pink, Kathleen Alsup, and Megan Moore for putting up with the "clavicle mess" for the past few years. I also extend my gratitude to everyone who helped scan the bones, including Dr. Mohamed Mahfouz, the radiology staff at the UT Medical Center, graduate student assistants, and Megan Moore. Dr. Mahfouz opened the door for me to begin the scanning project and provided me with the funds and software necessary to make the

project possible. Megan taught me how to construct boxes, load bones, scan the bones, and how to use the software to access my scans; her assistance with this project was invaluable.

I owe special thanks to the University of Tennessee's Anthropology Department for access to the William M. Bass Donated Collection and the William F. McCormick Collection. I also appreciate Lyman Jellema and the staff at the Cleveland Museum of Natural History for hosting my research visit to the Hamann-Todd Collection. I am grateful to Dr. Lyle Konigsberg for making the McKern and Stewart data accessible on the web and to Dr. William McCormick for his foresight in curating the McCormick Collection. Additionally, thanks to Dr. Steven Symes for loaning some of his archives for scanning. I would be remiss if I did not mention that Steve has been a generous friend, colleague and mentor to me throughout my graduate career. He has seen me through my thesis, preliminary exams, and my dissertation. His genuine interest in students provided me with opportunities that I would not have had otherwise. Thanks SO MUCH, Steve.

I am eternally indebted to Brandon Merkl for his help with the clavicle atlas, automated computer measurements, and the PCA-FDR statistical analyses. I am fortunate to have discovered in Brandon a brilliant, dedicated, and patient collaborator. Even though he was in the midst of preparing his own dissertation, Brandon always made time to meet with me and to explain complex engineering concepts in a manner that my anthropological brain could understand. His genius provided the basis for the innovative shape analyses and color maps in the pages to follow.

Three other special men in my life deserve my utmost appreciation and gratitude. First, thanks to my father, Dennis Langley, for believing in me and for never doubting that I could do this. I also owe everything to my husband, George Shirley, for enduring two years of marriage to a doctoral candidate. It cannot be exciting watching your wife spend all hours of the day and night at a computer, but George never complains and offers nothing but love and support. I do not take it for granted that I am very lucky to have George by my side. And finally, thanks to my advisor and mentor, Dr. Richard Jantz. Dr. Jantz is not merely an advisor to his students, he is also a friend and colleague. His unique mentoring has helped me to learn and grow past my conceived capabilities. It was his suggestion that somebody should "do something with those clavicles" that led to this exciting project that combined both of our research interests. I hope to be able to enjoy collaborations with Dr. Jantz for many years to come. I came to UT as a student, but I am leaving an anthropologist. And I should add—an anthropologist who knows that I really know nothing at all. Without statistics, that is. And Dr. Jantz, I'll never call you 'Richard'. Sorry.

Last, but certainly not least, thanks to all my friends, family, and colleagues who have endured boring clavicle conversations and still offered inspiration and support—Franklin Damann, Dr. Mariateresa Tersigni-Tarrant, my mom, Parin and Aunt Shirley, Nino, Bruce and Diane Van Horn, Chris Pink, Kathleen Alsup, Megan Moore, Anne Kroman, Greg Berg, Liz DiGangi, Jon Bethard, and anyone else I may have forgotten. Thank you … thank you very much … thank you.

Abstract

This study investigates skeletal maturation and sexual dimorphism in the human clavicle in the American population. Specifically, traditional methods of age and sex estimation were compared to novel approaches. Transition analysis, or probit regression, was used to study secular trends in epiphyseal union of the medial clavicle. A Bayesian approach was used to develop age-at-fusion ranges that are less sensitive to the effects of developmental outliers and age mimicry. Linear regression was used to evaluate the sexing accuracy of three commonly-used clavicular measurements and six newly developed measurements. Additionally, a statistical clavicle atlas was used to explore size and shape dimorphism.

For the epiphyseal union study, clavicles from 1,289 individuals from cohorts spanning the 20th century were scored with two scoring systems. A simple 3-phase scoring system proved the least subjective, while retaining accuracy levels. Significant secular trends were apparent in the onset of skeletal maturation, with modern Americans transitioning to fusion approximately 4 years earlier than early 20th century Americans and 3.5 years earlier than Korean War era Americans. Consequently, the secular trend towards earlier maturation appears to have occurred primarily during the latter half of the 20th century. These results underscore the importance of using modern standards to estimate age in modern individuals.

For the sexual dimorphism study, all analyses were performed on three dimensional models of CT scanned clavicles. Linear discriminant analysis was performed on nine computer-automated measurements from 1,414 clavicle models. Cross-validated accuracy rates of the best models hovered around 92%. Two new measurements of the lateral end proved to be

useful sex estimators, whereas accuracy rates from the medial end were low. Additionally, sex-specific statistical atlases were used to visualize areas of highest dimorphism. A statistical treatment combining Principal Components Analysis (PCA) and Fisher's Discriminant Ratio (FDR) showed high magnitudes of curvature difference between males and females, particularly in the anterior and superior curvature of the midshaft and the posteriorly oriented curvature of the lateral end. The areas highlighted by the PCA-FDR method show promise as new sexing criteria for the human clavicle.

Table of Contents

Chapter 1: Literature Review 1

The Clavicle: Development, Growth, and Maturation 2

Maturation 7

Adult Sex Estimation 16

Chapter 2: Materials and Methods 28

Medial Clavicular Epiphyseal Fusion 28

Size and Shape Properties of the Clavicle 38

Surface Model and Atlas Creation 38

Principal Components Analysis and Fisher's Discriminant Ratio 44

Linear Discriminant Analysis 45

Chapter 3: Results 49

Medial Clavicular Epiphyseal Fusion 49

Sex Estimation 62

Linear Discriminant Analysis 62

Clavicle Atlas: Size and Shape Dimorphism 64

Chapter 4: Discussion 74

Secular Change in Skeletal Maturation 74

Sexual Dimorphism 777

Chapter 5: Conclusions—Looking Ahead 81

References Cited 87

Appendices 101

Appendix A 102

Appendix B 104

Appendix C 105

Vita 108

List of Tables

Table Page

1. Observer error 50
2. McCormick individuals descriptive statistics 51
3. Todd individuals descriptive statistics 52
4. McKern and Stewart males descriptive statistics 53
5. Transition ages 53
6. T-tests for sexual dimorphism 56
7. T-tests for secular trends 58
8. Male age ranges 59
9. Female age ranges 59
10. 5-Phase test of PDF age ranges..... 61
11. 3-Phase test of PDF age ranges..... 61
12. LDA accuracy rates 63

List of Figures

Figure Page

1. No fusion 30
2. Beginning fusion 31
3. Active fusion 31
4. Active fusion (lateral view..... 32
5. Recent fusion 32
6. Complete fusion 33
7. FDB survivorship 37
8. FDB age-at-death distributions 37
9. Packed clavicles 40
10. Scanned bones 40
11. DICOM image slice 41
12. 3-D surface generations 42
13. Traditional clavicle measurements 46
14. Measurements from bone models 46
15. Age-at-transition distributions 55
16. Three sample comparisons 56
17. A-P color maps of PC1 65
18. S-I color maps of PC1 65
19. A-P color maps of PC2 67
20. S-I color maps of PC2 67
21. PC2-PC10 inferior view 68
22. PC2-PC10 superior view 70
23. PC2-PC10 anterior view 71
24. PC2-PC10 posterior view 72
25. PC2-PC10 medial and superior view 73
26. PC2-PC8 inferior view 80
27. Medial axis representation 85
28. Medial axis representation with 32 ellipses 86

Chapter 1: Literature Review

Skeletal biologists apply their knowledge of human skeletal variation to answer questions about skeletal age, sex, ancestry, stature, occupation, medical history, personal habits, health status, and cause of death. Skeletal analyses encompass prehistoric, historic, and modern human populations and are made on the population level, as well as on the individual level. A primary component of any skeletal analysis is determining age and sex, and anthropologists frequently must accomplish this task in the face of incomplete or fragmentary skeletons. For this reason, developing aging and sexing criteria from various skeletal elements has been a primary research focus in skeletal biology.

Estimates of sub-adult skeletal age are typically based on long bone length, epiphyseal fusion, and dental development or eruption sequences and are more precise and accurate than adult age estimates based on degenerative changes at the pubic symphysis, auricular surface, sternal rib ends, cranial sutures, or dental attrition. However, sex estimation is more accurate with adult skeletons than with sub-adults. Accuracy rates for sexing infant skeletons hover around chance, although rates improve as puberty approaches and sexual dimorphism becomes more apparent.

This dissertation will explore age and sex estimation from the clavicle in modern Americans. Age estimation will focus on the medial clavicular epiphysis, and adult sex estimation will be investigated in terms of size and shape dimorphism. Maturation of the medial clavicular epiphysis will be examined for secular trends, as well. Accordingly, this review

of the literature is multi-faceted and will discuss major achievements in age estimation from the clavicle, secular trends in maturation, and sex estimation from the human skeleton.

The Clavicle: Development, Growth, and Maturation

The clavicle is considered a long bone, as it has a medullary cavity and an epiphysis at either end to permit growth. The clavicle is the first fetal bone to undergo primary ossification, and its medial epiphysis is the last to fuse (Humphrey 1998; Scheuer and Black 2000). However, whereas other long bones undergo initial endochondral ossification, the clavicle ossifies via intramembranous ossification with no prior endochondral ossification (Kreitner et al. 1998). The two primary ossification centers appear by the 6th week intrauterine and fuse together about one week later (Ogata and Uhthoff 1990; Scheuer and Black 2000). After the osteoid matrix is laid down, cartilage appears at the acromial and sternal ends of the bone, at which point growth becomes a combination of endochondral and membranous ossification. The medial cartilaginous mass contributes more to clavicular growth in length than does the lateral mass, perhaps as much as 80% of the bone length (Ogata and Uhthoff 1990; Scheuer and Black 2000). The combination of the spatial location of the two ossification centers at either end of the bone and endochondral ossification at these sites gives the clavicle its unique s-shape by 8-9 prenatal weeks (Ogata and Uhthoff 1990; Scheuer and Black 2000). The bone attains its adult form by 11 prenatal weeks (Ogata and Uhthoff 1990). Growth slows after birth until the growth spurt between 5 and 7 years, then slows again until the pubertal growth spurt (Black and Scheuer 1996).

During adolescence secondary ossification centers appear at the medial and lateral ends of the clavicle. Todd and D'Errico (1928) documented fusion of the lateral epiphysis in the early 20th century, but it has since been given little attention in the literature. Sometimes the lateral epiphysis appears as a separate flake of bone, whereas in other cases the epiphysis simply glazes over to assume a quiescent appearance; this typically occurs during the 20th year (Todd and D'Errico 1928). Far more attention has been devoted to studying the development and fusion of the medial epiphysis (Black and Scheuer 1996; Ji et al. 1994; Jit and Kulkarni 1976; Kreitner et al. 1998; Li et al. 2001; McKern and Stewart 1957; Schmeling et al. 2004; Schulz et al. 2005; Schulz et al. 2008; Stevenson 1924; Todd and D'Errico 1928; Webb and Suchey 1985). Medial epiphyseal ossification begins at the onset of puberty, but the medial epiphysis does not fuse to the shaft completely until some 10 years after its initial appearance (Kreitner et al. 1998; Scheuer and Black 2000). Consequently, medial epiphyseal fusion has proven useful in estimating skeletal age in young adults. The epiphysis appears initially as a small speck of bone in the center area of the sternal end and spreads until it nearly covers the entire medial surface. Scheuer and Black (2000) offer the following timeline for clavicular maturation: a well-defined medial flake appears between 16 and 21 years; the flake covers the majority of the medial surface between 24 and 29 years; complete fusion occurs between 22 and 30 years. An Austrian study offered a similar age summary, but without overlapping ages: commencement of fusion between 18 and 20 years, active fusion between 21 and 25 years, and complete fusion between 26 and 30 years (Szilvassy 1977).

Among the earliest mention of the medial clavicular epiphysis in the American population was Stevenson's (1924) documentation of epiphyseal union in the arms, legs, and girdles. Stevenson's (1924) observations were based on the Hamann-Todd Collection (then called the Western Reserve Collection). He noted the commencement of union as early as age 22 and completed union in all cases by age 28, though he did not mention sex differences in epiphyseal union. Several years later, Todd and D'Errico (1928) published a more extensive study of the medial and lateral clavicular epiphyses. They scored the Western Reserve clavicles according to a 4-phase system: (1) no union, (2) beginning union, (3) recent union with a scar, and (4) complete union (with loss of all trace of the site of union). They discovered that union occurred between 18 and 29 years. The ossifying epiphysis typically begins to unite around age 21, and union is practically complete by age 25. They did not report any significant sex or race differences in their sample.

During the 1950's McKern and Stewart (1957) reported their findings on epiphyseal union in the Korean War dead; a portion of their report discussed the medial clavicle. They used a 5-phase scoring system, adding a phase for "active" union to Todd and D'Errico's (1928) system : (1) no union, (2) beginning, (3) active, (4) recent, and (5) complete. McKern and Stewart (1957) reported that union begins at age 18, possibly as early as 17, and the majority fuse by age 25. They found clavicles with unattached epiphyses in individuals up to age 22; however, complete union was not evident prior to age 23, and all individuals had reached the final stage of union by age 30. Since the sample was comprised of males only, they could not investigate sex differences. They did, however, offer a mechanism for epiphyseal fusion.

According to their results, union begins at the center of the epiphyseal face and proceeds superiorly, where it may progress either anteriorly or posteriorly; the inferior margin is the final site of fusion.

In order to provide the forensic community with standards from a modern reference sample, Webb and Suchey (1985) published their findings on an autopsy sample from the Los Angeles area. The sample consisted of skeletal material from over 800 autopsies from the late 1970's. Scoring was done on the bones after extraction, processing, and cleaning. Their 4-phase scoring system differed from either of the previous systems in that it placed more emphasis on the nonunion stages: (1) nonunion with no epiphysis, (2) nonunion with separate epiphysis, (3) partial union, and (4) complete union. Phase 1 lasted to age 25 in males and 23 in females. Phase 2 was observed in males from 16-22 years and in females from 16-21 years. Phase 3, which marks the onset of fusion, was noted in males from 17-30 years and in females from 16-33 years. Complete union was observed as early as age 21 in males, with all males completely fused by age 31. The earliest age at which complete union was noted in females was age 20, and all were fused by age 34.

Black and Scheuer (1996) documented medial clavicular fusion on a compilation of skeletal material from the 18th-19th century Spitalfields, St Bride's, and St. Barnabus documented collections, as well as individuals from the 20th century Museu Bocage Portuguese collection in Lisbon. Their 5-phase scoring system is similar to Webb and Suchey's (1985) in that it gives considerable attention to the nonunion phases: (1) distinctive ridges and furrows on the metaphyseal surface; no epiphyseal flake attached, (2) less ridges and furrows on the

metaphyseal surface; no epiphyseal flake attached, (3) flake commencing fusion, (4) flake expanded across the metaphyseal surface, (5) complete fusion with no trace of a fusion line. The authors claim that the first two phases offer some resolution for distinguishing younger individuals around the age of 15. Their age ranges did not distinguish between males and females, and they found considerable overlap between samples. Phase 1 ranged from 11-17 years, phase 2 from 15-22 years, phase 3 encompassed 19-23 years, phase 4 individuals were between ages 23-28 years, and phase 5 individuals were 25+ years. The sample size was too small to make inferences about genetic/population differences and/or secular trends.

Although CT scans have been used to study epiphyseal fusion, most of these studies use slice thicknesses that are too thick to detect meaningful details about the various stages of fusion (i.e. thicknesses greater than 1 mm). For example, in one study slice thickness ranged from 1mm-8mm (Kreitner et al. 1998). They divided union into 4 stages: (1) nonunion without ossification of the epiphysis, (2) nonunion with detectable ossification of the epiphysis, (3) partial union, and (4) complete union of the epiphysis with the clavicular metaphysis. The ossified epiphysis appeared between 11 and 22 years. Partial union was documented from 16-26 years. Earliest age at complete union was 22 years, and union was complete in all individuals by age 27. They found no statistically significant differences in males and females, possibly due to small sample size. However, scoring CT-scanned bones from scans of varying slice thicknesses should be approached with caution because details apparent in 1 mm slices may not be visible in 8 mm slices. Furthermore, standards developed from CT scans should only be used to age CT scans and not dry bone.

In sum, the medial clavicular epiphysis has been established as a reliable aging method for young adults. However, the existing studies employ various scoring systems, and there is no standard scoring system with established error rates. Furthermore, the most up-to-date information for the American population is from 1970's autopsies (Webb and Suchey 1985). Accordingly, this dissertation aims to provide updated standards for the American population. Additionally, this study will evaluate the utility of 2 scoring systems (a 3-phase and a 5-phase scoring system) and provide observer error rates for each system. Secular trends in skeletal maturation of the medial clavicular epiphysis will be examined by comparing a modern autopsy sample to two earlier samples. A Bayesian approach will be used to derive robust age ranges for modern individuals, and these age ranges will be tested on a separate population.

Maturation

Among the primates, humans have a unique pattern of growth and development in that all stages of life history are relatively longer, and we possess a period of rapid skeletal and somatic growth just prior to adulthood—the adolescent growth spurt. Some argue that the adolescent growth spurt may have been in place as early as *Homo erectus*, as evidenced by the Nariokotome skeleton, but this remains open to debate (Clegg and Aiello 1999). Nonetheless, human growth prolongation has been attributed to increased encephalization in our species and corresponding patterns of brain development to facilitate learning, language acquisition, and balance acquisition for bipedality (Bogin 1999; Leigh and Park 1998; Penin et al. 2002). Post-cranial somatic growth and reproductive development are delayed, perhaps at the expense of early brain growth, as the human brain reaches 80% of adult weight and size by age

4 and nearly 100% by age 7 (Bogin 1999). This unique growth pattern has led to the suggestion that the adolescent growth spurt may provide a mechanism for the postcranial skeleton to "catch up" to the already mature cranium (Bogin 1999). Regardless, Bogin (1999) argues that the skeletal growth spurt in stature is unique to humans; some primates have a weight or body mass spurt, but none have a stature spurt.

During this growth spurt, the postcranial skeleton begins to assume adult form as the epiphyses unite to the diaphyses and bone growth is completed. This process occurs between 11 and 20 years of age in the majority of skeletal epiphyses, with females typically commencing fusion a couple of years earlier than males. The only exception is the clavicle, in which fusion occurs during the second decade and traces of its union sometimes extend into the late 20's (Scheuer and Black 2000). Although sex differences in skeletal maturation were not well-established until the 1920s, they have been documented in the literature since that time (Pryor 1928; Scheuer and Black 2000; Webb and Suchey 1985).

Efforts to understand human growth variation have explored factors such as health, diet, socioeconomic status, and ethnicity. Roche et al. (1974; 1978) investigated skeletal maturation in the context of race, geographic region, family income, and educational standards of parents in US children ages 6-17 years. They did not find consistent differences between African American and European Americans, as well as no regional or urban-rural differences. However, other researchers have attributed variation in skeletal maturation specifically to ethnic differences (Mora et al. 2001; Schaefer and Black 2005). Population-specific skeletal growth velocities were cited as the primary reason that Japanese sub-adults differ from English

sub-adults in terms of skeletal maturation (Murata 1992). Murata (1992) reasoned that this genetic difference in skeletal maturation tempo during puberty may explain the difference in adult stature between these populations. In a separate investigation, measurements from the radius, ulna, and wrist short bones revealed that Japanese children reached skeletal maturity 1-2 years earlier than European *and* Chinese children (Murata 1997).

Still, ethnicity is cited less frequently than socioeconomic status as the most influential variable in maturational differences (Abioye-Kuteyi et al. 1997; Alberman et al. 1991; Bagga and Kulkarni 2000; Bodzsar 2000; Cardoso 2008a; Kim et al. 2008; Laska-Mierzejewska et al. 1982; Low et al. 1982; Malina 1979; Prado 1984; Rimpela and Rimpela 1993; Todd 1937). Schmeling and co-workers (2000) argue that skeletal maturation occurs in stages that are the same for all ethnic groups; the critical factor that brings about differences in ossification rates is the socioeconomic status of a given population. They maintain that, although a genetically-determined potential of skeletal maturation may exist, this potential does not depend on ethnicity. Instead, growth potential is realized under favorable environmental conditions, namely high socioeconomic status, and population affiliation has no effect on skeletal age (Schmeling et al. 2005). In fact, regression analysis has shown a positive relationship between ossification rates, medical modernization, and economic progress; low modernization levels are associated with delayed ossification (Schmeling et al. 2005; Schmeling et al. 2006).

These findings have practical implications, as well, particularly in instances where accurate age estimates are of vital importance. Specifically, applying skeletal maturation standards to individuals of lower socio-economic status than the reference population will

underestimate age in most instances (Schmeling et al. 2005; Schmeling et al. 2006). This phenomenon was illustrated by a comparison between the above-average socioeconomic status US90 Project Heartbeat! Study children and the Tanner-Whitehouse UK standards, wherein the US90 children matured significantly earlier than the UK children (Tanner et al. 1997). Frisancho and colleagues (1970) documented 36-38% delayed growth in early childhood in poor Central American children compared to US children. The delay decreased to 5-9% during adolescence, which still represents a 0.75-1.8 year delay in epiphyseal union (Cardoso 2008a). Unfortunately, neither of these studies effectively controlled for ethnicity or socioeconomic differences in their analysis. Interestingly, Cardoso (2008a) interprets the significant difference in childhood versus adolescent maturational delay as an indication that skeletal maturation during early childhood is more affected by nutritional upsets than during adolescence. A study by Dreizen and coworkers (1967) supports this hypothesis. They found that chronic under-nutrition retarded skeletal growth and delayed menarche, but the retardation was greater prior to menarche. Furthermore, post-menarcheal undernourished females experienced an extended growth period during which they attained the same adult stature as the well-nourished females. In other words, the developmental stage at which populations are compared to one another may impact the results of growth studies.

Another matter of concern in making developmental status comparisons is the method used to establish stage of union. Many skeletal maturation standards are based on radiographic studies of living individuals, and the appearance of the epiphyses on radiographs are not necessarily the same as they appear in dry bone (Cardoso 2008a; Todd 1937).

Additionally, modern studies are using computed tomography (CT) scans to establish age ranges for epiphyseal fusion. Consequently, biological anthropologists and archaeologists must contend with the fact that the commencement of fusion can be detected significantly earlier with radiographs and CT scans than with dry bone observations. One study of the medial clavicular epiphysis found that radiography and CT produced the same age ranges for fusion 98% of the time (Schulz et al. 2008). However, they were not able to score 13% of the radiographic images properly because the epiphyseal line was obstructed by other structures. Nonetheless, these authors warn that forensic age estimates based on conventional radiographs should use standards developed from radiographs, and age estimates based on CT scans should refer to CT-based standards (Schulz et al. 2008). Similarly, Cardoso (2008a) asserts that "radiographic standards are better suited to estimate bone age in radiographs, and skeletal collections are more appropriate to estimate bone age in dry bone material" (169). Nonetheless, Cardoso (2008a) surmises that eliminating methodological problems would reveal socioeconomic status as the primary cause of maturational differences. Consequently, information regarding socioeconomic status and modernization levels should not be dismissed as trivial when making forensic and/or bioarchaeological age estimates (Cardoso 2008a; Klepinger 2001).

Klepinger (2001) appropriately poses 2 questions related to the reports just discussed: (1) "To what extent do skeletal indicators of maturation follow those of the soft tissues?" and (2) "Is attaching a chronological age to a stage of skeletal maturation becoming increasingly unreliable?" (789). With regard to the first question, at least some research has suggested that

menarcheal onset is more closely related to skeletal age than to chronological age (Dreizen et al. 1967; Simmons 1944). However, with regard to the second question, Klepinger (2001) argues that individuals from lower socioeconomic strata are more likely to be overrepresented as unidentified forensic cases because they live more dangerous lives. Moreover, she maintains that these individuals are equally likely to be advanced or retarded in terms of average growth. In order to address such highly unpredictable variation in individual cases, Klepinger (2001) advises that age estimates should encompass a minimum of 2 standard deviations and should use multiple skeletal and dental indicators of developmental age. Likewise, the Study Group on Forensic Age Diagnostics (Arbeitsgemeinschaft für Forensische Altersdiagnostik) recommends that sub-adult age estimates be based several indicators, including a physical examination, an X-ray of the left hand, a dental examination, and, when appropriate, radiographic or CT assessment of the medial clavicular epiphysis (Schulz et al. 2008). Certainly, using multiple age indicators to arrive at an age estimate is a good idea, especially since individual areas of a given skeleton may be advanced or delayed relative to other areas. However, using a range greater than two standard deviations may create unreasonably wide age ranges.

Another possible error source in age estimation arises from secular trends in maturation. As used by human biologists, the term “secular trends” refers to biological changes that occur over a long period of time (i.e. several decades or generations) as a result of environmental factors (Roche 1979). According to Malina (1979), secular changes are not necessarily caused by introducing growth-stimulation factors, but rather by eliminating growth-inhibiting factors (i.e. nutritional stress, environmental stresses, and disease). Secular changes

may be positive, as with increases or accelerations, or they may be negative (decreases or decelerations) (Roche 1979). The issue posed by secular trends cause with regard to age estimates is that most skeletal aging standards were derived from documented collections of individuals born during the 19th and early 20th centuries. Since this time, improvements in socioeconomic status, nutrition, and health care in these populations have contributed to increases in weight and stature, as well as accelerated maturation. Hence, age estimates based on reference populations that have undergone significant secular changes may be biased.

Perhaps the most extensively documented positive secular trend in maturation is decreasing menarcheal age, which has been studied for more than a century in populations worldwide. In Europe, decrease in age at menarche was documented during the early 19th century (Malina 1979). Malina (1979) reported a delay in maturation during industrialization due to disease, nutritional, and social stresses associated with over-crowded cities during the Industrial Revolution. Shortly thereafter, however, improvements in environmental and nutritional quality led to accelerated growth. Since the 1920's Europeans have experienced rather stable caloric intakes and reduced caloric expenditures and, consequently, increases in body weights. Other factors contributing to body weight increases, overall health improvements, and accelerated maturation are: increased calcium intake, introducing cereals at early age in infant diet, increased consumption of processed sugars and fats, improved socioeconomic status, improved health status, improvements in water and sanitation, elimination of infectious diseases, reduction in infant mortality, increased life expectancy, and reduction in family size (Malina 1979). The only reported negative secular trend in menarcheal

age was during World War II, but this was minor and temporary, and acceleration resumed once social and economic conditions were restored to pre-war levels.

Overall, age at menarche in western industrialized populations has decreased over the past 4-5 decades by around 4-6 months per decade (Fredriks et al. 2000; Malina 1979). In the Dutch population, for example, median age at menarche is 13.15 years (Fredriks et al. 2000). In Austria, mean menarcheal age is between 12.8 and 14.4 years (Kralj-Cercek 1956). Such decreases in age at menarche have been documented in Japan (Hoshi and Kouchi 1981), South Korea (Hwang et al. 2003), China (Huen et al. 1997; Leung et al. 1996; Lin et al. 1992; Low et al. 1982; Low et al. 1981; So and Yen 1992), India (Bagga and Kulkarni 2000), Venezuela (Farid-Coupal et al. 1981), Poland (Laska-Mierzejewska et al. 1982), Finland (Rimpela and Rimpela 1993), Belgium (Vercauteren and Susanne 1985), the Netherlands (Fredriks et al. 2000; Wellens et al. 1990), Britain (Cameron 1979), Spain (Prado 1984), Sweden (Liu et al. 2000), and France (La Rocherbrochard 2000). Although the precise age of pubertal onset is more difficult to detect in males, a marked decrease in the age of voice-breaking in males has been noted, as well (La Rocherbrochard 2000). Data from Johann Sebastian Bach's boys in choirs in Leipzig between 1732 and 1750 indicate that voices broke at age 17, whereas 20th century estimates report this change at age 13 (Daw 1970; Taranger et al. 1976). Furthermore, earlier pubertal onset in males has been reported in China (Leung et al. 1996; Lin et al. 1992; Wong et al. 1996), Peru (Gonzales et al. 1982), Sweden (Liu et al. 2000), and France (La Rocherbrochard 2000).

Recent studies in the United States suggest that obesity has a significant effect on menarcheal age (Wattigney et al. 1999). In obese African American and European American

females, age at menarche occurred between 10 and 13 years compared to 11-14 years for normal weight females (Wattigney et al. 1999). Moreover, the obese cohort girls were significantly more likely to reach menarche before age 12. Apparently, African American adolescents are especially obese among the American population, as evidenced from triceps skin fold thicknesses above the National Health and Nutrition Examination Survey's 95th percentile for the past two decades (Gordon-Larsen et al. 1997). This trend is contributing to earlier maturation in African American females compared to their European American counterparts, as shown by mean ages for onset of breast development (9.96 years for European American females and 8.87 years for African American females) (Herman-Giddens et al. 1997). Furthermore, breast development had commenced by age 8 in 48% of African American females versus 15% of European American females. These estimates are significantly earlier than a report from a decade earlier, in which breast development in European American girls began at an average age of 11 years (Bogin 1988). Similar obesity correlates have been reported in populations worldwide (Eveleth and Tanner 1990; Vignolo et al. 1988). Moreover, Komlos and Breitfelder (2007) propose that obesity is not only responsible for earlier maturation in the American population, but also for shorter adult heights than their European counterparts, specifically the Dutch. Apparently, the earlier American growth spurt comes at a price, namely shorter adult stature as a result of less growth during adolescence. Other possible contributing factors to the stature differences between Western European and American populations include a sub-optimal American diet and/or the superiority of the European health care system (Komlos and Lauderdale 2007).

Since sexual maturation is closely related to skeletal maturation, acceleration in pubertal onset may be indicative of acceleration in skeletal maturation (Maresh 1972). Indeed, a study comparing Poland's rate of skeletal maturation from hand and wrist data showed acceleration of 0.22-0.66 years/decade in ages of fusion (Himes 1984). Likewise, a significant secular trend was detected in Southern Chinese girls in Hong Kong using radiographic data from the hand and wrist (So and Yen 1990). The authors cited hygiene, nutrition, and socioeconomic improvements as the primary contributing factors. Additionally, Crowder and Austin (2005) report that contemporary North American adolescents have advanced union of the tibia and fibula compared to earlier studies. Investigations of more skeletal epiphyses may reveal similar trends. Presumably, dental development is under stricter genetic control than epiphyseal union, and secular trends may not be detectable in dentition, though this remains speculative (Bernhard 1995). What is certain is that many industrialized countries have realized increases in stature and earlier sexual maturation due to a positive secular growth change (Fredriks et al. 2000; Kim et al. 2008).

Adult Sex Estimation

Sex estimation can be approached metrically or non-metrically. Non-metric techniques include visual assessments of pelvic traits such as the greater sciatic notch, ischio-pubic ramus, sub-pubic concavity, and ventral arc (Arsuaga and Carretero 1994; Phenice 1969). Additionally, cranial features such as the supra-orbital ridges, nuchal area, mastoids, and chin are used frequently in sex estimation (Buikstra and Ubelaker 1994; Walker 2008). These methods typically rely on a size and/or shape continuum that ranges from "definitely male" to

“ambiguous” to “definitely female”. However, the most sexually dimorphic non-metric traits may vary from one population to another (Patriquin et al. 2003; Walker 2008). An investigation of pelvic dimorphism in South African whites and blacks revealed differences between these two ethnic groups (Patriquin et al. 2003). In whites pubic bone shape and sub-pubic concavity were the most accurate features in discriminating sex (88% average accuracy); whereas greater sciatic notch form was the best sex estimator (87.5% accuracy) in blacks, and pubic bone shape fell close behind (84.5%). Nonetheless, pubic bone shape was the most consistently reliable sex indicator across groups in both sexes.

Although cranial non-metric traits are not as accurate as pelvic traits for sex determination, they can provide successful estimates, particularly in the absence of postcranial elements. Several studies have shown the supraorbital region to be the most dimorphic region of the skull. For example, a crude examination of supra-orbital margin shape indicated that this feature alone is 70% accurate at sex estimation in adults (Graw et al. 1999). Williams and Rogers (2006) identified supraorbital ridge, zygomatic extension, nasal aperture, mastoid size, gonial angle, and overall size and architecture as “high-quality traits,” meaning that intra-observer error is less than 10% and accuracy is greater than 80%. They were able to determine sex with 96% accuracy and 92% precision using a combination of 20 skull traits. In an earlier investigation they achieved 88% accuracy from 17 cranial traits and ranked the traits in terms of classifying ability (Rogers 2005). Supraorbital ridge, nasal aperture, zygomatic extension, and malar size are of primary importance. Chin form and nuchal crest ranked as secondary contributors; mastoid size ranked of tertiary value; nasal size and mandibular symphysis/ramus

size ranked fourth; and, finally, forehead shape ranks fifth, and palate size and shape rank sixth (Rogers 2005). Generally cranial non-metric analyses yield accuracy rates between 80% and 90% (Konigsberg and Hens 1998; Walker 2008). Cranial and pelvic non-metric traits are successful sex estimators, but concern for considerable inter-observer error and subjectivity in non-metric sex assessment led to the development of metric techniques for sex estimation.

The landmark study in metric sex estimation by Giles and Elliott (1963) used 9 cranial measurements from European Americans and African Americans to develop race-independent discriminant functions for sex determination. These functions proved between 82% and 89% accurate at sex estimation, which is comparable to the 88%-90% obtained by Fordisc's discriminant functions on modern Americans (France 1998; Jantz and Ousley 2005; Komar and Buikstra 2008). Discriminant analyses of craniometric data from other populations have produced similar accuracy rates. Franklin and colleagues (2005a) reported accuracies between 77% and 80% for South Africans using 8 cranial measurements extracted from three-dimensional data. Uytterschaut (1986) achieved between 81% and 89% accuracy with race-independent functions on populations from Amsterdam, Zulu, and Japan. These discriminant functions were developed with 4 measurements (glabello-occipital length, bizygomatic breadth, nasal height and nasal breadth), suggesting that just a few highly dimorphic measurements are adequate to determine sex from the cranium. Steyn and Iscan's (1998) discriminant functions for South African whites used a combination of 12 cranial and 5 mandibular measurements. Bizygomatic breadth alone classified sex with 80% accuracy, and bigonial breadth was the most dimorphic mandibular dimension. Overall, accuracies ranged from 80% to 86%, and cranial

measurements outperformed mandibular measurements (Steyn and Iscan 1998). On the other hand, Calcagno's (1981) mandibular discriminant functions for European Americans, African Americans, and American Indians proved greater than 90% accurate at sex estimation. Nonetheless, their discriminatory performance decreased considerably with tests on different populations. Calcagno (1981) attributed this decrease in accuracy to differences in size between the target and test populations.

Spradley and colleagues (2008) reported similar size-related issues with Hispanic post-cranial measurements, wherein European American reference data for humeral head diameter, humeral epiphyseal breadth, and femoral head diameter failed to sex Hispanics satisfactorily. Typically, humeral biepicondylar diameter alone discriminates sex in European Americans with 90% accuracy, but in the smaller-bodied Hispanic population, the small males are sexed as females, and, consequently, accuracy rates drop to chance levels for males. In fact, as early as the 1950s Thieme and Schull (1957) noted that the applicability of discriminant functions depends mainly on 2 factors: the degree of skeletal sexual dimorphism in different populations (i.e. *between* populations), and the degree to which sexual dimorphism is the same *within* populations. They contend that inter-population size variation within a single sex is often more pronounced than intra-population sexual dimorphism (Thieme and Schull 1957). This last point is precisely what Spradley et al. (2008) reported with European American versus Hispanic males. In short, most researchers agree that population-specific discriminant functions should be developed for sex estimation among various populations (Calcagno 1981; Cowal and Pastor 2008; Frutos 2002; Murphy 2002; Spradley et al. 2008; Steyn and Iscan 1998).

Population-specific dimorphism aside, postcranial dimensions discriminate sex more accurately than cranial or mandibular measurements, as they typically give classification rates in excess of 90% from a minimal number of measurements (Spradley and Jantz 2003). Thieme and Schull (1957) tested a series of measurements from the femur, humerus, clavicle, and pelvis on Terry Collection African Americans. At 96.5% accuracy, ischium-pubis (I/P) index was the single best discriminator, followed by femoral head diameter (95%); a discriminant function using all 7 measurements proved to be 98% accurate. A test of these discriminant functions on the Terry Collection African Americans and European Americans plus African American specimens from the Howard University Medical School Collection gave slightly less accurate discrimination (88-91%). However, as with the earlier study, pubis length and femoral head diameter were the best variables for sex estimation (Richman et al. 1979). Similarly, Washburn (1948; 1949) reported above 90% accuracy with the I/P index on Bantu, Bushmen, and Hamann-Todd individuals. On the other hand, a more extensive investigation of the femur found that epicondylar breadth discriminates better than head diameter(Van Gerven 1972). Likewise, a test of the distal breadths of the femur and tibia in South African whites showed that femoral epicondylar breadth nearly as accurate as the combination of both measurements (90.5% versus 91.4%, respectively) (Steyn and Iscan 1997). Recently, a collaboration between anthropologists and biomedical engineers at the University of Tennessee has improved patellar sex estimation from 74-78% (Kemkes-Grottenthaler 2005) and 85% (Dayal and Bidmos 2005) to 90.3% using linear discriminant classification and to 93.5% with feed-forward back propagation neural networks (Mahfouz et al. 2007a).

The upper limb is equal to—if not greater than—the lower limb in its discriminating capabilities. For example, accuracy rates as high as 95.8% were obtained with humeral epicondylar breadth alone in South African white females; head diameter was the next best classifier (Steyn and Iscan 1999). A Guatemalan study reported accuracy rates of 95.5% from maximum humeral head diameter and 98.2% from a stepwise discriminant analysis with six variables (maximum length, maximum head diameter, midshaft circumference, maximum diameter at midshaft, minimum diameter at midshaft, and epicondylar breadth) (Frutos 2005). The high accuracy rates from humeral head diameter intimate a significant degree of sexual dimorphism in the shoulder region, which has direct implications for this dissertation.

Indeed, studies thus far suggest that the bones of the shoulder girdle discriminate sex with a high degree of accuracy (Frutos 2002; McCormick et al. 1991; Murphy 1994; Murphy 2002; Steel 1966). Parsons (1916) studied clavicular dimensions in a British population in the early 20th century. He developed sectioning points for length, circumference at midshaft, and for an index of the medial articular surface (height/width) using average measurements. He noted considerable overlap in clavicular length between males and females and reported a 22% error rate using length alone. However, midshaft circumference proved a more valuable indicator of sex, with an error rate of 16%, and the medial articular surface index had an associated error rate of 17%. Steel (1966) derived discriminant functions from Parson's data and attained an 87% accuracy rate. More recently, McCormick and co-workers (1991) examined the efficacy of clavicle length and circumference in a contemporary European American population from East Tennessee. They derived single cutoff values that gave a 93%

overall accuracy rate, although their study did not incorporate discriminant analysis or cross-validation. A discriminant analysis on a Guatemalan forensic sample using a combination of clavicle and scapula measurements (clavicle length, midshaft circumference, and height and width of the glenoid fossa) reported accuracies between 86% and 95% (Frutos 2002). These population-specific discriminant functions fared up to 60% better than classification rates using a North American forensic reference sample, probably because of the issue with smaller body size in Hispanic populations discussed earlier (Spradley et al. 2008). Similarly, Murphy (2002) used the height and breadth of the glenoid fossa and the diameters of the acromial and sternal ends of the clavicle to develop a function for New Zealand Polynesians. Interestingly, diameter of the sternal end is significantly more accurate than diameter of the acromial end (82% and 65%, respectively) (1994). Unfortunately, functions that incorporate multiple measurements from various skeletal elements can be problematic when dealing with fragmentary remains. Nonetheless, these studies underscore the dimorphic properties of the shoulder region and highlight the potential for highly accurate sex estimation with just a few measurements.

Recently, researchers began using methods such as Fourier analysis and geometric morphometric techniques to explore the utility of shape information for sex estimation. Simply defined, shape is "all the geometric information that remains when location, scale, and rotational effects are filtered out from an object"(Kendall 1977). This trend grew out of the realization that, although size accounts for a large portion of the skeletal variation between males and females, differing selection pressures have contributed to considerable differences in form, as well. Isolating scale may contribute to effective discrimination for a variety of skeletal

elements (Arsuaga and Carretero 1994; Kimmerle et al. 2008a; Rosas and Bastir 2002; Tanaka et al. 2000; Uytterschaut 1986; Van Gerven 1972). For example, Van Gerven (1972) found that 51% of the variation in the femur between males and females is related to size (as captured by 17 standard anthropomorphic measurements of length and robusticity), but including shape information improves sex discrimination. He attributed the shape differences primarily to pelvic differences, citing the functional relationship between the femur and pelvis as the major factor. Simply put, females have a broader pelvis and a more anteriorly projected acetabular angle, which in turn creates differences in femoral shape and orientation in order to facilitate efficient bipedal locomotion (Van Gerven 1972). Although his analysis lacked the technology offered by digitizers, computed tomography scans, and statistical programs specially designed for geometric morphometric analyses, Van Gerven (1972) recognized the interface between form and function and its implications for discriminating sex from skeletal elements. Indeed, in the past several years this insight has significantly impacted knee replacement surgery as biomedical engineers have developed sex-specific knee implants to address the different functional demands placed on male and female knees (Mahfouz et al. 2008; Merkl and Mahfouz 2005a).

Sex-related shape differences have been documented in other areas of the skeleton, as well. The exploration of shape differences in pelvic anatomy has focused largely on highly dimorphic non-metric traits. For example, an investigation using a combination of landmarks and semi-landmarks quantified statistically significant differences between males and females in the shape of the greater sciatic notch and ischio-pubic ramus, with sciatic notch being more

dimorphic than the ramus (Gonzalez et al. 2007). Furthermore, this method has low inter- and intra-observer error rates compared to morphoscopic scoring methods. Not unexpectedly, humeral shape is highly dimorphic, as well. Elliptical Fourier analysis of the proximal humeral outline indicated that males have a more anteriorly placed lesser tubercle and a more postero-medially oriented greater tubercle than females (Tanaka et al. 2000). Discriminant functions based on a combination of shape characteristics of the outline and the bounded area classified sex with 93-96% accuracy. Researchers have elucidated significant differences in vertebral body shape, too, with female vertebral bodies exhibiting more gracile characteristics as early as age 8 (Taylor and Twomey 1984). This difference is likely due to growth differences between male and female vertebrae, including an earlier female growth spurt in vertebral height and more rapid growth in transverse diameter in males.

Shape analyses of crania have shown promise over traditional metric sexing techniques. A study focusing on the supraorbital ridges, nuchal crest, and frontal sinuses reported 100% classification accuracy with a modern Taiwanese sample (Hsiao et al. 1996). In general, males were found to have greater glabella projection and larger frontal sinuses. These authors used non-traditional measurements taken from lateral radiographs and foreshadowed the findings of later geometric morphometric analyses, namely that the supraorbital region increases discrimination when measured precisely. Furthermore, this finding correlates well with the findings from traditional metric analyses regarding the supraorbital region.

In effort to determine precise the effects of size and sex on craniofacial shape in European Americans and African Americans, Kimmerle et al. (2008a) conducted an analysis of

shape variables and centroid size. They found that size does not have an effect on shape in the American population. In other words, different sized individuals of the same sex are similar in shape. However, sex does have a significant influence on shape. They were able to discriminate sex with approximately 77% accuracy in African Americans and European Americans with shape variables alone. Combining shape variables and centroid size increased accuracy rates to 90% for African Americans and 87% for European Americans. Kimmerle and colleagues (2008a) conclude that explicit shape differences contribute to cranial sexual dimorphism, and these differences may have distinct patterns in various ancestral groups. Three-dimensional geometric morphometric methods are particularly useful in elucidating these shape differences.

Likewise, Rosas and Bastir (2002) assert that removing size in studies of cranial sex dimorphism offers a way to tease out its influence in sex discrimination. Thin-plate spline analysis of the skull revealed sexual dimorphism in the angulation of the posterior cranial base, the occipital region, and the nasopharyngeal region. Additionally, mandibular ramus flexion, anterior symphysis curvature, and pre-angular notch development were significantly dimorphic. However, the supraorbital ridges were more dimorphic than any other cranial feature. Allometry accounted for 54% of the total variance; 37% of the variation was attributed to the "sex specific factor" (i.e. differences between male and female cranial anatomy not related directly to size) (Rosas and Bastir 2002). In an similar analysis of cranial shape dimorphism in 5 modern South African populations, aspects of the frontal bone and zygomatic arches effectively discriminated sex among all populations (Franklin et al. 2004). Isolating scale has shown

potential in discriminating sex from the mandible, as well. An elliptic Fourier analysis of the mandible in lateral view revealed that 97% of males and 92% of females showed significant sexual dimorphism (Schmittbuhl et al. 2002; Schmittbuhl et al. 2001). Consequently, these authors maintain that shape contributes more than previously recognized to mandibular sexual dimorphism. Nonetheless, shape-related patterns of sexual dimorphism may vary among populations and, consequently, analyses should take population variation into account (Franklin et al. 2004; Gonzalez et al. 2007).

What emerges from these studies is the importance of shape in sexual dimorphism and the capacity to improve discriminatory performance with non-traditional features and innovative discovery methods. Our understanding of shape variation in the human skeleton has grown over the last several years, and novel data collection methods and statistical programs have been developed to explore this aspect of skeletal variation. Thus far, shape has proven a promising sex discriminator, particularly when combined with information about size. Recently, collaboration between anthropologists and biomedical engineers has proven invaluable in understanding skeletal size and shape dimorphism for the fields of biological anthropology and medical technology. By utilizing a newly developed statistical treatment that combines Principal Components Analysis (PCA) and the Fisher Discriminant Ratio (FDR), Mahfouz et al. (2007b) have been able to detect the most sexually dimorphic locations on the distal femur and explore how shape differences between the sexes change relative to size. Furthermore, an algorithm designed to detect bony landmarks automatically and propose measurements that most effectively capture the highest degree of sexual dimorphism promises

to improve upon our ability to discriminate sex from skeletal remains with 3-D coordinate and/or linear measurement data (Mahfouz et al. 2007b). Moreover, tests of the comparability of 3-D coordinate data with traditional linear measurement data have revealed that 3-D landmark coordinates can be transformed mathematically into 2-D linear measurements with high reliability (Decker and Hilbelink 2008; Franklin et al. 2005b). Hence, using a combination of size and shape information is accomplished easily and imparts advantages over using either alone. This analysis of sexual dimorphism will explore the metric and geometric morphometric properties of the human clavicle using a combination of traditional and novel statistical methods (i.e. linear discriminant analysis and statistical atlases, respectively).

Chapter 2: Materials and Methods

Medial Clavicular Epiphyseal Fusion

Two skeletal collections were scored for epiphyseal fusion (the William F. McCormick Clavicle Collection and the Hamann-Todd Osteological Collection). Additionally, McKern and Stewart's (1957) raw data from the Korean War males is included as part of this analysis. The William F. McCormick clavicles are housed at the University of Tennessee in the Anthropology Department. This documented autopsy collection consists of approximately 2,000 clavicle pairs from 1986-1998 autopsies in East Tennessee. The sample is 95% European American, 4% African American, and 1% Latino, Asian, and Native American, which roughly reflects the demographic composition of that region (Census 2000). A subset of 594 McCormick individuals aged 11-33 years was scored for medial epiphyseal fusion (448 males, 146 females). This broad age range was chosen to ensure that ample numbers of individuals were included from all stages of fusion (i.e. unfused, fusing, and fused).

The Hamann-Todd Human Osteological Collection is housed at the Cleveland Museum of Natural History. The collection consists of skeletal remains from approximately 3,000 cadavers autopsied between 1912 and 1938. The demographic composition of the Hamann-Todd Collection is 62% European American and 38% African American. A subset of 354 individuals aged 11-30 years were scored from this collection (255 males, 99 females). This sample was also used to test for ethnic differences in epiphyseal union. Additionally, 341 males aged 16-33 years from 1950-1952 Korean War fatalities are included in this analysis. These clavicles were

scored as part of McKern and Stewart's report *Skeletal Age Changes in Young American Males* (1957), and the raw data is available online (http://konig.la.utk.edu/paleod.htm). In sum, the total study sample is comprised of clavicles from 1,289 individuals from cohorts spanning the 20th century. The composition of this sample offers a way to control for population differences and examine secular trends within the American population.

As discussed in the literature review, several scoring systems are available for use. Webb and Suchey's (1985) system, however, was unsuitable for these collections because it was impossible to score phases 1 or 2 of this system (i.e. nonunion with no epiphysis or nonunion with a separate epiphysis). Since the clavicles were processed, dried, and stored in bags, I could not ascertain with certainty whether an unfused bony epiphysis was present at any point during this curation process. Black and Scheuer's (1996) system was suitable for use but differentiating between the first 2 phases proved challenging and subjective (Phase 1: no fusion with distinctive ridges and furrows on the metaphyseal surface; Phase 2: no fusion with less ridges and furrows on the metaphyseal surface); consequently, this system was omitted from the study. In the end, the clavicles were scored with McKern and Stewart's (1957) 5-phase system and with the 3-phase system used by most forensic anthropologists, as well as for coding entries in the Forensic Anthropology Databank. Apparently, the 3-phase system was suggested initially by Larry Angel in a casual conversation and later adopted by the authors of *Data Collection Procedures for Forensic Skeletal Material* (Moore-Jansen et al. 1993) as a standard for scoring epiphyseal union in the human skeleton.

The 5 phases of McKern and Stewart's system are (1) no union, (2) beginning union, (3) active union, (4) recent union with a scar, and (5) complete union with no scar. Each of these phases is shown in Figures 1-6. "No union" was scored if there was no remnant of a flake fused to the shaft. Typically, the medial articular surface had a coral-like appearance (Figure 1). A bone was scored as "beginning union" if the epiphyseal flake had commenced fusion to the medial articular surface, but less than 50% of the surface was covered by the flake (Figure 2). "Active union" was scored if 50% or more of the surface was covered by epiphyseal flake and fusion was actively occurring (Figure 3). The epiphysis was considered actively fusing if the flake clearly appeared as a separate entity with some space between the edges of the flake and the bone surface (Figure 4). "Recent union" was scored when the flake was completely fused to the bone but a trace of the fusion event remained in the form of a fusion scar and/or as small bony nodules on the outer rim of the medial surface (Figure 5). A fusion scar should not be confused with the line left from the joint cartilage; this line is usually on the edge of the bone and not on the medial surface where fusion occurs. Finally, "complete fusion" was scored when no trace of the fusion event remained and the articular surface was quiescent (Figure 6). The 3-phase system scores clavicles as unfused, fusing, or fused. In this system, phases 2-4 in the McKern and Stewart system (beginning, active, and recent union) are lumped into the "fusing" stage.

In order to determine the best scoring system, a random sample of 50 clavicles was selected for intra- and inter-observer error tests. For the intra-observer error test, the 50 clavicles were scored by the principal researcher one year after the initial scoring. To calculate

Figure 1. No fusion.

Note the coral-like appearance of the medial articular surface. From left: 12 year old European American male, 10 year old European American male, 12 year old European American female.

Figure 2. Beginning fusion.

Flake commencing fusion; less than 50% of surface covered by flake. From left: 16 year old European American male, 17 year old European American male, 20 year old European American female.

Figure 3. Active fusion.

Greater than 50% of the surface is covered by flake; fusion actively occurring. From left: 22 year old European American male, 25 year old European American male, 20 year old European American male.

Figure 4. Active fusion (lateral view).

Note the visible gap between the epiphyseal flake and the bone surface.

Figure 5. Recent fusion.

Arrows from left to right: fusion scar, bony nodule, bony nodule. From left: 23 year old European American male, 23 year old European American male, 25 year old European American male.

Figure 6. Complete fusion.

From left: 36 year old European American male, 49 year old European American male, 37 year old European American female.

three independent observers with different experience levels scored the clavicles with each scoring system. Observers waited one week between scoring the clavicles with the 5-phase system and the 3-phase system. Spearman's rank correlation coefficients (r_s) were calculated to evaluate the consistency between observers' scores. Wilcoxin signed ranks tests were used to test for significant scoring differences using a 95% confidence interval. Additionally, a Spearman's rank correlation coefficient (r_s) was calculated to test the relationship between age and phase in each scoring system. Spearman's rank correlations were performed in NCSS (Number Cruncher Statistical System), and the Wilcoxin signed ranks tests were done with SAS version 9.1.

Transition analysis was conducted on the data to determine the average age at which the transition from one phase to the next occurs. The transition analysis was performed with Konigsberg's Nphases2 computer program, which is available at http://konig.la.utk.edu/nphases2.htm. The Fortran-based program performs a logistic regression wherein the intercept and slope are converted to the mean and standard deviation, respectively (Boldsen et al. 2002). Nphases gives an age at transition, which is the maximum likelihood estimate of the likelihood function provided by the transition analysis. This estimate represents the average age at which an individual is most likely to transition from one phase to the next. Transition analysis has been used in the anthropological literature to study senescent changes in bone such as those at the auricular surface and pubic symphysis of the pelvis and the sternal rib ends (Alsup 2007; Boldsen et al. 2002; Kimmerle et al. 2008b; Konigsberg et al. 2008). However, the phase systems used to score age-related changes in the medial clavicle

meet the assumptions of transition analysis, and thus the method is applicable in this context. Namely, the fusion event can be represented in a series of invariant phases, and the morphological change is unidirectional with respect to those phases (Boldsen et al. 2002). One attractive feature of transition analysis is that it gives a robust estimate of age-at-fusion that is less sensitive to outliers than the percentile method. In order to examine secular trends in skeletal maturation, separate transition analyses were done on each skeletal sample (Hamann-Todd, McKern and Stewart, and McCormick). Furthermore, males and females were analyzed separately to investigate sexual dimorphism in epiphyseal union.

A cumulative probit model on log age, or proportional odds model with a probit link, was used for the transition analysis. The cumulative probit model assigns the same standard deviation to each transition, and the natural log scale assures that the transition distribution is log normal (Konigsberg et al. 2008). A Bayesian approach was used to obtain useable age ranges from the transition likelihood estimates. (Boldsen et al. 2002). Bayes' Theorem can be stated as:

$$\Pr\,(a|c_j) = \frac{\Pr(c_j|a)f(a)}{\int_0^\infty \Pr(c_j|x)f(x)\mathrm{dx}} \qquad (1)$$

where $\Pr(a|c_j)$ is the probability that a skeleton died at age a given it has characteristics c_j (in other words, the probability of age given phase); $\Pr(c_j|a)$ is the probability that a skeleton with characteristics c_j died at age a (in other words, the probability of phase given age); and $f(a)$ is a probability density function for age (Boldsen et al. 2002; Kimmerle et al. 2008b). $\Pr(c_j|a)$ is obtained from the transition analysis performed in Nphases and is mathematically represented as:

$$\Pr(c_j|a) = \int_{-\infty}^{a} f\left(a\middle|\mu_{j-1}, \sigma\right)da - \int_{-\infty}^{a} f(a|\mu_j, \sigma)da \qquad (2)$$

where $f(a|\mu,\sigma)$ is the normal probability function of age a with mean μ and standard deviation σ; this function pertains to the transition distributions, and is independent of the *f(a)* from the informative prior in equation 1. The probability density function *f(a)* (equation 1) is obtained by fitting a Gompertz hazard model to an informative prior distribution to obtain hazard parameters. The Gompertz hazard model is:

$$h(t) = \alpha_3(\exp\beta_3 t) \qquad (3)$$

where h is the hazard rate, t is age shifted by 11 years, α_2 is the baseline mortality, and β_3 is the senescent component. For this analysis, the Forensic Data Bank (FDB) was selected for the prior age-at-death distribution because of its similarity to the target population. By using an informative prior, we can select appropriate values for *f(a)* that are independent of the skeletal samples (Boldsen et al. 2002). Konigsberg et al. (Konigsberg et al. 2008) recommend that the prior is a "reasonable [guess] at what the possible age should be for an individual case prior to an osteological analysis" (544). The FDB provides such an approximation for forensic anthropology cases. Figure 7 contains a Kaplan-Meier survivorship curve of the FDB and the Gompertz curve using parameters from the FDB. Note that the Gompertz distribution is an appropriate hazard model for FDB mortality. Figure 8 shows the differences between the male and female distributions in the FDB. Consequently, separate hazards were run on males and females. The statistical package "R" was used to model the Gompertz distribution and to do

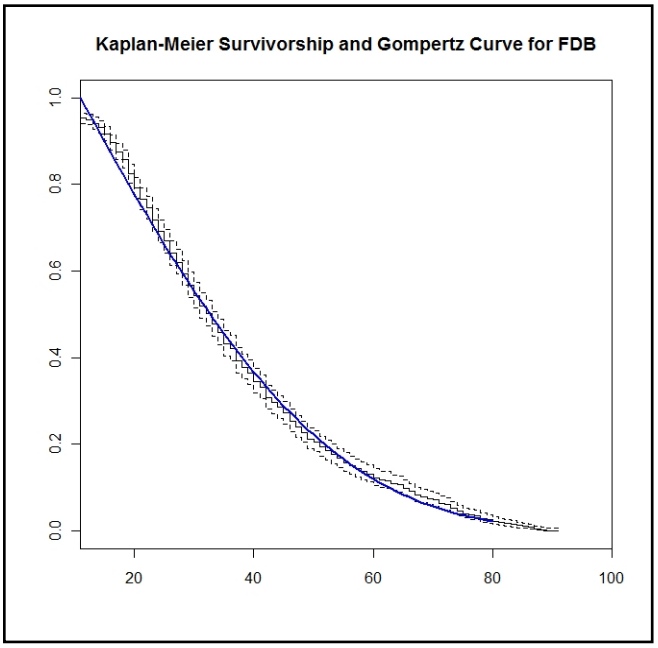

Figure 7. FDB survivorship.

Kaplan-Meier Survivorship with a 95% confidence interval (stepped lines) and Gompertz curve (solid blue line).

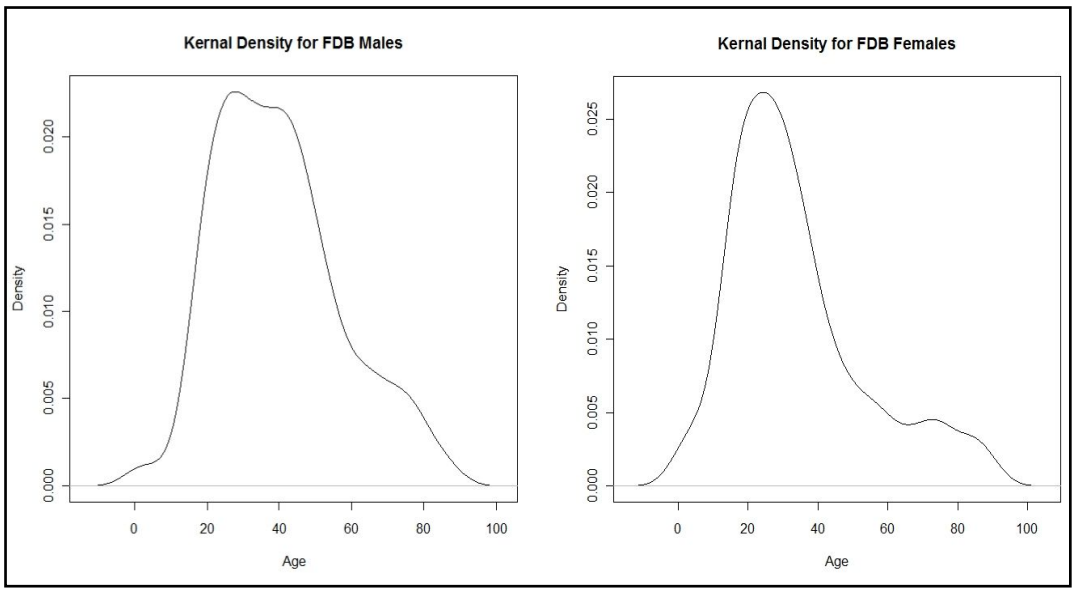

Figure 8. FDB age-at-death distributions.

the Bayesian analysis (http://www.r_project.org). "R" scripts for Gompertz hazards and for estimating the highest posterior density regions were adapted from scripts available at Dr. Lyle Kongsberg's webpage (http://konig.la.utk.edu).

The transition analysis likelihood estimates were combined with the hazard parameters from the Gompertz-modeled FDB to obtain the posterior density regions for each phase, which are equivalent to the most likely ages at death in each phase. Note that these age ranges are not confidence intervals around mean ages, but instead probability estimates of the most likely age at death (Kimmerle et al. 2008b). These age ranges were tested on 12 individuals from the University of Tennessee's William M. Bass Donated Collection. Finally, a Student's t-statistic was used to evaluate differences between the sexes and between birth cohorts. The t-statistic was calculated using the maximum likelihood estimates of the transitions and the corresponding standard errors from the log-normal transition ages:

$$t = \frac{\bar{x}_2 - \bar{x}_1}{\sqrt{se_2^2 + se_1^2}} \qquad (4)$$

Size and Shape Properties of the Clavicle

Surface Model and Atlas Creation

Size and shape data was collected from computed tomography (CT) scans of the McCormick Clavicles. A total of 2145 clavicles were scanned, but only individuals ≥28 years of age were used in this study, as growth is completed by this age. In addition, fractured or otherwise pathological clavicles were omitted. Transforming the CT scanned images into a usable format for statistical analyses involves several processes: (1) manually segmenting the

DICOM images, (2) generating triangular mesh surface models from the segmented images, (3) making 3-D models from the surface generations, and (4) creating a clavicle atlas. The clavicles were packed in layers into foam board boxes and scanned with a GE Lightspeed 16 slice computed tomography scanner using 0.625 mm X 0.625 mm X 0.625 mm cubic voxels (Figures 9 and 10). Because radiation levels are not a concern when scanning skeletal material (as opposed to living tissue), the slice thickness can be adjusted to obtain detailed representations of the bones. The result is high resolution, three dimensional radiographs in the form of DICOM image slices (Figure 11).

This stacked image data was uploaded to the commercially available software Amira (Mercury Computer Systems, San Diego) for manual image segmentation and surface model generation. In manual segmentation, each DICOM slice is examined individually, and bone is selected using a variety of tools similar to those available in Adobe Photoshop. With a sample of this size, manual image segmentation is more time consuming than automated image segmentation, but it has been found to be a reliable process with negligible inter- and intra-observer error (Mahfouz et al. 2007b). In order to delineate the regions of interest, maximum and minimum threshold voxel grayscale values were specified, and any voxels that fell between these values were selected. The voxel values correspond to the various densities of different materials. Figure 11 shows the process of manual image segmentation and differing voxel values. Note that cortical bone is the densest material with the highest voxel value, and it shows up as European American; trabecular bone is gray; the least dense materials are black

Figure 9. Packed clavicles.
Clavicles packed into foam board box for scanning.

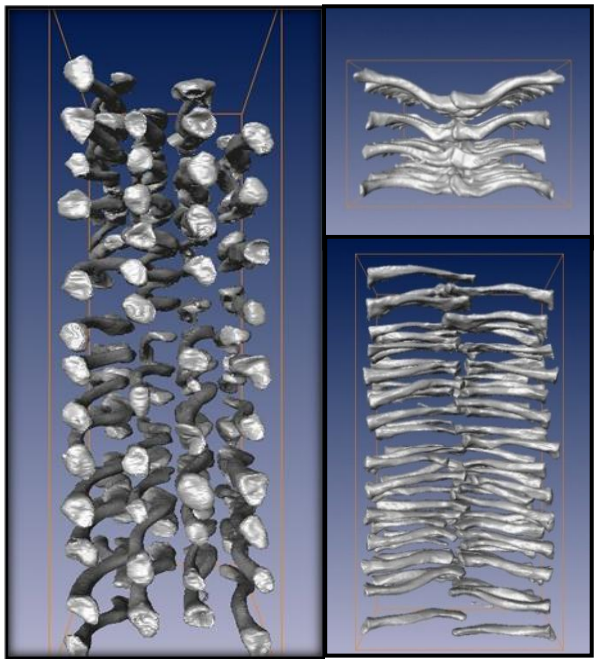

Figure 10. Scanned bones.
Three-dimensional surface generations showing *in virtro* scanned clavicles.
Left: y-z axis view. Top right: x-y axis view. Bottom right: x-z axis.

Figure 11. DICOM image slice.
This DICOM slice viewed in Amira shows varying voxel values. Voxel values representing bone are selected for manual image segmentation.

(i.e. foam and air). After all of the voxels representing clavicles were selected, each clavicle was assigned a separate material and label to differentiate it by side and case number. Once the labeled materials were created, a smoothing algorithm was applied to assist with making the surface generations and models (Figure 12). Next, triangulated surface mesh generations were created from the segmented images, and these surface generations were imported into Matlab in ASCII format in order to generate 3-D models of each bone (Figure 12). The bone models of the adult clavicles contain between 30,000 and 50,000 triangular faces, depending on the size of the bone. Both adult and sub-adult clavicles were modeled, but the bone surface area of the child clavicles was too small to obtain a sufficient number of triangular surfaces for addition to the clavicle atlas. Consequently, sub-adult clavicles were omitted from this analysis.

The surface models were used to create the sex-specific statistical clavicle atlases. A bone atlas is an average mold, or template mesh, that captures the primary shape variation of a

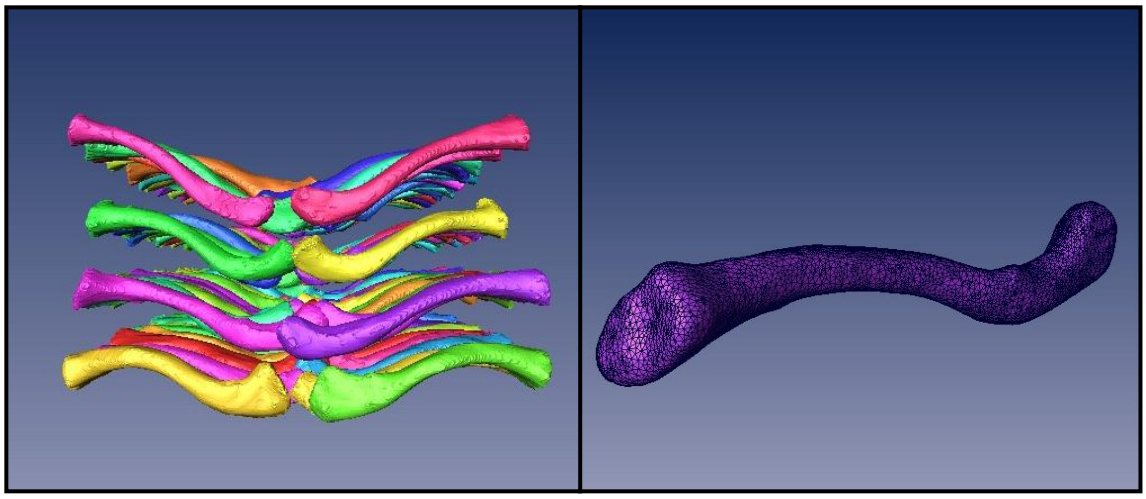

Figure 12. 3-D surface generations.
Left: Surface models in situ in the foam box. Right: surface model with super-imposed triangular faces.

bone and allows for the comparison of global shape differences between males and females (Mahfouz et al. 2007b). Bone atlases were developed initially by biomedical engineers as a way to digitally recreate a bone and conduct statistical shape analyses (Mahfouz et al. 2008; Mahfouz et al. 2006; Merkl and Mahfouz 2005b), but they have proven useful in biological anthropology as a means of studying sexual dimorphism (Mahfouz et al. 2007a; Mahfouz et al. 2007b) and for reconstructing hominid fossils and making shape comparisons among fossil species (Merkl et al. 2007; Sylvester et al. 2007) .

The following descriptions detailing atlas construction were adapted directly from Mahfouz and colleagues (2007b). To begin the atlas construction, several clavicles with average shape characteristics were selected to serve as guidelines for constructing the template mesh. Once the bone atlases was created, individual bone models were added to the atlas using an iterative closest point (ICP) algorithm and Mutual Correspondence Warping (MCW) to match

the points of the training models to the corresponding points in the template mesh so that all of the bones have the same number of vertices and the same triangular connectivity. Prior to applying the ICP algorithm, the centroids of the template mesh from the atlas and the new mesh from the models were aligned. Next, vertex-to-vertex ICP was used to attain rigid alignment of the template mesh to the new mesh, and MCW was used to prevent misalignments. The MCW uses a combination of the closest vertex-to-vertex correspondences from the template mesh to the new mesh and those from the new model to the template model to move the points from the template mesh towards the new points on the model. An iterative smoothing algorithm is then applied to the deformed template mesh in order to eliminate any discontinuities that may have resulted from the ICP and MCW. The smoothing algorithm uses the properties of the surrounding triangles to determine the smoothing vector. The end result is a triangular mesh in which adjacent triangles have similar geometric properties, although they vary dimensionally. After the atlas and model are rigidly aligned, an affine transformation without iteration was used to align the template and new mesh; this step comprises the rotation, translation, scaling, and shear of traditional geometric morphometric methods in order to remove the effects of size. Finally, surface-to-surface matching was applied to create new points on the surface of the new model; these points have local spatial characteristics similar to the template model. The final result is a sample of clavicles that all contain the same number of points and share the same spatial relationship, thus satisfying the requirements for the Principal Components Analysis.

Principal Components Analysis and Fisher's Discriminant Ratio

After the male and female atlases were created, a statistical treatment combining Principal Components Analysis (PCA) and Fisher's Discriminant Ratio (FDR) was used to determine the most sexually dimorphic areas of the clavicle. This method was first employed by Mahfouz and co-workers (2007b) in order to investigate sex differences in the distal femur for medical implant and forensic purposes. The PCA-FDR method combines the data-reduction capabilities of PCA with the discriminatory power of FDR to locate the points of greatest difference between male and female clavicles (Mahfouz et al. 2007b).

First, PCA was performed on the atlases as an initial step in preparing the data for further analysis[1]. PCA averages the corresponding points across all of the clavicle models and computes the mean clavicle shape. The resulting principal components consist of orthogonal eigenvectors that define a new set of coordinates with reduced dimensionality when the original features of the models are projected onto the eigenvectors scaled by the inverse of the singular values. These new coordinates can be used to compare the differences between males and females by using Fisher's Discriminant Ratio to weight the PCA coordinates, and then summing the weighted vectors such that consistent differences between the two classes are retained. Fisher's Discriminant Ratio is the ratio of the between-class to within-class covariance matrices, and it is used frequently in pattern classification. The resulting vector magnitudes

[1] Descriptions of statistical treatments adapted directly from Mahfouz M, Merkl B, Fatah E, Booth RJ, and Argenson J. 2007b. Automatic methods for characterization of sexual dimorphism of adult femora: distal femur. Computer Methods in Biomechanics and Biomedical Engineering 10:477-456.

were reinterpreted as a 3D deviation vectors for each of the points on the models in the atlas. The deviation vector magnitudes were applied to a color map in order to visualize the areas of highest dimorphism. These maps were created using various combinations of the first 10 principal components. The first principal component was used to visualize differences primarily due to scale, and the second through tenth principal components were used to visualize shape differences.

Linear Discriminant Analysis

Several automated measurements were taken directly from the bone models, including three traditional and six non-traditional measurements. Previous work with femoral bone models and atlases have shown the process of obtaining measurements from manually segmented 3-D images to be reliable and accurate (Mahfouz et al. 2007b). Furthermore, computer automated measurements minimize observer and instrumental error. In order to ensure that growth was completed, the sample was truncated to include only those individuals 28 years of age and older (n=1,414). The traditional measurements were maximum length, sagittal diameter at midshaft, and vertical diameter at midshaft (Moore-Jansen et al. 1993) (Figure 13). New measurements included maximum midshaft diameter, minimum midshaft diameter, maximum and minimum diameters at the medial end of the shaft, and maximum and minimum diameters at the lateral end of the shaft (Figure 14).

Maximum length was taken by locating the most medially and laterally projecting points on the clavicles and taking the Euclidean distance between these points. Midshaft diameters

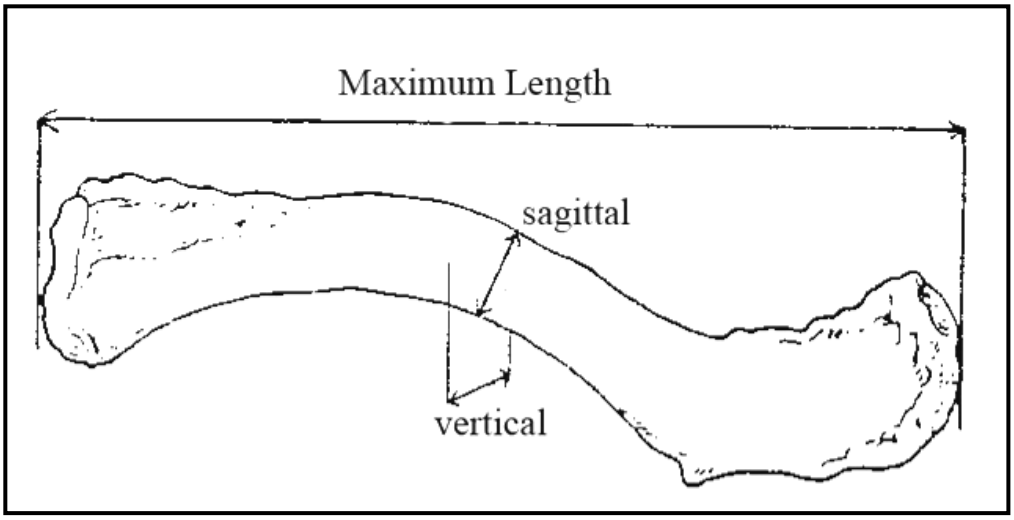

Figure 13. Traditional clavicle measurements.
Adapted from *Data Collection Procedures for Forensic Skeletal Material* (Moore-Jansen and Jantz 1990).

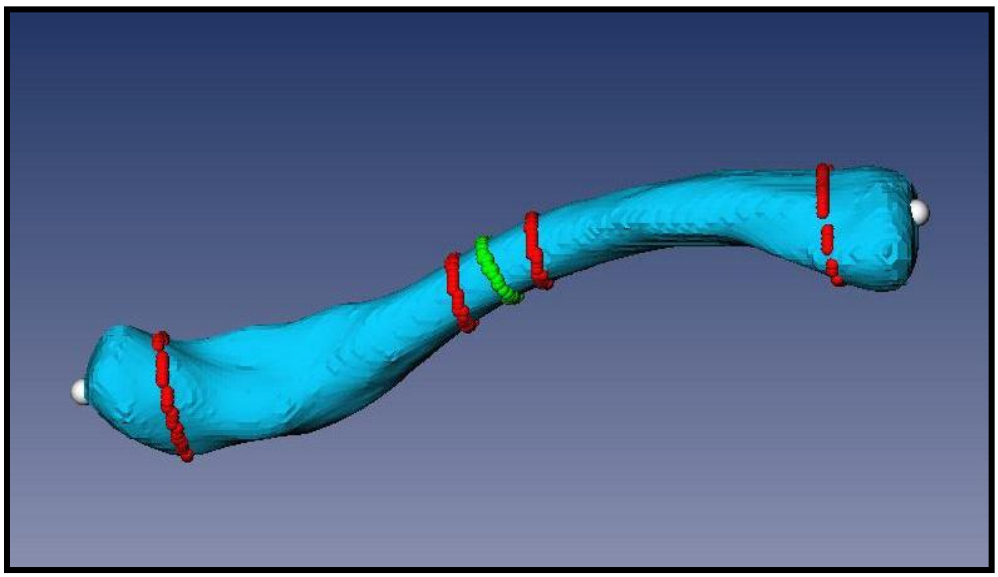

Figure 14. Measurements from bone models.
White dots represent the most medial and lateral projecting points. The red ellipses closest the ends were used to measure the medial and lateral end diameters. The red ellipses near the midshaft were used for geometric projection of the green ellipse for measuring sagittal and vertical diameters.

were taken by first finding the midpoints of the medial and lateral endpoints. Next, 2 additional points were created by moving 5% of the maximum length in the medial direction and 5% in the lateral direction; ellipses were constructed around these points, and the centers of these ellipses were connected with a line. Next, a perpendicular plane was placed at the midpoint of this line. The intersection of the plane with the bone shaft represents another ellipse; the maximum and minimum midshaft diameters were derived from the major and minor axes of this ellipse. Sagittal and vertical diameters at midshaft were measured by using the lateral end ellipse as a reference to approximate the anterior-posterior (AP) and superior-inferior (SI) anatomical orientations. The major axis of the ellipse is oriented in the AP direction, a unit vector in the SI direction is orthogonal to the AP direction, and both are orthogonal to a unit vector extending from the medial end to the lateral end (the ML direction). Based on these three unit vectors, or directions, geometrical projection was used to translate points from the x, y, and z coordinates to AP, SI, and ML coordinates. Because an ellipse drawn around the midshaft is located in a curved portion of the bone, it is not likely parallel to the AP-SI plane and, therefore, must be projected onto a two dimensional AP-SI plane. Once this is accomplished, the maximum diameters in the AP and SI directions can be measured (i.e. sagittal and vertical diameters, respectively). Diameters at the medial and lateral ends were taken by constructing ellipses at points representing 90% of the clavicle length in either direction, then measuring the major and minor axes of these ellipses.

Linear discriminant analysis with cross-validation was performed on the clavicle measurements in SAS Version 9.1. McHenry's algorithm for multivariate variable selection was

used to aid in determining the best variable models for the discriminant analysis. Whereas stepwise variable selection uses the significance values of the variables to derive the best model, McHenry's algorithm uses the R-squared values, which tends perform better in pinpointing the predictive power of the variables (i.e. variable significance does not equate to superior discriminatory capability). Separate discriminant analyses were run for African Americans (*n*=53) and European Americans (*n*=1354) in order to provide population-specific discriminant functions; all other ancestries were omitted from the discriminant analysis due to inadequate sample size (n=7). The right clavicle was used in these analyses because of the large number of right clavicles in the collection; although sample size was not a concern with the European American individuals, all efforts were made to obtain a statistically significant African American sample.

Chapter 3: Results

Medial Clavicular Epiphyseal Fusion

The Spearman's rank correlation coefficient between phase and age for the McCormick, Hamann-Todd, and McKern and Stewart samples are 0.88, 0.80, and 0.81, respectively, for the 5-phase scoring system and 0.83, 0.75, and 0.75 for the 3-phase scoring system. All of these coefficients indicate a strong positive relationship between age and phase in both scoring systems. The observer error tests of both systems show that the 3-phase system is less subjective than the 5-phase system (Table 1). The 3-phase scoring system resulted in fewer absolute differences between observers than the 5-phase system (44 differences versus 85 differences, respectively). Nonetheless, only 4% of the total differences involved scoring discrepancies greater than one phase. Whereas the Spearman's rank correlations were slightly higher with the 5-phase system, the p-values from the Wilcoxin signed ranks test were more favorable with the 3-phase system. Additionally, conversations with observers who scored the clavicles revealed that all observers found the 3-phase system easier to use. However, the 5-phase system can provide a means of fine-tuning age estimates, particularly estimates based on multiple age indicators. For example, beginning fusion is likely to represent an individual at the younger end of the age range. In fact, all individuals in the beginning fusion stage were under 25. Likewise, recent fusion is likely to represent an individual on the older end of the age range; no individuals in the recent fusion stage were under 20 years of age.

Table 1. Observer error. Observer 1 was used to test intra-observer error.

Observer	1		2		3	
5 Phase System	**# diff* (r_s)†**	**S‡ (p)§**	**# diff* (r_s)†**	**S‡ (p)§**	**# diff* (r_s)†**	**S‡ (p)§**
1	1 (.990)	-0.5 (1.00)				
2	11 (.929)	9 (0.55)				
3	5 (.962)	5 (0.31)	13 (.930)	-3.5 (0.69)		
4	16 (.889)	53 (0.003)	20 (.884)	-55 (0.03)	19 (.882)	-47.5 (0.05)
3 Phase System						
1	1 (.963)	-0.5 (1.00)				
2	2 (.921)	0 (1.00)				
3	4 (.861)	2.5 (0.63)	6 (.790)	-3.5 (0.69)		
4	11 (.683)	9 (0.55)	11 (.683)	-9 (0.54)	9 (.746)	1.5 (1.00)

*Actual number of scoring differences between observers, †Spearman's rank correlation coefficient (r_s), ‡Wilcoxin Signed Rank statistic (S), §p-value associated with the Wilcoxin signed rank statistic

A Student's t-test with a 95% confidence interval was used on the Hamann-Todd sample to test for significant ethnic differences in the transitions. None of the transitions were significantly different for African American males versus European American males in either the 5-phase or the 3-phase system. The small female sample size for some of the transitions precluded testing each transition in the 5-phase system for females. Nonetheless, there were no significant differences between African American and European American females in the 3-phase system. Consequently, ethnicities were pooled in this analysis.

Tables 2-4 present the descriptive statistics from each sample with both scoring systems. The mean, standard deviation, and a 95% confidence interval of the mean are not given for the first and last phases, as these numbers are artifacts of the upper and lower truncation ages for the sample (i.e. 11 and 30 years, respectively) and offer no interpretive value. Age ranges derived from the raw data are included in the last column; these broad age

Table 2. McCormick sample descriptive statistics. Mean ages, 95% upper and lower confidence levels, and observed age ranges from raw data.

Phase	*n*	Mean Age	Std. Dev	95% CI	Observed
5 Phase System					
Males					
1	34	-	-	-	≤18
2	72	19.3	2.31	18.7-19.8	13-24
3	83	22.4	2.26	22.0-22.9	17-29
4	52	25.9	2.66	25.1-26.6	22-32
5	207	-	-	-	≥19
Females					
1	11	-	-	-	≤19
2	22	17.8	1.89	17.0-18.7	15-23
3	26	21.0	2.54	19.9-22.0	17-26
4	16	25.4	2.94	23.8-26.9	20-31
5	70	-	2.47	28.8-29.9	≥24
Total					
1	45	-	-	-	≤19
2	94	18.9	2.29	18.5-19.4	13-24
3	109	22.1	2.40	21.6-22.5	17-29
4	68	25.8	2.72	25.1-26.4	20-32
5	277	-	-	-	≥19
3 Phase System					
Males					
1	34	-	-	-	≤18
2	207	22.2	3.47	21.7-22.7	13-32
3	207	-	-	-	≥19
Females					
1	11	-	-	-	≤19
2	64	21.0	3.77	20.0-21.9	15-31
3	70	-	-	-	≥24
Total					
1	45	-	-	-	≤19
2	271	21.9	3.57	21.5-22.3	13-32
3	277	-	-	-	≥19

Table 3. Todd sample descriptive statistics. Mean ages, 95% upper and lower confidence levels, and observed age ranges from raw data.

Phase	*n*	Mean Age	Standard Deviation	95% CI Mean Age	Observed Age Range
5 Phase System					
Males					
1	48	-	-	-	≤25
2	24	22.7	2.38	21.7-23.7	17-28
3	63	23.4	2.39	22.8-24.0	18-29
4	45	27.1	2.59	26.4-27.9	22-30
5	75	-	-	-	≥23
Females					
1	21	-	-	-	≤25
2	3	20.7	3.06	13.1-28.3	18-24
3	18	22.7	2.59	21.4-24.0	18-26
4	14	25.1	2.77	23.5-26.7	21-29
5	43	-	-	-	≥22
Total					
1	69	-	-	-	≤25
2	27	26.7	2.66	21.6-23.7	17-28
3	81	23.3	2.44	22.7-23.8	18-29
4	59	26.7	2.75	25.9-27.4	21-30
5	118	-	-	-	≥22
3 Phase System					
Males					
1	48	-	-	-	≤25
2	132	24.6	3.08	24.1-25.1	17-30
3	75	-	-	-	≥23
Females					
1	21	-	-	-	≤25
2	35	23.5	2.99	22.5-24.5	18-29
3	43	-	-	-	≥22
Total					
1	69	-	-	-	≤25
2	354	24.4	3.09	23.9-24.8	17-30
3	118	-	-	-	≥22

Table 4. McKern and Stewart sample descriptive statistics. Mean ages, 95% upper and lower confidence levels, and observed age ranges from raw data.

Phase	*n*	Mean Age	Standard Deviation	95% CI Mean Age	Observed Age Range
5 Phase System					
1	118	-	-	-	≤24
2	45	19.7	1.60	19.3-20.2	16-23
3	60	21.4	2.03	20.9-21.9	18-29
4	48	24.1	2.45	23.4-24.9	19-30
5	70	-	-	-	≥20
3 Phase System					
1	118	-	-	-	≤24
2	153	21.8	2.69	21.3-22.2	16-30
3	70	-	-	-	≥20

ranges encompass the full range of variation in the data sets, including outliers. Table 5 presents the results obtained from the transition analysis; the ages in this table are the anti-logged maximum likelihood estimates of the age at which an individual transitions from one phase to the next (i.e. the average age at which the transition is most likely to occur). Since the transition analysis was applied with a cumulative probit model, all transitions are assigned the same standard deviation.

Figure 15 shows the age-at-transition distributions from the 5-phase system for the McCormick and Todd samples. Note the greater separation between the phases in the McCormick sample versus the Todd sample, particularly between the first two transitions (no fusion-beginning fusion and beginning-active). Figure 16 shows the age-at-transition distributions for each of the three samples using the 3-phase system. These graphs illustrate that the onset of fusion in the modern McCormick individuals occurs significantly earlier than in

Table 5. Transition ages. Anti-logs of the maximum likelihood estimate of age at transition (± standard error) and standard deviations (S.D.) from the cumulative probit on log-age transition analysis.

Transition	McCormick Males	McCormick Females	Todd Males	Todd Females	McKern & Stewart
5 Phase System					
None-Beginning	16.19 ± 1.02	15.18 ± 1.03	20.59 ± 1.02	19.02 ± 1.03	19.70 ± 1.01
Beginning-Active	20.39 ± 1.01	18.87 ± 1.02	21.86 ± 1.01	19.78 ± 1.03	20.94 ± 1.01
Active - Recent	23.98 ± 1.01	23.10 ± 1.02	25.24 ± 1.01	23.22 ± 1.02	23.24 ± 1.02
Recent-Complete	26.09 ± 1.01	25.58 ± 1.02	27.94 ± 1.01	25.35 ± 1.02	25.99 ± 1.01
S. D.	1.10 ± 1.01	1.10 ± 1.01	1.12 ± 1.01	1.12 ± 1.01	1.11 ± 1.01
3 Phase System					
None-Fusing	16.15 ± 1.02	15.42 ± 1.03	20.50 ± 1.02	19.24 ± 1.03	19.73 ± 1.01
Fusing-Fused	26.03 ± 1.01	25.39 ± 1.02	27.91 ± 1.01	25.35 ± 1.02	26.01 ± 1.01
S.D.	1.11 ± 1.01	1.11 ± 1.01	1.12 ± 1.01	1.13 ± 1.02	1.12 ± 1.01

the Todd and McKern and Stewart individuals. Additionally, the McKern and Stewart Korean War males are more closely aligned with the Todd males for the first transition, indicating that the majority of the secular trend towards earlier onset of skeletal maturation appears to have occurred during the last several decades.

Table 6 presents the t-test results for sexual dimorphism in epiphyseal union. Significant levels of sexual dimorphism were present in the majority of the transitions in the Hamann-Todd sample, but in only one transition for the McCormick sample. Nonetheless, age differences in epiphyseal union are detectable in the modern sample even though these differences do not register as statistically significant across all transitions. The magnitude of sexual dimorphism is not constant across all stages of growth and development, and since sex differences in skeletal maturation have been well-documented in the literature, it is not advisable to ignore these

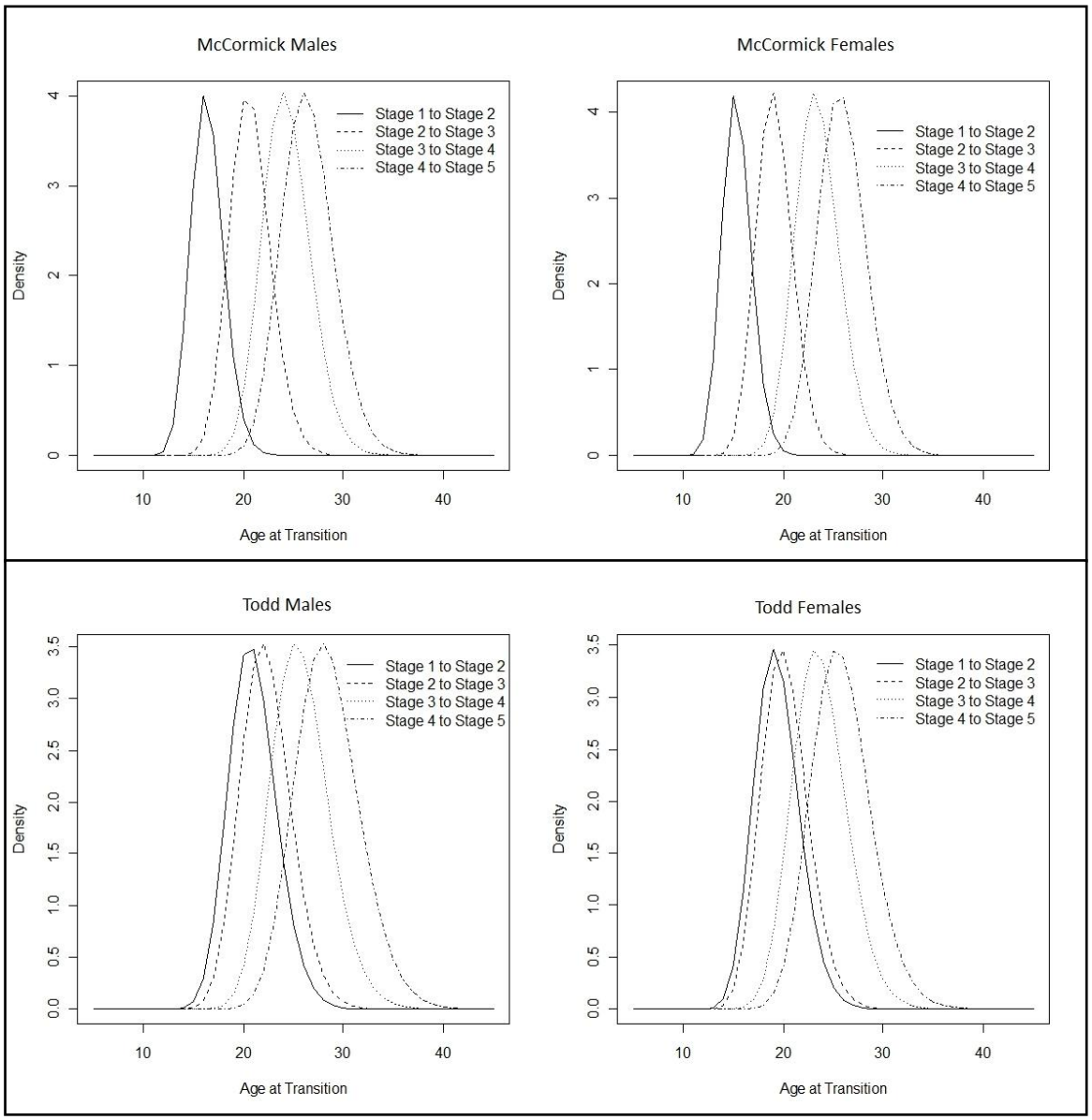

Figure 15. Age-at-transition distributions.

The age-at-transition distributions from the McCormick and Todd samples are represented for the 5-phase system.

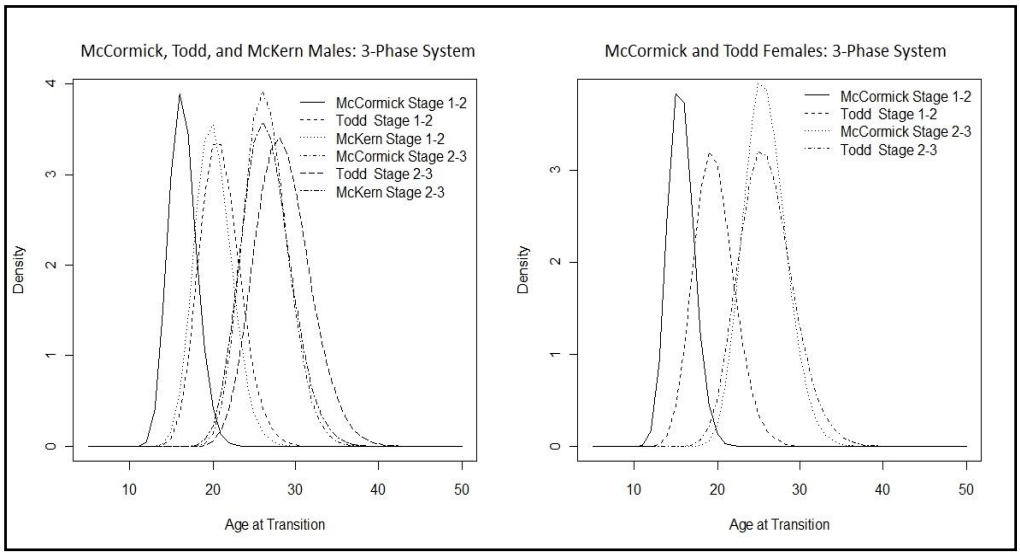

Figure 16. Three sample comparisons.
Probability density distributions from each sample using the 3-phase system.

Table 6. T-tests for sexual dimorphism. Highlighted rows are significant at α=.05. Degrees of freedom = ∞

Transition	t-statistic	p-value
McCormick: 5 Phase System		
No-Beginning	1.892	0.06
Beginning-Active	3.358	<0.001
Active-Recent	1.729	0.09
Recent-Complete	1.056	0.29
Todd: 5 Phase System		
No-Beginning	1.942	0.05
Beginning-Active	3.210	0.001
Active-Recent	3.426	0.001
Recent-Complete	4.270	<.0001
McCormick: 3Phase System		
No fusion-fusing	1.362	0.17
Fusing-Fused	1.247	0.21
Todd: 3 Phase System		
No fusion-fusing	1.802	0.07
Fusing-fused	3.790	<.001

differences (Pryor 1928; Scheuer and Black 2000; Webb and Suchey 1985). Consequently, age ranges are offered for males and females separately in the present study.

Table 7 shows the t-test results for cohort differences in epiphyseal union for males and females, respectively. Significant differences were noted in the all transitions between the McCormick and Todd males and between the McKern and Stewart and Todd males. The most significant differences in between the McCormick males and the McKern and Stewart males occur in the first transition. Likewise, the significant differences between the McCormick and Todd females occurred in the first transition (i.e. in the onset of epiphyseal union). The transition from no fusion to beginning fusion occurs approximately 4 years earlier in the modern McCormick individuals than in the turn of the century Hamann-Todd individuals (see also Table 3). Furthermore, McCormick males transition to beginning fusion an average of 3.5 years earlier than the McKern and Stewart Korean War males. Consequently, as mentioned above, the majority of the secular change in the onset of fusion appears to have occurred during the last several decades. Furthermore, the secular trend is more apparent in the onset of fusion than in the later stages, particularly in females; fusion commences significantly earlier in modern Americans but terminates at roughly the same age as in earlier Americans.

Tables 8 and 9 present the age ranges obtained with the Bayesian analysis. This analysis used the male and female hazard parameters obtained from the Gompertz-modeled FDB (males: α_3=0.01587542, β_3=0.03239691; females: α_3=0.03518139, β_3=0.01013983) and the transition maximum likelihood estimates (see Table 5) to obtain probability density functions and corresponding age ranges. Four different probabilities are given for each phase and for

Table 7. T-tests for secular trends. Highlighted rows are significant at α=.05. Degrees of freedom= ∞

	MALES		FEMALES	
Transition	**t**	**p**	**t**	**p**
McCormick-Todd: 5 Phase System				
No-Beginning	10.638	<.0001	5.308	<.0001
Beginning-Active	4.0873	<.0001	1.341	0.18
Active-Recent	3.378	<.001	0.185	0.85
Recent-Complete	4.531	<.0001	0.367	0.71
McCormick-McKern: 5 Phase System				
No-Beginning	10.265	<.0001	-	-
Beginning-Active	1.835	0.07	-	-
Active-Recent	2.187	0.03	-	-
Recent-Complete	0.258	0.80	-	-
McKern-Todd: 5 Phase System				
No-Beginning	2.547	0.01	-	-
Beginning-Active	2.761	0.006	-	-
Active-Recent	5.235	<.0001	-	-
Recent-Complete	4.130	<.0001	-	-
McCormick-Todd: 3 Phase System				
No fusion-fusing	10.191	<.0001	5.182	<.0001
Fusing-fused	4.490	<.0001	0.053	0.96
McCormick-McKern: 3 Phase System				
No fusion-fusing	10.053	<.0001	-	-
Fusing-fused	0.0250	0.98	-	-
McKern-Todd: 3 Phase System				
No fusion-fusing	2.084	0.04	-	-
Fusing-fused	3.807	<.001	-	-

Table 8. Male age ranges. Age ranges for modern males based on the highest posterior density regions from the McCormick sample.

Phase	HPD	50% CI	75% CI	90% CI	95% CI
5 Phase System					
1	12.7	≤13.7	≤15.2	≤16.6	≤17.4
2	18.2	16.8- 19.8	15.9-21.0	15.0-22.2	14.4-23.0
3	22.2	20.6-23.9	19.5-25.2	18.5-26.6	17.9-27.5
4	25.1	23.4-26.9	22.3-28.2	21.2-29.6	20.5-30.6
5	34.7	≥29.1	≥26.9	≥25.1	≥24.1
3 Phase System					
1	12.6	≤13.7	≤15.2	≤16.6	≤17.4
2	20.9	18.3-23.5	16.8-25.4	15.6-27.3	14.9-28.5
3	34.8	≥29.1	≥26.9	≥25.0	≥24.0

Table 9. Female age ranges. Age ranges for modern females based on the highest posterior density regions from the McCormick sample.

Phase	HPD	50% CI	75% CI	90% CI	95% CI
5 Phase System					
1	-	≤13.1	≤14.3	≤15.5	≤16.2
2	16.8	15.6- 18.2	14.8-19.2	14.0-20.3	13.5-21.0
3	20.7	19.2-22.4	18.2-23.6	17.3-24.9	16.7-25.7
4	24.1	22.6-25.8	21.5-27.0	20.5-28.4	19.9-29.3
5	30.1	≥25.9	≥24.3	≥23.2	≥22.5
3 Phase System					
1	-	≤13.2	≤14.6	≤15.8	≤16.5
2	18.8	16.8-21.9	15.7-24.0	14.6-25.9	14.1-27.1
3	30.1	≥25.7	≥24.1	≥22.9	≥22.1

both scoring systems (50%, 75%, 90%, and 95%). These prediction intervals represent the probability that an individual falls within an age range given that their clavicle exhibits the morphology of a given phase. Since this sample was truncated at 11 years on the lower end and 33 years on the upper end, the age ranges for first and last phases are expressed in terms of the oldest and youngest probable ages for those phases. Tables 10 and 11 show the results of the test on the William M. Bass individuals using both scoring systems. The correctly aged individuals are outlined with dotted lines. The test sample size is small and contains only one female, so statistical inferences should not be drawn from this test. Nonetheless, the test illustrates that aging accuracy is better with the 90% and 95% probabilities than with the 50% and 75% probabilities. The age ranges resulting from the Bayesian analysis were successful in aging the majority of the cases using the 90% and 95% probabilities; therefore, these probabilities encompass an adequate proportion of the individual variability in fusion.

The tables in this study can be used to arrive at an age estimate in several ways. Table 2 provides age ranges based on the raw data. In some instances, these age ranges are extended several years due to the presence of a single outlier. For example, of the 207 McCormick males exhibiting complete fusion, one individual is 19, no individuals are between 20 and 22, two individuals are 23, one individual is 24, and then the frequency increases with individuals 25 years old and beyond. In other words, the 19-year old outlier is not representative of typical fusion in that it constitutes a mere 0.48% of the variation. Nonetheless, some researchers argue that wide age ranges are necessary to address the unpredictable variation in individual cases (2001; Schaefer and Black 2005). In addition, Table 2 could be used to devise age ranges

Table 10. 5-Phase test of PDF age ranges. Test on 12 individuals from the William M. Bass Donated Collection using the 5-phase scoring system.

Sex	Age	Score	Predicted 50%	Predicted 75%	Predicted 90%	Predicted 95%
M	26	5	≥29.1	≥26.9	≥25.1	≥24.1
M	19	1	≤13.7	≤15.2	≤16.6	≤17.4
F	20	2	15.6- 18.2	14.8-19.2	14.0-20.3	13.5-21.0
M	27	4	23.4-26.9	22.3-28.2	21.2-29.6	20.5-30.6
M	26	4	23.4-26.9	22.3-28.2	21.2-29.6	20.5-30.6
M	27	5	≥29.1	≥26.9	≥25.1	≥24.1
M	27	5	≥29.1	≥26.9	≥25.1	≥24.1
M	29	5	≥29.1	≥26.9	≥25.1	≥24.1
M	25	1	≤13.7	≤15.2	≤16.6	≤17.4
M	27	4	23.4-26.9	22.3-28.2	21.2-29.6	20.5-30.6
M	23	3	20.6-23.9	19.5-25.2	18.5-26.6	17.9-27.5
M	25	5	≥29.1	≥26.9	≥25.1	≥24.1

Table 11. 3-Phase test of PDF age ranges. Test on 12 individuals from the William M. Bass Donated Collection using the 3-phase scoring system.

Sex	Age	Score	Predicted 50%	Predicted 75%	Predicted 90%	Predicted 95%
M	26	3	≥29.1	≥26.9	≥25.0	≥24.0
M	19	1	≤13.7	≤15.2	≤16.6	≤17.4
F	20	2	16.8-21.9	15.7-24.0	14.6-25.9	14.1-27.1
M	27	2	18.3-23.5	16.8-25.4	15.6-27.3	14.9-28.5
M	26	2	18.3-23.5	16.8-25.4	15.6-27.3	14.9-28.5
M	27	3	≥29.1	≥26.9	≥25.0	≥24.0
M	27	3	≥29.1	≥26.9	≥25.0	≥24.0
M	29	3	≥29.1	≥26.9	≥25.0	≥24.0
M	25	1	≤13.7	≤15.2	≤16.6	≤17.4
M	27	2	18.3-23.5	16.8-25.4	15.6-27.3	14.9-28.5
M	23	2	18.3-23.5	16.8-25.4	15.6-27.3	14.9-28.5
M	25	3	≥29.1	≥26.9	≥25.0	≥24.0

based on the mean ±1-2 standard deviations. However, age ranges obtained with simple descriptive statistics use what Konigsberg et al. (Konigsberg et al. 2008) refer to as a "hidden Bayesian approach" in that the reference sample also serves as the prior age-at-death distribution. This introduces the problem of age mimicry, whereby the age estimates of the target sample are influenced by the composition of the reference sample (Boldsen et al. 2002; Konigsberg and Frankenberg 1994). For these reasons, the authors have used a Bayesian approach to obtain conservative, statistically sound estimates that are less sensitive to the influences of age mimicry and developmental outliers (Tables 8 and 9). Of course, whenever possible age estimates in forensic cases should be based on multiple age indicators, particularly because individual areas of a given skeleton may be advanced or delayed relative to other areas (Klepinger 2001; Schulz et al. 2008).

Sex Estimation

Linear Discriminant Analysis

The McHenry's algorithm multivariate variable selection indicated that maximum length, maximum midshaft diameter, minimum midshaft diameter, and the maximum and minimum diameters at the lateral end offered high discriminatory capabilities. Table 12 presents the cross-validated classification rates from several models for African Americans and European Americans separately. The linear discriminant functions associated with these models are available in Appendix A. Classification rates are better for females than for males, presumably because of lower variance in the female sample. Maximum and minimum diameters at midshaft were included in most of the best models for both ethnic groups.

Table 12. LDA accuracy rates.

MODEL	EUROPEAN AMERICANS (*n*=1354)			AFRICAN AMERICANS (*n*=53)		
	Males	Females	Total	Males	Females	Total
Sagittal, Vertical	87.16	88.83	87.59	87.50	84.62	86.79
MaxMid, MinMid	86.67	87.68	86.93	87.50	100.00	90.57
MaxL, Sagittal, Vertical	91.34	91.40	91.36	82.50	92.31	84.91
MaxL, MaxMid, MinMid	91.24	92.55	91.58	82.50	100.00	86.79
MaxMedial, MinMedial	61.59	65.33	62.56	57.50	61.54	58.49
MaxLateral, MinLateral	76.32	83.95	78.29	70.00	84.62	73.58
MaxL, MaxLat, MinLat	87.66	90.54	88.40	80.00	100.00	84.91
MaxL, MaxMid, MinMid, MaxLat, MinLat	92.24	93.41	92.54	82.50	100.00	86.79
All Variables	92.34	93.70	92.69	80.00	100.00	84.91

Moreover, the lateral shaft diameters were promising discriminatory variables in both groups, but the medial shaft diameters performed the worst.

The best model for European Americans was the full model with all variables (92.69%). A five variable model with maximum length, maximum midshaft diameter, minimum midshaft diameter, maximum lateral shaft diameter, and minimum lateral shaft diameter was nearly as accurate (92.54%). However, a simple three variable model with maximum length and the maximum and minimum midshaft diameters yielded nearly 92% accuracy (91.58%). Interestingly, a simple model with maximum length and the two lateral diameters achieved 88.4% accuracy. These models agree with the McHenry's algorithm variable selection. Maximum length appears to be a critical measurement, and the maximum and minimum midshaft and lateral shaft diameters offer high discriminatory capacity, as well.

The results for the African American sample are similar. The best overall model was a two variable model: maximum and minimum midshaft diameters (90.57%). Adding maximum length to that model slightly decreased the accuracy rate (86.79%). In addition, substituting

maximum/minimum midshaft diameters in the two variable model with sagittal/vertical midshaft diameters gave similar results (86.79%). Likewise, the five variable model with maximum length, maximum midshaft diameter, minimum midshaft diameter, maximum lateral shaft diameter, and minimum lateral shaft diameter classified with 86.79% accuracy. Although these results are generally consistent with the results for the European American group, they should be considered preliminary because of the small African American sample size (n=53, with 40 males and 13 females). Appendix B presents a summary table of the means and standard deviations for each measurement evaluated in the linear discriminant analysis.

Clavicle Atlas: Size and Shape Dimorphism

The following results are based on the first 10 principal components of a 123 bone atlas. This preliminary atlas contains 87 principal components; 99.5% of the cumulative variance was used as the threshold for principal component selection. Appendix C gives the cumulative eigenvalue percentages for each of these principal components. The first 10 principal components accounted for 94.2% of the cumulative variance. Presumably, this percentage will increase as more bones are added to the atlas, but this is unlikely to alter the results of the PCA-FDR algorithm significantly.

Figures 17 and 18 show color maps of the first principal component. Red areas are the most dimorphic, followed by orange, yellow, green, light blue and dark blue. The first principal component accounted for 81.5% of the cumulative variance and reflects dimorphism primarily

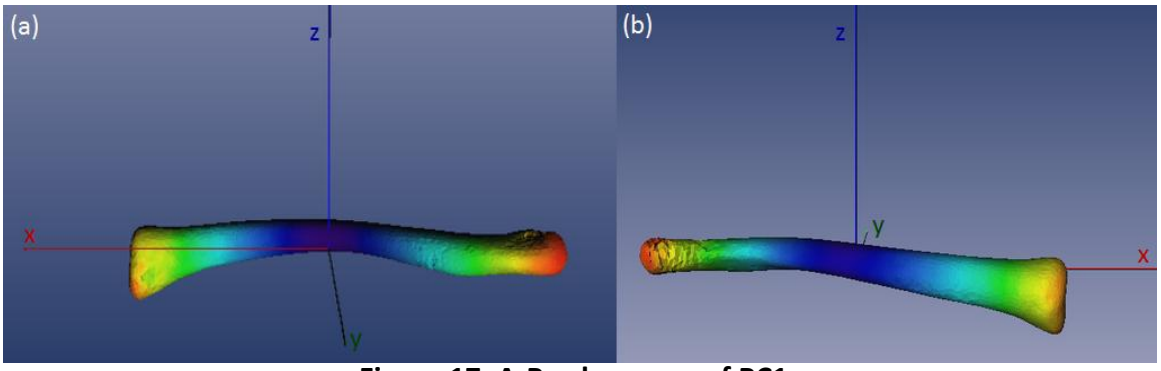

Figure 17. A-P color maps of PC1.
(a) posterior view (b) anterior view

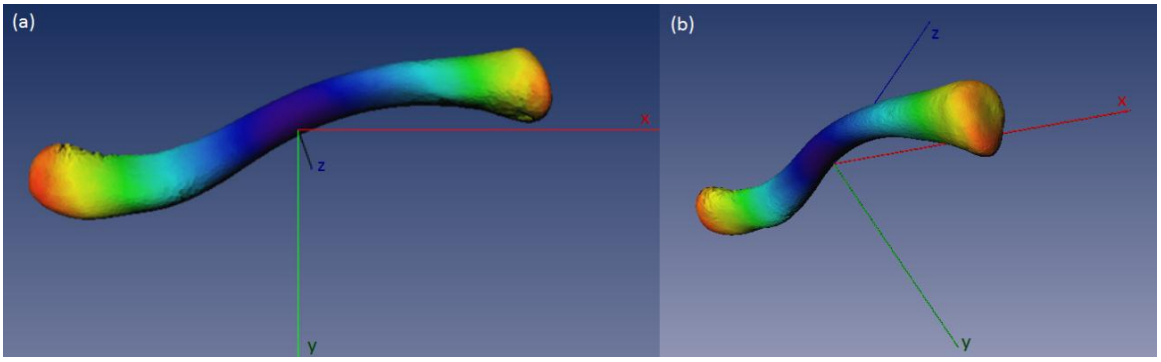

Figure 18. S-I color maps of PC1.
(a) superior view (b) inferior view

due to scale. These results show that the most size-dimorphic regions of the clavicle are on the medial and lateral ends, particularly the lateral end. The midshaft region appears less dimorphic because of the effects of scale in the first principal component. Because males are generally larger than females, the magnitudes of difference on the ends of the bone are greater than in the middle when scaling isometrically. In other words, the points on the ends of the bone "move" more than the points in the middle and therefore represent the greatest change on a global scale. This effect is eliminated with the remaining principal components because scale is removed.

Figures 19 and 20 are color maps of the second principal component. Since the first principal component has been removed, these magnitudes show sexual dimorphism in clavicular shape. Note the striking differences between these color maps and the maps of the first principal component; the primary differences in shape are due to the bone's unique curvature. The most dimorphic areas are the convex curvature of the anterior midshaft and the convex curvature of the posterior surface of the lateral end.

Figure 21 shows inferior views of various combinations of the second through tenth principal components. The inferior surface shows little shape dimorphism in the midshaft area, but pronounced dimorphism in the shape of the medial and lateral extremities (a-c). However, with addition of the sixth, seventh, and eighth principal components, deviation in curvature of the medial and lateral ends is apparent (d-f). Adding the ninth and tenth principal components captures additional shape dimorphism in the lateral curvature in the anterior, posterior, and inferior directions.

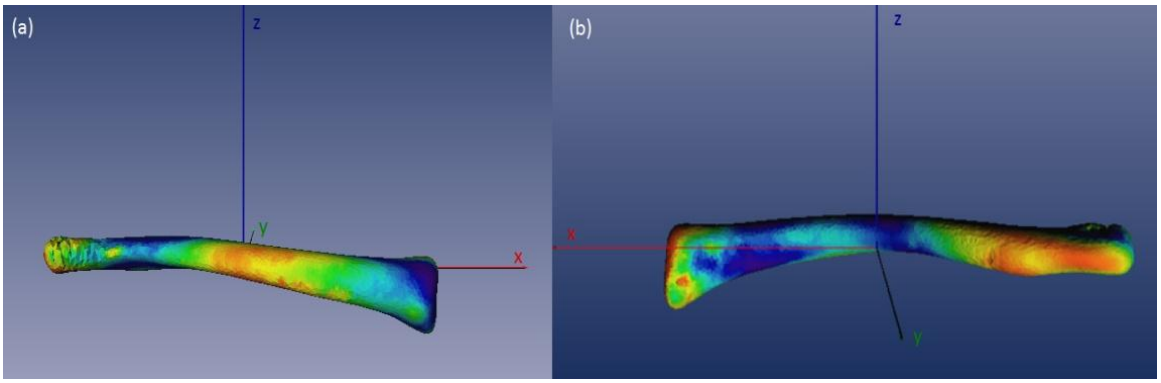

Figure 19. A-P color maps of PC2.
(a) anterior view (b) posterior view

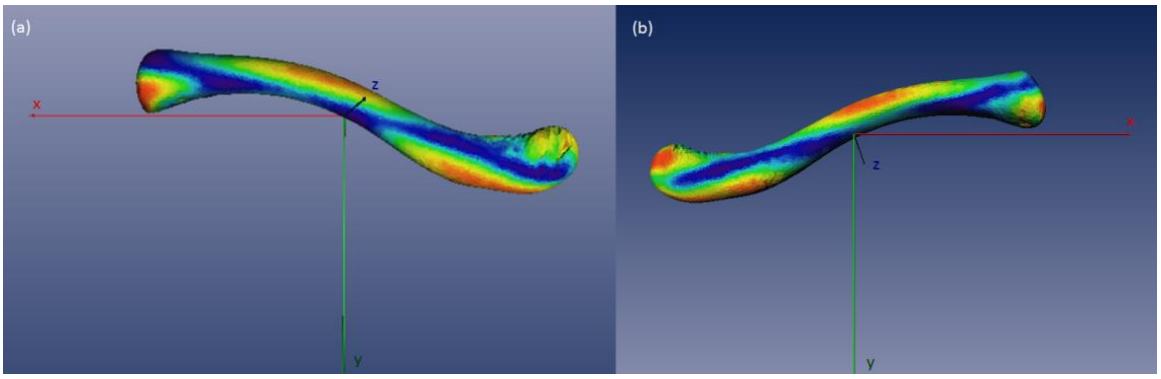

Figure 20. S-I color maps of PC2.
(a) superior view (b) inferior view

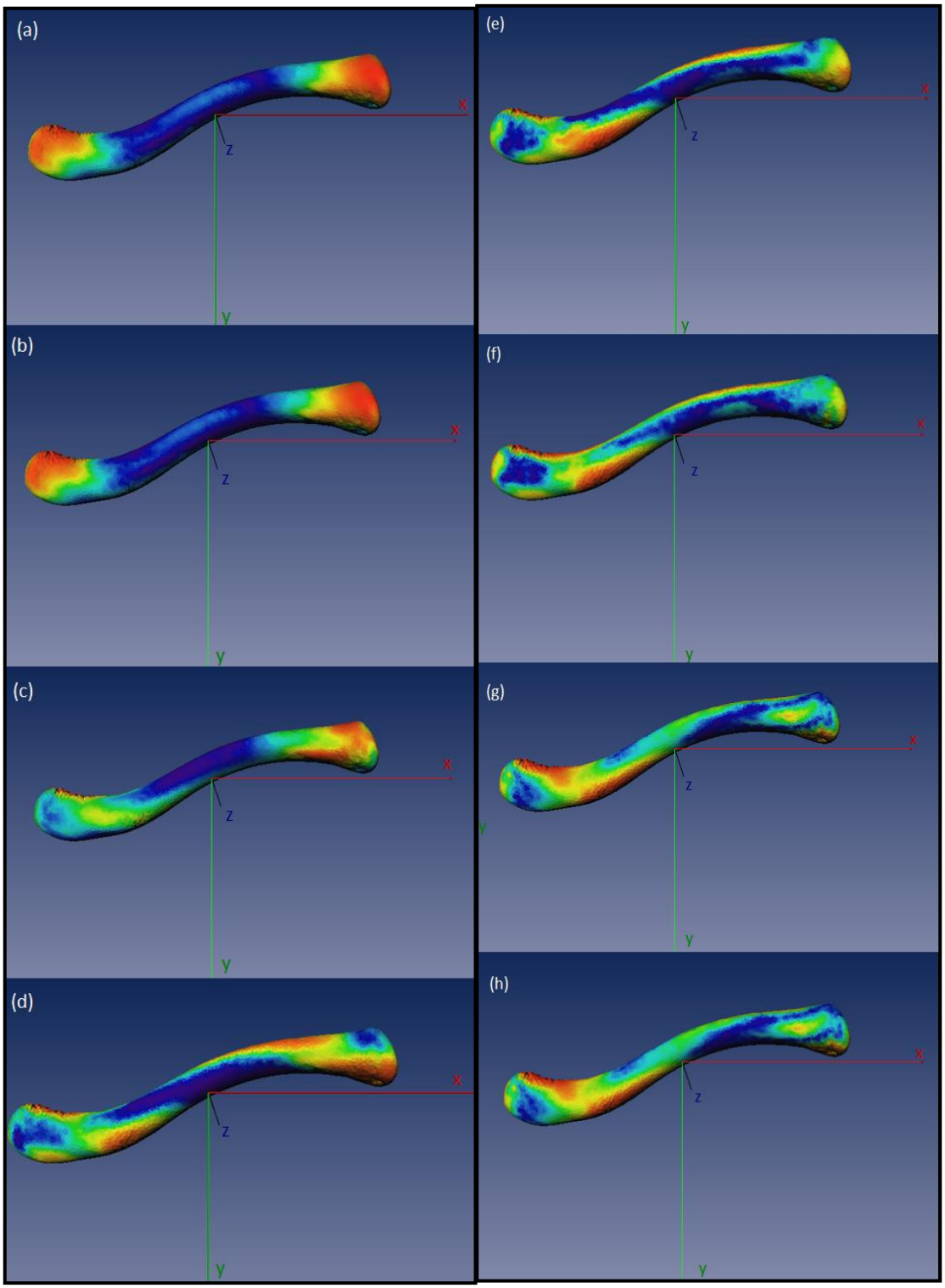

Figure 21. PC2-PC10 inferior view.
(a) PC2-3 (b) PC2-4 (c) PC2-5 (d) PC2-6 (e) PC2-7 (f) PC2-8 (g) PC2-9 (h) PC2-10

Figure 22 offers superior views of the second through tenth principal components. Unlike the inferior surface, the superior surface shows dimorphism in the midshaft region, thus indicating differences in the superior-inferior bending of the midshaft (Figure 22, a, b, g, h). Adding the fourth through tenth principal components captures extreme magnitude differences in the antero-lateral end (Figure 22, c-h). Figure 23 depicts shape differences on the anterior surface of the clavicle. The concave curvature of the antero-lateral end is the area of highest dimorphism. Additionally, the sixth-tenth principal components highlight the anterior midshaft curvature as significantly dimorphic (Figure 23, d-h). Examination of the posterior surface of the bone reveals that the convex curvature of the postero-lateral portion of the shaft is highly dimorphic (Figure 24, d-h); additionally, there appears to be some dimorphism in the shape of the sternal end (Figure 24, c-e, g, h). Figure 25 provides supero-medial views of various combinations of the second through tenth principal components. The most dimorphic area indicated by the second principal component is on the posterior aspect of the medial surface, but this dimorphism is not picked up with any of the other principal components (Figure 25a). Adding the third and fourth principal components captures the curvature differences on the superior midshaft surface mentioned earlier (Figure 25, a-b). These views also show the high magnitudes of dimorphism in the lateral end, both in the concave anterior curvature and in the convex posterior curvature (Figure 25, d-h). Additionally, dimorphism in the anterior bowing of the midshaft region is apparent (Figure 25, e-h).

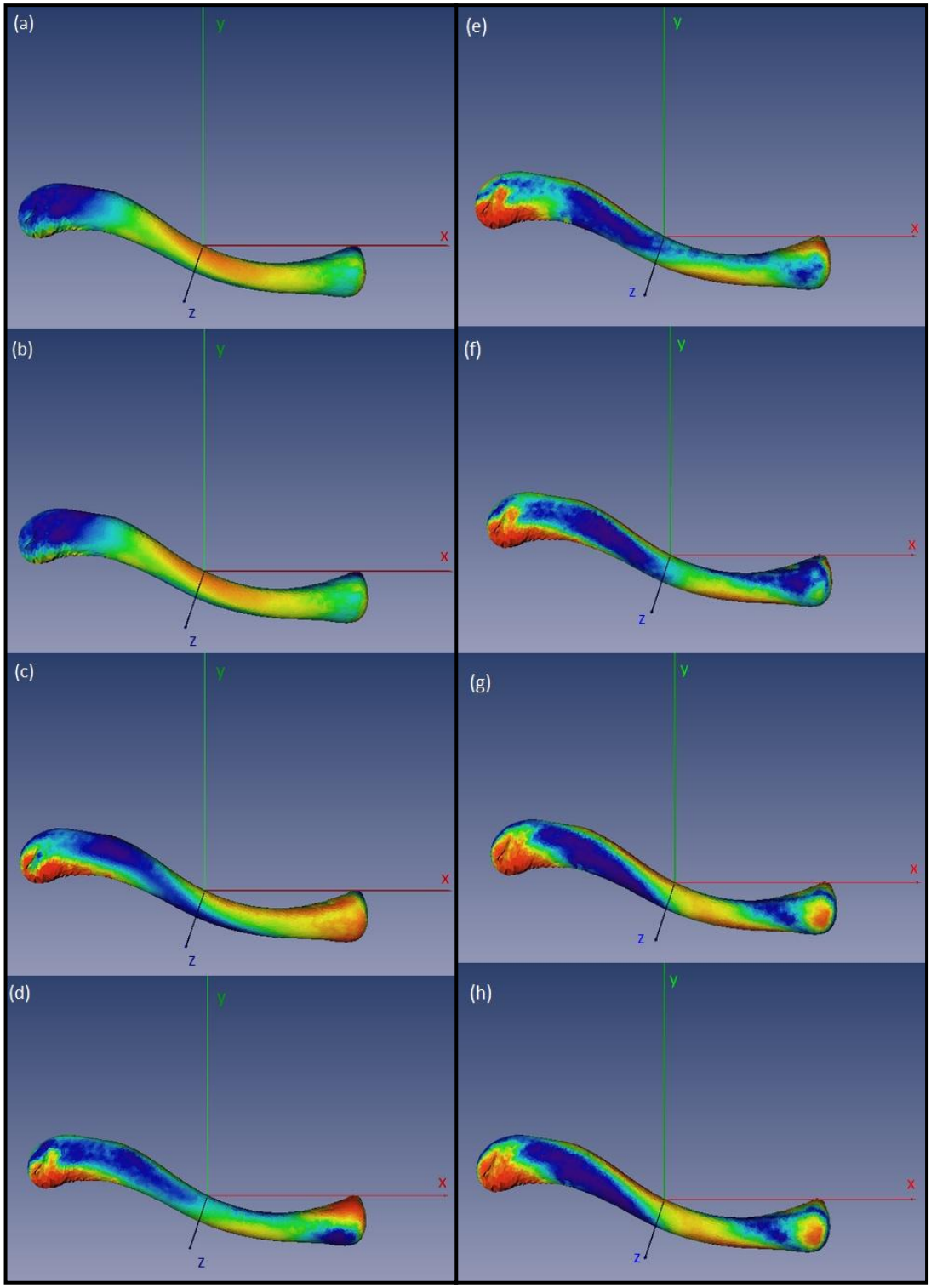

Figure 22. PC2-PC10 superior view.
(a) PC2-3 (b) PC2-4 (c) PC2-5 (d) PC2-6 (e) PC2-7 (f) PC2-8 (g) PC2-9 (h) PC2-10

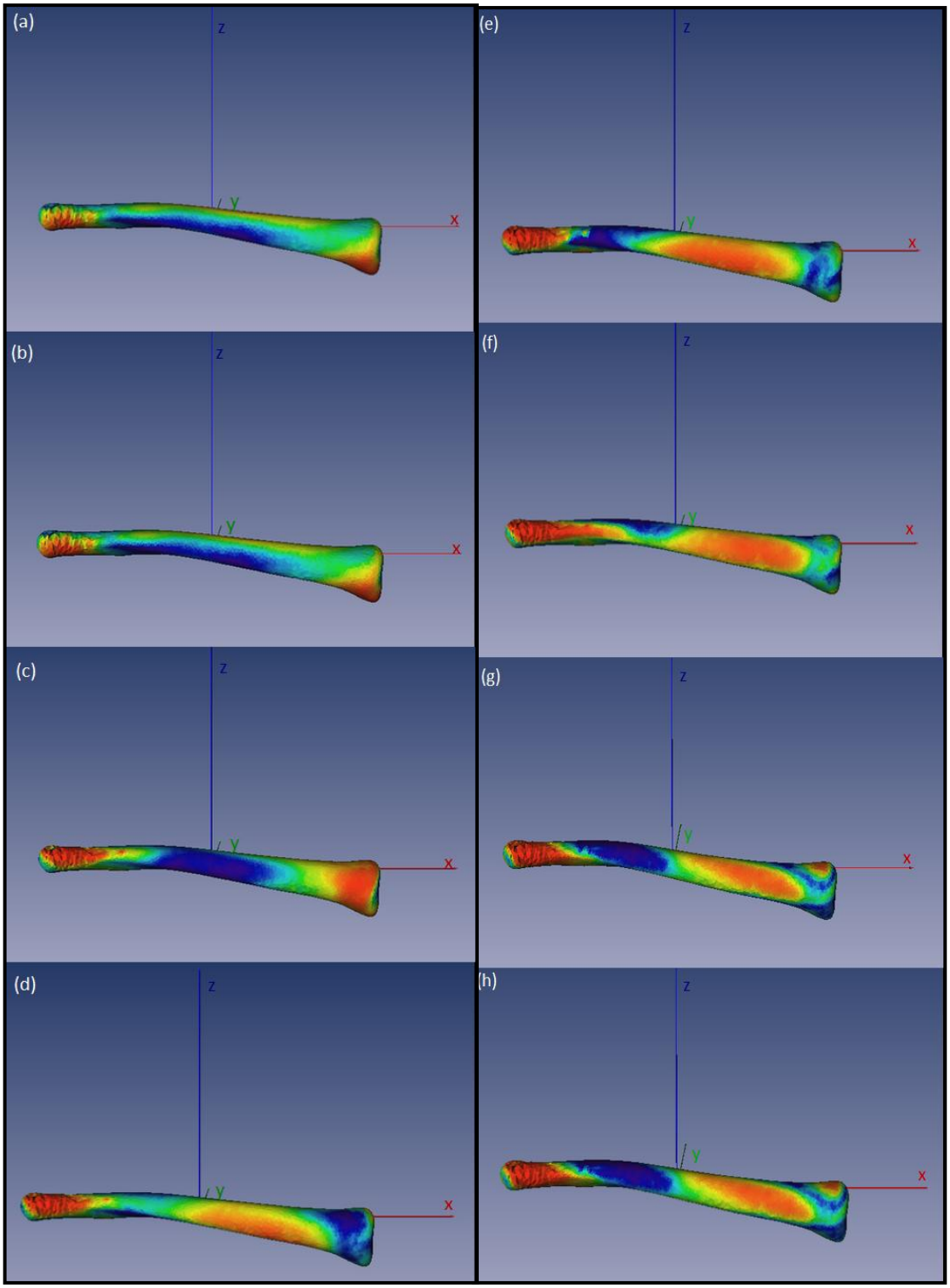

Figure 23. PC2-PC10 anterior view.
(a) PC2-3 (b) PC2-4 (c) PC2-5 (d) PC2-6 (e) PC2-7 (f) PC2-8 (g) PC2-9 (h) PC2-10

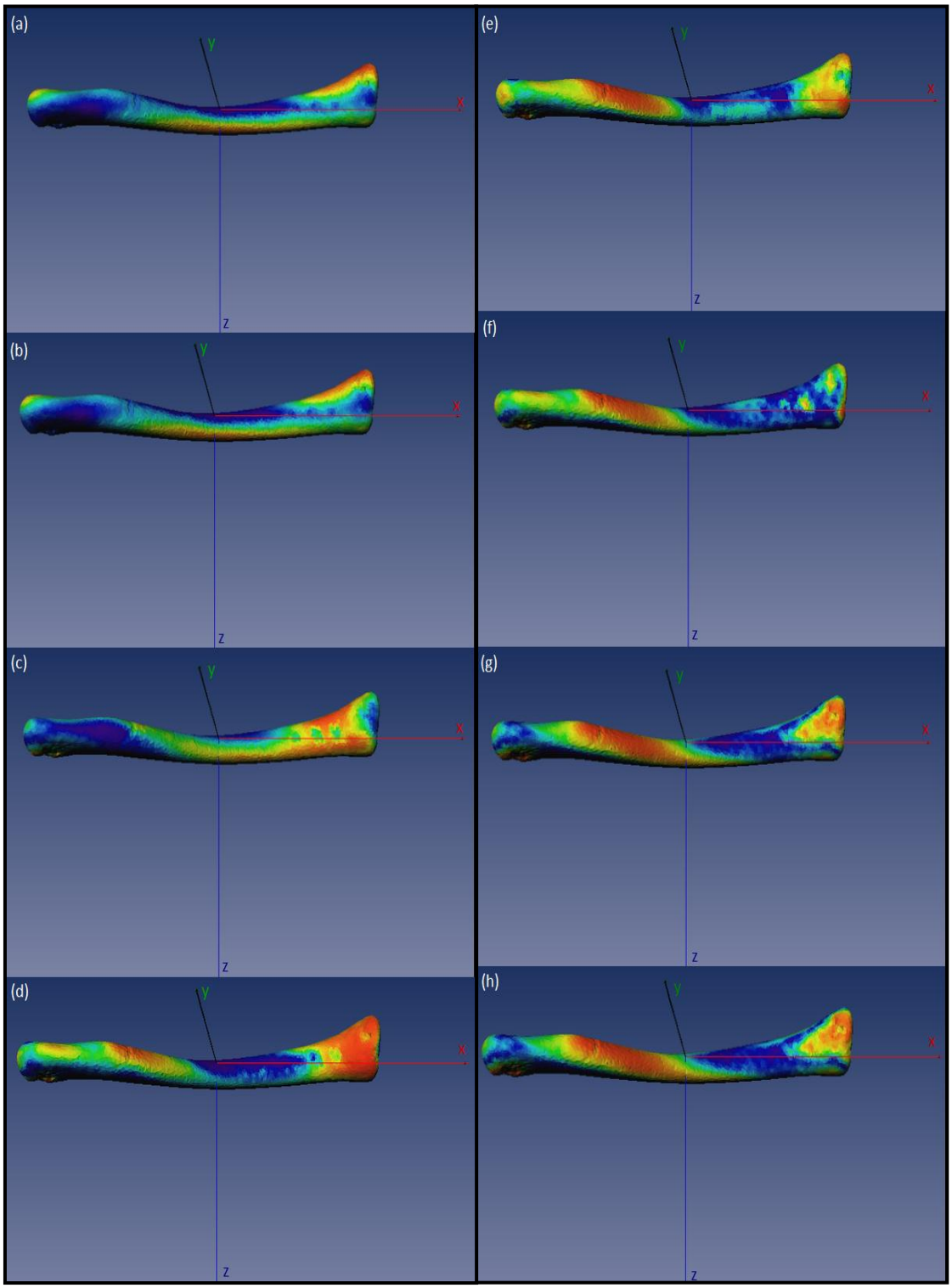

Figure 24. PC2-PC10 posterior view.

(a) PC2-3 (b) PC2-4 (c) PC2-5 (d) PC2-6 (e) PC2-7 (f) PC2-8 (g) PC2-9 (h) PC2-10

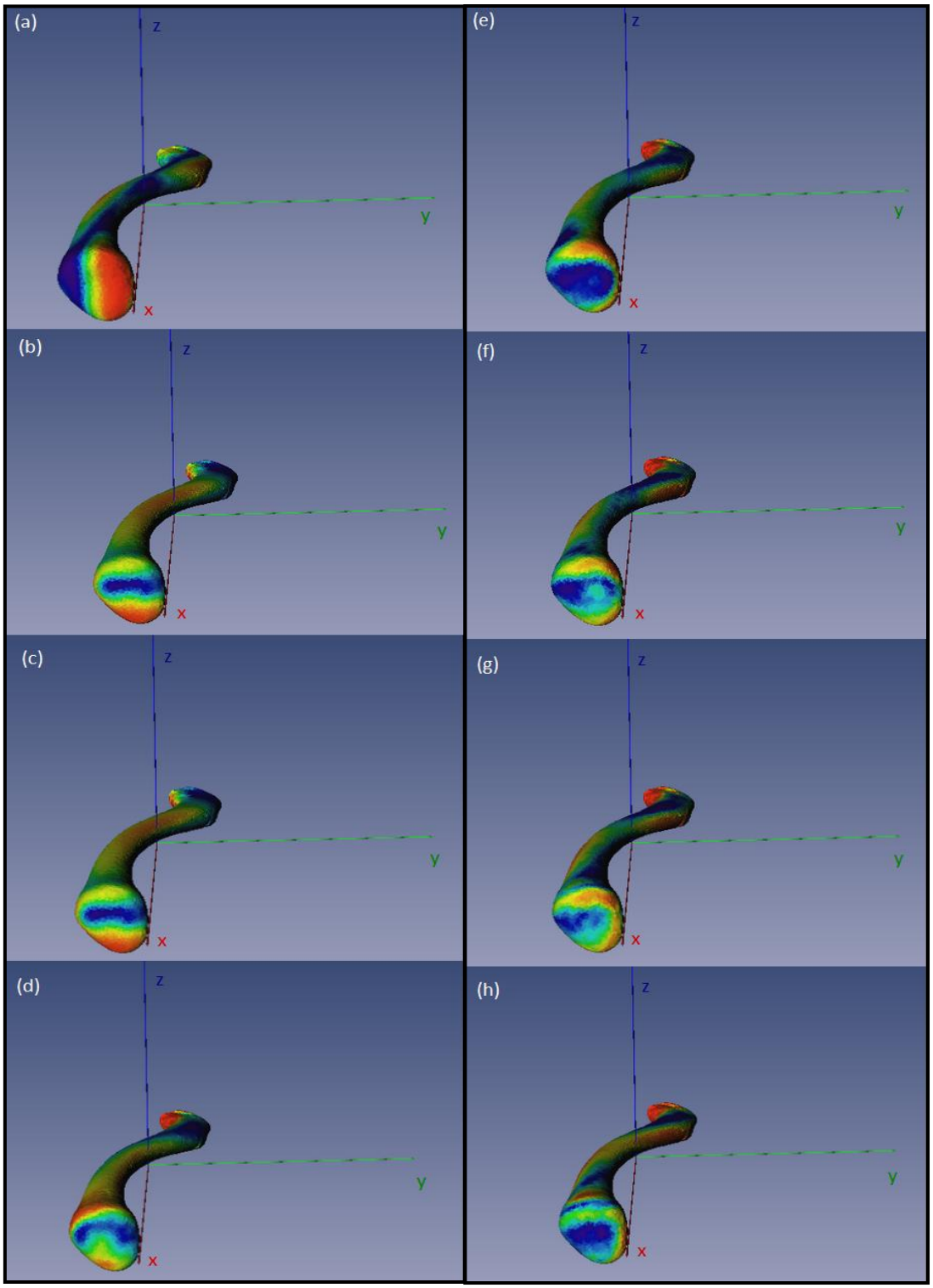

Figure 25. PC2-PC10 medial and superior view.
(a) PC2 (b) PC2-3 (c) PC2-4 (d) PC2-5 (e) PC2-6 (f) PC2-7 (g) PC2-8 (h) PC2-10

Chapter 4: Discussion

Secular Change in Skeletal Maturation

Although maturational secular change has been well-documented in terms of somatic changes (i.e. earlier menarche and development of secondary sex characteristics), changes in skeletal maturation have received little attention, primarily because of the lack of sizable sub-adult skeletal samples available for cross-cohort comparisons. In fact, age at menarche in western, industrialized populations has decreased over the past five decades by around 4-6 months per decade, and pubertal onset in American females occurs as early as 8-10 years of age (Fredriks et al. 2000; Herman-Giddens et al. 1997; Malina 1979; Morrison et al. 1994). Consequently, we might expect to see related changes in skeletal maturation. The results of this study indicate that the average age at which epiphyseal union commences in the medial clavicle is 4 years earlier in modern Americans than in the American population from the early 20th century. Furthermore, comparisons between modern American males and American males from the Korean War era suggest that much of this acceleration has occurred during the latter part of the 20th century.

Reasons for secular changes in height, weight, and maturation are numerous; however, these changes are not necessarily caused by introducing growth-stimulation factors, but rather by eliminating growth-inhibiting factors (i.e. nutritional stress, environmental stresses, and disease) (Malina 1979). Several factors contributing to body weight increases, overall health improvements, and accelerated maturation are stable caloric intake, reduced caloric expenditure, reduction in physical activity levels, increased calcium intake, introducing cereals

at early age in infant diet, increased consumption of processed sugars and fats, improved socioeconomic status, improved health status, improvements in water and sanitation, elimination of infectious diseases, reduction in infant mortality, increased life expectancy, and reduction in family size (Malina 1979). Furthermore, recent studies suggest that obesity has a significant effect on menarcheal age. Obese American females reach menarche earlier than normal weight females and are more likely to reach menarche before age 12 (Wattigney et al. 1999). Similar obesity correlates have been reported in the European population (Vignolo et al. 1988) and in populations worldwide (1990). In a forensic context, secular trends can affect age estimates derived from epiphyseal union because many skeletal aging standards were derived from documented collections of individuals born during the 19th and early 20th centuries. As indicated by the results in this study, age estimates based on reference standards from populations that have undergone significant positive secular change may be significant overestimates.

The standards proposed in this study were derived using skeletal samples from the American population. Significant ethnic differences between African American and European American Americans were not detected in this analysis, possibly because the factors affecting growth are similar for the American population as a whole, regardless of ethnicity. Indeed, socioeconomic status is cited more frequently than ethnicity as the most influential variable in maturational differences (Abioye-Kuteyi et al. 1997; Alberman et al. 1991; Bagga and Kulkarni 2000; Bodzsar 2000; Cardoso 2008a; Kim et al. 2008; Laska-Mierzejewska et al. 1982; Low et al. 1982; Malina 1979; Meijerman et al. 2007; Prado 1984; Rimpela and Rimpela 1993; Todd 1937).

Schmeling and co-workers (2000) argue that skeletal maturation occurs in stages that are the same for all ethnic groups; the critical factor that brings about differences in ossification rates is the socioeconomic status of a given population. Although a genetically-determined potential of skeletal maturation may exist, this potential does not appear to depend on ethnicity. Instead, growth potential is realized under favorable environmental conditions (namely high socioeconomic status), and population affiliation appears have no appreciable effect on skeletal age (Schmeling et al. 2005). In fact, regression analysis has shown a positive relationship between ossification rates and medical modernization, as well as economic progress (Schmeling et al. 2005; Schmeling et al. 2006). Nonetheless, the results presented here should not be applied to populations outside of the United States, especially those in which different growth environments may influence development and maturation. Furthermore, recent evidence suggests that Hispanic individuals in the United States may be experiencing differential maturation due to ethnic differences, lower socioeconomic status, or a combination of these factors (Crowder and Austin 2005). Careful controls should be used in future research regarding the effects of ethnicity and socio-economic status in order to ascertain the individual effects of these factors on growth and maturation.

Another matter of concern in making developmental status comparisons is the method used to establish stage of union. Many skeletal maturation standards are based on radiographic studies of living individuals, and the appearance of the epiphyses on radiographs are not necessarily the same as they appear in dry bone (Cardoso 2008a; Todd 1937). Additionally, many studies are beginning to use CT scans to establish age ranges for epiphyseal

fusion. Commencement of fusion can be detected earlier with radiographs and CT scans than with dry bone observations (Meijerman et al. 2007). Consequently, forensic age estimates based on conventional radiographs should use standards developed from radiographs, whereas age estimates based on CT scans should refer to CT-based standards, and estimates based on dry bone observations should use standards developed from dry bone (Cardoso 2008a; Cardoso 2008b; Schulz et al. 2008). Accordingly, the standards proposed in this study should be applied to dry bone observations only.

The results of this study call attention to the importance of using modern standards to assess the age of modern individuals. As the American population continues to change, particularly with the current obesity epidemic, forensic anthropologists will be charged with the task of evaluating how these changes in human biology affect our understanding of human skeletal variation. The dynamic nature of human populations necessitates constant research in order to ensure the most rigorous standards and current practices are available for use. Finally, age ranges derived using Bayesian statistics circumvent the issues of age mimicry and developmental outliers. Traditionally, these methods have been employed to establish standards for adult age estimation. However, they have proven useful in this study of sub-adult age estimation. Consequently, a Bayesian approach should be considered in future evaluations of sub-adult skeletal aging.

Sexual Dimorphism

Traditional approaches to sex estimation from skeletal elements focus on establishing linear discriminant models from a combination of osteometric variables. The discriminating

capability of these models depends on the degree of size dimorphism captured by the measurements. Given a set of easily reproducible measurements, discriminant functions are straightforward and yield reasonably accurate sex estimations if applied to the appropriate populations. The bone measurements used by anthropologists today were adapted from Martin (1956; 1957) and commonly include measurements of length, width, diameter, and circumference. Finding the measurements with the highest discriminatory capacity is usually a matter of extensive data collection and trial and error.

This study explored 2 approaches to discovering high degrees of sexual dimorphism in the clavicle. The first was a traditional metric approach using linear discriminant analysis. This analysis used a combination of established and newly developed computer-automated measurements from 3-D models. The traditional measurements were found to be highly accurate sex discriminators; in particular, maximum length is a critical variable for most linear discriminant functions. Several of the new measurements proved useful, too. The maximum and minimum midshaft diameters performed just as well as, if not slightly better than sagittal and vertical diameters at midshaft. Moreover, finding maximum and minimum diameters with a set of calipers or on computer models is easier than approximating the sagittal and vertical diameters, because the latter require a rough estimation of the bone's anatomical orientation. In many instances sagittal diameter is greater than vertical diameter, so these measurements may function similar to maximum and minimum diameters. However, some clavicles are "twisted" or unusually torqued, and the opposite scenario is true. In these instances, the maximum and minimum diameters will give a more consistent sex estimate than the sagittal

and vertical diameters. Consequently, maximum and minimum midshaft diameters are capable substitutes for the traditional midshaft diameters in that they capture the true robusticity of the clavicle midshaft and are more easily measured. In addition, the diameters of the lateral end proved to be accurate sex discriminators, particularly in combination with maximum length. On the other hand, the diameters of the medial end were not highly dimorphic and are not recommended for sex estimation.

The second approach was similar to geometric morphometric techniques in that it used 3-D statistical atlases to capture the primary shape variation in the clavicle. The PCA-FDR statistical treatment revealed striking differences between areas of greatest dimorphism due to scale and those attributable to shape. Like the linear discriminant analysis, the statistical atlas results suggested that measurements of the lateral end are promising criteria for sex estimation. The shape analyses showed that there are significant differences in curvature between males and females (Figure 28). The primary differences are in the anterior and superior curvature of the midshaft and in the posteriorly-oriented curvature of the lateral end. In particular, the greatest magnitudes of difference in the lateral clavicle were in the concave curvature of the anterior surface and in the corresponding convex curvature of the posterior surface. The most useful feature of the PCA-FDR statistical treatment is that it functions as an automated feature generation technique by highlighting areas of highest discrimination.

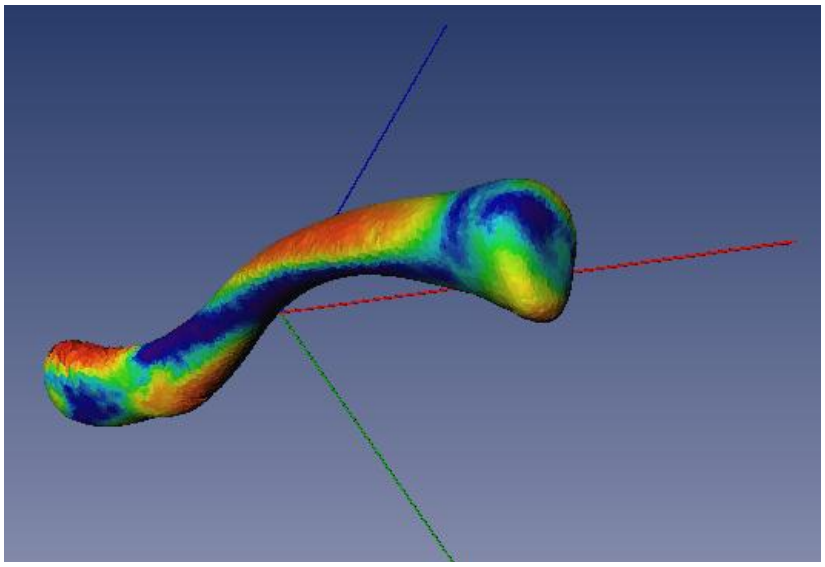

Figure 26. PC2-PC8 inferior view.
Note areas of significant dimorphism highlighted in red.

According to the authors of the algorithm, landmarks located in the areas of greatest magnitude differences (i.e. the red areas) should be able to stand alone as highly accurate sex discriminators (Mahfouz et al. 2007b). In other words, this methodology provides an exploratory technique for finding the most dimorphic areas in any skeletal element with relation to scale and/or shape. Landmarks and/or measurements generated in these areas should capture the highest levels of dimorphism in a given skeletal element. This provides the potential of estimating sex with a minimal number of highly accurate measurements. Additionally, this research indicates that effective quantification of shape may be able to enhance our sex discrimination capabilities from the clavicle, as well as from other skeletal elements.

Chapter 5: Conclusions—Looking Ahead

This dissertation focused on a single skeletal element, but the research findings have far-reaching implications for our understanding of modern human skeletal biology and skeletal variation in past populations. In terms of skeletal maturation, indications from the medial clavicle provide evidence for earlier onset of skeletal maturation. This secular trend is in line with the trend towards earlier somatic maturation--particularly earlier menarcheal onset--and may have implications for secular increases in stature. Undoubtedly, there is a link between somatic and skeletal maturation, and the next logical step for this line of research would be the exploration of maturational secular trends in additional skeletal elements.

Although it is difficult to control for all variables that influence skeletal maturation, research projects of this type should aim to make time-depth comparisons of populations from the same geographic region. Socioeconomic status has the most significant impact on maturation rates, and the secular trends we are observing in the American population are likely the result of gradual increases in socioeconomic status and its associated precipitates (i.e. improved healthcare, better nutrition, reduction in infectious diseases, etc.). Many populations worldwide are experiencing similar changes, so we might expect to document analogous trends. Nonetheless, extrapolating data from one population to another is not advisable when comparing skeletal or somatic maturation data. While it is possible that ethnic differences in maturation rates exist, the more significant problem is that all populations are not at the same socioeconomic level. Furthermore, as globalization continues some populations (i.e. those in

developing countries) will undergo a more drastic socioeconomic transition than others. Therefore, the magnitude of biological change may be greater in these populations than in westernized nations. For these reasons, future research concerning secular change in skeletal maturation should control for population differences to ensure that the same socioeconomic climate has been present throughout time.

In terms of sub-adult age estimation, this study used a Bayesian approach to develop standards for the modern American population. Recently, Bayesian statistics have been applied to adult age estimation from documented senescent changes in various skeletal regions (the pubic symphysis, auricular surface, and sternal rib ends). A Bayesian approach is superior to the percentile method typically used for developing skeletal-based age ranges in that the demographic variables of the test population do not affect the age estimates of the target population. Many of the sub-adult aging standards used by forensic anthropologists were developed on archaeological or early 20th century populations. The obvious problem with applying these standards to modern individuals is the issue of secular changes in skeletal maturation; the less obvious problem is the issue of age mimicry. The age-at-death distributions of archaeological samples and anatomical collections are not the same as that of a forensic sample. Consequently, this study used a forensic prior (the Forensic Data Bank) to model the age ranges from the forensic test population (the McCormick Collection). Although age mimicry is a greater concern in adult age estimation where age ranges are wider and have larger error terms, we should seek to avoid its affects across the board. Transition analysis and Bayesian statistics provide robust age estimates that are less susceptible to the influences of

outliers and age mimicry, and this approach should be explored further in establishing age ranges from epiphyseal union.

In terms of sex estimation, this research used both traditional and novel methods to explore sexual dimorphism in the human clavicle. The construction of statistical atlases from 3-D bone models and the use of the PCA-FDR algorithm proved highly effective in exploring sexual dimorphism. This type of analysis is the key to developing an in depth understanding of sex differences in the human skeleton. The PCA-FDR algorithm has the capability of pinpointing areas of highest dimorphism and thereby automatically detecting landmarks and new measurement criteria. Once a bone atlas is developed, it can be used to study the primary shape variation within and between past and present populations. The statistical atlas also offers the ability to reconstruct full bones from fragmentary remains. Moreover, the atlas can improve sexing accuracy from skeletal remains and reduce the number of measurements needed to obtain highly accurate sex estimates. Also, bone atlases would be useful tools for visualizing evolutionary and secular changes in the human skeleton.

The next step in this study is to pinpoint landmarks along the clavicle in the most dimorphic areas and to develop a way to quantify and/or qualify clavicular shape. One way to approach this is to use a method called medial axis representation. In order to represent the medial axis of the clavicle, the medial and lateral endpoints are located, and a line is constructed between these two points. This line is then bisected with a perpendicular plane, thereby representing the middle of the shaft. The contour of the cross-section where the perpendicular plane bisects the shaft is an ellipse. The center of the ellipse is located by using

the points that comprise the ellipse (note: the center of the ellipse is not necessarily near the line connecting the medial and lateral endpoints). Next, two lines are drawn from the center of the ellipse to the medial and lateral endpoints, respectively (Figure 26a). This series of steps is repeated with the two resulting lines, and again with the resulting four lines, etc. so that the bone is divided into smaller and smaller subsections (Figure 26, b-d). The final product is a series of cross-sections, each with its own center point; connecting these center points creates a medial axis that approximates the medial axis of the clavicle and provides a geometric representation of curvature (Figure 27). Measuring the angles between each segment provides a way to quantify the degree of curvature at various points on the shaft. These angles can be used as variables in a discriminant analysis to establish the discriminatory capabilities of clavicular curvature. Medial axis representation may be a viable method for assessing curvature differences in sub-adults, as well. Conducting separate analyses of males and females at one or two year intervals would yield information about the development of sexual dimorphism in clavicular shape and may offer more accurate sub-adult discrimination than traditional measurements.

The digital preservation of skeletal elements using CT or laser scans is an invaluable research tool. The small expense and time that is involved with digital preservation creates limitless possibilities. First, the images are preserved in the event of repatriation or accidental damage or destruction. Second, digital images can be stored on the worldwide web, thereby permitting access to this information by any interested party. In many cases, this would

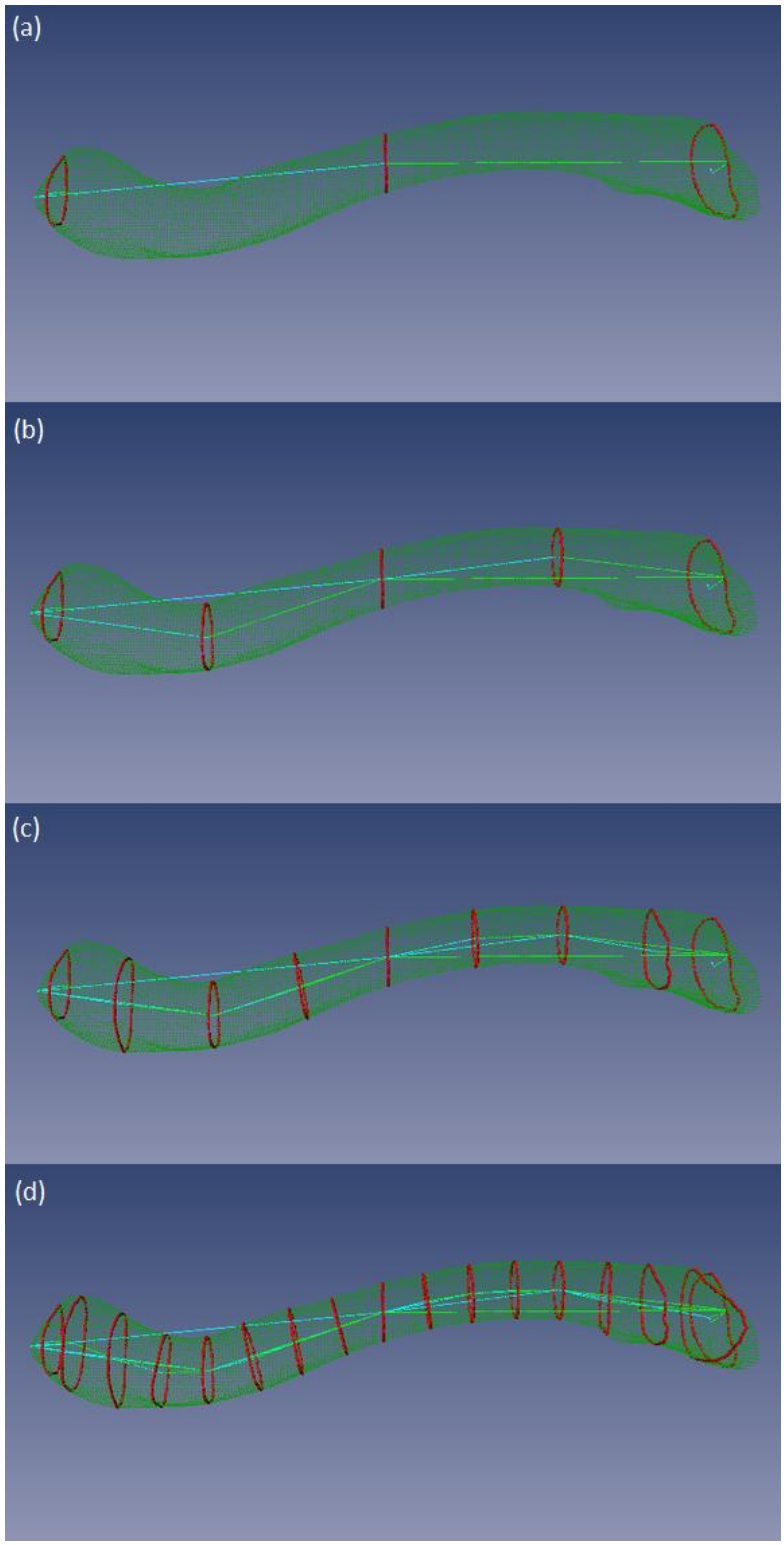

Figure 27. Medial axis representation.

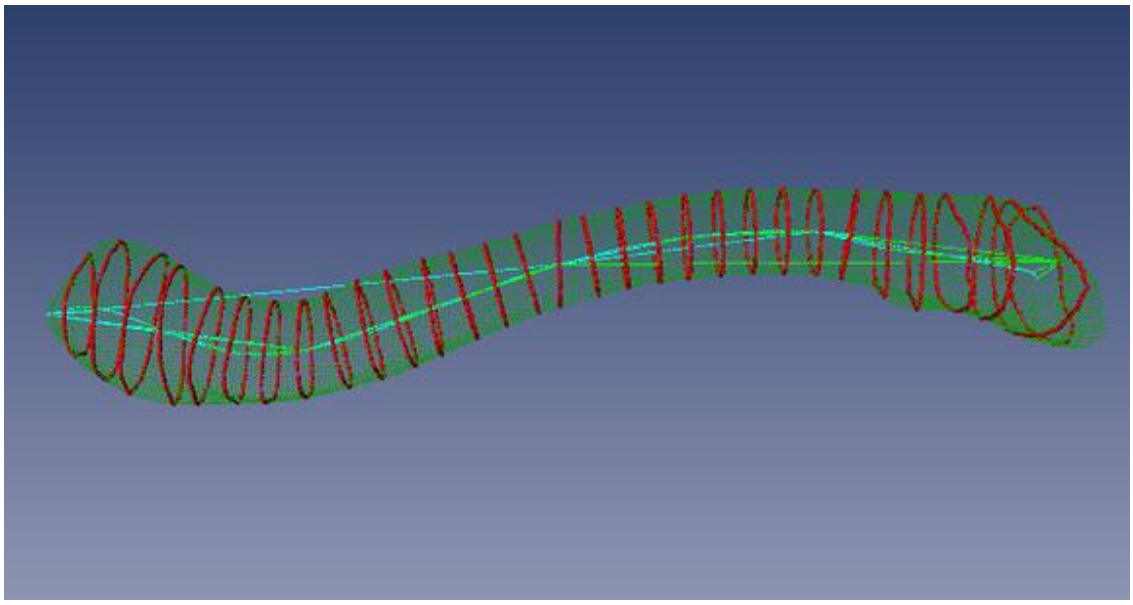

Figure 28. Medial axis representation with 32 ellipses.

eliminate research-associated travel, thus saving time and money that could be invested elsewhere in the research budget. Furthermore, the ability to perform computer-automated measurements reduces the time and observer and/or instrumental error associated with caliper measurements. Finally, digital images provide a non-destructive means of obtaining cross-sectional data for biomechanical studies. In the end, the research possibilities and ease of data access outweigh any extra costs associated with 3-D imaging.

References Cited

Abioye-Kuteyi EA, Ojofeitimi EO, Aina OI, Kio F, Aluko Y, and Mosuro O. 1997. The influence of socioeconomic and nutritional status on menarche in Nigerian school girls. Nutrition & Health (Bicester) 11(3):185-195.

Alberman E, Filakti H, William S, Evans S, and Emanuel I. 1991. Early influences on the secular change in the adult height between the parents and children of the 1958 birth cohort. Ann Hum Biol 18:127-136.

Alsup BK. 2007. Investigation of second, fourth and eighth sternal rib end variation related to age estimation [thesis]. Knoxville (TN): University of Tennessee. 48 p.

Arsuaga J, and Carretero J. 1994. Multivariate analysis of the sexual dimorphism of the hip bone in a modern human population and in early hominids. Am J Phys Anthropol 93:241-257.

Bagga A, and Kulkarni S. 2000. Age at menarche and secular trend in Maharashtrian (Indian) girls. Acta Biologica Szegediensis 44(1-4):53-57.

Bernhard W. 1995. New studies on the secular trend of acceleration of the permanent dentition. Zeitschrift fuer Morphologie und Anthropologie 81(1):111-123.

Black S, and Scheuer L. 1996. Age changes in the clavicle from the early neonatal period to skeletal maturity. International Journal of Osteoarchaeology 6:425-434.

Bodzsar E. 2000. A review of Hungarian studies on growth and physique of children. Acta Biologica Szegediensis 44:139-153.

Bogin B. 1988. Patterns of Human Growth. Cambridge, UK: Cambridge University Press.

Bogin B. 1999. Evolutionary perspective on human growth. Annu Rev Anthropol 28:109-153.

Boldsen JL, Milner GR, Konigsberg LW, and Wood JW. 2002. Transition analysis: a new method for estimating age from skeletons. In: Hoppa RD, and Vaupel JW, editors. Paleodemography: age distributions from skeletal samples. Cambridge: Cambridge University Press. p 73-106

Buikstra J, and Ubelaker DH. 1994. Standards for Data Collection from Human Skeletal Remains. Fayetteville: Arkansas Archaeological Survey.

Calcagno JM. 1981. On the applicability of sexing human skeletal material by discriminant function analysis. Journal of Human Evolution 10(2):189-198.

Cameron N. 1979. The growth of London England UK school children 1904-1966: An analysis of secular trend and intra-county variation. Ann Hum Biol 6(6):505-526.

Cardoso H. 2008a. Epiphyseal union at the innominate and lower limb in a modern Portuguese skeletal sample, and age estimation in adolescent and young adult male and female skeletons. Am J Phys Anthropol 135(2):161-170.

Cardoso HFV. 2008b. Age estimation of adolescent and young adult male and female skeletons II, epiphyseal union at the upper limb and scapular girdle in a modern Portuguese skeletal sample. Am J Phys Anthropol 137(1):97-105.

Census U. 2000. http://www.census.gov/index.html.

Clegg M, and Aiello L. 1999. A comparison of the Nariokotome *Homo erectus* with juveniles from a modern human population. Am J Phys Anthropol(110):81-93.

Cowal L, and Pastor R. 2008. Dimensional variation in the proximal ulna: evaluation of a metric method for sex assessment. Am J Phys Anthropol 135(4):469-478.

Crowder C, and Austin D. 2005. Age ranges of epiphyseal fusion in the distal tibia and fibula of contemporary males and females. J Forensic Sci 50(5):1001-1007.

Daw S. 1970. Age of boys' puberty in Leipzig, 1727-49, as indicated by voice breaking in J. S. Bach's choir members. Hum Biol 42(1):87-89.

Dayal M, and Bidmos M. 2005. Discriminating sex in South African blacks using patella dimensions. J Forensic Sci 50(6):1294-1297.

Decker S, and Hilbelink D. Virtual skull anatomy: Three-dimensional computer modeling and measurement of human cranial anatomy; 2008 February 18-23; Washington, D.C. p 312.

Dreizen S, Spirakis CN, and Stone RE. 1967. A comparison of skeletal growth and maturation in undernourished and well-nourished girls before and after menarche. Journal of Pediatrics 70(2):256-&.

Eveleth P, and Tanner J. 1990. Worldwide variation in human growth. Cambridge: Cambridge University Press.

Farid-Coupal N, Contreras ML, and Castellano HM. 1981. The age at menarche in Carabobo, Venezuela with a note on the secular trend. Ann Hum Biol 8(3):283-288.

France D. 1998. Observational and metric analysis of sex in the skeleton. In: Reichs K, editor. Forensic Osteology. Springfield, IL: Charles C. Thomas. p 163-186.

Franklin D, Freedman L, and Milne N. 2005a. Sexual dimorphism and discriminant function sexing in indigenous South African crania. Homo 55(3):213-218.

Franklin D, Freedman L, and Milne NH-. 2005b. Three-dimensional technology for linear morphological studies: a re-examination of cranial variation in four southern African indigenous populations. HOMO 56:17-34.

Franklin D, Milne N, and Freedman L. 2004. A geometric morphometric study of cranial sexual dimorphism in selected indigenous populations of South Africa. Am J Phys Anthropol 123(38):96.

Fredriks A, Van Burren S, and Burgmeijer R. 2000. Continuing positive secular growth change in the Netherlands 1955-1997. Pediatric Research 47(3):316-323.

Frisancho A, Garn SM, and Ascoli W. 1970. Unequal influence of low dietary intakes on skeletal maturation during childhood and adolescence. Am J of Clinical Nutrition 23(9):1220.

Frutos L. 2002. Determination of sex from the clavicle and scapula in a Guatemalan contemporary rural indigenous population. Am J Forensic Med Pathol 23(3):284-288.

Frutos LR. 2005. Metric determination of sex from the humerus in a Guatemalan forensic sample. Forensic Science International 147(2-3):153-157.

Gonzales G, Crespo-Retes I, and Guerra Garcia R. 1982. Secular change in growth of native children and adolescents at high altitude I. Puno, Peru (3800 meters). Am J Phys Anthropol 58:191-195.

Gonzalez PN, Bernal V, Perez SI, and Barrientos G. 2007. Analysis of dimorphic structures of the human pelvis: its implications for sex estimation in samples without reference collections. Journal of Archaeological Science 34(10):1720-1730.

Gordon-Larsen P, Zemel BS, and Johnston FE. 1997. Secular changes in stature, weight, fatness, overweight, and obesity in urban African American adolescents from the mid-1950's to the mid-1990's. Am J Hum Biol 9(6):675-688.

Graw M, Czarnetzki A, and Haffner H. 1999. The form of the supraorbital margin as a criterion in identification of sex from the skull: investigation based on modern skulls. Am J Phys Anthropol 108:91-96.

Herman-Giddens ME, Slora EJ, Wasserman RC, Bourdony CJ, Bhapkar MV, Koch GG, and Hasemeier CM. 1997. Secondary sexual characteristics and menses in young girls seen in office practice: A study from the Pediatric Research in Office Settings Network. Pediatrics 99(4):505-512.

Himes J. 1984. Early hand-wrist atlas and its implications for secular change in bone age. Ann Hum Biol 11(1):71-75.

Hoshi H, and Kouchi M. 1981. Secular trend of the age of menarche of Japanese girls with special regard to the secular acceleration of the age at peak height velocity Hum Biol 53(4):593-598.

Hsiao T, Chang H, and Liu K. 1996. Sex determination by discriminant function analysis of lateral radiographic cephalometry. J Forensic Sci 41:792-795.

Huen KF, Leung SSF, Lau JTF, Cheung AYK, Leung NK, and Chiu MC. 1997. Secular trend in the sexual maturation of Southern Chinese girls. Acta Paediatrica 86(10):1121-1124.

Humphrey L. 1998. Growth patterns in the modern human skeleton. Am J Phys Anthropol 105(1):57-72.

Hwang J, Shin C, Frongillo E, Shin K, and Jo I. 2003. Secular trend in age at menarche for South Korean women born between 1920 and 1986: the Ansan Study. Ann Hum Biol 30(4):434-442.

Jantz R, and Ousley S. 2005. FORDISC 3.0: Computerized Forensic Discriminant Functions. Knoxville: University of Tennessee.

Ji L, Terazawa K, Tsukamoto T, and Haga K. 1994. Estimation of age from epiphyseal union degrees of the sternal end of the clavicle. Hokkaido Igaku Zasshi 69(1):104-111.

Jit I, and Kulkarni M. 1976. Times of Appearance and Fusion of Epiphysis at the Medial End of the Clavicle. Indian Journal of Medical Research 64(5):773-783.

Kemkes-Grottenthaler A. 2005. Sex determination by discriminant analysis: reliability of patella measurements. Forensic Sci Int 147(2-3):129-133.

Kendall D. 1977. The diffusion of shape. Advances in Applied Probability 9:428-430.

Kim JY, Oh IH, Lee EY, Choi KS, Choe BK, Yoon TY, Lee CG, Moon JS, Shin SH, and Choi JM. 2008. Anthropometric changes in children and adolescents from 1965 to 2005 in Korea. Am J Phys Anthropol 136(2):230-236.

Kimmerle E, Ross A, and Slice D. 2008a. Sexual dimorphism in America: geometric morphometric analysis of the craniofacial region. J Forensic Sci 53(1):54-57.

Kimmerle EH, Konigsberg LW, Jantz RL, and Baraybar JP. 2008b. Analysis of age-at-death estimation through the use of pubic symphyseal data. J Forensic Sci 53(3):558-568.

Klepinger L. 2001. Stature, maturation, variation and secular trends in forensic anthropology. J Forensic Sci 46(4):788-790.

Komar D, and Buikstra J. 2008. Forensic Anthropology: Contemporary Theory and Practice. New York, Oxford: Oxford University Press.

Komlos J, and Breitfelder A. 2007. Are Americans shorter (partly) because they are fatter? A comparison of US and Dutch children's height and BMI values. Ann Hum Biol 34(6):593-606.

Komlos J, and Lauderdale BE. 2007. The mysterious trend in American heights in the 20th century. Ann Hum Biol 34(2):206-215.

Konigsberg L, and Frankenberg S. 1994. Paleodemography: “not quite dead". Evol Anthropol 3:92-106.

Konigsberg L, and Hens S. 1998. Use of ordinal categorical variables in skeletal assessment of sex from the cranium. Am J Phys Anthropol 107:97-112.

Konigsberg LW. October 19, 2007. http://konig.la.utk.edu/nphases2.htm. Nphases2.

Konigsberg LW, Herrmann NP, Wescott DJ, and Kimmerle EH. 2008. Estimation and evidence in forensic anthropology: age-at-death. J Forensic Sci 53(3):541-557.

Kralj-Cercek L. 1956. The influence of food, body build, and social origin on the age at menarche. Hum Biol 28(4):393-406.

Kreitner K, Schweden F, Riepert T, Nafe B, and Thelen M. 1998. Bone age determination based on the study of the medial extremity of the clavicle. Eur Radiol 8:1116-1122.

La Rocherbrochard E. 2000. Age at puberty of girls and boys in France: measurements from a survey on adolescent sexuality. Population: An English Selection 12:51-79.

Laska-Mierzejewska T, Milicer H, and Piechaczek H. 1982. Age at menarche and its secular trend in urban and rural girls in Poland. Ann Hum Biol 9(3):227-234.

Leigh S, and Park P. 1998. Evolution of human growth prolongation. AmJ Phys Anthropol 107:331-350.

Leung S, Lau J, Xu Y, and Tse L. 1996. Secular changes in standing height, sitting height and sexual maturation of Chinese—the Hong Kong growth study, 1993. Ann Hum Biol 23(4):297-306.

Li S, Li Z, Tao L, and Xu K. 2001. Bone age of the medial extremity of the clavicle: Determination on CT data from 695 individuals in northeast of China. Journal of China Medical University 30 (Suppl):34-37.

Lin W, Chen A, Su J, Xiao J, and Ye G. 1992. Secular change in the growth and development of Han children in China. Ann Hum Biol 19(3):249-265.

Liu YX, Wikland KA, and Karlberg J. 2000. New reference for the age at childhood onset of growth and secular trend in the timing of puberty in Swedish. Acta Paediatrica 89(6):637-643.

Low W, Kung L, and Leong J. 1982. Secular trend in the sexual maturation of Chinese girls. Hum Biol 54(3):539-552.

Low W, Kung L, Leong J, and Hsu L. 1981. The secular trend in the growth of southern Chinese girls in Hong Kong. J Morph Anthrop 72(1):77-88.

Mahfouz M, Abdel Fatah E, Merkl B, and Mitchell J. 2008. Automatic and manual methodology for 3-dimensional measurements of distal femoral gender differences and femoral component placement. Journal of Knee Surgery accepted with revisions.

Mahfouz M, Badawi A, Merkl B, Fatah E, Pritchard E, Kessler K, Moore M, and Jantz R. 2007a. Patella sex determination by 3D statistical shape models and nonlinear classifiers. Forensic Sci Int 173(2-3):161-170.

Mahfouz M, Booth R, Argenson J, Merkl B, Abdel Fatah E, and Kuhn M. 2006. Analysis of variation of adult femora using sex-specific statistical atlases. 7th Intl Symp on Computer Methods in Biomechanics and Biomedical Engineering. Antibes, Cote d'Azur, France.

Mahfouz M, Merkl B, Fatah E, Booth RJ, and Argenson J. 2007b. Automatic methods for characterization of sexual dimorphism of adult femora: distal femur. Computer Methods in Biomechanics and Biomedical Engineering 10:477-456.

Malina R. 1979. Secular changes in size and maturity: causes and effects. Society for Research in Child Development 44(3-4):59-102.

Maresh MM. 1972. A 45 year investigation for secular changes in physical maturation. Am J Phys Anthropol 36(1):103-110.

Martin R. 1956. Lehrbuch der Anthropologie. Revised Third Edition, Volume 3. Saller K, editor. Stuttgart: Gustav Fischer Verlag.

Martin R. 1957. Lehrbuch der Anthropologie. Revised Third Edition, Volume 4. Saller K, editor. Stuttgart: Gustav Fischer Verlag.

McCormick W, Stewart J, and Greene H. 1991. Sexing of human clavicles using length and circumference measurements. Am J Forensic Med Pathol 12(2):175-181.

McKern T, and Stewart T. 1957. Skeletal age changes in young American males. Analysed from the standpoint of age identification. Natick, MA: Quartermaster Research and Development Center, Environmental Protection Research Division. 89-97 p.

Meijerman L, Maat GJR, Schulz R, and Schmeling A. 2007. Variables affecting the probability of complete fusion of the medial clavicular epiphysis. International Journal of Legal Medicine 121(6):463-468.

Merkl B, and Mahfouz M. Unsupervised three-dimensional segmentation of medical images using an anatomical bone atlas; 2005a; Singapore.

Merkl B, and Mahfouz M. 2005b. Unsupervised three-dimensional segmentation of medical images using an anatomical bone atlas. 12th International Conference on Biomedical Engineering. Singapore.

Merkl B, Sylvester A, and Mahfouz M. 2007. A three-dimensional shape comparison of AL129-1a and modern human distal femora. Annual Meeting of American Association of Physical Anthropologists. Philadelphia, PA.

Moore-Jansen P, and Jantz R. 1990. Data collection procedures for forensic skeletal material. Knoxville: The University of Tennessee Department of Anthropology and Forensic Anthropology Center.

Moore-Jansen PM, Ousley SD, and Jantz RL. 1993. Data collection procedures for forensic skeletal material. Knoxville: The University of Tennessee Department of Anthropology and Forensic Anthropology Center.

Mora S, Boechat MI, Pietka E, Huang HK, and Gilsanz V. 2001. Skeletal age determinations in children of European and African descent: Applicability of the Greulich and Pyle standards. Pediatric Research 50(5):624-628.

Morrison JA, Barton B, Biro FM, Sprecher DL, Falkner F, and Obarzanek E. 1994. Sexual maturation and obesity in 9-year-old and 10-year-old black and white girls: the national heart and lung and blood institute growth and health study. Journal of Pediatrics 124(6):889-895.

Murata M. 1992. Characteristics of pubertal growth in Japanese children from the standpoint of skeletal growth. Acta Paediatr Jpn 34:236-242.

Murata M. 1997. Population-specific reference values for bone age. Acta Paediatr Suppl 423:113-114.

Murphy A. 1994. Sex determination of prehistoric New Zealand Polynesian clavicles. J Archaeol 16:85-91.

Murphy A. 2002. Articular surfaces of the pectoral girdle: sex assessment of prehistoric New Zealand Polynesian skeletal remains. Forensic Sci Int 125(2-3):134-136.

Ogata S, and Uhthoff H. 1990. The early development and ossification of the human clavicle--an embryologic study. Acta Orthop Scand 61(4):330-334.

Parsons F. 1916. On the Proportions and Characteristics of the Modern English Clavicle. J Anat 51(1):71-93.

Patriquin ML, Loth SR, and Steyn M. 2003. Sexually dimorphic pelvic morphology in South African whites and blacks. Homo-Journal of Comparative Human Biology 53(3):255-262.

Penin X, Berge C, and Baylac M. 2002. Ontogenetic Study of the Skull in Modern Humans and the Common Chimpanzees: Neotenic Hypothesis Reconsidered With a Tridimensional Procrustes Analysis. Am J Phys Anthropol 118:50-62.

Phenice TW. 1969. A newly developed visual method of sexing os pubis. Am J Phys Anthropol 30(2):297-&.

Prado C. 1984. Secular change in menarche in women in Madrid. Ann Hum Biol 11(2):165-166.

Pryor J. 1928. Difference in the ossification of the male and female skeleton. J Anat 1928 Jul;62(Pt 4):499-506 62(4):499-506.

Richman E, Michel M, Schulter-Ellis F, and Corruccini R. 1979. Determination of sex by discriminant function analysis of postcranial skeletal measurements. J Forensic Sci 24(1):159-167.

Rimpela AH, and Rimpela MK. 1993. Towards an equal distribution of health? Socioeconomic and regional differences of the secular trend of the age of menarche in Finland from 1979 to 1989. Acta Paediatrica 82(1):87-90.

Roche A. 1979. Secular trends in human growth, maturation, and development. Monogr Soc Res Child Dev 44(3-4):1-120.

Roche A, Roberts J, and Hamill P. 1974. Skeletal maturity of children 6-11 years, United States. Vital Health Stat 11(140):1-62.

Roche A, Roberts J, and Hamill P. 1978. Skeletal maturity of youths 12--17 years racial, geographic area, and socioeconomic differentials. United States, 1966-1970. Vital Health Stat 11(167):1-98.

Rogers T. 2005. Determining the sex of human remains through cranial morphology. J Forensic Sci 50(3):493-500.

Rosas A, and Bastir M. 2002. Thin-plate spline analysis of allometry and sexual dimorphism in the human craniofacial complex. Am J Phys Anthropol 117:236-245.

Schaefer MC, and Black SM. 2005. Comparison of ages of epiphyseal union in north American and Bosnian skeletal material. J Forensic Sci 50(4):777-784.

Scheuer L, and Black S. 2000. Developmental juvenile osteology. New York: Academic Press.

Schmeling A, Olze A, Reisinger W, and Geserick G. 2005. Forensic age estimation and ethnicity. Legal Medicine 7(2):134-137.

Schmeling A, Reisinger W, Loreck D, Vendura K, Markus W, and Geserick G. 2000. Effects of ethnicity on skeletal maturation: consequences for forensic age estimations. Int J Legal Med 113:253-258.

Schmeling A, Schulz R, Danner B, and Rösing F. 2006. The impact of economic progress and modernization in medicine on the ossification of hand and wrist. Int J Leg Med 120:121-126.

Schmeling A, Schulz R, Reisinger W, Muehler M, Wernecke K-D, and Geserick G. 2004. Studies on the time frame for ossification of the medial clavicular epiphyseal cartilage in conventional radiography. International Journal of Legal Medicine 118(1):5-8.

Schmittbuhl M, Le Minor JM, Schaaf A, and Mangin P. 2002. The human mandible in lateral view: elliptical Fourier descriptors of the outline and their morphological analysis. Annals of Anatomy-Anatomischer Anzeiger 184(2):199-207.

Schmittbuhl M, Le Minor JM, Taroni E, and Mangin P. 2001. Sexual dimorphism of the human mandible: demonstration by elliptical Fourier analysis. International Journal of Legal Medicine 115(2):100-101.

Schulz R, Muehler M, Mutze S, Schmidt S, Reisinger W, and Schmeling A. 2005. Studies on the time frame for ossification of the medial epiphysis of the clavicle as revealed by CT scans. International Journal of Legal Medicine 119(3):142-145.

Schulz R, Mühler M, Reisinger W, Schmidt S, and Schmeling A. 2008. Radiographic staging of ossification of the medial clavicular epiphysis. Int J Legal Med 122:55-58.

Simmons K. 1944. The Brush Foundation study of child growth and development, II: Physical growth and development. Monogr Soc Res Child Devel 9(1).

So L, and Yen P. 1990. Secular trend in skeletal maturation in Southern Chinese girls in Hong Kong. Zeitschrift fuer Morphologie und Anthropologie 78(2):145-154.

So L, and Yen P. 1992. Secular trend of menarcheal age in Southern Chinese girls. Zeitschrift fuer Morphologie und Anthropologie 79(1):21-24.

Spradley M, and Jantz R. Skull vs. postcranial elements in sex determination; 2003; Chicago, IL. American Academy of Forensic Sciences. p 239.

Spradley M, Jantz R, Robinson A, and Peccerelli B. 2008. Demographic change and forensic identification: problems in metric identification of Hispanic skeletons. J Forensic Sci 53(1):21-28.

Steel F. 1966. Further observations on the osteometric discriminant function. The human clavicle. Am J Phys Anthropol 25(3):319-322.

Stevenson P. 1924. Age order of epiphyseal union in man. Am J Phys Anthropol 7:53-93.

Steyn M, and Iscan M. 1998. Sexual dimorphism in the crania and mandibles of South African whites. Forensic Sci Int 98:9-16.

Steyn M, and Iscan MY. 1997. Sex determination from the femur and tibia in South African whites. Forensic Science International 90(1-2):111-119.

Steyn M, and Iscan MY. 1999. Osteometric variation in the humerus: sexual dimorphism in South Africans. Forensic Science International 106(2):77-85.

Sylvester A, Merkl B, and Mahfouz M. 2007. Reconstructing the AL 288-1 femur using three-dimensional computer models. Annual Meeting of American Association of Physical Anthropologists. Philadelphia, PA.

Szilvassy V. 1977. Altersschaetzung an der sternalen gelenkflaechen der schluesselbeine. Beitr Gerichtl Med 35:343-345.

Tanaka H, Lestrel PE, Uetake T, Kato S, and Ohtsuki F. 2000. Sex differences in proximal humeral outline shape: Elliptical Fourier functions. J Forensic Sci 45(2):292-302.

Tanner J, Oshman D, Babbage F, and Healy M. 1997. Tanner-Whitehouse bone age reference values for North American children. Journal of Pediatrics 131(1 Part 1):34-40.

Taranger J, Engström I, Lichtenstein H, and Svennberg- Redegren I. 1976. Somatic pubertal development. Acta Paediatr Scand Suppl 258:121-135.

Taylor J, and Twomey L. 1984. Sexual dimorphism in human vertebral body shape. J Anat 138(2):281-286.

Thieme F, and Schull W. 1957. Sex determination from the skeleton. Hum Biol 29(3):242-273.

Todd T. 1937. Atlas of skeletal maturation. St. Louis: Mosby.

Todd T, and D'Errico J. 1928. The clavicular epiphyses. Am J Anat 41:25-50.

Uytterschaut H. 1986. Sexual dimorphism in human skulls. A comparison of sexual dimorphism in different populations. Hum Evol 1(3):243-250.

Van Gerven D. 1972. The contribution of size and shape variation to patterns of sexual dimorphism of the human femur. Am J Phys Anthropol 37(1):49-60.

Vercauteren M, and Susanne C. 1985. The secular trend of height and menarche in Belgium: are there any signs of a future stop. European Journal of Pediatrics 144(4):306-309.

Vignolo M, Naselli A, Di Battista E, Mostert M, and Aicardi G. 1988. Growth and development in simple obesity. Eur J Pediatr 147:242-244.

Walker P. 2008. Sexing skulls using discriminant function analysis of visually assessed traits. Am J Phys Anthropol 136(1):39-50.

Washburn S. 1948. Sex differences in the pubic bone. Am J Phys Anthropol 6(2):199-208.

Washburn S. 1949. Sex differences in the pubic bone of Bantu and Bushmen. Am J Phys Anthropol 7(3):425-432.

Wattigney WA, Srinivasan SR, Chen W, Greenlund KJ, and Berenson GS. 1999. Secular trend of earlier onset of menarche with increasing obesity in black and white girls: the Bogalusa heart study. Ethnicity & Disease 9(2):181-189.

Webb P, and Suchey J. 1985. Epiphyseal union of the anterior iliac crest and medial clavicle in a modern multiracial sample of American males and females. Am J Phys Anthropol 68(4):457-466.

Wellens R, Malina R, Beunen G, and Lefevre J. 1990. Age at menarche in Flemish girls: current status and secular change in the 20th century. Ann Hum Biol 17(2):145-152.

Williams B, and Rogers T. 2006. Evaluating the accuracy and precision of cranial morphological traits for sex determination. J Forensic Sci 51:729-735.

Wong GWK, Leung SSF, Law WY, Yeung VTF, Lau JTF, and Yeung WKY. 1996. Secular trend in the sexual maturation of southern Chinese boys. Acta Paediatrica 85(5):621-622.

Appendices

Appendix A

Linear Discriminant Functions for African Americans:

(1.62537***MaxLength***) – (4.20468***Sagittal***) – (27.40155***Vertical***) + (23.23126***MaxMidShaftDiam***) + (23.06525***MinMidShaftDiam***) + (1.64699***MaxLateralShaftDiam***) + (3.89571***MinLateralShaftDiam***) - (2.71267***MinMedialShaftDiam***) + (0.46371***MaxMedialShaftDiam***) - 49.0729

(1.62946***MaxLength***) + (5.69315***MaxMidShaftDiam***) + (6.57740***MinMidShaftDiam***) + (2.31137***MaxLateralShaftDiam***) – (0.29912***MinLateralShaftDiam***) - 44.45691

(1.56322***MaxLength***) + (6.99691***MaxMidShaftDiam***) + (8.84981***MinMidShaftDiam***) - 42.53862

(1.49322***MaxLength***)+ (14.32338***Sagittal***) + (2.57876***Vertical***) - 43.94224

(1.23495***MinMedialShaftDiam***)+ (1.14245***MaxMedialShaftDiam***) - 4.55263

(2.01025***MaxLength***)+ (3.62114***MaxLateralShaftDiam***) + (0.90788***MinLateralShaftDiam***) - 39.97778

(12.83068***MaxMidShaftDiam***)+ (6.68862***MinMidShaftDiam***) - 24.66408

(16.87301***Sagittal***)+ (3.65575***Vertical***) - 25.79818

(3.91442***MaxLateralShaftDiam***)+ (0.46918***MinLateralShaftDiam***) - 9.70075

Linear Discriminant Functions for European Americans:

(1.76363***MaxLength***)+ (7.15023***Sagittal***) + (6.97469***Vertical***) + (1.20028***MaxMidShaftDiam***) + (2.38921***MinMidShaftDiam***) + (1.02573***MaxLateralShaftDiam***) + (3.22045***MinLateralShaftDiam***) – (0.84286***MinMedialShaftDiam***) - (0.79395***MaxMedialShaftDiam***) - 49.55772

(1.81213***MaxLength***) + (7.97156***MaxMidShaftDiam***) + (6.75265***MinMidShaftDiam***) + (0.81928***MaxLateralShaftDiam***) + (3.02102***MinLateralShaftDiam***) - 49.77214

(1.75465***MaxLength***) + (9.69371***MaxMidShaftDiam***) + (8.33896***MinMidShaftDiam***) - 47.29460

(1.74066***MaxLength***) + (9.41708***Sagittal***) + (10.15932***Vertical***) - 48.39731

(2.07187***MaxLength***) + (1.98119***MaxLateralShaftDiam***) + (5.24167***MinLateralShaftDiam***) - 41.83148

(11.47731***Sagittal***) + (10.61207***Vertical***) - 25.34004

(11.28001***MaxMidShaftDiam***) + (8.97450***MinMidShaftDiam***) - 23.77848

(1.57114***MinMedialShaftDiam***) + (1.24635***MaxMedialShaftDiam***) - 5.41107

(2.85147***MaxLateralShaftDiam***) + (4.14743***MinLateralShaftDiam***) - 11.75199

Appendix B

Means and standard deviations of computer-automated measurements (in centimeters).

MEASUREMENT	European Am. Males (n=949)		European Am. Females (n=292)		African Am. Males (n=43)		African Am. Females (n=13)	
	Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
MAXIMUM LENGTH	15.96	0.87	14.26	0.87	16.13	0.95	14.44	0.49
SAGITTAL MIDSHAFT DIAM.	1.29	0.13	1.06	0.10	1.38	0.13	1.17	0.11
VERTICAL MIDSHAFT DIAM.	1.17	0.13	0.96	0.11	1.20	0.15	1.02	0.09
MAX. MIDSHAFT DIAMETER	1.40	0.13	1.15	0.11	1.46	0.15	1.25	0.09
MIN. MIDSHAFT DIAMETER	1.09	0.12	0.90	0.10	1.13	0.13	0.97	0.08
MIN. MEDIAL DIAMETER	2.20	0.35	1.98	0.34	2.17	0.45	1.92	0.39
MAX. MEDIAL DIAMETER	1.83	0.29	1.62	0.26	1.87	0.30	1.68	0.27
MAX. LATERAL DIAMETER	2.67	0.38	2.18	0.32	2.58	0.42	2.01	0.28
MIN. LATERAL DIAMETER	1.22	0.20	0.99	0.16	1.22	0.22	0.99	0.18

Appendix C

Principal Component	Cumulative Eigenvalue Percentage
1	81.4632
2	85.1737
3	88.0707
4	89.9897
5	91.3051
6	92.187
7	92.8703
8	93.3843
9	93.8178
10	94.1759
11	94.4864
12	94.7557
13	95.0008
14	95.2342
15	95.4388
16	95.623
17	95.8006
18	95.9594
19	96.1104
20	96.2471
21	96.3746
22	96.4985
23	96.6185
24	96.7275
25	96.8326
26	96.9288
27	97.0239
28	97.115
29	97.2037
30	97.2875
31	97.3658
32	97.4416
33	97.515
34	97.5854
35	97.6533
36	97.7187

37	97.7829
38	97.8458
39	97.9068
40	97.9673
41	98.0262
42	98.0834
43	98.137
44	98.1894
45	98.2406
46	98.2897
47	98.3377
48	98.3844
49	98.4293
50	98.4726
51	98.5156
52	98.5569
53	98.5968
54	98.6362
55	98.6743
56	98.7118
57	98.7475
58	98.7827
59	98.8166
60	98.8494
61	98.8818
62	98.9137
63	98.9449
64	98.9756
65	99.0053
66	99.0346
67	99.0629
68	99.0901
69	99.1169
70	99.143
71	99.1686
72	99.1936
73	99.2182
74	99.2424
75	99.2664
76	99.2898
77	99.3126

78	99.3351
79	99.3572
80	99.3787
81	99.3993
82	99.4199
83	99.4404
84	99.4604
85	99.48
86	99.4994
87	99.5182

Vita

Natalie Shirley was born Natalie Langley in Kinder, Louisiana on May 16, 1975. She is the daughter of Dennis Langley and Rita Davenport. She received her high school diploma from the Louisiana School for Math, Science, and the Arts in 1993. Natalie attended Tulane University for two years and then transferred to Louisiana State University to complete her Bachelor of Arts in German and Anthropology in December 1998. During her undergraduate career, Natalie spent one year as a direct exchange student at the Universitaet Bonn in Germany. Natalie entered the Louisiana State University Master's program in Anthropology in January 1999 and graduated in May 2001 with a Master of Arts. She began her doctoral work at the University of Tennessee in August 2003 and received her Doctor of Philosophy in May 2009.