THESIS

SEX DETERMINATION USING THE FIRST THORACIC VERTEBRA IN 19TH CENTURY AMERICAN AND ANCIENT NUBIAN HUMANS

Submitted by

Charles Cundiff

Department of Anthropology

In partial fulfillment of the requirements

For the Degree of Master of Arts

Colorado State University

Fort Collins, Colorado

Fall 2010

COLORADO STATE UNIVERSITY

August 23, 2010

WE HEREBY RECOMMEND THAT THE THESIS PREPARED UNDER OUR SUPERVISION BY CHARLES CUNDIFF ENTITLED SEX DETERMINATION USING THE FIRST THORACIC VERTEBRA IN 19^{TH} CENTURY AND ANCIENT NUBIAN HUMANS BE ACCEPTED AS FULFILLING IN PART REQUIREMENTS FOR THE DEGREE MASTER OF ARTS.

Committee on Graduate Work

J. LaBelle

Jason LaBelle

Phillip J. Chapman

Phillip Chapman

Ann L. Magennis

Advisor: Ann Magennis

Diane L. France

Co-Advisor: Diane L. France

Kathleen Sherman

Department Head: Kathleen Sherman

ABSTRACT OF THESIS

SEX DETERMINATION USING THE FIRST THORACIC VERTEBRA IN 19TH CENTURY AMERICAN AND ANCIENT NUBIAN HUMANS

Many metric sex determination methods exist and have proven to be useful (Albanese 2003; Albanese et al. 2008; France 1998; Freiman et al. 2008; Frutos 2002; Frutos 2005; Gapert et al. 2009; Giles and Elliot 1963; Ozer and Katayama 2008; Ozer et al. 2006; Pastor 2005; Phenice 1969; Yu et al. 2008). Most metric sex determination methods rely on differences in stature and musculature between the sexes. The first thoracic vertebra is of interest because of its ease of identification and location at the boundary of many muscle groups. Linear measurements were taken on 161 T1s from two osteological collections housed in Colorado. The first population, housed at CU Boulder, is derived from cemeteries excavated in Kulubnarti, Sudan. Burials in from this sample range from c.a. 550 AD to c.a. 1500 AD. The second population, housed at CSU Fort Collins, are remains from the cemetery of an asylum locate in Pueblo Colorado in the late 19th century.

A linear function was used to determine the best classifying features of the T1. From the 4 best classifiers (length of the transverse process, length of the spinous process, body diameter and coronal breadth of the vertebral foramen) a discriminant function was created for purposes of classification. Cross-validated results for the entire

population give an accuracy of 86.76% for females and 89.25% accuracy for males. For the CSU (American) population 92.31% of females were classified correctly while 95.56% of males were correctly classified. For the CU (Nubian) population 92.59% of females were correctly classified as female with 85.71% being correctly classified as male.

Charles Raymond Cundiff
Department of Anthropology
Colorado Sate University
Fort Collins, CO 80523
Fall 2010

CONTENTS

CHAPTER1: INTRODUCTION

The objective of this study is to use linear measurements of the first thoracic vertebra to differentiate sex in modern humans. The first thoracic vertebra is chosen because previous research indicates that vertebrae can be useful in sexing known populations, and because it is likely to be useful as a universal sex determination key. The ability to confidently assign sex to unidentified human remains is essential to archaeological and forensic investigations.

Vertebrae are essentially a mix of features which are affected by musculature and features that are not affected by musculature. The spinous process is an appendage entirely committed to muscle insertion while the transverse processes (of the thoracic vertebrae) split the role of muscle insertion center and articulation point for the ribs. These features can be described as plastic, in that they can change significantly during adult life. The body of the vertebra is not so plastic as the processes. While there are some muscle attachments ventrally, they are not as heavily muscled as the processes and the vertebral body has no muscle attachments on the superior or inferior sides. In at least the superior-inferior direction the vertebral body is metrically static during adult life, and is somewhat indicative of adult stature.

An advantage that comes from the presence of a mixture of plastic and static features is that shape becomes a factor in sexual dimorphism in addition to size. Shape is seen in the comparison between how the plastic features vary relative to the static features. In this way a male vertebra does not have to be larger than a female vertebra to be distinguished. The plastic elements of the male vertebrae simply have to be larger relative to the static elements of the vertebrae.

While all vertebrae are a generous mix of static and plastic elements, the first thoracic vertebra is uniquely placed to be both easily identifiable, and central to three large muscle groups. Muscle stress is therefore extremely high on all processes of the first thoracic vertebra. This heightened stress level acts as a magnifier to the plastic properties of the muscle insertion points. To accommodate the generally and relatively more muscular shoulders, backs and necks of male humans, the relative size differences between plastic and static features in male first thoracic vertebrae should be greater than the relative size differences between plastic and static features in female first thoracic vertebrae.

Previous research into differential morphology of lower thoracic and of lumbar vertebrae has shown that vertebrae can be useful in the classification of sex (Cheng et al. 1998; Freiman et al. 2008; Marino 1995; Pastor 2005; Yu et al. 2008). This will be discussed in Chapter 2, along with general information about vertebrae as well as methods to identify the first thoracic vertebra. Materials for this study include 161 first thoracic vertebrae from osteological collections housed at CU Boulder and CSU in Fort Collins, CO. The collections are described in Chapter 3 along with descriptions of each measurement and a brief discussion on the statistical analysis utilized. Chapter 4 involves the results of the primary study (the sexing of the first thoracic vertebra) and additionally includes incidental results gleaned from the collected data. The discussion (Chapter 5) is concerned with defining the importance of the results relative to the limitations of the study.

CHAPTER 2: METRIC METHODS OF SEX DETERMINATION AND VERTEBRAL ANATOMY

There is a long history of metric methods of sex determination (France 1998). Cranial measurements (France 1988; Giles & Elliot 1963) have also been used in addition to visual analysis of the cranium to determine sex. Dorsey (1897) used measurements of the articular surfaces of the long bones to determine sex on populations of Native Americans from the Northwest Coast of the United States, as well as from Ohio and Peru. He concluded "if the maximum diameter of the head of the humerus of any American skeleton measures 44mm, the chances are extremely great that it is a male…" (p. 82). Long bones have since proven to be useful within population groups.

Giles and Elliot (1962) we able to correctly sex 82.9% of the 1,022 subjects tested from the Terry Collection regardless of population affinity. The Terry Collection is made up of European and African Americans. Using a discriminant function for each population group (Giles & Elliot 1963), better results were accomplished. Up to 86.6% accuracy was accomplished for European Americans and up to 87.6% accuracy for African Americans within the study group. It was later shown that these methods broke down completely outside of the test population (Birkby 1966).

There are several other postcranial methods of metric sex determination. Recent methods include measurement of the scapula, clavicle, various tarsals, metatarsals, and metacarpals. Of particular relevance to this thesis, discriminant functions of measurements of scapula have been used (Ozer et al. 2006) to achieve up to 94% accuracy using a medieval Anatolian population. In 2010, Dabbs reported a similar result with an overall accuracy of 95.7% on the Hamann-Todd collection with cross validation, though when tested on a different sample (also from the Hamann-Todd collection) the accuracy was 92.5%. Testing the method on a collection from Wichita State University Biological Anthropology Laboratory further dropped the accuracy of the test to 84.4%. Dabbs (2010) demonstrates through this study that it is necessary to validate results with multiple test populations. Even with cross-validation, the result of 95.7% accuracy is likely optimistic.

In order to utilize the difference in musculature between the sexes, features directly affected by muscular development must be rated against those that are not. The vertebral body is a load bearing structure with size determined genetically, similar to the long bones (assuming adequate nutrition during growth). The pedicles and the spinous process are, in part, shaped by the power of the musculature of the back. Because one of these variables is fixed (the depth, breadth and height of the vertebral body) and the other is plastic, some ratio between them may indicate a difference in musculature, and therefore differentiate between the sexes.

The relative robustness of muscle insertions on the skeleton is directly affected by the forces of the muscles which pull on them (Benjamin et al. 2006). The human spine is a region with multiple large muscle groups exerting extreme forces on the vertebrae in several directions (Gest & Schlesinger 1995; van Lopik & Acar 2007). Previous studies have shown features of various vertebrae to exhibit sexual dimorphism (Cheng et al. 1998; Freiman et al. 2008; Marino 1995; Pastor 2005; Yu et al. 2008).

Previous vertebral sex determination methods have utilized a large number of measurements to determine sex but have demonstrated that the combination of a few are best for distinguishing males from females (Marino 1995; Pastor 2005; Yu et al. 2008). Using a digital catalogue of Korean specimens, Yu et al. (2008) had greatest success using equations based on a few separate systems of the 12^{th} thoracic vertebra. Comparison of the superior and inferior coronal diameter of the end-plate of the vertebral body versus the maximum coronal diameter of the same measurement garnered an accuracy of over 80%. An equation correctly predicting sex with 90% accuracy utilized the coronal diameter of the end plate, the ratio of the middle height of the body versus the posterior height of the body, and the length of the mammillary (transverse) process and the pedicle along the axis of the pedicle. Pastor (2005) utilized the junction of T12 and L1 to garner 90% accuracy using ratios between the vertebral body and the pedicle, and the vertebral body and the length of the spinous process.

Yu et al. (2008) used the Digital Korean database at the Catholic Institute for Applied Anatomy to develop a metric sex-determination system using the 12^{th} thoracic vertebra. Twenty two measurements of 102 vertebrae were taken. These measurements were used to create 23 discriminant function equations. Initial results were likely clouded by too much information as sex was predicted 62.7%-85.3% of the time. Using a stepwise method of discriminant function analysis the number of variables was reduced from 22 to 4. This simplification resulted in sex prediction with increased accuracy at 90.0%.

Replication of the precise measurement technique of Yu et al. (2008) is not possible as the study material consists of a digital database. Precision is suspect in the Yu et al. (2008) study as the topographic images used have a 1-mm axial thickness. The calipers used in the present study are precise to 0.001mm. While the hundredths place is outside the tolerance for measurement error for the present study, the loss of the tenths place may reduce the resolution of the previous study.

Marino (1995) examined the first cervical vertebra as a means of sex determination. Seven regression and seven discriminant function equations were created from eight measurements taken from the superior and inferior articular regions of 100 first cervical vertebrae from the Terry Collection. The test was controlled using 100 first cervical vertebrae from the Hamann-Todd collection. The regression equations correctly predicted sex on the study sample with 77-85% accuracy and 75-85% accuracy was achieved for the discriminant functions. When applied to the control (Haman-Todd) sample only 60-84% accuracy was achieved.

Pastor (2005) demonstrated an accuracy of 88.9% in white males and females from the Terry collection, using the coronal diameter of the end plate versus the length of the spinous process of the first lumbar vertebra. Pastor's findings are only listed in an abstract of proceeding, so his methods are not known.

The First Thoracic Vertebra

The first thoracic vertebra is of specific interest due to its ease of identification and connection between the muscle systems of the thoracic musculature and that of the head and neck. Considering success determining sex using other parts of the vertebral column (Marino 1995; Pastor 2005; Yu et al. 2008) the first thoracic vertebra is a strong candidate for the addition of a new and accurate sex determination method.

The first thoracic vertebra is identified in two ways. If the seventh cervical vertebra is present it will occlude perfectly with the superior side of the first thoracic.

Without the seventh cervical vertebra the first thoracic can be easily confused with the second thoracic. However, it often can be distinguished from other thoracic vertebrae by the presence of a single, relatively large articular facet for the first rib. Presented below is a basic anatomy of the vertebral column in order to illustrate how the first thoracic is identified.

Anatomy of the Vertebrae (Bass 1979)

Characteristics common to all vertebrae

- Body- The load bearing portion of the vertebrae, it is the oval to bean shaped structure located anterior to the vertebral foramen. It is connected to the rest of the vertebra by a pedicle on each (posterior-lateral) side.
- Spinous Process- A posterior protruding process.
- Transverse Process- Laterally facing processes.
- Superior articular processes- Apophyseal processes pointing superiorly and facing posteriorly.
- Inferior articular processes- Apophyseal processes pointing inferiorly and facing anteriorly

Differentiating Characteristics of the Vertebrae

Cervical Vertebrae-

The seven cervical vertebrae define the neck region of the spine. The first two cervical vertebrae are known as the *atlas* and the *axis* and have very unique morphologies designed to hold the cranium. The cervical vertebrae do not connect to ribs.

- Body- Oval-shaped and small
- Spinous Process- All cleft at end except that of C7
- Transverse Foramen- A superior-inferior directed hole in each transverse process

- Transverse Processes- Small, growing more robust towards C7.
- C7- Very similar in shape and size to T1, with large body and prominent transverse processes. Differentiated from T1 by presence of transverse foramen

Thoracic Vertebrae-

The 12 thoracic vertebrae make up the body of the spine. They connect to the ribs transversely and increase in size from 1 to 12.

- Body- Large, bean-shaped to round
- Spinous Process- Elongate; Points from slightly inferiorly (T1) to nearly parallel to the superior-inferior line (T12). Superior edge starts rounded (T1), becoming sharper towards T12.
- Transverse Processes- Prominent and robust (except T11 and T12).

Lumbar Vertebrae-

The lumbar vertebrae make up the lower spine and do not connect to any ribs. They are broad in order to carry the bulk of the weight of the body. The 5th lumbar articulates with the superior side of the sacrum which resembles several fused vertebrae.

- Body- Very large, roundish.
- Spinous process- Broad, project horizontally in the posterior direction.
- Transverse Processes (Mammillary Process)- Short and extending posterior-laterally.

The 1st Thoracic

The 1st thoracic is differentiated from the rest of the thoracic vertebrae by a body that is wide laterally and short dorso-ventrally. The body of the 1st thoracic is the most "bean-shaped" of the thoracic vertebrae.

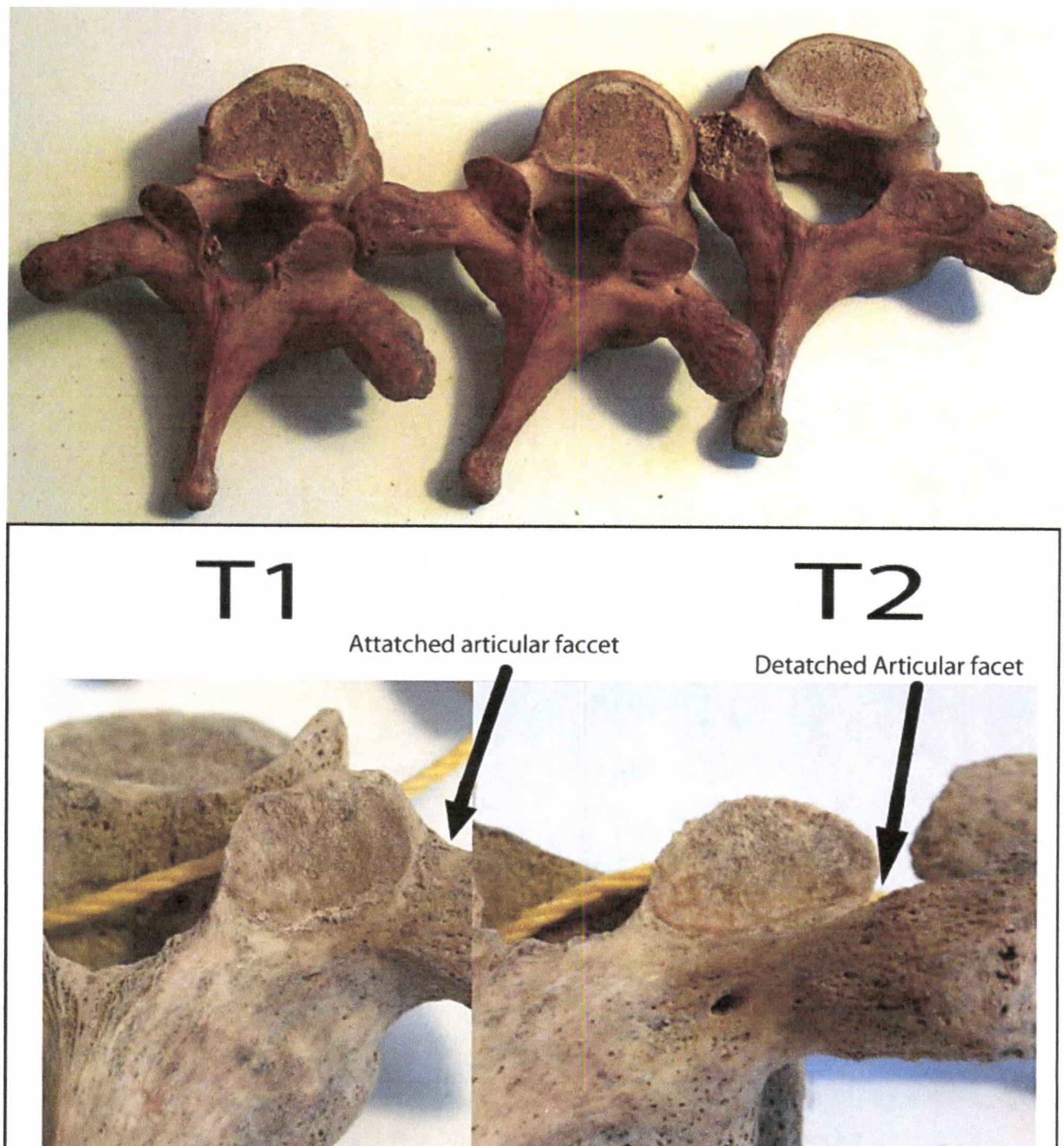

Figure 2.1 T3, T2 and T1 as viewed from above and the superior articular facets of T1 and T2

The spinous process is the least inferior-pointing of the thoracic vertebrae, often angled nearly flush with the superior side of the vertebral body.

Figure 2.2 Side view of Thoracic Vertebrae T5-T1 and C7.

A basic knowledge of human vertebrae allows for easy identification of the first thoracic vertebra. Additionally it is important that the specific measurements used to create the sex identification method are easily identifiable to future users of the method. Therefore only unambiguous measurements of the first thoracic vertebra are included in this study. While it is likely that space-and-shape-based measurements of the facets may also be useful in the determination of sex, the use of such measurements would require extreme familiarity with the vertebra in question, as well as specialized and expensive equipment.

CHAPTER 3: MATERIALS AND METHODS

Materials used for this study consisted of skeletal remains from two collections housed at Colorado Universities. A total of 161 first thoracic vertebrae were measured. Fifty eight first thoracic vertebrae from the Pueblo Colorado Mental Health Institute housed at Colorado State University in Fort Collins, Colorado were used in the study. Of those, 45 first thoracic vertebrae came from males and 13 from females. Ninety six first thoracic vertebrae from the Nubian collection from the University of Colorado, Boulder were also used. The Nubian collection is comprised of two sub-populations from the R and S burial sites. Sixty nine first thoracic vertebrae (35 males and 34 females) came from the R site while 37 first thoracic vertebrae (17 males and 20 females) were measured from the S site.

The Nubian collection, housed at the University of Colorado, Boulder, contains a total of 418 individuals from the site of Kulubnarti, Sudan (Sheridan 1992). The sample was exhumed in 1979 from two cemeteries dating from early Christian times (ca. 550-750 AD) and the Christian era of this region (ca. 750-1500 AD). Both sexes are represented by individuals aged from five months *in utero* to 51+ years of age. Preservation is excellent due to natural mummification. Sexing of many of the individuals is extremely reliable due to the preservation of soft tissue.

The CSU collection consists of remains from forgotten graves excavated in 1992 and 2000 on the site of the Pueblo Colorado Mental Health Institute (Bower et al. 2007). This cemetery was active between 1879 and 1899, and was likely used to inter individuals who either had no families to claim them, or which would not claim them.

More than 95% of the inmates were immigrants to Colorado with nearly a third having been born in Europe. Males comprise 77% of the skeletons from the cemetery. The medical record from the institution reports that 505 inmates died during the years between 1880 and 1900. The majority were likely buried in regular cemeteries. Preservation of the specimens used in the study is excellent.

Fifteen linear measurements of the first thoracic vertebra were chosen for model selection. Several of the measurements selected were shown to be useful in sex determination by Yu et. al (2008) on the 12th thoracic vertebra. The remainder was selected to test possible variation in robusticity of the spinous process. All measurements were taken with digital Mitutoyo calipers except for the height of the middle body which was taken by analog calipers modified to take measurements of center-points. Only those features which have been shown to be useful in model calibration are pictured for clarification.

Figure 3.1. Body diameter of the first thoracic vertebra (**1 BD**).
The diameter of the vertebral body measured in the transverse plane. Measurement is taken in the apparent center with caliper tines on the anterior side and the edge of the vertebral foramen.

Figure 3.2. Height of the middle of the vertebral body **(2 HM)**
The height of the center of the vertebral body measured in the superior-inferior axis.

Figure 3.3 Height of the posterior vertebral body **(3 HP)**
The height measured at the center of the posterior edge (that edge which borders the vertebral foramen) measured in the sagittal plane.

Figure 3.4 Pedicle length **(4 Pli)**
The width of the pedicle measured as the smallest distance between the vertebral body and the posterior side of the pedicle. Measurement is taken in the transverse plane.

Figure 3.5 Transverse Process **(5 M)**
The length of the transverse process as measured from the adjacent edge of the vertebral foramen to the end of the transverse process. Measurement is taken in the coronal plane.

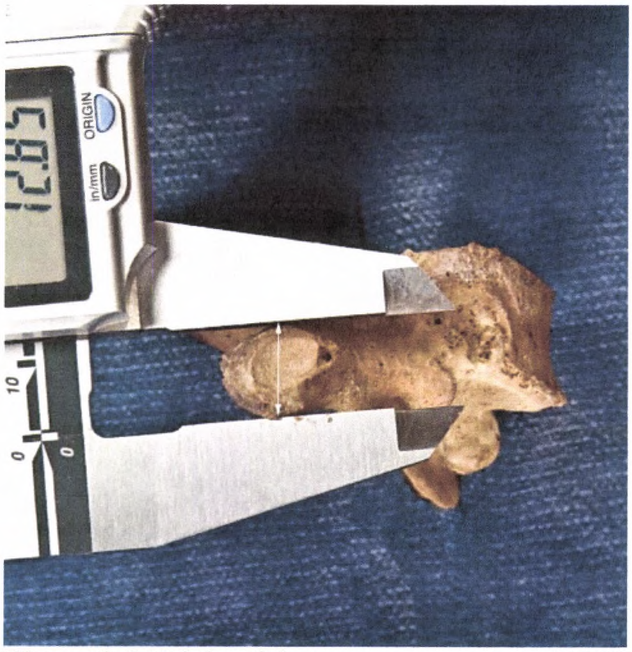

Figure 3.6 Height of the transverse process (**6 MH**)
The *maximum* height of the transverse process in the sagittal plane. To ensure consistency in measurement the caliper tines are held parallel to the top edge (the transverse plane) of the vertebral body.

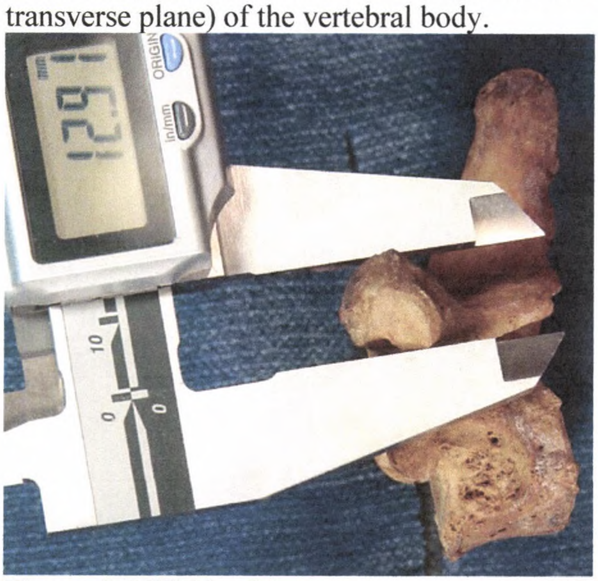

Figure 3.7 Width of the transverse process (**7 MW**)
The *maximum* width of the transverse process measured in the sagittal plane. To ensure consistency in measurement the caliper tines are held perpendicular to the top edge (the transverse plane) of the vertebral body.

Figure 3.8 Length of the spinous process (**8 SL**)
The *linear distance* of the posterior end of the spinous process from the posterior edge of the vertebral foramen. Caliper tines are kept perpendicular to the transverse plane of the vertebral body. The tine is laid inside the vertebral foramen flush to the side of the posterior edge of the vertebral foramen. Calipers are aligned in the sagittal plane with measurement along both the sagittal and transverse planes.

9 SW Width of the spinous process
The *maximum* width of the posterior tubercle of the spinous process measured in the coronal plane.

Figure 3.9 Sagittal width of the vertebral foramen (**10 FDs**)
Width of the vertebral foramen measured in the apparent center of the foramen in the sagittal plane.

Figure 3.10 Coronal width of the vertebral foramen (**11 FDc**)
Width of the vertebral foramen measured in the apparent center of the foramen in the coronal plane.

Figure 3.11 Pedicle Height (**12 PH**)
Superior-inferior height of the pedicle measure in the coronal plane.

13 SCw Center width of the spinous process
Minimum width of the spinous process measured in the coronal plane.

14 SCh Center height of the spinous process
Minimum height of the spinous process measured in the coronal plane that is *not* on the tubercle at the end of the spinous process.

15 SH Height of the spinous process
Maximum height of the tubercle at the end of the spinous process measured in the coronal plane.

Statistical Analysis

Model Development

The statistical analysis utilized SAS software (SAS 2008) for model selection and evaluation. A stepwise-linear model selection procedure (Proc Logistic) was used to select variables which should be the most helpful in the discrimination between male and female. Once the variables with the best selective properties were determined, a quadratic discriminant function was applied (Proc Discrim). The discriminant function is a formula based on a set of means and covariance matrices estimated for each classification group. Using all the data, Proc Discrim creates a generalized squared distance function and a posterior probability is output for each; mis-classification rates are given based on a probability greater than 50%. These classification rates tend to be optimistic, as each point is used in the estimation of the function that classifies itself. This problem is mitigated by a cross-validation process in which the discriminant function is re-estimated for each data point, with that point removed from the data set.

To evaluate the sensitivity of the final model to minor variations in the data set as a whole, the model was reselected 10 times with a random 10% of the data removed each time. The stability of the modes selected by the stepwise method was observed over these 10 runs. For some of these 10 runs additional models that performed well in Proc Discrim were also considered. A cross-validated error rate, generated by Proc Discrim, was compared for each of the models created within and between each run.

The same statistical procedure was used to evaluate the ability to classify the American and Nubian populations. The stepwise procedure, Proc logistic, was used to select variables which would likely be most helpful in the discrimination between 19^{th} century Americans and the Medieval-aged Nubians. These variables were then used by Proc Discrim to estimate a discriminant function to classify each specimen into one of the two sub-populations. Models were then created using each sub-population to test for sex. First, the model created by each sub-population was used to classify itself utilizing the cross-validation function. Then the model created by each sub-population was used to classify the other sub-population. For the second case there was no cross-validation as the sub-population that was classified contained no corresponding point in the classifying data set. Following the same statistical method an attempt was also made to classify samples into age classes to check for model interference caused by age.

Discriminant functions have been used for many years to process metric data in sex determination applications (Falsetti 1995; Frutos 2002; Giles & Elliot 1963; Marino 1995; Ozer et al. 2006). With the advent of high speed computers discriminant analysis has become much more accessible. Complex quadratic discriminant equations utilizing large amounts of data can be quickly created. In this study most of the runs consisting of trials with multiple groups of variables, well over 10,000 discriminant functions were created and solved by the SAS software (2008).

CHAPTER 4: MODEL CREATION AND EVALUATION

The first thoracic vertebra is a strong indicator of sex. Using the entire data set the Proc Logistic procedure selected the best four classifiers, including M (transverse process), SL (spinous process length), FDc (coronal width of the vertebral foramen), and BD (body diameter). The non cross-validated classifications gave a 92.64% accuracy for females, and a 90.32% accuracy for males. Cross-validation estimated an accuracy of 86.76% for females and 89.25% accuracy for males. Using five variables the results were slightly lower: 91.18% for females, 91.40% for males (not cross-validated), and 86.76% for females and 88.17% for males (cross-validated).

Model Stability

To examine stability of the model given for the entire data set, the data set was randomized. A new data set was then created with ten percent of the data removed. This was repeated ten times with the resulting ten data sets each missing 10% of the data. The model was then re-estimated for each of the ten data sets. In Table 4.1 the classifying variables for each of the data sets are listed in the order recommended by the stepwise procedure, Proc logistic. The percentages on the right are the rate at which each was classified in to its group correctly (F% = the percentage of females that are correctly identified as female, M% = the percentage of males correctly identified as males). The four classifiers given by the whole model are in bold, and the best result is underlined. The cross-validated results are reported.

TABLE 4.1—Correct classification percentages between male and female of each decimated sub-population. The variables which were selected by the entire data set are in bold. The best overall result for each run are underlined.

Run	Classifying Variable					F%	M%
1st	**BD**	**M**	**SL**	**FDc**		93	90.4
2nd	BD	MW	SL	FDs		85.71	89.1
	BD	**M**	**SL**	**FDc**		87.3	89.16
	BD	SL	FDc	PH		88.89	85.45
3rd	**BD**	**M**	**SL**	**FDc**		86.44	90.59
4th	**BD**	**M**	**SL**	**FDc**		86.44	89.53
	BD	M	SL	FDs	FDc	86.44	91.86
5th	**BD**	**M**	**SL**	**FDc**		86.89	89.29
	BD	M	SL	FDC	SCw	85.25	90.48
6th	BD	SL	FDs	FDc		87.10	86.75
	BD	**M**	**SL**	**FDc**		85.48	89.16
	BD	M	SL	FDs	FDc	85.42	89.16
7th	**BD**	**M**	**SL**	**FDc**		89.66	88.51
	BD	M	MW	SL	FDc	86.21	91.95
8th	**BD**	**M**	**SL**	**FDc**		87.30	89.02
	BD	M	MH	SL	FDc	84.13	91.46
9th	**BD**	**M**	**SL**	**FDc**		93.22	93.02
10th	**BD**	**M**	**SL**	**FDc**		87.93	87.21

For each four-variable selection the original four variables gave the best result. Twice a five-variable model gave the better result. The stability of these results give confidence in the model selected with these four variables (BD, M, SL and FDc).

Each sub-population of the data set was classified separately into male and female. The late 19th century American data classification rates were much higher than the ancient Nubian data, and much higher than the data set as a whole. The Nubian data set gave a slightly lower classification rate than did the entire data set. In the Nubian data set the variables recommended by the stepwise procedure (Proc Logistic) estimated classification rates that were exactly the same as the rates for a different set of variables (see Table 4.2). This was computed multiple times to make sure no error was made. The combination of variables used for the final model estimation was not the one recommended by Proc Logistic, but testing showed it to give as good a result as the combination that was recommended by the stepwise procedure, Proc Logistic. For the 19th century American data set the preferred model was better overall than the model recommended by Proc logistic. All estimated classification rates are cross-validated.

TABLE 4.2—Correct classification percentages for each sub-population as classified by itself.

Sub-Population	Classifying Variable					F%	M%
CSU	BD	M	MW	SL		69	97
	BD	**M**	**SL**	**FDc**		92.31	93.33
	BD	MW	SL	PH	SH	92.31	91.11
CU	M	MH	SL	FDc		88.89	87.76
	Pli	M	SL	FDc		79.63	85.71
	BD	**M**	**SL**	**FDc**		88.89	87.76*
	M	MH	SL	FDc	SCw	88.89	83.67

*BD M SL FDc was not recommended by *proc logistic* for the Nubian data

Each sub-population was used to classify the other using the standard variables. The 19th century American model classified the Nubian females very well (94.44%) but classified Nubian males poorly (71.13%). The Nubian model classified 19th century American males very well (95.56%) but classified 19th century American females very poorly (61.54%). These results are likely due to the general greater robusticity of members of the 19th century American population. Overall, the model created by the larger American population was inclined to classify several of the more gracile Nubian males as female. Likewise all but the smallest American males will be selected by the Nubian model to be male, and many of the more robust American females will also be sorted into the male category.

TABLE 4.3—Correct classification percentages for each sub-population as classified by the other sub-population. CSU is the American population, CU is the Nubian population.

Classifying pop vs. pop Classified	Classifying variable				F%	M%
CSU vs CU	**BD**	**M**	**SL**	**FDc**	94.44	71.43
CU vs CSU	**BD**	**M**	**SL**	**FDc**	61.54	95.56

The model created by the whole set was then used to classify each sub-population. These results will likely be optimistic because each point in the sub-population was used, in part, to classify itself. Only the previously selected model including BD M SL FDc was used. This combination is used to create the functions described later in this section.

TABLE 4.4—Correct classification percentages for each sub-population as classified by the entire data set. CSU is the American Population, CU is the Nubian population.

Population	Classifying Variable				F%	M%
CSU	**BD**	**M**	**SL**	**FDc**	92.31	95.56
CU	**BD**	**M**	**SL**	**FDc**	92.59	85.71

To test for possible discrepancies between the two populations the same statistical procedure was applied to the American and Nubian populations. Proc Logistic recommended a completely different set of variables than those used for sex determination. The combination of HM, MW, FDs and SCw show that the American population can be correctly identified nearly 90% of the time while the Nubian population can be correctly identified 84% of the time. Whether this difference is due to genetic differences between the two population, or if it is due to differences in lifetime activity is not clear. The width at the center of the spinous process, the sagittal diameter of the vertebral foramen and the middle height of the vertebral body are used in every recommended classification. Notably absent from the classification is the length of the spinous process (SL) which has been key in nearly every experiment thus far.

TABLE 4.5—Correct classification percentages between the American and Nubian Populations.

Classifying Variable					American%	Nubian%
Hm	SW	FDs	SCw		86.21	85.44
Hm	M	FDs	SCw		86.21	85.44
Hm	MW	FDs	SCw		89.66	84.31*
Hm	M	SW	FDs	SCw	87.93	86.41

*Underlined are results that represent the highest classification rate

The use of only male data gives a slightly better result selecting for American specimens (91.11%) and it is slightly worse at selecting for Nubian specimens (83.67%) using either of two, five variable models.

TABLE 4.6—Correct classification percentages between American and Nubian male populations.

Classifying Variable					American	Nubian
HM	SW	FDs	SCw		86.67	83.67
BD	Hm	FDs	SCw		88.89	79.17
Hm	FDs	PH	SCw		86.67	81.25
BD	HM	SW	FDs	SCw	91.11	83.67*
HM	SW	FDs	SCw	SCh	91.11	83.67*

*Underlined results represent the highest classification rates

Because the Nubian population is represented by populations of two different antiquities a comparison of variation between the R and S burial sites was performed. The same statistical procedures (Proc Logistic and Proc Discrim) were used to create the specific model for this test. The best results given utilized 5 or 6 variables. The selected variables again differ from those selected to differentiate between the American and Nubian (whole) populations. The height of the center and the end of the spinous process as well as the length of the spinous process is used in every classification. However the use of only these three variables gives hardly better than a 50% classification rate. The result of over 70% classification for both cemeteries demonstrates that there is some physical difference between the two Nubian populations. (R% = the percentage of R burials correctly identified as R burials, S% = the percentage of S burials correctly identified as S burials)

TABLE 4.7—Correct classification percentages for between the Nubian sub-populations R burial and S burial sites.

Classifying Variable						R%	S%
SL	SCh	CH				52.17	55.88
Hp	SL	SCh	SH			71.01	66.67
M	SL	SCh	SH			53.62	47.06
HM	SL	SCh	SH			68.12	60.61
Hp	M	SL	SCh	SH		71.01	72.73
Hp	SL	FDs	SCh	SH		71.01	60.61
Hp	MH	SL	SCh	SH		73.91	66.67
Hp	M	MH	SL	SCh	SH	76.81	66.67
Hp	MH	SL	FDs	SCh	SH	72.46	63.64

(R% = the percentage of R burials correctly identified as R burials, S% = the percentage of S burials correctly identified as S burials)

Age

An attempt was made, using the data collected, to determine whether age at death had an effect on the outcome of the model. Initially five age classes were created: 17-25, 26-33, 33-41, 41-49 and 50+. The best result classified its age group correctly only 20% of the time. Narrowing the classification to 3 age groups (17-32, 33-49 and 50+) gave a better but still useless result. Sixty three percent of age-group one were classified into that age-group. However 60.61% of 50+ year old specimens were also classified into that (17-32 years old) age group. Age therefore has a negligible effect on sex discrimination.

The final equation

The discriminant function used by Proc Discrim outputs a D^2_j equation for each group to be classified. In the case of this study there are two: one for males and one for females. The results of these two equations are then used in the posterior probability equation to describe either the probability that a specimen is male, or the probability that a specimen is female. Posterior probabilities are useful in that they allow the user to know how confident to be in the assignment of sex.

Equation 4.1 The generalized squared distance function for the j^{th} group is:

$$D^2 j1(X) = (X - \overline{X}(x)_j)^T \operatorname{cov}^{-1}(x)j\,(X - \overline{X}(x)_j) + \ln|\operatorname{cov}(x)_j|$$

X is the 4 by 1 vector of input variables (M, SL, FDc and BD) for the specimen being classified. $\overline{X}_j$ is the vector averages of the four variables in the j^{th} group of the training data set. Cov_j is the covariance matrix for the j^{th} group of four variables in the training data set. Cov^{-1} is the inverse of the covariance matrix, and $|cov_{(x)j}|$ is the matrix determinant. Listed below are the female and male means and their corresponding covariance matrices.

Table 4.8—Female and male means and covariance matrices

	Female Mean	Female Covariance Matrix			
M	24.71603	4.966842	1.450916	-0.09503	0.939594
SL	27.48206	1.450916	5.750849	-0.14357	0.646361
FDc	19.74529	-0.09503	-0.14357	1.680876	0.123939
BD	15.36882	0.939594	0.646361	0.123939	1.853894
	Male Mean	Male Covariance Matrix			
M	28.66796	5.915427	3.92504	0.366937	1.007077
SL	32.17778	3.92504	12.02871	0.012352	1.490272
FDc	21.02441	0.366937	0.012352	2.691355	0.028997
BD	17.39516	1.007077	1.490272	0.028997	2.485538

Equation 4.2 The estimated posterior probability for the classifying observation with input value x into the jth group is:

$$\Pr_j(x) = \frac{\exp(-0.5\, D^2{}_j(X))}{\sum_k \exp(-0.5\, D^2{}_k(X))}$$

For female posterior probabilities, $D^2{}_1(X)$ is the result of the generalized squared distance from the *female* mean function applied to the inputs for the specimen being classified while $D^2{}_2(X)$ represents the *male* generalized squared distance function. For the male posterior probability the values are reversed.

These procedures are too difficult to do by hand and require covariance matrices to be included. A simple input-output Excel function has been created for easy use. Inputs are the X array (M, SL, FDc and BD) with the outputs being the male and female posterior probabilities. In this example the subject is very likely female. The probability that it is female is 99.99% while the probability that it is male is less than 0.00005%.

Table 4.9—Example of inputs and outputs from Excel file created for easy classification using the model created using the entire data set.

Variable	Input
M	29.9
SL	35.54
FDc	24.17
BD	16.88

Prob F	Prob M
0.999954	4.6E-05

Out of 161 specimens, 19 were incorrectly classified. With the confidence levels being 86% for females and 89% for males, the error rate is then given to be 14% for females and 11% for males. Below are the specimens which are incorrectly assigned. For the females four are assigned as male with over 90% confidence while one is nearly 100% male. The other five border on the 50% mark and are sufficiently ambiguous to be considered a minor error. Half of the male errors fall in this sufficiently ambiguous range with five being over 70% confidence as female, and four near or above 90%. Two have a 100% chance of being female according to the predicting model. The results listed below have been cross-validated.

Table 4.3—Individual errors from cross-validated data with male and female posterior probabilities.

Specimen	Sex	F	M	INTO[1]	Specimen	Sex	F	M	INTO[2]
D20	F	0.0001	0.9999	M	R180	M	0.5092	0.4908	F
S104	F	0.0257	0.9743	M	C20	M	0.5261	0.4739	F
S1	F	0.0386	0.9614	M	R15	M	0.5902	0.4098	F
A10	F	0.0640	0.9360	M	R56	M	0.6080	0.3920	F
R33	F	0.4002	0.5998	M	R141	M	0.6473	0.3527	F
R37	F	0.4088	0.5912	M	S5	M	0.7581	0.2419	F
S118	F	0.4230	0.5770	M	S107	M	0.8806	0.1194	F
E12	F	0.4532	0.5468	M	S100	M	0.9285	0.0715	F
R70	F	0.4799	0.5201	M	C14	M	0.9808	0.0192	F
					R158	M	0.9999	0.0001	F

1 Incorrectly classified into males. 2 Incorrectly classified into females

CHAPTER 5: DISCUSSION AND CONCLUSIONS

It is apparent from the results of this study that the first thoracic vertebra is useful in sex determination. For a known and narrowly defined population, classification rates can exceed 95%. In a cross-temporal, cross-racial, and cross-cultural setting the first thoracic vertebra classifies with between 86.76% and 89.25% accuracy. This is consistent with preliminary results for many metric sex determination methods (Dabbs & Moore-Jansen 2010; Giles & Elliot 1963). Even with the benefits of cross-validation and quarantining a portion of a data set for testing, the preliminary tests usually give an overly optimist result. More testing using this method will need to be done to determine the actual classification rates. Based on testing of other metric methods, classification rates will likely be in the mid to upper 80% range. This tendency is illustrated well in Dabbs and More-Jansen (2010) where the calibration model gave an accuracy of 95.7%, a test on a sample left out of the calibration was 92.5% accurate, and a test on an outside sample was 84.4% accurate.

Because the model for this study utilized two disparate populations, the difference between the model's predicted accuracy and the accuracy applied to real world samples may be smaller than that of single-sample studies. This is due to the greater variation gained using two populations. Additionally one of those populations contains a great deal of variability due to the fact that it contains remains from over a thousand year period.

Given the classifying variables (length of the transverse process, coronal breadth of the vertebral foramen, length of the spinous process and the diameter of the centrum) it is unclear whether the hypothesis that musculature would affect the sexability of the first thoracic vertebra has been supported. Due to the location of T1 relative to a large number of muscle systems it was posited that musculature would affect morphology. Only two of the features used in the classification contain a large number of muscle insertions and in each case it is the length of the feature (not breadth or width) that is measured for its selective capacity. It is therefore apparently the size of the first thoracic as it relates to the skeleton in general that affects its classification as to one sex or the other. Musculature does not appear to be among the best indicators of sex affecting the first thoracic vertebra.

Of some concern is the fact that one member of each sex was classified to nearly 100% certainty as the opposite sex. Specimen R158, a Nubian male that was classified with 99.99% certainty as female, is a particularly small individual with all measurements being well below the averages for females. In this case it would appear that R158 was classified by the model as female simply because of his size. This appears to be the case with many of the more gracile males in the study. All but two of the incorrectly classified males are from the Nubian population.

D20, an American female that was classified with 99.99% certainty as male is not statistically a large female. Only the measurement for the coronal diameter of the vertebral foramen is larger than the female average for that variable, and is smaller than the male average. The body diameter measurement is nearly identical to the female average with the length of the spinous process being quite a bit smaller than the female average. The measurement for the transverse process is the smallest in the entire data set. This demonstrates the power of the discriminant function to assign sex not only based on size, but on the relative shape of the vertebra. In this case D20 happens to exhibit statistically male morphology. For S104, a large Nubian female, it is apparent that the incorrect classification is a function of both size and shape. S104's values for every variable closely resemble those of the male averages.

One would expect that most of the male errors would come from the Nubian population, since this population is overall more gracile than the American population and the majority of the female errors would come from the American population. However this is not entirely the case: in both instances most of the errors come from Nubian specimens. This may be due to much greater diversity present in the Nubian population. The specimens from the Colorado Asylum represent a veritable snap-shot of Colorado demography at the end of the 19th century. The Nubian sample represents a long exposure (over 1000 years) to a genetically diverse population and a transitory culture.

Morphological differences between the American and Nubian populations are significant enough that American specimens are correctly classified as American nearly 90% of the time and Nubian are correctly classified 84% of the time. Additionally a discernable difference was found within the Nubian population when broken up by cemetery. Each burial site was correctly identified approximately 70% of the time. These results can have multiple implications. First, it is possible that the overall shape of T1 is genetically determined. This would account for the larger variation between the American population and the Nubian population. The smaller yet still significant difference between the two Nubian cemetery sites can be accounted for by hypothesizing a difference in genetics. There was a large demographic shift in the population around Kulubnarti (Sheridan 1992) during the centuries dividing the cemeteries. This gene-flow event would account for a recognizable difference in the two populations.

Conversely a life-history hypothesis to account for differences between the populations is just as viable as a genetics-based hypothesis. An individual's life history is the product of many factors. Ontological nutrition deficiencies can have a major effect on the adult stature of an individual. Other factors that can affect the skeleton are disease and muscle use. The human skeleton is a dynamic system that responds to multiple stimuli. Hard work can affect musculature and therefore affect musculature's effect on the skeleton. It can be assumed that industrial age Americans led a very different life than that of 5th to 15th century Africans. Additionally it can be assumed that there was a significant difference in living between the two represented Nubian samples. However, the model differentiating the two Nubian populations uses features which may all be plastic (three measurements including the spinous process, the transverse process and the height of the pedicle). Plastic features are those which can change during adult life in response to changes in muscle use. The features mentioned above are all features with multiple muscle insertion points. It is therefore likely that life history was a major factor affecting the differences between the two Nubian groups.

The ability to discern between the two possibilities is obscured by the fact that the variables used to create the best fitting model are equally split between features that are presumed to be genetically determined (structure and support oriented) and those which *may* be affected by musculature (muscle insertion points). The middle height of the vertebral body and the sagittal diameter of the vertebral foramen are primarily structural (load bearing, and not tension bearing and therefore not likely to change greatly in morphology during adult life) in nature and therefore have a morphology which is likely determined by genetics, but also perhaps by gestational and developmental nutrition. The transverse process is an insertion point for several muscle groups and its width may be affected by musculature and other life forces. In addition the center-width of the spinous process is also *possibly* affected by musculature.

Because the specific combination of variables chosen for the best-fit model represents both structural and plastic skeletal features, the best hypothesis to describe the variation between the populations is that a combination of genetics and life history are affecting the result.

Limitations

There are several considerations which will limit the accuracy of this study. As with all studies which rely in skeletal morphology, ethnicity and lifestyle can have a large effect on size and shape. Because of the use of two disparate data sets, neither of which are representative of any modern population, any results gleaned from this study can not necessarily be applied precisely to modern forensic cases. Instead results will be an indication of whether or not the 1st thoracic vertebra is indeed useful for sex classification. Unknown pathological anomalies may also disrupt results of analysis, though these hopefully will be relegated to the tail ends of the statistical results or recognized and discarded from the study.

Variation in the human musculature is also a concern. France (1988, 1990) cautions the use of musculature-affected features of the cranium in sexing cross-cultural groups. While males tend to be larger and more robust than females in the general human populace, there is considerable overlap in nearly all statistical differences between human males and human females. This will undoubtedly be reflected in the results of this, or any other study concerning sex determination. Age may also be a factor, though preliminary analysis showed that age was not an important factor in this study.

Sample size is another consideration which may complicate the replicability of results. As is evident in the results section, the removal of only 10% of the data can shift the classification results a statistically significant amount. With a larger training data set, the model sensitivity to outliers (anomalous data points) is reduced. The training data set used in this study is small and therefore sensitive to outliers. In order to assess the true accuracy of this method an outside population will need to be examined.

The implications for future research from this study are twofold. Within a single known population the first thoracic vertebra proved to be an excellent indicator of sex, and may be useful across populations.

In order to maximize the effectiveness of a first thoracic vertebra-based method in single populations, a new discriminant function would be calculated using data from any population of interest. This tactic could be very useful in archaeological populations for which there is a catalogued contingent, but for which some specimens are found without sufficient cranial or pelvic remains due to scavenging or burning.

As a universal sex determination key, the first thoracic vertebra may give up to and beyond 80% correct classification rates. To test this hypothesis a new discriminant function would be calculated using data from an ethnically and culturally very broad data set. This universal key would be useful in modern forensic investigations, as well as in archaeological settings for which the population of origin is unknown.

REFERENCES CITED

Albanese, J. 2001. A metric method for sex determination using the hipbone and the femur. Paper presented to the Journal of Forensic Sciences 29th Annual Meeting of the Canadian Association for Physical Anthropology, Winnepeg, Canada, 2001.

Albanese, J., G. Eklics & A. Tuck. 2007. A metric method for sex determination using the proximal femur and fragmentary hipbone. Paper presented to the The Anual Meeting of the American Association of Physical Anthropology, Philedelphia, PA, 2007.Bass, W. 1979. Developments in the identification of human skeletal material (1968-1978). *American Journal of Physical Anthropology* **51**, 555-562.

Benjamin, M., H. Toumi, J.R. Ralphs, G. Bydder, T.M. Best & S. Milz. 2006. Where tendons and ligaments meet bone: attachment sites ('entheses') in relation to exercise and/or mechanical load. *Journal of Anatomy*, 471-490.

Birkby, W.H. 1966. An evaluation of race and sex identification from cranial measurements. *American Journal of Physical Anthropology* **24**, 21-&.

Bower, N.W., S.A. McCants, J.M. Custodio, M.E. Ketterer, S.R. Getty & J.M. Hoffman. 2007. Human lead exposure in a late 19th century mental asylum population. *Science of the Total Environment* **372**, 463-473.

Cheng, X.G., Y. Sun, S. Boonen, P.H.F. Nicholson, P. Brys, J. Dequeker & D. Felsenberg. 1998. Measurements of vertebral shape by radiographic morphometry: sex differences and relationships with vertebral level and lumbar lordosis. *Skeletal Radiology* **27**, 380-384.

Dabbs, G.R. & P.H. Moore-Jansen. 2010. A method for estimating sex using metric analysis of the scapula. *Journal of Forensic Sciences* **55**, 149-152.

Dabbs, J.M. 1992. Testosterone measurements in social and clinical-psychology. *Journal of Social and Clinical Psychology* **11**, 302-321.

Dorsey, G.A.A. 1897. A sexual study of the size of the articular surfaces of the long bones in Aboriginal American skeletons. *Boston Medical and Surgical Journal*, 80-82.

Falsetti, A.B. 1995. Sex assessment from the metacarpals of the human hand. *Journal of Forensic Sciences* **40**, 774-776.

France, D.L. 1988. Osteometry at Muscle Origin and Insertion in Sex Determination. *American Journal of Physical Anthropology* **76**, 515-526.

—. 1990. A Cautionary Note in the Cross-cultural Use of the Mastoid Process in Sex Determination. *Journal of Forensic Medicine Istanbul* **6**, 149-456.

—. 1998. Observational and metric analysis of sex in the skeleton. In *Forensic Osteology* (ed.) K.J. Reich. Springfield, IL: CC Thomas.

Freiman, H.C., C.J. Zambrano & A.B. Falsetti. 2008. Use of metric analysis on human thoracic vertebrae as a means to determine sex and to aid in the seriation of isolated vertebrae, 2008.

Frutos, L.R. 2002. Determination of sex from the clavicle and scapula in a Guatemalan contemporary rural indigenous population. *American Journal of Forensic Medicine and Pathology* **23**, 284-288.

—. 2005. Metric determination of sex from the humerus in a Guatemalan forensic sample. *Forensic Science International* **147**, 153-157.

Gapert, R., S. Black & J. Last. 2009. Sex Determination from the Occipital Condyle: Discriminant Function Analysis in an Eighteenth and Nineteenth Century British Sample. *American Journal of Physical Anthropology* **138**, 384-394.

Gest, T.R. & J. Schlesinger. 1995. *MedCharts Anatomy*. New York: ILOC.

Giles, E. & O. Elliot. 1962. Race identification from cranial measurements. *Journal of Forensic Sciences* **7**, 147-157.

—. 1963. Sex determination by discriminant function analysis. *American Journal of Physical Anthropology* **21**, 53-68.

Listi, G.A. & H.E. Bassett. 2005. Test of an alternative method for determining sex from the os coxae: Applications for modern Americans. Paper presented to the 57th Annual Meeting of the American-Academy-of-Forensic-Sciences, New Orleans, Louisiana, 2005.

Marino, E.A. 1995. Sex estimation using the first cervical vertebra. *American Journal of Physical Anthropology* **97**, 127-133.

Ozer, I. & K. Katayama. 2008. Sex determination using the femur in an ancient Japanese population. *Collegium Antropologicum* **32**, 67-72.

Ozer, I., K. Katayama, M. Sagir & E. Gulec. 2006. Sex determination using the scapula in medieval skeletons from East Anatolia. *Collegium Antropologicum* **30**, 415-419.

Pastor, R. 2005. Sexual dimorphism in vertebral dimensions at the T12/L1 junction. Paper presented to the Proceedings of the 57th Annual Meeting of the American Academy of Forensic Sciences, New Orleans, LA. Colorado Springs, CO, 2005.

Phenice, T.W. 1969. A newly developed visual method of sexing os pubis. *American Journal of Physical Anthropology* **30**, 297-&.

Rosenberg, K. & W. Trevathan. 2002. Birth, obstetrics and human evolution. *International Journal of Obstetrics and Gynaecology* **109**, 1199-1206.

SAS. 2008. SAS. In *SAS* (ed.) S.I. Inc. Cary, NC, USA: SAS Institute Inc.

Sheridan, S.G. 1992. Minor and trace element distributions in bone: reconstruction of diagenetic, dietary, and disease patterns in an ancient Nubian population. Dissertation: University of Colorado.

Sutherland, L.D. & J.M. Suchey. 1991. Use of the ventral arc in pubic sex determination. *Journal of Forensic Sciences* **36**, 501-511.

Van Damme, R., P. Entin, B. Vanhooydonck & A. Herrel. 2008. Causes of sexual dimorphism in performance traits: a comparative approach. *Evolutionary Ecology Research* **10**, 229-250.

van Lopik, D.W. & M. Acar. 2007. Development of a multi-body computational model of human head and neck. *Journal of Multi-Body Dynamics* **221**, 175-197.

Yu, S.B., U.Y. Lee, D.S. Kwak, Y.W. Ahn, C.Z. Jin, J. Zhao, H.J. Sui & S.H. Han. 2008. Determination of sex for the 12th thoracic vertebra by morphometry of three-dimensional reconstructed vertebral models. *Journal of Forensic Sciences* **53**, 620-625.