To the Graduate Council:

I am submitting herewith a thesis written by Jaymelee Kim entitled "Weaver's Method Reassessed: the Relationship between Age and the Estimation of Sex in Juveniles Using the Auricular Surface." I have examined the final electronic copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Arts, with a major in Anthropology.

Murray K. Marks, Major Professor

We have read this thesis
and recommend its acceptance:

Tricia Redeker Hepner

Walter Klippel

Accepted for the Council:

Carolyn R. Hodges, Vice Provost and
Dean of the Graduate School

(Original signatures are on file with official student records)

WEAVER'S METHOD REASSESSED: THE RELATIONSHIP BETWEEN AGE AND THE ESTIMATION OF SEX IN JUVENILES USING THE AURICULAR SURFACE

A Thesis
Presented for the
Master of Arts Degree
The University of Tennessee, Knoxville

Jaymelee Kim
May 2008

Acknowledgments

I wish to thank all those who have helped me to complete my Master of Arts degree in Biological Anthropology. First, I would like to thank my advisor, Dr. Murray Marks, for his advisement, encouragement and support. I would also like to thank my committee members, Dr. Tricia Redeker Hepner and Dr. Walter Klippel for their continuous feedback and honesty. I would like to thank Dr. Richard Jantz for his review of my statistical methods, and the University of Tennessee's Department of Anthropology for presenting me with the William Bass Endowment to fund my research travels. I would also like to thank Kevin Hufnagl for his assistance in completing the statistics portion of my research, and Suzanne Johnson for her helpful suggestions. Finally, I would like to thank the Cleveland Museum of Natural History for access to the Hamann-Todd Collection, Washington University for the use of the Trotter Fetal Collection, and Dr. David Hunt for his consultation and access to skeletal materials at the Smithsonian Institution.

Abstract

The purpose of this research was to test Weaver's method of sex estimation using the auricular surface on neonatal skeletal remains of varying ages. The age of the individual was used to examine whether or not there existed a stronger correlation between age and elevation than sex and elevation. By assessing these relationships, it could be determined if Weaver's method is more appropriate for individuals of certain ages more than individuals of other ages. Juvenile skeletons (n = 167) ranging in age from fourteen weeks post-conception to twenty years of age were assessed from two skeletal collections, and the results were analyzed using 2x2 contingency tables and logistic regressions. Weaver's method yielded a 45.16% accuracy for determining sex in females and a 56.14% accuracy for estimating sex in males. These accuracies, close to chance, do not compare to Weaver's original study which generated a 75% accuracy for determining sex of female fetuses and a 92% accuracy for determining sex in males fetuses. When age was incorporated into the analyses, unlike suggestions made by previous researchers, there was no correlation between it and auricular surface elevation.

Table of Contents

Chapter One: Introduction 1

Assessing Sex in Adult Skeletal Remains 2

Assessing Sex in Juvenile Skeletal Remains 12

Chapter Two: Materials & Methods 18

Materials 18

Methods 19

Visual Assessment. 19

Statistical Assessment 21

Chapter Three: Results 27

Chapter Four: Discussion 38

Research Possibilities 46

Conclusion 51

References Cited 53

Appendix A 57

Vita 63

List of Tables

3.1 Contingency Table of Sex versus Elevation 27

3.2 Contingency Table of Sex versus Elevation in the Hamann-Todd Collection 28

3.3 Contingency Table of Sex versus Elevation in the Trotter Fetal Collection 28

3.4 Contingency Table of Collection versus Correctly Sexed Individuals 29

3.5 Contingency Table of Ancestry versus Correctly Sexed Individuals 29

3.6 Logistic Regression Results 35

List of Figures

2.1 Elevated Auricular Surface 20

2.2 Unelevated Auricular Surface 20

2.3 Example of a 2x2 Contingency Table 22

2.4 Example of Logistic Regression Output 25

4.1 Histogram of Correctly Sexed Individuals by Age 47

4.2 Histogram of Incorrectly Sexed Individuals by Age 47

Chapter One: Introduction

In 2004, the National Center for Juvenile Justice reported that 1,551 individuals, composing 10.4% of homicide victims that year, were younger than eighteen years old (2006). The following year, the Federal Bureau of Investigation's 2005 Uniform Crime Report stated that 12% of all homicide victims were younger than nineteen years old. Unfortunately the instances of crimes against children are not limited to the United States; rather, incidents of children as victims can be seen worldwide. With forensic anthropologists' growing role in investigations concerning homicide, human rights violations, and war crimes, continuous efforts on improving identification methods remain a priority. Complications in victim identification surround each recovered set of juvenile human skeletal remains. In order to begin the process of identifying the children and resolving the circumstances surrounding their deaths, a thorough biological profile and examination of the skeleton must be completed. Regrettably, due to the fragile and underdeveloped nature of the juvenile skeleton, establishing a biological profile can prove difficult.

To further the complex nature of evaluating immature skeletal remains, few methods used to determine the sex of young victims are considered reliable. Due to the scarcity of large subadult skeletal collections of known sex and the limited representation of each age group, developing adequate methods proves difficult. Additional hindrances stem from biological anthropology's nature of being a young discipline that has not yet established standards for juvenile sex estimation or parameters for acceptable accuracies. Despite these obvious obstacles, researchers have attempted to develop both metric and morphological methodologies that can be used in evaluating the sex of juveniles. Hoping

to contribute to the development of sex estimation standards, research presented here examines Weaver's method for determining sex in fetal and neonatal remains when applied to juveniles of varying ages. Using an array of ages allows for a complex assessment of Weaver's theory and illustrates the relationships between auricular surface elevation, age and sex of an individual.

Assessing Sex in Adult Skeletal Remains

Researchers such as Thomson (1899) began to look at the pelvic girdle as a tool for determining sex in the adult skeleton in the late nineteenth century. Pelvic bones, historically known as useful elements in sex estimation in adults, include two innominates and the sacrum. Innominates, or os coxae, begin as three separate bones that fuse together during growth and development: the ilium, ischium and pubis. Washburn (1948) focused his 1940s research on the study of the ischium and pubis in macaques as tools for sex estimation; he noted that the pubic bones in females were notably longer than those in males. In order to compare samples, he analyzed the ischium-pubis index and hoped to apply this method of study to adult human skeletal remains as well. Traditional, observable sex characteristics could be derived from this elongation of the pubic bone: formation of the subpubic angle, the triangularity of the obturator foramen, and the height of the pubic symphysis (Washburn, 1948). Examining 300 human individuals from the Hamann-Todd Collection, Washburn (1948) calculated the ischium-pubis indices in hopes of finding the primary cause of morphometric sex differences in the pelvis. Encouragingly, he discovered that this measurement correctly determined sex

90% of the time if the ancestral group was known. Ancestry needed to be identified since he found an overlap of indices for black females and white males.

Though Washburn's method for sex determination was accurate, it had its shortcomings. Only with knowledge of an individual's heritage was accuracy maintained, and Washburn emphasized that using ischium-pubis indices for sex estimation remained limited to adult skeletons. Up to age one, no differences in pubic bone length between males and females were found (Reynolds, 1945; Washburn, 1948). Moreover, the average difference between pre-pubescent males and females was minute at about 2-3% (Washburn, 1948). Because of this small difference between the sexes in juveniles, Washburn's method proved to be inappropriate for sex estimation in children.

Continuing his research with Hanna into the early 1950s, they looked at an Eskimo population in the U.S. National Museum collection and analyzed the pelves in order to determine sex (Hanna and Washburn, 1953; Stewart, 1954). Instead of using only the ischium-pubis index, the sciatic notch angle and the interiliac index were considered as well. Separating the population into males and females, Hanna and Washburn concluded that an ischium-pubis index of ninety-two or greater and a sciatic notch angle of sixty degrees or more indicated that the individual was female. Correspondingly, an ischium-pubis index of less than ninety-two and a sciatic notch angle of less than seventy degrees was male (Hanna and Washburn, 1953; Stewart, 1954). Since the sex of the skeletal remains in the population was unknown, Hanna and Washburn compared their results to those of Hrdlicka's when he determined sex using cranial characteristics. Of the 173 individuals, only eight did not match Hrdlicka's assigned sex (Hanna and Washburn, 1953).

Following their success, Hanna and Washburn encouraged their method to be taught in place of traditional sex estimation techniques. They believed that their method was easily replicable and could be taught to beginners who did not have past experience in sex estimation (Stewart, 1954). Stewart criticized this perspective. Rather, Stewart supported a multi-factorial approach, and the use of various techniques in combination with one another. Like his contemporary researcher, Phenice, Stewart emphasized the use of nonmetric, morphological traits that had high accuracies and were quicker than methods that require careful measurement (Stewart, 1954).

Washburn's work and Stewart's commentary were followed by the criticism of Phenice in the 1960s. While Phenice agreed that the pelvis was an ideal skeletal element for sex estimation, he also stated that Washburn's technique required both time and fully intact skeletal material (Phenice, 1969). Often, especially in historical contexts, skeletal remains in varying states of preservation are studied under strict time constraints, making Washburn's method lacking. Knowing this, Phenice proposed the use of three characteristics of the pelvis which were both clear to the inexperienced and more expedient than Washburn's ischium-pubis measurements and calculations (Phenice, 1969). Phenice suggested that researchers look at the ventral surface of the pubis; in females there was a clear ventral arc and in males there was only a slight ridge. Additionally, Phenice stated that on the medial aspect of the ischio-pubic ramus, females had a pronounced ridge, and males had a broad, flat surface. The third trait that was to be used in combination with the other two could be seen from the dorsal view of the pubis and ischio-pubic ramus. Apparently, females had a concave ischio-pubic ramus while those of males were convex (Phenice, 1969).

Testing his methodology on 275 individuals from the Terry Skeletal Collection, Phenice found that only eleven were sexed incorrectly. More importantly, his technique worked equally well on both black and white individuals. The accuracy of sex estimation in blacks was 95% and the accuracy in whites was 95.62% (Phenice, 1969). Disappointingly, like Washburn's method, Phenice emphasized that his technique was only effective with adult skeletal remains and was not appropriate for juvenile remains "…for there are indications that the ventral arc and the subpubic concavity are not well developed until the female has reached about 20 years of age" (Phenice 1969: 300).

Throughout the late 1970s and the 1980s, anthropologists continued their focus on the use of discriminant function analyses and other metric methods of sex estimation (Black, 1978; Dittrick and Suchey, 1986; MacLaughlin and Bruce, 1985). Black (1978) studied 114 prehistoric skeletons in Ohio from the Libben site, a late Woodland ossuary. Theorizing that because males laid cortical bone faster than females, they would have a greater long bone circumference than females, Black decided to use this hypothesis to try to estimate sex in his sample. First, Black applied Phenice's successful method of sex estimation using the three traits of the pubic bone in order to assign sex to his sample. Because the individuals did not have known age and sex, he tested to discern whether his method had results that were consistent with results from the Phenice method. Then he categorized individuals into sex categories based on circumference of the femur at midshaft; the midpoint of eighty-one millimeters which fell between the female mean and the male mean was used as the deciding factor. After completing his analyses, Black (1978) found that the mean circumference of the males was significantly different than that of females using a t-test. Furthermore, he completed several regression analyses and

discovered that the slope for male circumference and age differed from zero, but the females' did not. With this, it must be noted that the R^2 value for the male regression equation was .12, and this low value implied that age only explained a small portion of femur circumference variability. Consulting other discriminant function analyses, Black concluded that his method had an 85% consistency with Phenice's method. Similar to many other researchers in this field, Black emphasized the uncertainty of the effects that age, population and historical timeframe had on the success of his sex estimation method (Black, 1978).

Subsequent to the work of Black, using prehistoric skeletal remains from central California, Dittrick and Suchey (1986) compiled eighteen measurements of 207 individuals in an effort to establish a highly accurate multivariable method for sex estimation. Sex for both their Early Horizon sample and their Middle/Late Horizon sample was established using Phenice's three traits of the os pubis as well as pubic bone width and shape of the ventral rampart. Ready to begin their research, they measured nine aspects of the femur: maximum length, physiological length, maximum diameter of the head, subtrochanteric anterior-posterior diamenter, subtrochanteric medio-lateral diameter, midshaft circumference, midshaft anterior-posterior diameter, midshaft medio-lateral diameter, and bicondylar width. Similar measurements were taken on the humerus: maximum length, physiological length, vertical diameter of the head, transverse diameter of the head, least circumference, midshaft maximum diameter, midshaft minimum diameter, bicondylar width, and deltoid tuberosity diameter. Afterwards, the authors completed univariate analysis and multivariate analyses. Surprisingly, they found

that the use of multiple variables produced similar results to those of several of the variables analyzed independently (Dittrick and Suchey, 1986).

For the Early Horizon sample, the diameter of the humeral head correctly sexed 85.7% of the individuals and the physiological length of the femur correctly sexed 85% of the individuals. Middle and Late Horizon samples had similar results; the maximum diameter of the femur correctly categorized 90.6%, and the humeral head diameter correctly determined sex in 90.3% (Dittrick and Suchey, 1986). Regardless of these promising results, the accuracies did not supercede the traditional examinations of the pelvis. Even Dittrick and Suchey (1986: 9) deferred to the pelvis stating that only "in individuals lacking the os pubis, the next best area would seem to be one of the femoral end dimensions or head of the humerus." Once more it is important to note that many of these methods would not be practical to test in juveniles because the proximal ends of their long bones have not fully developed, nor fused.

Studies like this continued in the 1980s as MacLaughlin and Bruce (1985) attempted to identify a univariate method for estimating sex in skeletal remains. Using a prehistoric Scottish population, the authors examined the morphometric cranial and pelvic traits developed by Krogman (1973) and Stewart (1979) to classify the remains into fifty-two males and twelve females. By testing the maximum anterior-posterior diameter of femur inferior to the midshaft point, they found that males' femoral diameters averaged to 29.8 $\pm$ .4 millimeters, and the females' averaged 24.2 $\pm$ .7 millimeters. Of the forty-three individuals analyzed, four individuals' metric-based sex assignments were inconsistent with the morphologic-based sex assignment, yielding a 90.6% accuracy if one assumes that the morphological assessment was correct. Along with the assumption

of correct morphological assessment also came the recognition of an extremely skewed sex ratio. MacLaughlin and Bruce (1985) attributed these results to differential burial practices. Others believed that skewed sex ratios in skeletal samples could be attributed to females disintegrating faster than males because of lower bone density (Walker et al., 1988). Contrastingly, Weiss (1972) showed that researchers tend to have bias in sex estimation, assigning male to more individuals than female. MacLaughlin and Bruce (1985) insisted that the latter was not the case in their research, and that their sex assignments were sound. Given that this was true, then their accuracies indicated a relationship between femoral diameter and sex in adult skeletons. Again, as with their contemporaries, they did not investigate comparable methodologies that could be used on juvenile skeletal remains.

The trend of diverging from the pelvis and turning to other effective means of sex estimation in the human skeleton continued into the mid-nineties with the research of Loth and Henneberg (1996). In contrast to many studies of the previous decade (Black, 1978; Dittrick and Suchey, 1986; MacLaughlin and Bruce, 1985), Loth and Henneberg turned their research back to the skull, and like many anthropologists before them (Giles, 1964; Iscan and Ding, 1995), chose to re-examine the mandible for use in sex estimation. Because they believed that skeletal elements used for sex estimation should have morphological sex differences, be robust in the archaeological record and be consistent across populations and timeframes, the mandible was the likely candidate.

Loth and Henneberg (1996) examined 300 mandibles from the Dart Collection at Witwatersrand Medical School and the University of Pretoria Anatomy Department, all of which were of known sex and at least twenty years old. Primarily of South African

heritage, only nine were white, and 100 others that displayed abnormal bony growths or were missing more than two teeth were parceled out into a pathological subsample. Hypothesizing that the posterior border of the mandibular ramus varied in shape between males and females, they tested the Dart Collection and 247 individuals of European, African and Native American decent who were housed at the Smithsonian Institute in Washington, D.C., most of which were in the Terry Collection. Interestingly, the researchers discovered that in the collections in question, the posterior border of the mandibular ramus did appear to vary in shape between males and females. Results showed that the portion of the ascending ramus which is level with the occlusal surface flared in males, while it remained straight in females. Acknowledging that if flexure was observed in females, then it occurred higher on the ramus near the mandibular condyles, Loth and Henneberg carefully scored the mandibles (Loth and Henneberg, 1996). Creating a +1 to -1 scale, with +1 indicating bilateral flexure and -1 indicating no flexure, each side of the mandible was scored separately. When the examinations were finished, Loth and Henneberg (1996) analyzed their results using 2x2 contingency tables.

Ramus shape, as established by Loth and Henneberg, was 99.1% effective for discerning males and 98.8% effective for females with an overall accuracy of 99% in the normative South African sample. Pathological remains which were tested separately from the normative group had a lowered overall accuracy of 91%. Differences in accuracy between the two African groups, normative and pathological, were significant with a p-value of $<.05$. Optimistically, an overall accuracy for all mandibles was 94.2% (Loth and Henneberg, 1996).

Loth and Henneberg's apparent success with sex estimation using the mandible, quickly came under critique by Koski (1996) who stated that "on living white children and young adults of female sex showed that a flexure between the dorsal surfaces of the ramus and condylar process is a very common feature indeed," (Koski, 1996: 545). Moreover, he disagreed with Loth and Henneberg's reasoning on the physiological causes of differences in flexure between the sexes (Koski, 1996). Unwavering, Loth and Henneberg responded to Koski's critique, emphasizing that they only recommended using ramus flexure at the occlusal plane for sex estimation. Reviewing their use of multiple populations, the researchers noted the differences in accuracy that occurred between groups, and they encouraged further testing of their method on different samples and populations (Loth and Henneberg, 1998).

Their contemporaries took that advice, but when testing of their method continued other researchers did not have such successful results. Rather, Donnelly et al. (1998) only earned a 62.5% accuracy in their study. Ninety-six mandibles were tested blindly; sixteen were of known sex from the University of Tennessee's Forensic Collection and the other eighty mandibles were sexed based on pelvic examinations. All of the known samples were over twenty-five years old, and the Native Americans were said to be adults since epiphyseal fusion was complete and the third molars had erupted. Using two observers, the first observer was told the scoring scale without knowing which trait was considered masculine or feminine. They were then asked to score one side of a mandible at a time while the other side remained concealed by a sleeve. The scoring was done over four days with two practice days followed by one day of scoring for each side. A second observer then scored each side of the mandibles on separate days. When these

assessments were complete, Donnelly et al. (1998) found an accuracy of 80.4% for males and 50.0% for females.

Two years later, Hill (2000) published the results of her test of Loth and Henneberg's method. Taking 158 mandibles from the Hamann-Todd Collection, Hill had sixty-five American Blacks and ninety-three American Whites. Scoring each side as either +1, 0, or -1, the total score for a single mandible ranged from +2 to -2. Sex estimation for the sample had a success rate of 79.1% overall with males touting a 91.3% accuracy, but females revealing a 56.4% accuracy. Like Donnelly et al., Hill found that males had a better sex estimation accuracy than females. This trend was seen in both the American Blacks and American Whites with American Black males having an accuracy of 94.7% and females having an accuracy of 51.9%. Similarly, American Whites had an accuracy of 89.2% for males and 60.7% for females (Hill, 2000). Though Loth and Henneberg's technique had varying results, unlike the work many of her fellow researchers in the arena of adult sex estimation methodologies, their technique showed promise in identifying sex in both adults and juveniles.

Over the decades, researchers have shifted their ideologies regarding sex estimation in the human skeleton. Research began with primary interest in the skull and pelvis, and many methods of analyzing these components developed. Following from this, trends shifted from metric to nonmetric observations, and other skeletal elements were scrutinized for their effectiveness in separating males from females. Unfortunately, while many of these methodologies appeared to have successful results on the adult skeleton, they were inappropriate for the juvenile skeleton. Discouragingly, this inability to transfer popular adult techniques to subadult remains led many anthropologists to

believe that juvenile remains could not be sexed, and they reiterated that "…the secondary characteristics do not manifest themselves until puberty; thus it is impossible to judge the remains of children and adolescents because the means available relate to adult traits" (Bass, 1995: 25). Whether or not sex estimation in subadults is possible or not continues to be a debate among anthropologists and researchers alike, but it does not change the need to test whether or not techniques work in children. Acknowledging this, many have attempted to identify such methodologies over the years.

Assessing Sex in Juvenile Skeletal Remains

Since the inception of physical anthropology, biological anthropologists and bioarchaeologists have sought to tease information from skeletal remains which may be used in endeavors such as identifying victims or establishing paleodemographics. Knowing this, estimating the sex of skeletal material has remained a key component of understanding ancient populations as well as discovering the identities of unknown forensic victims. As it has been discussed, the estimation of sex in the adult or mature skeleton is generally accepted as an accomplishable task through the use of common metric and nonmetric methodologies. However, anthropologists have yet to come to a consensus on sex estimation in juvenile and fetal remains; some even have taken the stance that sex estimation in prepubescent individuals is impossible.

A lack of confidence in the ability to estimate sex of subadult populations has not eradicated the need to develop a demographic profile of juvenile remains found in both ancient and modern contexts. Because of this, since the nineteenth century, researchers have been studying juvenile growth and development in an attempt to define sexually

dimorphic, osseus characteristics. Skeletal components considered sexually dimorphic in adults have become the focus of study in the immature skeleton with an assumption that there is a predisposition toward these traits that is observable before puberty and even in infancy. While studies that encompassed many elements of the body developed, including the skull, mandible, and cervical vertebrae, the ilium has remained a promising indicator of sexual dimorphism (Fehling, 1877; Thompson, 1899).

Regardless of the fact that methodologies for sex determination in juveniles have extended back for several hundred years, or the acknowledgment that physical anthropologists have been preoccupied with cranial characteristics, the skulls of juveniles have not been as deeply studied for sexual dimorphism as those of adults. Only after the mid-1900s have anthropologists analyzed cranial, mandibular and vertebral components of the subadult skeleton that may be sexually dimorphic (Bailit and Hunt, 1964). Instead, initial studies into the realm of sex estimation in subadults began with a focus on what is currently recognized as the primary skeletal component in assessing the sex of adult skeletal remains; the pelvic girdle (Fehling, 1876; Washburn, 1948).

Exploring pelvic morphometric features for subadult sex estimation grew popular with Fehling's study of 130 fetal pelves that led him to the conclusion that after the fourth gestational month observable sex differences had begun to develop (Fehling, 1876; Reynolds, 1945). This conclusion was further supported by Thomson when he conducted a similar study on a much smaller sample of eight fetal pelves (Thomson, 1899; Reynolds, 1945). Both Fehling and Thomson inspired research into subadult sex estimation and presented the idea as an accomplishable task for biological anthropologists to undertake.

Shortly following their success, a surge of research contradicted their findings and many researchers, such as Yamamura and Morton (Reynolds, 1945), openly disagreed; they concluded that the pelvis could not be used as a source of juvenile sex estimation. The open debate led Reynolds to execute his own research concerning the development of the pelvic girdle from birth to one year. Using radiographs, Reynolds took a series of measurements six times on ninety-five children between birth and twelve months of age. From the radiographs Reynolds (1945) found that females had greater bi-ischial breadth, sciatic notch breadth, inlet breadth, and pubic length while males experienced greater pelvic height, ischio-iliac space and ilium breadth. He also noted that female measurements were more variable than male measurements which later researchers also believed to be true (Reynolds, 1945; Weaver, 1980). Reynolds' findings provided the basis for further research into using features of the innominates as estimators of sex.

One researcher who determined that these correlations were relevant was Barbara Boucher who in 1957 proceeded to look for sexual dimorphism in fetal pelves not only between males and females, but between populations (Boucher, 1957; Choi and Trotter, 1970). Her research encompassed thirty-three American White and ninety-six American Black fetuses, and she found that in both populations the subpubic angle was larger in females than in males. Interestingly enough, she noted that though the sciatic notch was larger in American Black females, it was not larger in American White females when compared to males (Boucher, 1957).

By the mid-1950s researchers had delved into a strong debate about the presence or absence of sexually dimorphic features of the subadult skeleton. Yet, many of their primary foci were on the pelvic girdle and metric analyses of breadths or indices of

various features. The topic had even expanded its exploration of juvenile sexual dimorphism into the realm of population differences and the effects of ancestry. Expansion of variable usage and reevaluation of colleagues' work opened the field of prepubescent sexual dimorphism to include dental development and growth factors (Hunt and Gleiser, 1955).

In 1955 Hunt and Gleiser observed the possibility of juvenile sex estimation based on dental and bone development. Examining radiographs, they measured root completion, compared dental age to age based on bone development, and even created different standards for males and females. According to Hunt and Gleiser (1955), if the two ages corresponded for male standards, but diverged by female standards, then the subject was male, and vice versa. At ages two, five and eight the sexing accuracies were 73%, 76% and 81% respectively (Hunt and Gleiser, 1955). Hoping to recreate his apparent success, Edward Hunt then further tested this method in 1964 with Howard Bailit. Together they discovered that out of a sample of fifty children and using the development of mandibular teeth, the accuracy was 58%, little more than chance, minimizing the utility of his method (Bailit and Hunt, 1964).

With this as an obvious deviation from the traditional studies of the innominates, focus by professionals in the field still often turned to the pelves as a source of sexual dimorphism. This is further attested to by the 1980 publications of David Weaver; however, unlike his predecessors he also turned his research toward the possibility of nonmetric sexually dimorphic characteristics in the ilia. In Weaver's research, he completed both metric analyses and a nonmetric analysis of the fetal innominates, and the latter produced patterned results (Weaver, 1980; Hunt, 1990). Like researchers before

him, Weaver measured sciatic notch width and depth, ilium anterior and posterior length and iliac height and width. In addition to this, Weaver visually observed the auricular surface elevation on the ilia. According to Weaver, the only significant results were produced by the auricular surface analysis in which elevated edges of the auricular surface indicated that the subject was female, while unelevated edges indicated that it was male. Weaver's (1980) study generated an accuracy of 92% for males and 75% for females within his fetal sample.

Since then, Weaver's nonmetric method, which has opened the door for the study of nonmetric traits, has been tested and retested. David Hunt used Weaver's auricular surface analysis on a separate sample in order to verify that Weaver's findings were not a population-specific trait. In 1990, Hunt used a sample of 275 Arikara subadult ilia recovered from three separate sites (Hunt, 1990). Using children of unknown age and sex whose epiphyseal fusion had not begun, placing them at less than twelve years of age, Hunt's results determined that the "raised trait is 5.6 times more prevalent than the nonraised trait in newborns" (Hunt 1990, 884). Because of this vast deviation from a 1:1 sex distribution, Hunt (1990) theorized that surface elevation correlated to age, and he argued that though Weaver had success with fetal remains, the method may not appropriate for other subadult samples.

Hunt's analyses of Weaver's method was then expounded upon by Mittler and Sheridan in their 1992 research on a sample of fifty-eight juveniles that ranged in age from birth to eighteen years old. Like Hunt, they sought to retest Weaver's nonmetric method on their subadult population in an effort to determine its accuracy for ancient population studies as well as forensic applications (Mittler and Sheridan, 1992). Contrary

to Hunt's vastly skewed results, Weaver's method correctly determined the sex of 74.1% of the juveniles. Interestingly enough, his method proved to be almost 30% more effective in distinguishing males than females with an 85% accuracy for male sex determination and a 58% accuracy for female sex estimation. In addition to this variation in accuracy, Mittler and Sheridan (1992), like Hunt, suggested that Weaver's method is least effective in young, prepubescent juveniles and that if the subadult is over ten years in age and has an elevated auricular surface, then there is a 99% probability that it is a female. These findings, like Hunt's, alluded to an increasing differentiation between the sexes as children age.

Following Mittler and Sheridan's study, research expanded to other skeletal components as biological anthropologists began to study features such as chin prominence, ramus shape, or dental arcade width in subadults (Loth and Henneberg, 1996; Schutkowski, 1993), but others maintained a steady focus on characteristics of the innominates (Holcomb and Konigsberg, 1995). With the hope of testing Weaver's hypothesis that the auricular surface elevation can distinguish between male and female juvenile remains, and in an effort to establish the relationship between age and auricular elevation, in this research the long-standing tradition of analyzing the pelves to test methods of sex estimation in juveniles has been continued.

Chapter Two: Materials & Methods

Materials

In order to accumulate a large sample composed of skeletal remains that varied in age from infant to adult, two skeletal collections were used. Due to the sensitive nature of a child's death, juvenile remains are not predominantly found in skeletal collections of known age and sex. Typically individuals who die at a young age are not donated to collections, but are buried by surviving relatives; finding a collection with a significant number of fetal remains became a primary concern.

After contacting several curators, it became apparent that the Trotter Fetal Collection, which is housed at Washington University's School of Medicine in St. Louis, Missouri, met the needs of the project. Gathered in the 1950s and 1960s from American Black and American White stillbirths, this collection was comprised of the ilia, ischia, pubii and femora of 147 individuals of known age and ancestry. Of these, based on the mother's date of last menstruation, eighty-nine individuals also had an estimated age-at-death and eighty-two of those had known sex. Composed of the necessary osteological elements, with known ages and sexes, the Trotter Fetal Collection proved to be a promising sample for the examination and application of Weaver's sex estimation method.

Aspiring to analyze Weaver's method on juveniles that had surpassed the fetal and infant ages, a second skeletal collection was needed. The Hamann-Todd Collection met the criteria. Amassed in the early 1900s, remains that otherwise would have went to Potter's Field, were turned over to Western Reserve Medical School. Thomas Wingate Todd became the director of the Hamann Museum of Comparative Anatomy and

Anthropology, where the skeletal remains from the unidentified cadavers were stored. Todd continued the collection of both adult and subadult skeletal remains of known age and sex. The skeletal remains are now part of the Hamann-Todd Collection, housed at the Cleveland Museum of Natural History in Cleveland, Ohio (Wilton, 1939). From this collection, eighty-five subadults from one to twenty years of known age and sex were used.

Methods

Visual Assessment. The data from each individual was recorded in a Microsoft Excel spreadsheet (Appendix A). Identified only by the curation number, the left innominate of each individual was analyzed, and if the left element was missing, the right was used. A "Y" was typed into the spreadsheet to indicate that the auricular surface was elevated and an "N" was entered to indicate an unelevated auricular surface. After completing the analysis of a collection, the first ten percent of elements that had been studied were revisited. This reassessment was done to reduce intraobserver error which would diminish naturally with repeated analysis as a researcher becomes more accustomed to the observation process. If an observation from the reassessment differed from the initial assessment, then the second observation was used.

When assessing the Trotter Fetal Collection, the left ilium of an individual was removed from its protective case, and only the identification number on the case was noted; the known age and sex were disregarded at this time. Following this, the ilium was placed on its dorsal side, and an auricular surface received an elevated score (Figure 2.1 and 2.2) if "the sacro-iliac surface was elevated from the ilium along its entire length



Figure 2.1: An elevated auricular surface. (Photo from Trotter Fetal Collection, housed at Washington University in St. Louis, Missouri.)



Figure 2.2: An unelevated auricular surface. (Photo from Trotter Fetal Collection, housed at Washington University in St. Louis, Missouri.)

and along both the anterior and posterior edges of the sacro-iliac surface" (Weaver 1980). If any portion of the rim appeared to be unelevated, then it received a score of unelevated. After each ilium was studied, it was placed into its case and the next individual was examined. Some of the individuals were lacking the innominates, and thus they were excluded from the data collection process.

The process for examination of the juveniles in the Hamann-Todd Collection was the same as that for the Trotter Fetal Collection. Each individual was housed in a removable drawer at the Cleveland Museum of Natural History. Left innominates from several individuals were removed from their respective drawers at once, making them only identifiable by their curation numbers. They were examined, and if the entire auricular border appeared elevated, then they were noted as such, but if it appeared unelevated, that was recorded before the innominates were returned to their storage drawers. Upon the completion of the juvenile examination, the first ten percent were also re-evaluated to reduce error.

Statistical Assessment. All observations were taken from the Microsoft Excel spreadsheet and entered into NCSS. Because the infant and fetal ages were recorded in weeks since conception, the Trotter Fetal Collection ages did not correspond to the units of the Hamann-Todd Collection. In order to rectify this, thirty-eight weeks, or the human gestation period since conception, was added to each of the ages which were recorded in years. Furthermore, the ages recorded in weeks were converted to their applicable age in years. These two changes converted all ages to years, making age zero equivalent to conception. In the instances in which a second observation had been taken, only the second observation was used for the statistical analysis. With the categorical variables of

elevation and sex, and the continuous variable of age, contingency tables and logistic regressions were used to study the variables and their relationships with one another.

Contingency tables were used for each independent variable and the dependent variable. To assess the significance of the results found in the contingency tables, Chi-squared tests are used. A Chi-squared p-value of less than .05 indicates a significant result, further supporting that a relationship did in fact exist between the variables. The use of crosstabulation or 2x2 contingency tables can be illustrated in an example using eye color and handedness. By asking 100 individuals if they are left handed or right handed, and noting their respective eye color, a contingency table can be used to analyze the data. If forty-two, blue-eyed people are left handed, and eight are right handed, while ten brown-eyed people are left handed and forty are right handed, then the contingency table would look like the one in Figure 2.3. With this, the significance can be seen by the

The FREQ Procedure
Table of Handedness by Eye_Color

Handedness	Eye_Color		
Frequency	Blue	Brown	Total
Left	42	10	52
Right	8	40	48
Total	50	50	100

Figure 2.3: Example of a 2x2 Contingency Table

Chi-squared p-value. The contingency table in Figure 2.3 has a Chi-square p-value of <.0001. Being considerably less than .05, the p-value indicates a high correlation between eye color and handedness in the sample. Contingency tables were used in the same manner to look for correlations between sex and elevation or correct identification of sex using Weaver's method and skeletal collection.

In addition to utilizing contingency tables to establish relationships, logistic regressions were used in NCSS to study the correlations between the remainder of the variables. A logistic regression model needed to be used when the dependent variable was binary, or categorical such as male and female or elevated and unelevated. The output of logistic regression models indicates if there is a relationship between a dependent variable and one or more independent variables, and it shows the direction and strength of the relationship.

While contingency tables only show the significance of one variable at a time, logistic regressions illustrate the interactions of independent variables and their influence on the dependent variable. A logistic regression model provides a beta (β) value, or regression coefficient, for the y-intercept as well as each independent variable. The regression coefficient, which is either positive or negative, implies the direction of the relationship between the independent variable and the dependent variable for the model. Moreover, if the respective p-value is less than alpha, then the relationship is statistically significant and a correlation exists. If an independent variable is not statistically significant, then it can be determined that it has no effect on the dependent variable if the other independent variables are held constant. This suggests that any other independent

variables in the model have more influence on the dependent variable than the statistically insignificant variable. In addition, if the regression coefficient is positive and significant for an independent variable, then it indicates that as the value of that independent variable increases by one unit, the likelihood of the dependent variable having the value of one increases. Likewise, if the regression coefficient is negative and significant for an independent variable, then it implies that as the independent variable increases by one unit, the likelihood of the dependent variable having the value of the one decreases. The value of one is the same as the reference group, while zero indicates the opposite of the reference group. Furthermore, if the p-value is greater than alpha, then the coefficients are not statistically significant and no relationship, or only a weak relationship, exists.

Moreover, the output of a logistic regression function also produces an odds ratio. The odds ratio illustrates the strength of relationship of the independent variable and the dependent variable in question. Assuming all other independent variables are constant and that there is a one unit change in the independent variable in question, the odds ratio provides a percentage. This is the odds of the action illustrated by the model occurring with each one unit increase in the independent variable. However, unlike a probability which is based between one and zero, odds ratios are based around the value of one. For instance, an odds ratio greater than one shows an increase in the relationship, and an odds ratio of less than one indicates a decrease in the relationship.

Finally, the output also provides an R^2 value. The R^2 value is derived from the difference between the observations and the expected values. The expected values are the values derived from the logistic regression function generated by the logistic

regression model. From this, the R^2 value shows how close the actual observations are to the values from the logistic model. The closer R^2 is to one, the better the data fits the model and vice versa.

Figure 2.4 displays an example of the NCSS output when running the logistic regression model: Y = Elevation, X1 = Adjusted_Age, and X2 = Known Sex. These values are based on an unelevated auricular surface. The positive regression coefficient for Beta1: Adjusted_Age, indicates that as the individual gets older, the probability of having an unelevated auricular surface increases. Moreover, the odds ratio states that each year that age increases, the chances of being unelevated increase by 103%.

Run Summary Section

Parameter	Value	Parameter	Value
Dependent Variable	X_L__Elevated	Rows Processed	166
Reference Group	N	Rows Used	166
Number of Groups	2	Rows for Validation	0
Frequency Variable	None	Rows X's Missing	0
Numeric Ind. Variables	1	Rows Freq Miss. or 0	0
Categorical Ind. Variables	1	Rows Prediction Only	0
Final Log Likelihood	-112.52065	Unique Row Patterns	89
Model R-Squared	0.03422	Sum of Frequencies	166
Actual Convergence	5.014311E-12	Likelihood Iterations	4
Target Convergence	0.000001	Maximum Iterations	20
Model D.F.	3	Max Like Message	Normal Completion
Model	Age__yrs_|Sex		

rameter Significance Tests Section (Reference Group: X_L__Elevated = N)

Parameter	Regression Coefficient (B or Beta)	Standard Error	Wald Z-Value (Beta=0)	Wald Prob Level	Odds Ratio
B0: Intercept	-0.42551	0.29478	-1.443	0.14888	0.65344
B1: Age__yrs_	0.03104	0.01927	1.611	0.10716	1.03153
B2: (Sex="M")	-0.07797	0.32535	-0.240	0.81060	0.92499

Figure 2.4: Example of Logistic Regression Output - the dependent variable "Auricular Surface Elevation" and independent variables "Age" and "Sex".

Unfortunately, the p-value of .10716 is greater than .05, and thus is not statistically significant. The other independent variable, Beta2: Sex, indicates that males, when compared to females, are less likely to have an unelevated auricular surface. The odds ratio shows that the chances of males being unelevated decreases because it is only 92.5% of the original value. Once again, the p-value is greater than .05 at .81, indicating that there is no statistically significant correlation between elevation and sex. Finally the model R^2 value is .03, and this represents how well the model fits the data. With an R^2 value so far from one, it indicates that the model does not fit the data well. This further corroborates the lack of relationships between age and elevation and sex and elevation.

Chapter Three: Results

All individuals of unknown age were eliminated, leaving 167 juvenile skeletons from two collections to examine for this study. Upon reexamination of the first 10% of each sample to check for intraobserver error, only one individual from each of the collections was assigned a different sex. Satisfied that consistency in sex assignment for both collections occurred, 2x2 contingency tables (Tables 3.1-3.5) were created to assess the relationship between sex and elevation as well as to determine whether or not there was a significant difference in success between the two skeletal collections. Afterwards, logistic regression functions were used to determine the type of relationship, if any, between elevation and age. Employing the use of contingency tables, the known sex of the individuals was first placed against the independent variable "auricular surface elevation," based on Weaver's definition. These tables (Tables 3.1-3.3) show the number of individuals that are elevated and unelevated by sex. For each sex and elevation there

Table 3.1: Contingency table of total sample - Sex versus Elevation.

Sex	Elevation		
	No	Yes	Total
Female	34	28	62
Expected #	34.90	27.10	
% of Total	20.36	16.77	37.13
Row %	54.84	45.16	
Column %	36.17	38.36	
Male	60	45	105
Expected #	59.10	45.90	
% of Total	35.93	26.95	62.88
Row %	57.14	42.86	
Column %	63.83	61.64	
Total	94	73	167
	56.29	43.72	100.00
Chi-square	.0841	P-value	.7718

Table 3.2: Contingency table of the Hamann-Todd Collection – Sex versus Elevation.

Sex	Elevation		
	No	Yes	Total
Female	16	18	34
Expected #	18.40	15.60	
% of Total	18.82	21.18	40.00
Row %	47.06	52.94	
Column %	34.78	46.15	
Male	30	21	51
Expected #	27.60	23.40	
% of Total	35.29	24.71	60.00
Row %	58.82	41.18	
Column %	65.22	53.85	
Total	46	39	85
	54.12	45.88	100.00
Chi-square	1.137	P-value	.2863

Table 3.3: Contingency table of the Trotter Collection – Sex versus Elevation.

Sex	Elevation		
	No	Yes	Total
Female	18	10	28
Expected #	16.39	11.61	
% of Total	21.95	12.20	34.15
Row %	64.29	35.71	
Column %	37.50	29.41	
Male	30	24	54
Expected #	31.61	22.39	
% of Total	36.59	29.27	65.85
Row %	55.56	44.44	
Column %	62.50	70.59	
Total	48	34	82
	58.54	41.46	100.00
Chi-square	.5790	P-value	.4467

Table 3.4: Contingency table of the Collection versus Correctly Identified

Collection	Correctly Identified		
	No	Yes	Total
Hamann-Todd	48	37	85
Expected #	45.30	39.70	
% of Total	28.74	22.16	50.90
Row %	56.47	43.53	
Column %	53.93	47.44	
Trotter Fetal	41	41	82
Expected #	43.70	38.30	
% of Total	24.55	24.55	49.10
Row %	50.00	50.00	
Column %	46.07	52.56	
Total	89	78	167
	53.29	46.71	100.00
Chi-square	.7020	P-value	.4021

Table 3.5: Contingency table of Ancestry versus Correctly Identified

Ancestry	Correctly Identified		
	No	Yes	Total
Black	65	72	137
Expected #	63.99	73.01	
% of Total	38.92	43.11	82.04
Row %	47.45	52.55	
Column %	83.33	80.90	
White	13	17	30
Expected #	14.01	15.99	
% of Total	7.78	10.18	17.96
Row %	43.33	56.67	
Column %	16.67	19.10	
Total	78	89	167
	46.71	53.29	100.00
Chi-square	.1672	P-value	.6826

was the observed count of individuals, the expected number of individuals, the percent in that category of the total, the row percentage and the column percentage. Each contingency table also provided the associated Chi-squared p-value to determine if there was a significant relationship between the variables in question. First, Table 3.1 is comprised of the total sample looking at sex against the variable elevation. From this, it can be seen that 54.84% of the females were unelevated, and 45.16% of them were elevated. Comparatively, 57.14% of the males did not have raised auricular surfaces while 42.86% did have raised surfaces. Though it appeared that there was no correlation between known sex and auricular elevation, the Chi-squared p-value of .7718 confirmed the absence of a statistically significant relationship between the two variables.

Table 3.2 showed the eighty-five skeletons that were examined in the Hamann-Todd Collection. Here sixteen individuals, or 47.06% of the females had unraised auricular surfaces, whereas eighteen, or 52.94% of women had raised surfaces. Coinciding with this, thirty, or 58.82% percent of males, had unelevated auricular surfaces. Additionally, it can be seen that twenty-one, or 41.18%, of the males had raised auricular surfaces. With forty percent of the sample composed of female individuals the distribution between raised and unraised is close to even. While unraised auricular surfaces were more prevalent in males, the Chi-squared p-value of .2863 indicated that there was no significant relationship in the Hamann-Todd Collection between sex and elevation.

When examining the results for the individuals of the Trotter Collection (Table 3.3), of the eighty-two individuals, eighteen of the females, or 64.29%, had unelevated auricular surfaces, and only ten females, or 35.71%, had raised surfaces. Similarly,

55.56% of males were unelevated on the auricular border, and 44.44% of the males had raised auricular borders. Interestingly, though there were almost twice as many males as females in the sample, both sexes had more individuals with unelevated auricular surfaces. Once again, the Chi-squared p-value further supported the absence of a correlation between elevation and sex in the Trotter Fetal sample. With a value of .4467, far greater than .05, there was no relationship between elevation and sex in the sample.

The next contingency table (Table 3.4) analyzed collection versus the number of individuals who were correctly identified. The results of the table should have indicated whether there was a significant difference in the efficacy of Weaver's method between the two collections. Half of the Trotter Fetal Collection was correctly assigned sex using Weaver's method and half was not. Similarly, the Hamann-Todd Collection was close to equally divided; although there were more correctly identified in that collection than the Trotter Fetal Collection. Not unlike the previous tables, this contingency table produced a Chi-squared p-value of .4021, indicating that there was no strong significant difference between the collections in the efficacy of Weaver's technique.

Finally, as a matter of interest, a crosstabulation of ancestry against the number of correctly sexed individuals (Table 3.5) was completed. As with the basic premise of contingency tables, this table should have illustrated whether or not there was a significant relationship between ancestry and the number of individuals that were correctly sexed. This calculation illustrated that 47.45% of the American Black individuals were incorrectly sexed using Weaver's auricular surface elevation, and the remaining 52.55% were assigned the correct sex. Similarly, 43.33% of the skeletons of American White ancestry were incorrectly identified, while 56.67% of them were

correctly assigned their reported sex. Being close to equally divided regardless of ancestry suggests that there is no significant relationship between the ability to correctly estimate sex using Weaver's method and ancestry. This conclusion was further supported by the high Chi-squared p-value of .6826, which established that there was no statistically significant relationship.

In order to assess possible relationships between age and elevation, multiple logistic regressions were completed. Chiefly, logistic regressions were deemed appropriate because the variables sex and ancestry were binary. Additionally, the continuous variable of age can be assessed with logistic regression functions. Logistic regression allows the researcher in this instance to control for the independent variable of sex, while assessing the influence that age has on the dependent variable of auricular surface elevation.

Normally, when performing a regression analysis, a Pearson correlation table is created in order to test for multicollinearity. Multicollinearity occurs in a multiple regression model when there is a significant amount of linear correlation amongst two or more variables. If multicollinearity was present, it could be identified in a Pearson correlation table which showed the significance of the relationship between two or more variables. However, because there was only one continuous variable, there were no other independent variables that could have had a statistically significant relationship with age, thus no testing for multicollinearity was required.

After finding no significant correlations in the contingency tables, logistic regressions were executed. Initially, the logistic regression model, Model 1, was created for all males from both skeletal collections, using elevation as the dependent variable,

and the adjusted age as the independent variable. Correspondingly, for Model 2 all females were analyzed in the same manner as the males. Following this, a regression model, Model 3, was created using sex as the dependent variable with elevation and age as the independent variables. Ultimately, Model 4, like Model 1 and 2, has elevation as the dependent variable and age and sex as the independent variables, as well as an interaction variable of age with sex. Summaries of the results from these models can be found in Table 3.6

Modeling a dependent variable (Y) of Elevation for only the males in the sample, Adjusted_Age was made the independent variable (X1). For this first model, the reference group was Elevation = N, and stated that researchers should consider an unelevated surface to be the default, or one value, in the situation described by the regression (Table 3.6). The regression coefficient for this model was .0216 and positive. Because the regression coefficient was positive, it showed that as a male grew one year older, he was more likely to have an unelevated auricular surface. Furthermore, the odds ratio of 1.0218 illustrated a 2% percent increase of those odds. Unfortunately, this correlation was not statistically significant with a p-value of .3329. If the p-value was less than .05, the correlation would have been considered statistically significant. In contrast, the Model R^2 would have been close to the value of one if the data had fit the model well. However, the regression established that the data did not fit the model with an R^2 value of .0256.

Subsequently, in Model 2, the dependent variable (Y) of elevation was plotted with the adjusted age (X1) for all of the females. The regression coefficient (Table 3.6) β_1 was .0500 with a reference group of Elevation = N. The positive regression coefficient

indicated that with each year that females got older, it was more likely that their auricular surface was unelevated. More importantly, the odds ratio of approximately 1.05 showed that with each year, a female's chances of having an unelevated surface increased by 5%. Despite this relationship, the correlation proved weak with a p-value of about .14, and confirmed that the relationship was not statistically significant. Lastly, the model R^2 value of .0535 supported the fact that the data did not fit the model well. In Model 2, no significant correlation between elevation and age in females existed.

After completing these two initial logistic regressions, a third model used a dependent variable (Y) for sex, as this would be a likely scenario in a forensic anthropology or bioarchaeological setting. The independent variables, X1 and X2, were Adjusted_Age and Elevation, respectively. With Elevation = Y, the regression had a raised auricular surface as a frame of reference. Likewise, the reference group of Sex = F for the dependent variable referred to the probability of the sex being female. When the NCSS software completed the regression model, the coefficient for Adjusted_Age was -.0012, and it showed that while controlling for elevation, as one grew older, the chances of being female diminished. The corresponding odds ratio of .9989 could be described as the chance of being female reduced by .001% with each year of increased age. With this, the p-value for X1, Adjusted_Age, was .9521, and demonstrated that there was no statistically significant relationship between the variable and the model presented.

Additionally, the second independent variable X2, or Elevation, had a regression coefficient of -.0911. A negative regression coefficient revealed that controlling for age, when the surface was elevated, the chances of being female were reduced.

Table 3.6: Table of logistic regression results

Model	Dependent Variable (Y)	IndependentVariable	Reference Group	Regression Coefficient (β)	P-value	Odds Ratio	Model R^2
1	Elevation	Adjusted_ Age	Elevation = N	.0216	.3329	1.0218	.0256
2	Elevation	Adjusted_ Age	Elevation = N	.0500	.1396	1.0513	.0535
3	Sex	Adjusted_ Age	Sex = F	-.0012	.9521	.9989	.0008
		Elevation=Y		-.0911	.7795	.9130	
4	Elevation	Adjusted_ Age	Elevation = N	.0500	.1396	1.0513	.0415
		Sex=M		.1338	.7702	1.1432	
		Adjusted_ Age*Sex		-.0285	.4823	.9719	

Like X1, the odds ratio of .9130 which was less than one, indicated that the chances were reduced to .91% of the original value. Additionally, the p-value of .7795 proved that the variables' relationship with the model was not statistically significant. Finally, the model R^2 of .0008 demonstrated that the model did not fit the data well, which corresponded to the insignificant p-values.

Lastly, Model 4 was put into a logistic regression function. In this fourth model, the probability of having an elevated auricular surface (Y) was modeled by three independent variables. Adjusted_Age and Sex acted as the independent variables X1 and X2, while a third variable was an interaction variable Adjusted_Age*Sex, or X1X2. The reference for Sex was male and that meant the regression function described a situation in which males were more or less likely to behave in a certain manner when compared to females. With this, the reference group for the logistic regression model was Elevation = N, like that of Model 2. The first independent variable X1 had a regression coefficient of .0500, which implied that as a person grew a year older, the odds of having an unelevated auricular surface increased. Likewise, the odds ratio of 1.0513 indicated a 5.13% increase in the odds of that situation occurring. Lamentably, the p-value of .1396 was not significant and showed that this variable did not have statistically recognizable influence on the model.

Then the second independent variable X2, or Sex, had a regression coefficient of .1338. As with the independent variable X1, a positive regression coefficient indicated that at any given age, males were more likely than females to have an unelevated auricular border. In support of this, the odds ratio for X2 of 1.1432, illustrated a 14.32%

increase in the chance of males having unelevated auricular surfaces. Though this corresponded with Weaver's statement, the p-value was not significant at .7702 and again showed that this variable did not fit the model at a statistically significant level. The last variable, X1X2 shows how sex and age interact with each other in the model. Unlike X1 and X2, the regression coefficient for X1X2 is negative at -.0285. With this, the provided odds ratio of .9719 was for any given age. To calculate the odds ratio of males having an unelevated auricular surface with each increasing year of age the product of the odds ratios of X1 and X1X2 was computed. This totaled 1.0218.

The odds ratios described an increase in odds of 2.18%, that for each year of age, males would have an unelevated auricular surface. Interestingly, the odds of females having an unelevated auricular surface with each year of age also increased, but by 5.13% instead. Finally, adjusting for any given age, the odds of an unelevated auricular border occurring were 14% greater for males than females. Contrary to Weaver's method, males appeared to be more likely to have an elevated auricular surface longer than females, and the odds of being unelevated increased faster for females. Regrettably, like the other models and variables, the p-value is .4823, and it was not significant. The poor fit of the model to the data also showed in the model R^2 value of .0415, which was not close to one.

Chapter Four: Discussion

Determining sex in the skeletal remains of juveniles has posed a challenge to forensic anthropologists and biological archaeologists since the advent of anthropology. Many researchers even believe that sex determination in prepubescent individuals is not possible since sex markers in adult skeletons develop after puberty. Knowing this, one would imagine that finding features that are not affected by puberty or other biological development would be key in discovering juvenile sex estimation methods. Regardless of this, the most prominent elements used for identifying the sex in adult skeletons remain a starting point for anthropologists in the efforts for determining sex in subadults.

Practicing Weaver's technique of judging auricular surface elevations to estimate sex, subadults of a variety of ages were studied. By taking the independent variables into consideration, the efficacy of Weaver's method was analyzed, as was its success across populations, collections and ages. It has been suggested that age has a larger effect on elevation than sex does. Through the use of several statistical methods the relationships between these variables was tested.

The 2 x 2 contingency tables used in the analysis of the data (Tables 3.1-3.5) all included binary, independent variables. The tables revealed the relationships between overall sex and elevation, sex against elevation of the individual collections, the collections and accuracy of identification, and ancestry and accuracy. The first (Table 3.1) showed no correlation between sex and elevation. If sex and elevation were supposed to be correlated, then Weaver's method for sex identification did not hold for either collection in the sample. Only 45.16% of females were correctly identified based

on auricular surface elevation. These results were worse than if sex had been arbitrarily assigned by chance. The males were identified with a slightly better accuracy than the females with a 57.14% accuracy. This slight favor toward the males may be due to the skewed sample ratio of males to females, with 105 males to 62 females. Regardless, 57.14% is little better than chance as well. Disappointingly, these results only mirror Weaver's results on the most basic level, with males having higher prediction accuracies than females. Weaver achieved accuracies of 95% for sexing fetal males and 75% for sexing fetal females (1980).

Results like this led to the investigation of sources of error, and the first issue to be addressed was the possibility of anthropological perspectives causing skewed conclusions in favor of males. Traditionally, in cases where a feature is either present or absent, bias tended to be in favor of males. For instance, Hooton believed that racial markers were most pronounced in males, supporting the idea that more variation can be seen amongst male specimens (Weiss, 1975). This allowed researchers to assume that when in doubt a specimen is male. This can also be seen in the research of Loth and Henneberg (1996: 475) where 0 was assigned "if the shape was neither clearly flexed nor straight." Because in their initial study, no one who received a score of 0 was female, they proposed that scores of 0, +1 and +2 be considered male (Loth and Henneberg, 1996). Later research however showed that females could in fact receive scores of 0.

Weaver's technique was set up in a similar manner. The auricular surface needed to be elevated along the entire length of the anterior and posterior border in order to be considered female. This description could be problematic and lead to error on several levels. First, the interpretation of this definition may have varied between observers.

Some may have interpreted this to include the central portion of the auricular surface while others may have considered only the outermost margin to pertain to elevation. Additionally, there are instances in which the elevation appeared to surround the entire auricular surface except for a small portion of the posterior border. At other times, portions of the border may have appeared minutely elevated, but considerably less elevated than other portions of the auricular edge. Weaver himself noted the difficulty of this standard for separating males from females when he stated that "complete elevation of this sort is not very common even in adult females. There is often a lack of elevation along the center of the posterior edge" (Weaver, 1980: 194). These complications in technique as well as interobserver error and variance in interpretation could contribute to the low and/or skewed accuracies when assigning sex based on elevation.

Another factor for poor results may be attributed to differences between population samples. With some features used for sex estimation, marked differences in success can be seen between populations due to the demography of the sample. Considering this, the second contingency table (Table 3.2) compared sex to elevation for just the individuals from the Hamann-Todd Collection. Comprised of eighty-five black and white individuals who ranged in age from one year to twenty years old, the sample was studied using crosstabulation tables. The results for the individuals in the Hamann-Todd collection were only marginally better than the overall results. 58.82% of the males were correctly identified and 52.94% of the females were correctly identified. Once again the method favored the males. This could be due to the presence of fifty-one males and thirty-four females showing an uneven distribution of the sexes. Apart from this, both the male and female prediction accuracies were only slightly better than chance.

Even while looking at this portion of the sample separate from the other collection, the efficacy of Weaver's method did not compare to the results he achieved in his original study.

Initially, Weaver used his sex estimation technique based on auricular surface elevation on the infants in the Hrdlicka Collection. Of these, Weaver (1980) received the best results on the fetal remains with 75% of the females and 91.7% of the males correctly sexed. For the purpose of testing whether or not Weaver's method works exclusively on Hrdlicka's fetal remains, or varies from sample to sample, a crosstabulation of sex and elevation in the Trotter Fetal Collection (Table 3.3) was completed. Surprisingly, the results for portion of the sample from the Trotter Fetal Collection appear to be worse than those from the older subadults in the Hamann-Todd Collection; 35.71% of the females and 55.56% of the males were correctly sexed. The sex estimation of the females was much worse than the 50% correct that would be expected from arbitrary sex assignment. Once again the males hovered around 56% accuracy in sex estimation which is still only 6% better than chance, but it is considerably more effective for males than females. These results still do not coincide with those of Weaver in his 1980 publication.

One would expect the fetal remains to perform well with this method of sex determination since it was first successfully tested on children of a similar age. However, differential effectiveness could be attributed not only to interobserver error and interpretation of methods, but to initial sample collection. The remains in the Trotter Fetal Collection were preserved in formalin for a short period of time after they were collected from refrigerated cadavers. The effects of the refrigeration as well as the

chemical reaction of the formalin may not have been realized. Fetal innominates are already small elements; the thermal and chemical conditions may have affected the slight elevations of the auricular surface. Though an unlikely source of error, it may have contributed minutely to the unfavorable results.

Knowing that both the juveniles of the Hamann-Todd Collection and the Trotter Fetal Collection did not perform well when using Weaver's method based on Table 3.2 and Table 3.3, the crosstabulation found in Table 3.4 was then completed. Here, the collections were correlated with the number of correctly sexed individuals. By doing this, one can observe whether or not there are any significant differences in the success of Weaver's method between the collections. While 56.47% of the Hamann-Todd Collection was correctly identified as either male or female based on auricular elevation, 48.79% of the Trotter Fetal Collection was correctly categorized. The difference in efficacy between these two collections does not appear significant, as both hover near 50%. This is further supported by the Chi-squared p-value calculation which yielded a value of .4467, as with the other results, being greater than .05 renders no significant relationship between the variables presented in Table 3.4. Determining sex based on Weaver's method of auricular surface analysis performed equally poor for both collections.

Lastly, a possible reason for ineffectiveness could be due to another component of sample demographics, ancestry. Though the Hamann-Todd Collection and the Trotter Fetal Collection are contemporary to each other, they are both composed of American Blacks and American Whites. It has been suggested that sex markers are not necessarily consistent between ancestral populations and have varying degrees of success because of

this. Taking ancestral populations into consideration, a crosstabulation for American Blacks and American Whites against the number of individuals correctly sex was created. The output shows that 52.55% of the American Blacks and 56.67% of the American Whites were accurately sexed. The American White population had a 4% higher success rate than the American Blacks, but both lingered around 50%. The fact that elevation did not appear to correlate with sex for either ancestral population is reflected in the statistically insignificant Chi-squared p-value .of .6826.

Following from this were several logistic regression models that used age as a variable as well. Model 1, with elevation acting as the dependent variable, used age as the dependent variable. This logistic regression function was completed on all males in the sample. It was determined that as males grew one year older, the chances of having an unelevated auricular surface in increased. Weaver's premise is based on males having an unelevated auricular surface border and females having an elevated auricular surface border. The conclusion of the logistic regression, though statistically insignificant, does not contradict Weaver's assumption. If the elevation of the auricular surface border is a sex characteristic, then it is possible that it would become more marked with age, especially as individuals reach and surpass puberty. Unfortunately one can not say that Model 1 supports Weaver's method either, as can be seen in Model 2.

The second logistic regression model was the same as the first, however, instead of using the males in the sample, the females were used. Again, using elevation as the dependent variable, and age as the independent variable, the results were the same as those found in Model 1. The model states that as a female grows a year older, the chances of becoming unelevated increase. This does contradict Weaver's original

premise. Using a sample that ranges from weeks since conception to twenty years old with this regression function implies, that as girls grow toward adulthood, their chances of having an unelevated auricular border increase. To make matters worse, Model 1 had an odds ratio of 1.0218, showing an increase in odds of 2.18% over the original value. Meanwhile, Model 2 touts an odds ratio of 1.0513, stating an increase in odds of 5.13% over the original value. Thus, according to Models 1 and 2, the females have greater odds of becoming unelevated along the auricular border with age than males have. This is in direct conflict with Weaver's theory. However, Model 2, like Model 1, also had an insignificant p-value of .1396 (Table 3.6). With both models proving to have statistically insignificant relationships, they must be disregarded, despite Model 1's agreement with Weaver's premises about the nature of elevated versus unelevated auricular surfaces.

The third logistic regression, unlike Model 1 and Model 2, uses sex as the dependent variable. This, rather than testing the nature of elevation and which variables affect it, tests the variables by which sex may be determined. In the field, anthropologists may encounter victims of violent crime and need to estimate the sex of the child, or archaeologists and skeletal biologists may find prehistoric remains that need to be sexed. In these situations sex of the child is unknown, while elevation is evident and age can be estimated based on physiological development. The model assesses this situation with the independent variables of elevation and age. Results from Model 3's logistic regression indicate that controlling for elevation, as age increases, the chances of sex being female is reduced. Looking at the composition of the sample, this could be attributed to the skewed sex ratio; in total there are 62 females and 105 males. With the lack of correlation between sex and elevation, it is understandable how the results of the

model came to be. In an actual situation, the chances of an individual being either male or female are equal. Regardless, the data from the sample does not fit the model well with a Model R^2 value of .0008.

Finally, the last logistic regression function uses elevation as the independent variable in order to try to determine which independent variables influence the raised and unraised auricular border. X1 and X2, two of the independent variables, represent age and sex respectively. Unlike previous models, Model 4 has an interaction variable, X1X2. The interaction variable allows researchers to see if the relationship between two variables is influencing the model. From the regression coefficients, the model states that controlling for X2 and X1X2, as an individual grows a year older, he or she is more likely to have an unelevated auricular surface. This has been seen in previous models as well. Furthermore, controlling for X1 and X1X2, males are more likely to be unelevated. The interaction variable of sex and age implies that as individuals get one year older, the chances of being unelevated increases faster for females than for males. This is probably again due to the skewed sex ratio of the sample.

All of the statistical methods employed indicate that there is no correlation between either sex and auricular surface elevation or age and auricular surface morphology. Unfortunately Weaver's method for assigning sex in juvenile skeletal remains doesn't appear to be as effective on the Trotter Fetal Collection and the Hamann-Todd Collection as the Hrdlicka Collection. Additionally, there does not appear to be a correlation between age and auricular elevation as was suggested by David Hunt (1990) and Mittler and Sheridan (1992).

In correlation with the lack of patterns in sex assignment, there also do not appear to be any patterns of misclassification. Looking at Figure 4.1 and Figure 4.2, the distribution for those correctly assigned sex is almost the same as the distribution for those incorrectly assigned sex. The youngest in the group faired slightly better than the oldest in the group, and ages seven to seventeen seemed to be correctly assigned sex more often than not. Though the difference is small, and no correlation with age was found to be statistically significant, it corresponds with Mittler and Sheridan's (1992) suggestion that accuracies in sex assignment were higher for ten to seventeen year old females and males greater than age ten. Unfortunately, the differences in counts of incorrect sex assignment and correct sex assignment are small and do not have any noticeable nor statistical pattern.

Research Possibilities

Despite the apparent lack of success, as mentioned before, others have obtained better results in the past when analyzing the auricular surface. Their success suggests the need for further research involving sex estimation and the auricular surface of the ilia. Moreover, with recent researchers moving away from the study of the pelvic girdle in juveniles to encompass other skeletal elements, a possibility of determining sex in subadults is still possible (Shutkowski, 1993; Franklin et al., 2007).

In 1993, Schutkowski had begun studying mandibular features in addition to an evaluation of several traits of the pelvis. Using the Coffin Plate Sample of Spitalfields, Schutkowski examined sixty-one juveniles of known age and sex. He analyzed the angle of the sciatic notch, the depth of the sciatic notch, "arch" criterion, curvature of the iliac

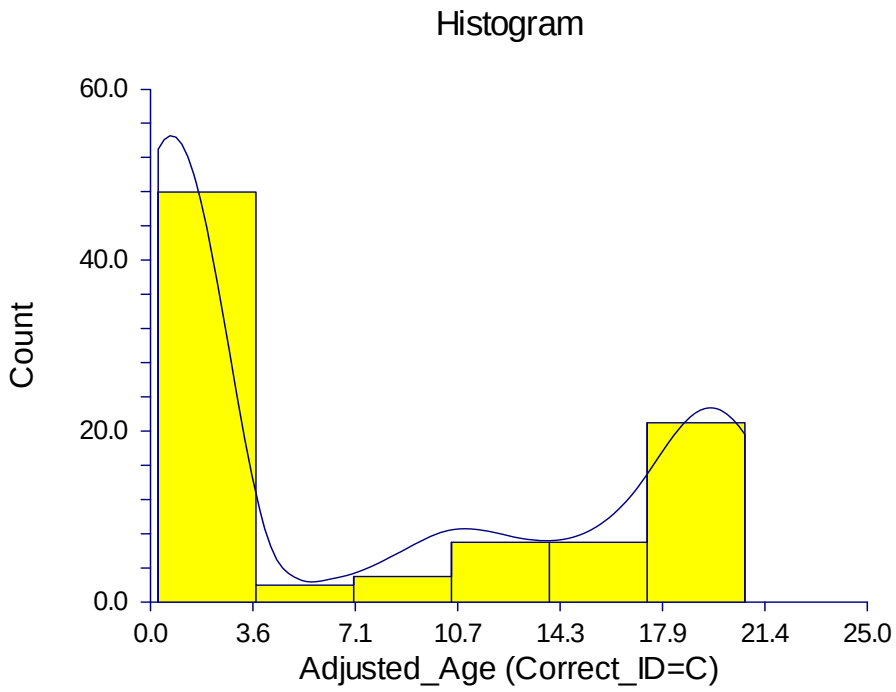


Figure 4.1 The counts of those correctly assigned sex grouped by age.

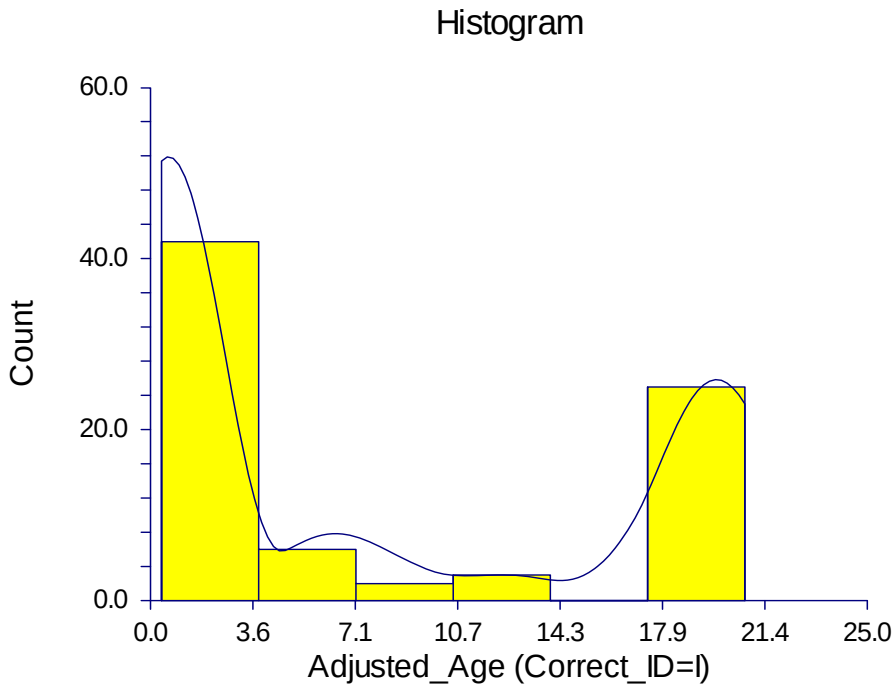


Figure 4.2 The counts of those incorrectly assigned sex grouped by age.

crest, protrusion of the chin, shape of the dental arcade and eversion of the gonion region (Schutkowski, 1993; Sutter, 2003). In these latter traits, he found that the mental eminence on female juveniles was relatively unelevated and gracile, while on males it was prominent, and when viewed superiorly appears angular and wide. Coinciding with this, the shape of the dental arcade appears parabolic in girls, but in boys is u-shaped. Lastly, Schutkowski studied the eversion of the gonion region, stating that in boys the gonion areas are everted and in girls this is not the case (Schutkowski, 1993; Sutter, 2003).

After completing his analysis, Schutkowski believed that children from birth to age five exhibit clear sex differences. Overall, he found that 94.1% of boys exhibit prominent and angular chin regions and 82.6% of them have wide anterior dental arcades. Gonion eversion is present in 73.9% of the males in the sample as well. With only a small portion of his sample exceeding age five, this distribution of sex traits is not seen. Schutkowski (1993) attributes this to small sample size. These results proved to be a promising step in using other skeletal elements aside from the pelvis in determining sex in subadult remains. Schutkowski was not the only researcher to begin looking at the mandible as an element for sex estimation.

Loth and Henneberg, previously mentioned for their work with sex estimation in adult skeletal remains, tested their methodology on juvenile skeletal remains as well. Sixty-two mandibles of known age and sex found in the Dart Collection, were used for their study (Loth and Henneberg, 2001). All individuals were either South African Blacks or South African Whites, but only the nineteen individuals with healthy mandibles between the age of seven months to four years were examined. These ages were chosen

based on eruption of central incisors and patterns between males and females that could be seen until age four (Loth and Henneberg, 2001). According to Loth and Henneberg (2001: 181), the anterior symphyseal region of a female, juvenile mandible is rounded and there is a "sharply angled transition from the symphyseal region to the lateral body in males, so different from the gradual transition in females." Because of these differences symphyseal base and body shape became the two qualifying traits to determine sex. Three researchers observed the nineteen mandibles individually, and all of the males were sexed correctly except one which was misclassified by the third observer. The females ranged in accuracy from 66.7% than 83.3%, depending on the observer. In addition to this, Loth and Henneberg used two other mandibles from South African forensic cases and 10 CT scans of living French children with an accuracy of 82% (Loth and Henneberg, 2001). Though Loth and Henneberg had varying results, the high accuracies are better than chance. This method appears to be a possible technique to improve an investigators chances when estimating sex on a juvenile skeleton.

The following year, Louise Scheuer tested Loth and Henneberg's method for sex estimation in juveniles and had less successful results. Scheuer utilized the Spitalfields Collection which had 36 mandibles of known age and sex. They ranged in age from five months to four years and seven months. The mandibles were evaluated twice in two separate sessions that were separated by two days. Scheuer (2002) states that each session had an overall accuracy of 75%, however the results for males differed than those for females. In session one 77.7% of the males and 66.6% of the females were accurately sexed, and in session two 85.1% of the males and 44.4% of the females were correctly assigned sex. This is likely in part due to the skewed sex ratio of 3:1 in favor of the

males (Scheuer, 2002). However, Scheuer's results do mirror the results of Loth and Henneberg with females being less accurately sexed than males.

Not only have scientists begun to look at mandibles but others have begun to study cranial features as well. For example, Molleson and Cruse (1998) have looked at several facial features, including the orbit shape, mentum and mandibular angle. The Spitalfields Collection, composed of 18th and 19th century skeletons, and a similar collection of 10th to 16th century skeletons from Wharram Percy, North Yorkshire were used. The researchers analyzed five features of three elements of the juvenile and adult skull. According to Molleson and Cruse (1998), females have a rounded eye orbit with sharp margins while males have angular orbits and robust margins. In addition to this, females have a rounded mandibular angle with even curvature and the male has an everted angle. Finally, females have a rounded, more gracile mental eminence than males have. These traits are based on sex differences found in mature skeletal remains.

Fifty-three skulls from the Spitalfields Collection were examined and 89% of them were correctly sexed using these traditional features. These results were found to be statistically significant with a Chi-squared p-value of $<.001$. When the juveniles of the same collection were scrutinized by the researchers, despite poor preservation of some features, 78% of them were correctly assigned sex. This, like the adult portion of the sample was also statistically significant with a Chi-squared p-value of $<.02$. Unlike the Spitalfields Collection, the Wharram Percy Collection did not have known sexes for the remains. Because of this, sex was determined using the pelvis, and the estimated cranial sex was consistent with the pelvic sex 89% of the time. For the subadults in the Wharram Percy sample, 15 individuals were assigned the sex of male, but only 13 were consistent

with the pelvic estimations. Furthermore, thirteen were assigned female, but one was male based on the pelvic sex estimation. Overall the methods were moderately successful (Molleson and Cruse, 1998).

Because of the fragile nature of the orbits in the juvenile skull, more robust elements such as the mandible or the pelvis are ideal for use in sex estimation. Over the last decade researchers have begun exploring other juvenile skeletal components for sexually dimorphic features. In addition to this, while there are still successful experiments using the innominates, investigators remain hopeful that standards using the pelvis can be established. Despite the ineffectiveness of Weaver's method in this study, other researchers have had more successful results

Conclusion

This study presents the results of examining the ilia of 167 subadults in order to test the efficacy of Weaver's method of sex estimation using the auricular surface and the relationship between this method and age. Using 2x2 contingency tables and logistic regression, no statistically significant correlations between age, sex and auricular elevation were found. There also do not appear to be patterns of misclassification, though the results are suggestive that further research should be conducted. With half of the sample comprised of fetal remains, and the other half encompassing ages from one to twenty years old, a larger sample of older subadults would be ideal. Furthermore, because other researchers have had higher accuracy rates using Weaver's method, an analysis of the age groups in their samples could be completed, as well as further testing on other samples and populations. More important, researchers should continue to study

juvenile growth and development in order to discover other sexually dimorphic traits that may appear at various stages of growth in subadults. Whether sexually dimorphic traits can be found in the ilia, mandible or cranium, any results that are statistically better than chance can aid in the important and controversial issue of juvenile sex estimation.

References Cited

Bailit H, and Edward E. Hunt J. 1964. The Sexing of Children's Skeletons from Teeth Alone and its Genetic Implications. American Journal of Physical Anthropology 22(2):171-173.

Bass W. 1995. Human Osteology: A Laboratory and Field Manual: Missouri Archaeological Society. Columbia, MO.

Black TK. 1978. A New Method for Assessing the Sex of Fragmentary Skeletal Remains: Femoral Shaft Circumference. American Journal of Physical Anthropology 48(2):227-231.

Boucher BJ. 1957. Sex Differences in the Foetal Pelvis. American Journal of Physical Anthropology 15(4):581-600.

Choi SC, and Trotter M. 1970. A Statistical Study of the Multivariate Structure and Race-Sex Differences of American White and Negro Fetal Skeletons. American Journal of Physical Anthropology 33(3):307-312.

Dittrick J, and Suchey JM. 1986. Sex Determination of Prehistoric Central California Skeletal Remains Using Discriminant Analysis of the Femur and Humerus. American Journal of Physical Anthropology 70(1):3-9.

Donnelly SM, Hens SM, Rogers NL, and Schneider KL. 1998. Technical note: A Blind Test of Mandibular Ramus Flexure as a Morphologic Indicator of Sexual Dimorphism in the Human Skeleton. American Journal of Physical Anthropology 107(3):363-366.

Federal Bureau of Investigation. Uniform Crime Reports. http://www.fbi.gov/ucr/ucr.htm.

Fehling H. 1876. Die Form des Beckens beim Foetus und Neugeborenen. Archives of Gynecology and Obstetrics 10:1-80.

Franklin D, Oxnard CE, O'Higgins P, and Dadour I. 2007. Sexual Dimorphism in the Subadult Mandible: Quantification Using Geometric Morphometrics. Journal of Forensic Sciences 52(1):6-10.

Giles E. 1964. Sex Determination by Discriminant Function Analysis of the Mandible. American Journal of Physical Anthropology 22(2):129-135.

Hanna RE, and Washburn SL. 1953. The Determination of the Sex of Skeletons, as Illustrated by a Study of the Eskimo Pelvis. Human Biology 25:21-27.

Hill C. 2000. Technical Note: Evaluating Mandibular Ramus Flexure as a Morphological Indicator of Sex. American Journal of Physical Anthropology 111(4):573-577.

Holcomb SMC, and Konigsberg LW. 1995. Statistical Study of Sexual Dimorphism in the Human Fetal Sciatic Notch. American Journal of Physical Anthropology 97(2):113-125.

Hunt DR. 1990. Sex Determination in the Subadult Ilia: An Indirect Test of Weaver's Nonmetric Sexing Method. Journal of Forensic Sciences 35(4):881-885.

Hunt EE, and Gleiser I. 1955. The Estimation of Age and Sex of Preadolescent Children from Bones and Teeth. American Journal of Physical Anthropology 13(3):479-487.

Iscan M, and Ding S. 1995. Sexual Dimorphism in the Chinese Cranium. International Association for Craniofacial Identification Program. p 24-25.

Koski K. 1996. Mandibular Ramus Flexure – An Indicator of Sexual Dimorphism? American Journal of Physical Anthropology 101(4):545-546.

Krogman, WM. 1973. The Human Skeleton in Forensic Medicine. Springfield, IL: C.C. Thomas.

Loth SR, and Henneberg M. 1996. Mandibular Ramus Flexure: A New Morphologic Indicator of Sexual Dimorphism in the Human Skeleton. American Journal of Physical Anthropology 99(3):473-485.

Loth SR, and Henneberg M. 1998. Mandibular Ramus Flexure IS a Good Indicator of Sexual Dimorphism. American Journal of Physical Anthropology 105(1):91-92.

Loth SR, and Henneberg M. 2001. Sexually Dimorphic Mandibular Morphology in the First Few Years of Life. American Journal of Physical Anthropology 115(2):179-186.

MacLaughlin SM, and Bruce MF. 1985. A Simple Univariate Technique for Determining Sex from Fragmentary Femora: Its Application to a Scottish Short Cist Population. American Journal of Physical Anthropology 67(4):413-417.

Mittler DM, and Sheridan SG. 1992. Sex Determination in Subadults Using Auricular Surface Morphology: A Forensic Science Perspective. Journal of Forensic Sciences 37(4):1068-1075.

Molleson T, and Cruse K. 1998. Some Sexually Dimorphic Features of the Human Juvenile Skull and their Value in Sex Determination in Immature Skeletal Remains. Journal of Archaeological Science 25:719-728.

National Center for Juvenile Justice. 2006. Easy Access to the FBI's Supplementary Homicide Reports: 1980-2004. http://ojjdp.ncjrs.org/ojstatbb/ezashr/.

Phenice TW. 1969. A Newly Developed Visual Method of Sexing the Os Pubis. American Journal of Physical Anthropology 30(2):297-301.

Reynolds EL. 1945. The Bony Pelvic Girdle in Early Infancy. A Roentgenometric Study. American Journal of Physical Anthropology 3(4):321-354.

Scheuer L. 2002. A Blind Test of Mandibular Morphology for Sexing Mandibles in the First Few Years of Life. American Journal of Physical Anthropology 119(2):189-191.

Schutkowski H. 1993. Sex Determination of Infant and Juvenile Skeletons: I. Morphognostic Features. American Journal of Physical Anthropology 90(2):199-205.

Stewart TD. 1954. Sex Determination of the Skeleton by Guess and by Measurement. American Journal of Physical Anthropology 12(3):385-392.

Stewart TD. 1979. Essentials of Forensic Anthropology. Springfield, IL: C.C. Thomas.

Sutter RC. 2003. Nonmetric Subadult Skeletal Sexing Traits: I. A Blind Test of the Accuracy of Eight Previously Proposed Methods Using Prehistoric Known-Sex Mummies from Northern Chile. Journal of Forensic Sciences 48(5):1-9.

Thomson A. 1899. The Sexual Differences of the Fetal Pelvis. Journal of Anatomy and Physiology 33:359-381.

Walker PL, Johnson JR, and Lambert PM. 1988. Age and Sex Biases in the Preservation of Human Skeletal Remains. American Journal of Physical Anthropology 76:183-188.

Washburn SL. 1948. Sex Differences in the Pubic Bone. American Journal of Physical Anthropology 6(2):199-208.

Weaver DS. 1980. Sex Differences in the Ilia of a Known Sex and Age Sample of Fetal and Infant Skeletons. American Journal of Physical Anthropology 52(2):191-195.

Weiss KM. 1972. On the Systematic Bias in Skeletal Sexing. American Journal of Physical Anthropology 37(2): 239-249.

Wilton, MK. 1939. Contributions of T. Wingate Todd to Anatomy and Physical Anthropology. American Journal of Physical Anthropology 25(2):145-186.

Appendix A

Individual	Elevated	Age (wks)	Age (yrs)	Sex	Correct ID	Ancestry	Collec-tion	Adj. Age
2	N	24	0.46	M	C	B	TF	0.46
3	N	20	0.38	M	C	B	TF	0.38
4	Y	20	0.38	F	C	B	TF	0.38
6	N	16	0.31	M	C	B	TF	0.31
7	Y	22	0.42	M	I	B	TF	0.42
8	N	26	0.50	M	C	B	TF	0.50
9	N	20	0.38	F	I	W	TF	0.38
10	Y	20	0.38	M	I	W	TF	0.38
13	N	38	0.73	M	C	B	TF	0.73
14	N	24	0.46	F	I	B	TF	0.46
15	Y	26	0.50	F	C	B	TF	0.50
17	N	16	0.31	M	C	B	TF	0.31
18	N	19	0.37	M	C	W	TF	0.37
19	N	34	0.65	F	I	B	TF	0.65
20	N	22	0.42	M	C	W	TF	0.42
21**	N	16	0.31	M	C	B	TF	0.31
23	N	16	0.31	M	C	B	TF	0.31
25	N	14	0.27	M	C	B	TF	0.27
26	N	24	0.46	F	I	B	TF	0.46
29	N	32	0.62	M	C	W	TF	0.62
30	Y	38	0.73	M	I	B	TF	0.73
34	N	20	0.38	M	C	B	TF	0.38
39	N	26	0.50	M	C	W	TF	0.50
41	Y	23	0.44	F	C	W	TF	0.44
42	Y	40	0.77	M	I	B	TF	0.77
46	Y	27	0.52	M	I	W	TF	0.52
47**	Y	32	0.62	M	I	B	TF	0.62
49	N	32	0.62	M	C	W	TF	0.62
56	N	40	0.77	M	C	B	TF	0.77
10B	Y	34	0.65	F	C	B	TF	0.65
11A	N	28	0.54	M	C	W	TF	0.54
12W	N	35	0.67	F	C	B	TF	0.67
13A	Y	26	0.50	M	I	B	TF	0.50
13B-1	Y	43	0.83	F	C	B	TF	0.83
14A	N	37	0.71	M	C	B	TF	0.71
14B	Y	37	0.71	M	I	W	TF	0.71
15B	N	37	0.71	F	I	B	TF	0.71
15W	N	30	0.58	M	C	B	TF	0.58

Individual	Elevated	Age (wks)	Age (yrs)	Sex	Correct ID	Ancestry	Collec-tion	Adj. Age
16B	N	27	0.52	M	C	B	TF	0.52
17A**	N	22	0.42	M	C	B	TF	0.42
17W	Y	38	0.73	M	I	B	TF	0.73
18B	N	39	0.75	M	C	B	TF	0.75
19A	N	40	0.77	F	I	B	TF	0.77
1B	N	40	0.77	F	I	B	TF	0.77
20B	Y	36	0.69	M	I	B	TF	0.69
20W	N	35	0.67	F	I	B	TF	0.67
21A	N	22	0.42	M	C	B	TF	0.42
21B	Y	40	0.77	M	I	B	TF	0.77
22A	Y	24	0.46	F	C	B	TF	0.46
22B	N	36	0.69	F	I	B	TF	0.69
23A	N	24	0.46	F	I	B	TF	0.46
24B	N	39	0.75	F	I	B	TF	0.75
26B	Y	42	0.81	M	I	W	TF	0.81
26W	N	28	0.54	M	C	B	TF	0.54
27W	N	32	0.62	F	I	B	TF	0.62
28A	N	27	0.52	M	C	B	TF	0.52
29A	Y	28	0.54	M	I	B	TF	0.54
29B	Y	40	0.77	M	I	B	TF	0.77
2B*	Y	30	0.58	M	I	B	TF	0.58
30W	N	32	0.62	F	I	B	TF	0.62
31A	N	24	0.46	M	C	B	TF	0.46
31W	Y	32	0.62	M	I	B	TF	0.62
32A	Y	24	0.46	M	I	B	TF	0.46
32B	Y	40	0.77	M	I	W	TF	0.77
33A	N	23	0.44	M	C	B	TF	0.44
33B	Y	28	0.54	M	I	B	TF	0.54
33W	N	20	0.38	M	C	B	TF	0.38
3A	N	28	0.54	F	I	B	TF	0.54
3B	Y	32	0.62	M	I	W	TF	0.62
3W	Y	40	0.77	M	I	B	TF	0.77
41-1	Y	31	0.60	M	I	W	TF	0.60
4A	N	32	0.62	F	I	B	TF	0.62
4B	Y	35	0.67	F	C	B	TF	0.67
5A	Y	24	0.46	M	I	B	TF	0.46
5B	Y	40	0.77	F	C	B	TF	0.77
6B	N	38	0.73	M	C	B	TF	0.73

Individual	Elevated	Age (wks)	Age (yrs)	Sex	Correct ID	Ancestry	Collec-tion	Adj. Age
6W	Y	20	0.38	M	I	B	TF	0.38
7B	N	30	0.58	F	I	B	TF	0.58
8A	N	32	0.62	M	C	W	TF	0.62
8W	N	35.5	0.68	F	I	B	TF	0.68
9A	Y	24	0.46	F	C	B	TF	0.46
9B	Y	33	0.63	F	C	B	TF	0.63
HTH 0017	Y		19	M	I	B	H-T	19.731
HTH 0098	N		18	M	C	W	H-T	18.731
HTH 0233	N		19	M	C	W	H-T	19.731
HTH 0291	Y		20	M	I	B	H-T	20.731
HTH 0376	N		20	M	C	W	H-T	20.731
HTH 0404	N		11	M	C	B	H-T	11.731
HTH 0437	Y		18	F	C	W	H-T	18.731
HTH 0485	Y		16	F	C	B	H-T	16.731
HTH 0526	Y		11	F	C	B	H-T	11.731
HTH 0527	Y		16	F	C	W	H-T	16.731
HTH 0548	Y		17	M	I	B	H-T	17.731
HTH 0563	N		19	M	C	B	H-T	19.731
HTH 0576	Y		16	F	C	B	H-T	16.731
HTH 0588	N		19	M	C	B	H-T	19.731
HTH 0598	N		20	M	C	B	H-T	20.731
HTH 0620	Y		19	M	I	B	H-T	19.731
HTH 0624	N		6	F	I	B	H-T	6.731
HTH 0632	N		10	F	I	B	H-T	10.731
HTH 0633	Y		14	F	C	B	H-T	14.731
HTH 0639	Y		20	M	I	B	H-T	20.731
HTH 0645	Y		12	F	C	W	H-T	12.731
HTH 0695	Y		18	M	I	B	H-T	18.731
HTH 0710	N		10	M	C	B	H-T	10.731
HTH 0721	N		18	M	C	B	H-T	18.731
HTH 0807	Y		20	M	I	B	H-T	20.731
HTH 0808	Y		20	M	I	B	H-T	20.731
HTH 0854	N		19	M	C	B	H-T	19.731
HTH 0872	Y		8	F	C	B	H-T	8.731
HTH 0955	Y		20	M	I	B	H-T	20.731
HTH 0964	Y		20	M	I	B	H-T	20.731
HTH 1012	N		18	F	I	B	H-T	18.731
HTH 1041	N		17	F	I	B	H-T	17.731

Individual	Elevated	Age (wks)	Age (yrs)	Sex	Correct ID	Ancestry	Collec-tion	Adj. Age
HTH 1074	N		4	F	I	B	H-T	4.731
HTH 1097	N		18	M	C	B	H-T	18.731
HTH 1098	N		5	F	I	B	H-T	5.731
HTH 1109	Y		20	M	I	B	H-T	20.731
HTH 1115	Y		5	F	C	B	H-T	5.731
HTH 1140	Y		18	M	I	B	H-T	18.731
HTH 1156	N		8	F	I	B	H-T	8.731
HTH 1168	N		1	M	C	B	H-T	1.731
HTH 1228	Y		20	M	I	W	H-T	20.731
HTH 1232	Y		16	F	C	B	H-T	16.731
HTH 1238	N		19	F	I	W	H-T	19.731
HTH 1240	Y		12	F	C	W	H-T	12.731
HTH 1328	Y		19	F	C	B	H-T	19.731
HTH 1339	Y		20	F	C	B	H-T	20.731
HTH 1379	N		1	M	C	B	H-T	1.731
HTH 1385	N		1	M	C	B	H-T	1.731
HTH 1441	N		10	M	C	B	H-T	10.731
HTH 1557	N		3	M	C	B	H-T	3.731
HTH 1575	N		20	M	C	B	H-T	20.731
HTH 1583	N		1	M	C	W	H-T	1.731
HTH 1590	N		18	F	I	B	H-T	18.731
HTH 1606	N		17	F	I	B	H-T	17.731
HTH 1688	N		10	M	C	B	H-T	10.731
HTH 1693	N		20	M	C	B	H-T	20.731
HTH 1704	N		20	M	C	B	H-T	20.731
HTH 1707	Y		20	M	I	B	H-T	20.731
HTH 1711	N		17	M	C	B	H-T	17.731
HTH 1768	N		1	M	C	B	H-T	1.731
HTH 1772	N		12	F	I	W	H-T	12.731
HTH 1784	Y		6	M	I	B	H-T	6.731
HTH 1834*	N		8	M	C	B	H-T	8.731
HTH 1894	N		1	M	C	B	H-T	1.731
HTH 1949	Y		19	F	C	B	H-T	19.731
HTH 1974	N		18	M	C	B	H-T	18.731
HTH 2036	N		7	F	I	W	H-T	7.731
HTH 2065	N		19	F	I	B	H-T	19.731
HTH 2074	Y		8	F	C	B	H-T	8.731

Individual	Elevated	Age (wks)	Age (yrs)	Sex	Correct ID	Ancestry	Collec-tion	Adj. Age
HTH 2075	Y		1	M	I	B	H-T	1.731
HTH 2104	N		20	M	C	B	H-T	20.731
HTH 2118	N		13	F	I	B	H-T	13.731
HTH 2135	Y		14	F	C	B	H-T	14.731
HTH 2141	N		4	F	I	B	H-T	4.731
HTH 2144	Y		6	M	I	B	H-T	6.731
HTH 2171	N		20	M	C	B	H-T	20.731
HTH 2370	N		1	M	C	B	H-T	1.731
HTH 2558	Y		19	M	I	W	H-T	19.731
HTH 2714	Y		1	F	C	B	H-T	1.731
HTH 2796	Y		20	M	I	B	H-T	20.731
HTH 2845	N		20	F	I	B	H-T	20.731
HTH 3112	N		15	M	C	B	H-T	15.731
HTH 3455	Y		18	M	I	B	H-T	18.731
HTH 3470	Y		17	M	I	B	H-T	17.731
HTH 3699	Y		19	F	C	W	H-T	19.731

*Changed when first 10% were reassessed

**Right element was used

Vita

Jaymelee Jane Kim was born in Findlay, Ohio, on November 11, 1983. She was raised in Arcadia, Ohio and attended Saint Wendelin Elementary School and Saint Wendelin High School in Fostoria, Ohio. She graduated from Saint Wendelin High School in 2002. She then enrolled at Rochester Institute of Technology in Rochester, New York and graduated in 2006 with a Bachelor of Science degree in packaging science and a minor in Japanese language and culture. From there, she went to the University of Tennessee, Knoxville and earned a Master of Arts degree in biological anthropology in 2008. Currently, she is pursuing her doctorate in biological anthropology at the University of Tennessee, Knoxville.