SEX ESTIMATION IN FORENSIC ANTHROPOLOGY: A TEST OF THE KLALES ET AL. (2012) METHOD WITH IMPLICATIONS OF ASYMMETRY

By

Sandra Call

RECOMMENDED:

Alexandra Klales, Ph.D.
Committee Member

Jamie Clark, Ph.D.
Committee Member

Patrick Druckenmiller, Ph.D.
Committee Member

Brian Hemphill, Ph.D.
Advisory Committee Chair

Ben Potter, Ph.D.
Chair, Department of Anthropology

APPROVED:

Todd Sherman, M.F.A.
Dean, College of Liberal Arts

John Eichelberger, Ph.D.
Dean of the Graduate School

27 July 2016
Date

SEX ESTIMATION IN FORENSIC ANTHROPOLOGY: A TEST OF THE KLALES ET AL. (2012) METHOD WITH IMPLICATIONS OF ASYMMETRY

A

THESIS

Presented to the Faculty

of the University of Alaska Fairbanks

in Partial Fulfillment of the Requirements

for the Degree of

MASTER OF ARTS

By

Sandra J. Call, B.A.

Fairbanks, AK

August 2016

Abstract

A sample of 204 American individuals was examined to assess the accuracy and reliability of the three non-metric traits described by Phenice (1969) and revised by Klales et al. (2012) for assigning sex. In addition, the bilateral stability of the three non-metric traits was assessed to determine if asymmetrical expression of the traits compromises the classification accuracy of the revised method, since a prior study found that application of Phenice's original technique yielded low classification accuracy when applied to the right innominate. Klales and colleagues claimed that expansion of the classification system from a dichotomous present/absent scale into five character states and the incorporation of logistic regression based on posterior probabilities vastly improves the accuracy rates for correct sex identification over the original method. Validity of the method developed by Klales and colleagues has not been tested by an external observer on a modern sample of American individuals (individuals who have died within the last 50 years). The current study tests the reliability and validity of Klales et al.'s (2012) technique for assigning sex of both the left and right innominate.

Validity was tested using the sample of innominates 204 individuals from the William Bass Skeletal Collection housed at the University of Tennessee, Knoxville. Intra- and inter-observer agreement was evaluated for Klales and colleagues' method. Intra-observer and inter-observer agreement was statistically evaluated with Cohen's weighted kappa and the intra-class correlation coefficient. A series of Wilcoxon matched-pairs signed-ranks tests were used to evaluate statistical differences in the trait scores between the left and right innominates.

Results show that the Klales et al. (2012) technique yields moderate to high levels of intra- and inter-observer agreement and yields correct sex identifications among individuals of known-sex in 93.6% of cases when all three traits are combined. Accuracy of correct sex

identification was further increased to 99% by re-calibrating the logistical regression equation to fit the sample obtained from the William Bass Skeletal Collection. A Wilcoxon matched-pairs signed-ranks test revealed a statistically significant difference in trait scores of the ventral arc between European and African Americans; however, this difference does not compromise the accuracy of the method for correct identification of sex in known-sex individuals.

Table of Contents

Page

Page

List of Figures

List of Tables

Acknowledgments

To my husband Jason Call, for his unwavering love and support; my family for inspiring me; Dr. Dawnie Steadman for allowing me access to the William Bass Collection; Dr. Michael Kenyhercz and Dr. Alexandra Klales for all their encouragement; my committee members for their mentorship.

Chapter 1 Introduction

1.1 The *Daubert* Criteria

Forensic anthropologists are called upon to assist law enforcement agencies in the identification of the skeletal remains of unknown individuals through estimation of the biological profile. In the last two decades, U.S. courts have presented forensic scientists, including forensic anthropologists, new challenges in the admissibility of scientific evidence and testimony as expert witnesses. In the past, admissibility of scientific evidence determined by a judge, relied solely on the general acceptance of a methodology or technique in the field from which the evidence derived. This procedure was known as the *Frye* standard (Delcarmen, 2010). Under the *Frye* standard, opposing experts were allowed to debate scientific opinion openly in court and without disclosing its scientific framework, allowing the jury to make judgments based on the strength of each party's argument (Steadman et al., 2006). However, a subsequent ruling issued by the United States Supreme Court in *Daubert vs. Merrell Dow Pharmaceuticals* (*Daubert vs. Merrell Dow Pharmaceuticals, Inc*., 1993) changed the way in which scientists and other expert witnesses contribute evidence (Steadman et al., 2006).

Daubert vs. Merrell Dow Pharmaceutical appeared before the U.S. Supreme Court in 1993. Jason Daubert, who suffers from serious birth defects, alleged that Bendectin, an anti-nausea drug manufactured by Merrell Dow Pharmaceuticals, had caused his injuries. The plaintiff's mother had been prescribed the medication to treat pregnancy-related symptoms while pregnant with the plaintiff. Expert epidemiologists testified that the drug, Bendectin had been linked to malformations and birth defects in an experiment conducted on laboratory mice. As pointed out by the state court, the epidemiological evidence had not provided sufficient proof that the drug had caused the birth defect using methods based on theory generally accepted in the epidemiological field. Due to the lack of scientific evidence as outlined by *Frye*, the testimony was ruled as inadmissible (*Daubert vs. Merrell Dow Pharmaceuticals, Inc*., 1993). The Supreme Court had overturned the state court's decision, implementing what is now known as the *Daubert* criteria.

The *Daubert* criteria that undergird the Daubert standards redefine a qualified witness as an expert by knowledge, skill, experience, training, or formal education, and may testify thereto in the form of an opinion. However, their expert testimony must be based on sufficient facts or data obtained through reliable principles and methods. The principles used to support sufficient facts must be applied to the facts of the case (Delcarmen, 2010). The *Daubert* decision demands that scientific conclusions are evaluated on four criteria: (1) conclusions must be based on firmly established theories; (2) the methods must be peer reviewed, (3) the methods must produce estimated error rates to assess validity, and (4) the underlying method must be generally accepted by the peer's community (Melnick, 2005; Christensen and Crowder, 2009).

More recently, the U.S. National Academy of Sciences (NAS) (2009) examined ways in which to strengthen forensic science in the United States in light of the *Daubert* decision. According to the 2009 report, the establishment of quantifiable measures of the reliability, validity, and accuracy of forensic analysis is needed. Furthermore, all results produced by scientific techniques and methods should reflect the degree of uncertainty, error, and bias associated with those results. Both the *Daubert* ruling and the NAS report resulted in a series of changes within forensic anthropology.

1.2 Response to *Daubert* in Forensic Science

In response to the ruling in *Daubert vs. Merrill Dow Pharmaceutical*, forensic anthropologists have been critically re-evaluating their techniques to ensure reliability and accuracy of the methods used for establishing demographic characteristics of the hard tissue remains of unknown individuals (Dirkmaat et al., 2008). The term reliability refers to the repeatability of a technique. Repeatability focuses both on the observations of a single observer at two separate points in time (intra-observer reliability) and the same observation made by two separate observers (inter-observer reliability). Intra- and inter-observer tests measure the level of observer agreement. If observers are found to be in agreement with each other that method is said to be repeatable or reliable. Tests of repeatability are important as they demonstrate the ability of a method to be applied by separate observers, or by the same observer at different times, and yield the same results. The term validity, as it will be used in this thesis, refers to the ability of a technique to render correct classifications. The concepts of both reliability and validity can be evaluated through a multitude of statistical tests. The way in which anthropologists ensure reliability and accuracy is through reliability and validation studies. Validation and reliability studies are critical to scientific credibility. Both types of studies have gained importance for uncovering weaknesses in methodologies and for identifying limitations of existing methods. These studies are instrumental in complying with the criteria outlined by *Daubert*. Discerning how well a method will perform beyond the sample from which it was developed is impossible without the use of validation and reliability studies.

Additional efforts are being made through the revision of existing techniques to bring them up to the *Daubert* standard. One such revision by Klales and colleagues (2012) aimed to improve the Phenice (1969) method which employs the three most widely used non-metric methods of the innominate for estimating biological sex. Phenice's technique and Klales et al.'s sub-sequent revision will be discussed in great detail in chapter two. This study has a two part focus. The first focuses on testing the performance of the Klales et al. (2012) revision on a modern (those who have died within the last 50 years) U.S. sample and the second focuses on bilateral expression of those traits used in the method and gaining optimal classification power.

1.3 Aim

In light of the recent changes to the admissibility of scientific evidence used in forensic cases and forensic anthropology's due diligence to uphold those standards, this current study has four primary objectives: 1) to test the validity of the Klales et al. (2012) method in order to firmly establish its accuracy as applied to a sample of modern inhabitants of the United States, 2) to test the repeatability of the method to ascertain its reliability, 3) to assess bilateral expression of Phenice's traits using Klales et al.'s technique to determine whether asymmetry impacts sex classification accuracy and, 4) to re-calibrate the original regression classification equation published by Klales and colleagues to establish a sample-specific equation to attain optimal classification accuracy for a modern U.S. sample, thereby mirroring the appropriate population for modern forensic cases to which the method may be applied,

1.4 Hypotheses

This research examines the validity and reliability of Klales et al.'s (2012) revised method of Phenice (1969) and tests whether asymmetry, if present, compromises the accuracy of the revised method. If asymmetry is found to exist, its pattern and relationship with biological sex will also be examined. The aims of this research will be addressed through consideration of the following four hypotheses.

H_0: Klales et al.'s (2012) revised method of Phenice (1969) is a valid technique, such that the rate of correct classification by sex among known-sex individuals in a specific sample of modern inhabitants of the United States is greater than 85%.

H_1: Klales et al.'s (2012) revised method of Phenice (1969) is not a valid technique, such that the rate of correct classification by sex among known-sex individuals in a specific sample of modern inhabitants of the United States is less than 85%.

H_0: Klales et al.'s (2012) revised method of Phenice (1969) is sufficiently reliable, such that error rates in repeatability occur in fewer than 5% of cases.

H_2: Klales et al.'s (2012) revised method of Phenice (1969) is not sufficiently reliable, for error rates in repeatability occur in greater than 5% of cases.

H_0: Klales et al.'s (2012) revised method of Phenice (1969) is not compromised by asymmetry between sides, for less than 5% of cases yield opposite sex identifications by morphological features.

H_3: Klales et al.'s (2012) revised method of Phenice (1969) is compromised by asymmetry between sides, for greater than 5% of cases yield opposite sex identifications by morphological features.

H_0: A re-calibration of Klales et al.'s (2012) ordinal LR equation to fit a specific sample of modern inhabitants of the United States increases the rate of correct classification by sex among known-sex individuals.

$H_{4:}$ A re-calibration of Klales et al.'s (2012) ordinal LR equation to fit a specific sample of modern inhabitants of the United States does not increase the rate of correct classification by sex among known-sex individuals.

The eighty five percent cut-off established in hypothesis one reflects the level of accuracy generally accepted in the anthropological community (Schmitt et al., 2006). The five percent cut-off in hypothesis two represents the generally accepted critical value for acceptance or rejection of a null hypothesis in statistical analysis (Bremer and Doerge, 2010). The five percent cut-off established in hypothesis three represents a pragmatic rejection rate for incorrect assignments by sex with Klales et al.'s (2012) methodology for the specific sample considered in this analysis. This threshold should not be considered as an "error rate" for generalization beyond this specific sample as the accuracy rates for acceptance of the alternative hypothesis require replication studies and/or a far lower critical value (e.g., $p < 0.0047$) (Sellke et al., 2001).

1.5 Significance

This study is significant for forensic anthropology, bioarchaeology, and paloeanthropology. The results of this research could possibly establish validity and reliability to a statistically sound method of sex estimation that meets the criteria outlined by *Daubert*. A re-calibration could demonstrate the flexibility of the method to conform to specific samples, attaining optimal classification accuracy beyond the sample of which the method was developed. The Klales et al. (2012) method has exhibited great potential for use in modern forensic cases and in samples outside of the United States; however, it has not been tested on a separate U.S. sample by an independent third party. Validation of the method, should it prove reliable, could result in its addition to standard operating procedures (SOP) within forensic laboratories throughout the U.S. In addition, the present research examines the symmetrical stability of these traits used in the Klales et al. (2012) method and investigates how asymmetry may affect this method for sex estimation.

Chapter 2 Forensic Anthropology, Estimation of Sex from Skeletal Remains, and the Relative Efficacy of Metric and Non-metric Variables for Sex Estimation

2.1 History of Forensic Anthropology

The historical development of forensic anthropology may be envisioned as encompassing three periods. The early 1800s to late 1930s marks the Formative Period. The Consolidation Period falls between the late 1930s to the early 1970s, and the Modern Period begins in the 1970s and represents the current state of forensic anthropology as we know it (Byers, 2011). The *Daubert* decision (*Daubert vs. Merrell Dow Pharmaceuticals, Inc.,* 1993) has introduced a paradigm shift within the discipline whereby forensic anthropologists are being called upon to substantiate their assertions with scientifically tested methods (Dirkmaat et al., 2008). Some suggest that this paradigm shift is signaling a new era of forensic anthropology.

During the Formative Period anatomists and law enforcement agencies began to recognize the value of the information provided by the study of the human skeleton to forensic cases (Byers, 2011). In 1849, one of the first forensic anthropological-like cases involved professors of anatomy Oliver Wendell Holmes Sr. and Jeffries Wyman who assisted in reassembling unknown human remains and provided law enforcement officials with the sex, ancestry, approximate height, and age of the decedent. The work of these anatomists led to the identification of the individual, and ultimately the conviction of the assailant. This case demonstrated the effectiveness of such methods in medico-legal investigations (Byers, 2011; Christensen et al., 2014). Thomas Dwight, credited as being the "father of anthropology" by T. Dale Stewart (Stewart, 1954), began publishing several articles and essays on the identification of human skeletal remains, which later gained attention from the FBI (Dwight, 1905).

Wilton Krogman's 1939 seminal publication of *Guide to the Identification of Human Skeletal Material* signaled the transition of forensic anthropology out of the Formative Period and into the Consolidation Period (Byers, 2011). Krogman's paper and similar publications consolidated what anthropologists knew about the identification of unknown human skeletal remains. These publications gained popularity within the FBI and became companions to those involved in medico-legal investigations. The involvement of the U.S. in World War II and in the Korean conflict significantly impacted the development of forensic anthropology. Service members killed and recovered from the battlefield were returned severely decomposed, thereby inhibiting traditional identification methods such as visual confirmation. As a consequence, the Central Identification Laboratory (CIL) was established by the U.S. Army Office of the Quartermaster to account for, identify, and lay to rest those service members. Charles Snow was selected to direct the laboratory and he actively solicited help from other anthropologists, such as Mildred Trotter and Goldine Gleser to create methods for estimating stature (Trotter and Gleser, 1952) and Thomas McKern and T. Dale Stewart to estimate age-at-death (McKern and Stewart, 1957). The CIL remains in operation today and is actively involved in efforts to identify and account for all those missing in action from past conflicts.

In 1972, Clyde Snow and Ellis Kerley managed to recruit enough of their colleagues to form a Physical Anthropology section within the American Academy of Forensic Sciences (AAFS) (Byers, 2011; Christensen et al., 2014). The formation of the new branch established forensic anthropology's relevance in medico-legal investigation, launching the field forward into the modern period. Quality assurance measures were needed for those practicing forensic anthropology in the United States leading the development of the American Board of Forensic Anthropology (ABFA). ABFA established a vigorous certification process in which qualified anthropologists holding specified prerequisites can become board certified diplomats. Today, the field of forensic anthropology is recognized as a sub-field of physical anthropology.

2.2 Methods of Biological Profile Estimation

When called upon to assist in the identification and recovery process, forensic anthropologists examine skeletal remains to estimate what is known as the "biological profile." The biological profile consists of the demographic details of the deceased; specifically their age, ancestry, sex, and stature. Many of the methods used to estimate each component of the biological profile derive from those developed in earlier periods of forensic anthropology. The estimation of age, sex, and stature has been expanded upon considerably as much of the early literature, particularly during the Consolidation Period, was based on skeletal material of male service members killed in action.

Estimating sex from skeletal remains involves the evaluation of traits that differentiate between males and females; these differences are referred to as sexual dimorphism. Sexual dimorphism is the consequence of sex specific hormones during growth and development (puberty), differential locomotive patterns, and reproductive function required in females (Byers, 2011). In bio-archaeological contexts, the accurate estimation of sex impacts the reconstruction of past populations based on demographic information. In a forensic context, the estimation of sex is one of the first steps in the identification process. The correct assignment of sex effectively eliminates approximately 50% of the population in search of missing person's records and databases. Furthermore, many methods of age, stature, and ancestry estimation are sex specific.

2.3 Metric Analysis

Techniques for sexing human skeletal remains are organized into two categories; metric and non-metric. Metric methods involve the collection and analysis of osteometric measurements based on landmarks, semi-landmarks, and sliding landmarks, while non-metric methods are based on the visual assessment of morphological traits. Tools used to collect measurements from skeletal remains include but are not limited to sliding calipers, spreading calipers, osteometric boards, measuring tapes and recently, two-dimensional digitizers and three-dimensional scanners. The points at which these measurements are collected are called landmarks. There are several types of landmarks that denote different points of origin on the skeleton. For example, type I landmarks are points defined by a juxtaposition of different tissues (Bookstein, 1991). Type I landmarks denote an origin on the skeleton where multiple bones meet and create a discrete point of convergence. Type II landmarks are points of maximum curvature, such as the apex of the greater sciatic notch. Distances between two landmarks can be described as the distance from point "*a*" to point "*b*" or as a ratio (such as mandible length: mandible height). The measurements collected are used to identify patterns of variation within and between individuals or between samples such as the difference in long bone length between males and females.

Metric methods are often praised as being objective as osteometric landmarks are said to be well defined, in most cases homogeneous, and are easier to locate on skeletal material than the visual descriptions often employed in non-metric methods. However, are not without their own limitations. Well-defined landmarks are not as easy to locate as many have assumed (Drew, 2013). Metric landmarks can be obscure, especially in cases dealing with highly eroded skeletal material. Furthermore, most metric techniques require a suite of measurements for at least two or more landmarks which often requires the material to be mostly complete and well preserved. In bio-archaeological and forensic contexts bones are often eroded, fragile, and highly fragmented.

Furthermore, metric methods sometimes require specialized equipment and extensive training on several soft and hardware systems to thoroughly analyze the metric data, such as two and three dimensional scanners and digitizers.

2.4 Metric Methods of Sexing the Adult Innominate

In his 1948 paper, *Sex Differences in the Pubic Bone*, Washburn states that despite numerous efforts, no method of determining the sex of unknown remains had yet been discovered. Piqued by his observations of sexual differences in the pelvic bones of monkeys, Washburn set out to search for similar differences in the human innominate. Using similar measurements developed by Schultz (1930) to compare animals of varying size, Washburn created an index (ischium length/pubis length) using a sample of 300 adult human skeletons. One of the landmarks used to measure the length of the ischium is the point at which the ischium and pubis join within the acetabulum. Historically, this landmark has been considered difficult to locate as it requires the observer to determine a transition in the thickness of bone and relies on the location of an irregularity represented by a notch (Stewart, 1954; Drew, 2013). To complicate matters further, Washburn does not offer an illustration to aid in locating this landmark. This ambiguous landmark consequently has been misinterpreted from the original description, leading analysts to use the center of the acetabulum (Seidler, 1980) as a base point instead. Nonetheless, Washburn claims that a classification accuracy rate in excess of 90% can be achieved using the ischium-pubis index.

Drew (2013) tested Washburn's 1948 index on a small sample from the Mary Rose, a 16^{th}-century English warship lost in a documented disaster. While the sample is small (54), it is believed to be comprised of all males. Of the 54 presumed males, 80% were correctly classified as male using Washburn's (1948) description. Additionally, Drew took measurements using the center of the acetabulum as the base point as described by Seidler (1980) and re-calculated the index. Classification dropped considerably with correct classifications in only 4% cases. Drew's study demonstrated some of the risks in misinterpretation of obscure landmarks when using metric techniques for data collection; however, Washburn's index has been applied successfully by other researchers (Sachdeva et al., 2014).

Sachdeva et al. (2014) tested Washburn's index on a North Indian sample of 100 individuals of known sex. The sample derived from the Department of Anatomy, Government

Medical College, of Punjab, India. The pubic length and height of the ischium were measured in accordance with Washburn (1948). Sachdeva and colleagues found that the pubis is significantly longer in females while the height of the ischium is consistently longer in males. The calculated mean ischio-pubic index was significantly different between the sexes ($p < 0.001$). The index correctly classified sex in excess of 98% of cases. The authors concluded that the ischio-pubic index is a valuable indicator of biological sex in the North Indian Population.

Using measurements of the width and height of the greater sciatic notch, Letterman (1941) found statistically significant differences between males and females. Letterman (1941) found that the mean greatest width of the sciatic notch is larger in females, while the greater depth is greater in males. However, Letterman also reports greater variation about the mean among males than among females. Singh and Potturi (1978) also measured width, depth, and length of the greater sciatic notch for the purpose of differentiating between males and females; however, contrary to Letterman (1941), the authors found the width and depth of the greater sciatic notch to be unsuccessful for differentiating the sexes. The authors instead propose that the posterior angle of the sciatic notch to be a better indicator of sex (Singh and Potturi, 1978).

Another study focused on the sciatic notch took a geometric morphometric approach to the examination of sexual dimorphism (Steyn et al., 2004). The aim of this study was to assess the greater sciatic notch of South African black and white males and females to ascertain if geometric morphometrics could discern whether differences in shape could be quantified using two-dimensional landmarks on photos. This study utilizes Relative Warp Analysis (RWA), which is similar to principle component analysis, to determine whether differences and/or similarities in shape are present in the four subgroups (black males, black females, white males, and white females). The sample used in this study derived from the Department of Anatomy, University of Pretoria, South Africa.

The authors found that both black and white South African females displayed a predominately wide shape (96% and 86%). Ninety-six percent of black males took on the typical narrow, deep male shape, while only 33% of white South African males displayed the typical male pattern. The authors attributed the wider sciatic notch in white males to stature or overall size. Steyn et al. (2004) states that South African white males have larger bones, and are thus taller than their black counterparts. Therefore, the author concluded that the shape of the sciatic notch is not a reliable sex estimator in the South African population at large.

Gonzalez et al. (2009) also analyzed the greater sciatic notch as well as the ischio-pubic complex using geometric morphometric analysis. Semi-landmarks, arbitrarily placed landmarks along a curve, were collected from a sample of 121 individuals of known sex from the Museum Anthropologico de Coimbra in Portugal. Each innominate was photographed in an identical position (auricular surface facing upward). Fourteen semi-landmarks were digitized along the margin of the sciatic notch, and two landmarks and semi-landmarks were digitized on the ischio-pubic region from the photographs. Generalized Procrustes Analysis was used to remove effects of landmark and semi-landmark configurations. Discriminant function (leave-one-out cross validation) and *k*-mean clustering was used to estimate sex. The authors found that the ischio-pubic complex (90%) outperformed the sciatic notch (91%) in classification accuracy. The authors also noted that females were misclassified more frequently than males, especially with regard to the ischio-pubic complex. Contrary to Steyn et al. (2004), Gonzalez et al. (2009) supports the utility of geometric morphometrics as a reliable methodology to discern pelvic shape differences between males and females.

Klales et al. (2009) also took a geometric morphometric approach to biological sex estimation. Klales et al. (2009) collected a suite of 23 type II landmarks from the entire innominate to ascertain their utility in discerning between male and female pelvic morphology and between different ancestry groups. The sample was derived from the Hamann-Todd Osteological Collection. The 253 inter-landmark distances were measured using a Microscribe G2 digitizer and 3Skull software. The inter-landmark differences were analyzed using FORDISC 3 (Jantz and Ousley, 2005). Stepwise selection was performed to isolate the best combination of smaller number of variables that produce the most accurate classifications. Five variables (inter-landmark differences) separated the sexes with 99% accuracy. The inter-landmark differences suggest that females have relatively larger pubic lengths and smaller ischial heights than males, which is supported by Sachdeva and colleague's (2014) study. The authors demonstrated the utility of the measurements to correctly classify biological sex of males (94%) and females (97%) of cases in a sample from the HTH.

Most recently, Bytheway and Ross (2010) selected 26 landmarks on the male and female innominate previously reported in the geometric landmark literature to differentiate between the sexes. Ten additional landmarks were selected by the authors to potentially capture the entire shape of the innominate. Males and females from the Terry Skeletal Collection housed at the

National Museum of Natural History were used in this study. Coordinates were collected directly from each individual innominate using a 3D digitizer. Generalized Procrustes Analysis was used to remove the effects of size. Discriminant function analysis (leave-one-out cross validation) was performed to for each group. Discriminant analysis revealed 100% correct classification. The authors conclude that the high classification accuracy is due to the inclusion of areas of variability between males and females that cannot be captured through traditional analysis.

2.5 Overview of Metric Methods

As these studies have demonstrated, metric methods are of great value to identifying sex of unknown human remains when landmarks are clearly defined and are identifiable by outside researchers. Each method has its own strengths and weaknesses; furthermore, each requires specialized equipment and extensive experience and training. The use of geometric morphometrics is beginning to become more commonplace as the utility of new landmarks are being discovered and proving useful in discriminating between anatomical shape and form of males and females.

2.6 Non-metric Analysis

A great deal of early research on non-metric traits focused on variants of the cranium. Influenced by evolutionary science of his time, Haeckel (1879) was convinced that many non-metric variants observed on skeletal remains were vestiges of the evolutionary stages through which the developing organism had passed (Saunders, 1989). Haeckel believed that over the course of time evolution had added new stages to produce life forms and non-metric traits were the evidence of an organism's descent (Haeckel, 1879). Many subsequent studies offered descriptive analyses of bony non-metric traits and their relationship with soft tissue structures, but added little to understanding the biological causation of such traits.

It was not until 1952 that Gruenberg designed an experiment to explore the genetic nature of non-metric skeletal variants. Observing skeletal variants on mice, Gruenberg found that single gene mutations could induce several minor skeletal variants as part of their syndromic effects (Berry and Berry, 1967; Saunders, 1989) in strains of inbred mice. He found that any single non-metric trait can be altered by mutational events; those non-metric traits were determined to be inherited. However, in cross bred strains of mice, incidences of non-metric traits had not followed a Mendelian pattern of inheritance; this meant that the presence or discontinuity of a trait is determined by a physiological threshold. As a consequence of Gruenberg's studies, a quasi-continuous model was devised proposing that a trait has an underlying continuity of visible expression; those above the threshold are affected, while those below the threshold remain unaffected. The continuous manifestation of a trait is both genetically and environmentally influenced.

Drawing on the work of Gruenberg, Berry and Berry (1967) found distinct ancestral differences between the incidences of non-metric traits in the cranium of Egyptians and Palestinians. Berry and Berry contributed a list of non-metric cranial traits along with a proposed statistical calculation to quantify biological difference between populations. The authors suggest that intrinsic factors such as age and sex have minimal effect on the appearance of non-metric traits observed human skeletal material. Berry and Berry concluded that the use of non-metric traits is superior to metric data in population studies. Since Berry and Berry's (1967) paper, non-metric traits have been confirmed to show statistically significant differences in sex, age, and side on which non-metric traits appear (Corruccini, 1974; Saunders, 1989). Many researchers who tested Berry and Berry's (1967) technique reported challenges in assessing the non-metric traits outlined in their paper (Corruccini, 1974). Berry and Berry acknowledged the inherent subjectivity in scoring non-metric traits and suggest that agreed upon criteria would help with classification.

Today, non-metric traits have been extended beyond population studies using the skull. Non-metric traits of the pelvis are arguably the best indicators of biological sex on skeletal remains (Letterman, 1941; Phenice, 1969; Walker, 2005; Klales et al., 2012). Non-metric methods for the purpose of estimating biological sex involve visual assessment of traits that may be difficult to quantify on a metric scale. While linear measurements can tell us how far or close two points are away from each other on a specimen only semi-landmarks, often utilized in geometric morphometrics, can tell us about its shape or structure. Often non-metric methods involve the assessment of presence/absence, degree of expression, or the overall morphology of a bone or specific trait. Non-metric techniques are often criticized as being subjective, less standardized, and prone to bias induced by inexperienced observers (Bruzek, 2002). However, despite these perceived challenges, non-metric techniques continue to be the preferred method of analysis by anthropologists as they are practical (Bass, 2005). Non-metric analyses can be performed quickly in the field, or on the scene, without the use of specialized equipment (Byers, 2011). Furthermore, difficult considerations in devising measurements that capture subtle, visually apparent morphologies aren't required. Lastly, and of great importance, such methods can often be applied to fragmentary remains, while many metric methods cannot.

An early approach to non-metric analysis used decision tables and charts to draw conclusions. While many tables and charts were devised to address a specific question, the application of the technique followed the same general protocol. Decision tables list the traits in question with corresponding columns providing descriptors of each trait such as present or absent, or small, intermediate, and large. The column with the most marks represents the decision considered most likely correct. For example, when analyzing traits of the skull for the purpose of assessing sex, more large traits are in favor of a male, while a majority of smaller traits are indicative of a female. In this context, the descriptors large and small are relative to the average male and female within a population. This does not include individuals suffering from pathological conditions causing abnormal or excessive bone growth. However, the terms large and small are ambiguous and mean almost nothing to the analyst with little experience, or comparative models. A drawback to this method lies in knowing whether all the pertinent traits are included on the table used to differentiate say between males and females. Another drawback lies with the descriptors, and capturing all possible characters for a specific trait. However, when paired with other lines of analyses decision tables have proven useful.

Non-metric techniques have made progress from those early days of obscure decision tables and have proven to be quite powerful when paired with statistical analysis. Recent studies (Walker, 2008; Hefner, 2009; Klales et al. 2012) have taken traditional non-metric traits of the skull and pelvis used by anthropologists to determine ancestry and sex and have re-calibrated them along an ordinal scale. Discriminant function and ordinal logistic regression analysis produce classification accuracies upward of 80 to 90%. These new methodologies also provide a means to determine the strength of estimation through posterior probabilities. These studies have demonstrated a way in which non-metric traits and methodologies, traditionally criticized as being void of statistical power, may be used, and have demonstrated their ability to make correct classifications, assess observations on a standardized scale, and produce quantitative evidence to support classification accuracy.

2.7 Non-metric Methods of Sexing the Adult Innominate

Several non-metric methods for estimating sex using bones of the innominate have been proposed. The most common methods assess features found on the innominate, the greater sciatic notch, and preauricular surface of the ilium.

Events related to parturition have been used in anthropological studies for the purpose of sex estimation. The pre-auricular sulcus has been credited as a means for sex estimation. Houghton (1974) reports two forms of the pre-auricular sulcus. The first form has been found on the ilium of both males and females and is usually wider and deeper. The second form, which is exclusive to females, is short and shallow. Using a sample of 119 pelves (54 males; 65 females), Houghton found that 81% of males and 71% of females were found to have the first form of the pre-auricular sulcus. What Houghton concluded was that the first form of the sulcus is a consequence of the attachment for the inferior portion of the ventral sacro-iliac joint. However, the second form of the sulcus, exclusive only to females, is the consequence of pregnancy. Only 23% of females were found to have the second form, while no males possessed a short shallow sulcus.

Kelley (1978) also examined pelvic traits associated with parturition. Kelley concluded that the pre-auricular groove is a sensitive indicator of parturition. However, this sulcus is occasionally found in nulliparous women. Kelley suggests that the combination of dorsal pitting and preauricular grooves yielded the highest degree of reliability for determining parity. Later, a radiologic investigation by Dee (1981) involving 100 men and 200 women revealed that while the sulcus is indeed a female trait, it can only be observed in approximately 25% of women. While these traits have been identified as reliable in the investigation of parity, the absence of these traits are not definitively indicative of a male.

A later test by Novak et al. (2012) of the preauricular sulcus observed 94 individuals from the William Bass Skeletal Collection (WBSC) and 104 individuals from the Terry Skeletal Collection (Novak et al. 2012). Observations of the sulcus were based on Buikstra and Ubelaker (1994) which illustrate a scale reflecting five possible character states the preauricular sulcus may manifest for standard data collection purposes. Novak and colleagues (2012) found that only 63% of females and 5% of males exhibited the trait. The authors concluded that while some males had a sulcus that it was the least developed of possible characters. While this trait correctly classified males in 94% of cases, correct classification of females occurred in only 63% of cases.

The sciatic notch in females is wide and shallow in women and narrow and deep in males (Letterman, 1941). While this observation is easily seen by the naked eye, lack of consensus over landmarks that thoroughly capture the differences have prevented metric assessment of this trait (Walker, 2005). A visual non-metric approach was proposed by Walker (2005) and modeled after a system created by Acsadi and Nemeskeri (1970). The system assigned scores ranging between -2 and +2 based on drawn illustrations of the sciatic notch, with a score of zero representing an androgynous form intermediate between male and female. According to Walker (2005), a common issue arising from the usage of Acasdi and Nemeskeri's (1970) scheme is that it was based on the distribution pattern of a European sample. Using samples from the Hamann-Todd, Terry, and St. Bride's collection, Walker aimed to find an equal representation of sciatic notch variation between adult males and females of African and European ancestry.

Walker's (2005) technique is valuable in situations where the anterior portion of the pelvis is eroded. Generally, in females, the sciatic notch is relatively wider than the typically deep narrow shape found in males. To address issues related to subjectivity, the sciatic notch is scored along a 5-point ordinal scale used in comparison to the skeletal remains in question. In applying the technique, sciatic notches are assigned scores corresponding with the shape and width of the sciatic notch pictured on the scale: -2 = definite female, -1 probably female, 0 androgynous, +1 probably male, and +2 definite male (Walker, 2005). Results revealed little overlap between the sexes on the extreme ends of the ordinal scale. Using 165 males and 131 females (n = 296) from the Hamann-Todd, Terry, and St. Bride's collection Walker (2005) used his ordinal scale to assign sex. Females were correctly sexed with 88% accuracy, while males were correctly sexed with 91% accuracy. Furthermore, when comparing Acsadi and Nemeskeri's scale, Walker found that a score of 2 in Acsadi and Nemeskeri's (1970) scheme representing a female form was more indicative of an intermediate form throughout his sample. Since then, Walker's adjusted ordinal scale has been used in forensic cases in the United States.

Based on previous studies of biological sex estimation, Bruzek (2002) combined several non-metric techniques and proposed a new holistic technique for sex determination using traits of the entire innominate. The aim of Bruzek's study was to address the issues of subjectivity and inexperience of new analysts, which are some of the drawbacks associated with non-metric techniques. Bruzek's solution involved using only three possible scores for five features of the innominate, which had previously been described as present or absent, or expanded on an extensive ordinal scale. Bruzek argues that the extensive ordinal scale divides each trait into several character states which makes evaluation difficult (Bruzek, 2002). The five features selected by Bruzek include the prearicular sulcus, greater sciatic notch, composite arch, inferior pelvis, and ischio-pubic proportions. The combined classification accuracy on two European samples ranged from 93 to 98% when using all five features (Bruzek, 2002). Listi and Bassett (2006) tested Bruzek's (2002) method using an American sample to ensure its applicability in the United States. Classification accuracy ranged from 90 to 92% correct depending on the experience of the observer with more experienced observers attaining higher accuracy than lesser experienced observers (Listi and Bassett, 2006).

The traits of Phenice (1969) are the most extensively used traits for estimating the sex of human skeletal remains. His visual method and subsequent revisions will be discussed in detail in the following sections.

2.8 Overview of Non-metric Methods

Non-metric analyses, however practical, are often criticized as being subjective (Rogers and Saunders, 1994). This is especially true when utilizing non-metric methods to determine demographic characteristics from unknown skeletal remains. However, it is often difficult to define non-metric traits and to devise measurements that adequately capture variations in shape that may be apparently visible to the eye (Walker, 2008). Observers of non-metric traits often report some confusion in identifying a prescribed trait or discerning a dividing line between its presence or absence (Christensen et al., 2014). Close examination of non-metric traits often recognize more than two states of trait manifestation which is a common critique of Phenice's method for identifying sex in skeletal remains. Phenice (1969) and the subsequent revision of his method by Klales et al. (2012) are the focus of the current study and will be discussed in detail below.

2.9 Phenice (1969)

Grant and Boilean (1965) suggested that the structure between the crus penis and crus clitoris, the ventral arc, and sub-pubic concavity are inherently variant between males and females. It is worth noting that the area of attachment of the crus penis and crus clitoris described by Grant and Boilean is found on the ischio-pubic ramus; however, the ischio-pubic ramus alone as an indicator of sex was not as discrete as the ventral arc and sub-pubic concavity (Grant and Boilean, 1965; Phenice, 1969). Grant and Boilean claimed that in almost every case, the ventral arc and sub-pubic concavity was either absent or present without intermediate forms. The claim of non-intermediate forms of the traits described by Grant and Boilean is highly significant. Such a bimodal distribution suggests that the traits are developmentally stable, completely homogenous when found and easily identifiable so that patterns of variation may be predicted easily and with great accuracy. This implies that even in the hands of researchers with little experience, sex can be reliably estimated using the three aforementioned traits.

Motivated by the research of Grant and Boilean (1965) T.W. Phenice (1969) proposed a new method for sexing human skeletal remains using these three traits of the innominate. The sample used in Phenice's study derived from the documented Terry Skeletal Collection which is dominated by black and white males and females. The cadavers were primarily collected from both local hospitals and institutional morgues throughout the state of Missouri, and are comprised of individuals born between 1822 and 1943. The age range within the collection is 16 to 102 years of age with an average of 45 (Hunt and Albanese, 2005).

Phenice tested his assertion that the three traits could accurately differentiate males from females using 275 adult individuals of known sex from the Terry Skeletal Collection. Phenice describes the traits and provides an illustration (Figure 1).

<u>The Ventral Arc (VA)</u>. The ventral arc is a slightly elevated ridge of bone found on the ventral surface of the pubis that extends from the pubic crest and bends laterally from the pubic symphysis as it trends caudally. The arc extends inferiorly to the subpubic concavity (Figure 1). While Phenice argued that the ventral arc is solely a female condition, it should be noted that that males may also possess a similar ridge of bone; however, it should not be confused with a ventral arc.

<u>The Subpubic Concavity (SPC)</u>. The subpubic concavity, which should be viewed from the dorsal surface of the pubis, is described as a lateral recurve, inferior to the lower margin of the pubic symphysis (Figure 1). According to Phenice, this recurve is also solely a female condition, although a slight concavity can be found in males as well.

<u>The Medial Aspect of the Ischio-pubic Ramus (MA)</u>. The ischio-pubic ramus is located caudal to the symphyseal surface of the pubis. Phenice describes this area to be broad and flat in

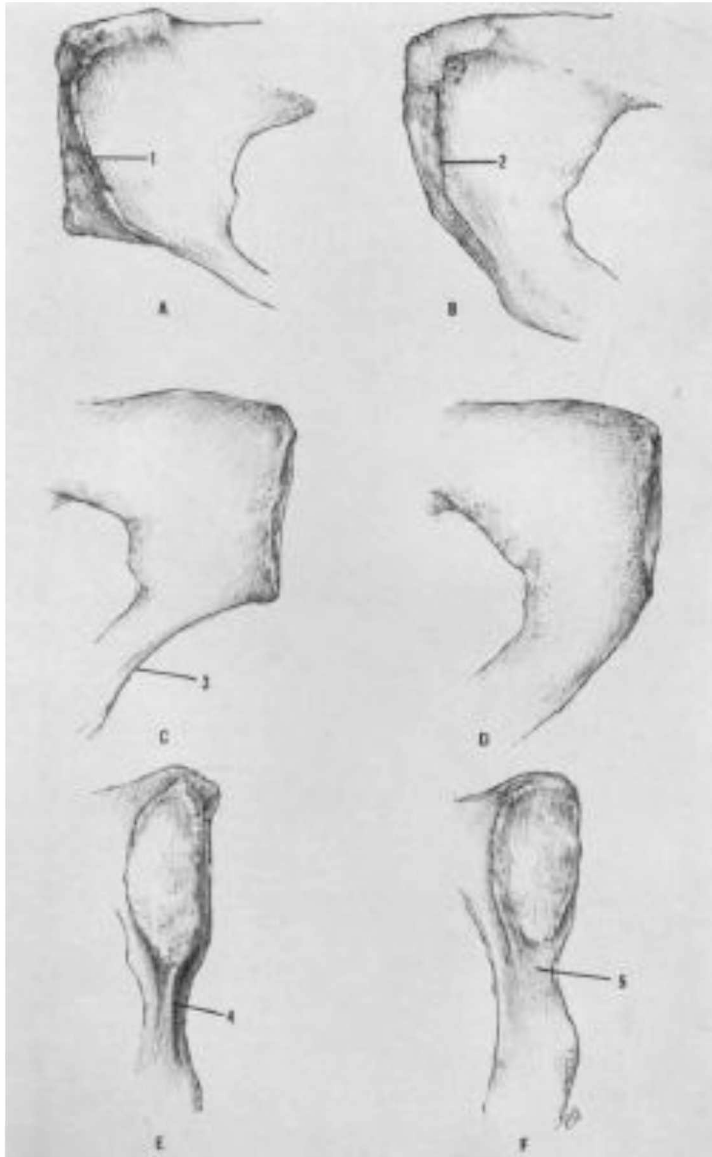

Figure 1. Phenice's (1969) non-metric traits of the female (left) and male (right) innominate. From top to bottom: ventral arc (1), the sub-pubic concavity (2), and the medial aspect of the ischio-pubic ramus. Adapted from "A newly developed visual method for sexing the os pubis," by TM Phenice, Am J Phys Anthropol, 1969, 30, 297-302. Copyright by Wiley, adapted with permission.

males, while in females lateral constriction results in a generally narrower ramus that includes a ridge of bone (Figure 1).

According to Phenice (1969) the presence of the three traits is the female condition while the absence of the traits is the male condition. Phenice recorded the presence and/or absence of the three traits on his sample. In the event that one trait may be too obscure to distinguish, the other traits were used to determine biological sex using the majority rule. With this approach, all three traits are weighted equally in their discriminatory power to discern males from females.

Of the 275 individuals tested by Phenice only 11 were sexed incorrectly (96%). Phenice's results were separated by sex and ancestry (Table 1). Importantly, and by his own admission, Phenice conceded that not all individuals were perfect males or females as Grant and Boilean (1965) previously state. When all three traits were not in agreement, Phenice suggested that at least one of the three traits would be obviously indicative of male or female and could be used to accurately estimate sex.

Table 1. Classification accuracies (%) of individuals correctly classified by biological sex Phenice (1969). (From Phenice, Am J Phys Anthropol, 1969, 30, 297-302.)

	Negro	White	Total
Males	95%	95%	95%
Females	94%	100%	96%
Total	94%	96%	96%

Since Phenice's publication, validation studies have rarely matched his success. Kelley (1978) tested Phenice's method using a prehistoric sample resulting in 90% accuracy. Kelley had applied Washburn's (1948) ischium-pubic index, width of the sciatic notch, and presence of the pre-auricular sulcus to establish "known sex" before applying Phenice's (1969) technique to estimate the biological sex of the sample. Lovell (1989) conducted a test of Phenice using 50 pubic bones of known sex individuals who died between the fifth and ninth decade of life resulting in 83% accuracy. Interestingly, Lovell found a moderate negative correlation between accuracy in estimating biological sex of a decedent and that of the individual's age. Lovell suggests that accuracy decreases as age increases ($r = -0.48$). Lovell concludes that perhaps the increased age is responsible for the difference in accuracies obtained in other studies. MacLaughlin and Bruce's (1990) study was based on three European skeletal series of known sex. In conjunction with MacLaughlin and Bruce (1990), Kelley (1978) included an intermediate expression between the male and female form, as Phenice's technique had not accounted for ambiguous expressions of each trait. Unlike other authors, MacLaughlin and Bruce (1990) focused their observations on the right innominate as opposed the left, producing the lowest classification accuracy at 59%. Investigation into the possibility of asymmetrical expression of the three traits had not been conducted. The ventral arc criterion was later tested by Sutherland and Suchey (1991) resulting in 96% accuracy without regard for the other traits. Sutherland and Suchey's (1991) study was also limited to younger individuals who died during the first and second decade of life. McBride et al. (2001) tested accuracy of Phenice's traits using an indicative computer algorithm (ID3) on a sample from the Terry skeletal collection. McBride et al. obtained a combined classification accuracy of 89.6%. Ubelaker and Volk (2002) tested Phenice's method using 198 individuals of known sex, also from the Terry collection and obtained 88.4% accuracy. The authors reported an increase in accuracy when other traits of the pelvis are used in conjunction with the three traits of Phenice.

Results of these validation studies are inconsistent, averaging around 84% for combined accuracy. As with most non-metric methods the consistencies at which the traits are perceived pose a challenge. Phenice's traits, however practical, fail to capture the full range of variation often encountered by analysts. Observations of the traits are limited to only the extreme expression of each trait. As MacLaughlin and Bruce (1990) found, the addition of intermediate forms had not improved classification accuracy which fell well below results obtained by Phenice. No consideration was given discriminatory power of each trait as they were all weighted equally by Phenice. Grant and Boilean (1965) indicate that some traits were not as discrete as others. Despite the short comings of the method, Phenice's traits are the three most cited non-metric features used by anthropologist for the determination of sex in both bioarchaeology and anthropology (Bass, 2005).

Unfortunately, Phenice's technique falls short in light of the changes set forth by *Daubert* and the recent recommendations made by the NAS (2009) to strengthen forensic science in the U.S. Results are heavily dependent on the experience of the observer, and all traits are weighted equally. Phenice also fails to support classifications within a statistical framework. The estimated sex classifications are not accompanied by posterior probabilities which provide a means to measure the degree of uncertainty of the results. Lastly, the simplistic presence/absence protocol does not account for the full range of variability of the three traits.

2.10 Klales and Colleagues (2012)

The Phenice method required modifications in light of the *Daubert* criteria. With the production of reliable and accurate methods now the focus of scientists involved in medico-legal investigations, prior methods are being reassessed (Steadman et al., 2006; Dirkmaat et al., 2008; Walker, 2008). Klales and colleagues (2012) aimed to investigate and improve Phenice's technique to meet the guidelines outlined by the *Daubert* criteria and the NAS (2009) report.

Reliability of Phenice's technique is questionable as results of sub-sequent validation studies are inconsistent. In addition, the aforementioned validation studies produced widely varying accuracy rates. Furthermore, the technique fails to produce posterior probabilities; a significant weakness pointed out by Klales et al. (2012). The posterior probability measures the probability of a membership of an unknown belonging in a group based on relative distance. Posterior probabilities allow for the degree of certainty to be established. To address the conflict of reliability and lack of posterior probabilities in Phenice's method, Klales and colleagues (2012) took into account the variation of expression in each of the three traits. Each of the three traits was expanded to five possible expressions without assumptions of male or female qualities. This revision accounted for ambiguity beyond the simple presence/absence dichotomy established by Phenice. The five expressions for each of the three traits were illustrated and paired with detailed descriptions establishing the criteria appertaining to each expression and the proper orientation from which to observe the trait (Figure 2).

> The proper orientation of the innominate to score the VA is to hold the innominate so that the ventral surface of the pubic bone is showing with the superior pubic ramus aligned horizontally. The five possible character states are: 1) Arc present at approximately or at least 40^{o} angle in relation to the symphyseal face with a large triangular portion of the bone inferiorly placed to arc, 2) arc present at approximately a $25\text{-}40^{o}$ angle in relation to the symphyseal face with a small triangular portion of the bone inferiorly placed to arc, 3) arc present at a slight angle (less than 25^{o}) to the symphyseal face with a slight, non-triangular portion of the bone inferiorly placed to arc, 4) arch present approximately parallel to the symphyseal face with hardly any additional bone present inferior to arc

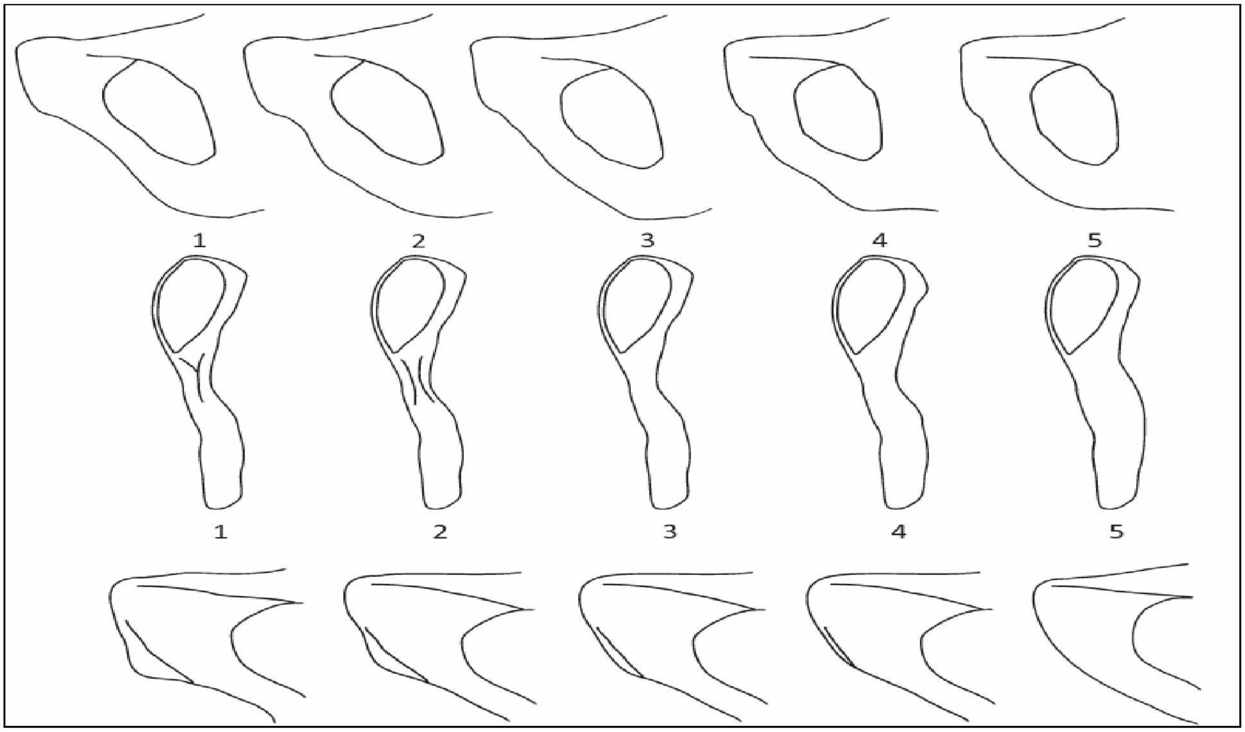

Figure 2. Ordinal scale for assessment of the subpubic concavity (top), conformation of the ischio-pubic ramus (center) and the ventral arc (bottom). Adapted from “A Revised Method of Sexing the Human Innominate Using Phenice’s Nonmetric Traits and Statistical Methods,” by Klales et al., Am J Phys Anthropol, 2012, 149, 104-114. Copyright by Wiley. Adapted with permission.

> and, 5) no arc present (therefore, no additional bone present inferior to the arc). . . . The proper orientation of the innominate to score the MA is to hold the innominate with the medial surface of the pubic bone showing with the symphyseal face aligned vertically. The five possible character states of are: 1) ascending ramus is narrow dorso-ventrally with a sharp ridge of bone present below the symphyseal face, 2) ascending ramus is narrow dorso-ventrally with a plateau/rounded ridge of bone present below the symphyseal face, 3) ascending ramus is narrow dorso-ventrally with no ridge present, 4) ascending ramus is medium width dorso-ventrally with no ridge present and, 5) ascending ramus is very broad dorso-ventrally with no ridge present. (Klales et al. 2012, pg. 108)

The samples used in Klales et al.’s study derived from two skeletal collections: the Hamann-Todd Human Osteological Collection (HTH) and the William Bass Donated Skeletal

Collection (WBSC). The HTH collection was amassed during the 20th century (Hunt and Albanese, 2005). The collection is comprised of over 3,100 individuals collected from morgues and institutions in possession of unclaimed remains. The individuals in the HTH collection are not as well documented as well as the Terry collection; however, it is primarily made up of black and white males and females born in the late 19th century to early 20th century (Hunt and Albanese, 2005). One hundred seventy innominates were selected from the HTH collection. Both blacks and whites were included as evenly as possible from the HTH collection. From the WBSC 140 innominates from black, white, Hispanic, Asian, Mexican, and Japanese individuals were included in the sample. All individuals from both samples are of documented age, sex, and ancestry.

The HTH sample was scored by Klales and Vollner (two of the original authors), individuals with considerable osteological experience, as well as by two individuals with limited experience. Additionally, the sample from the WBSC was scored by Klales as an independent validation of the method. The HTH was used to calibrate the classification function, which was applied to the WBSC. The classification accuracy produced by the WBSC was meant to represent the methods external validity; however, reliability of the method had not been tested using an outside researcher. Furthermore, the authors did not re-calibrate the ordinal logistical regression equation to test for optimal classification accuracy using a population-specific formula (Table 2).

Table 2. Classification accuracies (%) by trait for Klales et al. (2012).

Trait	Males	Females	Combined
MA	79.3	72.3	75.8
SPC	82.2	90.4	86.6
VA	80.5	96.4	88.5

Following the original scoring of the HTH sample, Vollner scored a sub-set of the sample again to assess intra-observer differences. Results of the intra- observer test revealed moderate to substantial agreement as outlined by the parameters of Landis and Koch (1977): ventral arc 0.645 (substantial agreement), sub-pubic contour 0.579 (moderate agreement) and, medial aspect

0.694 (substantial agreement). Since Phenice did not conduct such a test, a comparison of the reliability rates obtained by Klales and coworkers with that of Phenice is not possible. The scores of all four observers was assessed for inter-observer differences. Results of the inter-observer test using the Intra-Class Correlation (ICC), a descriptive statistic that describes how strongly units within a single group resemble each other (Landis and Koch, 1977); identified the degree of consistency in scores across the four observers for each trait in relation to each other using the revised character scores and detailed descriptions devised by Klales et al. (2012). The ICC yields values that range from a high of 0.9 to a low of 0. The four observers with varying levels of sexing experience were found to produce scores with high levels of agreement yielding values of 0.9 for the ventral arc and medial aspect, and nearly as high values for the sub-pubic contour (0.8). Results of both intra-and inter-observer tests demonstrate that the revised ordinal system devised by Klales and colleagues (2012) results in high levels of scoring consistency (Table 3).

Table 3. Inter- and intra-observer test results of Klales et al. (2012).

Trait	Cohen's Weighted Kappa	ICC
VA	0.645	0.9
SPC	0.579	0.8
MA	0.694	0.8

In addition to assessing the degree of scoring consistency across the four observers for left innominates encompassed by Hamann-Todd Human Osteological Collection (HTH), one of the authors (Klales) tested the degree of external validity of the revised ordinally-scaled method using with 140 innominates from the modern William Bass Skeletal Collection (WBSC) housed at the University of Tennessee. The classification accuracy was obtained using the ordinal logistic equation calibrated using the HTH sample. Combined classification accuracy of the historic HTH sample is 95.5% (males 99%, females 92%). The regression equation was then applied to the sample from the modern WBSC where combined classification accuracy decreased. The combined accuracy in correct estimation of sex of known-sex individuals from the WBSC sample was 86%, with an accuracy rate of 98% for females, but only 74% for males. Sex bias in accuracy, which is simply the difference in classification accuracy between males and females, is 24%. Sex bias is important because it measures how well a technique classifies males versus females. Obviously, the greater the sex bias the greater the difference in sex classification between males and females.

Though Klales et al.'s original study yielded accuracy rates that are somewhat lower than those claimed in Phenice's original study, observations were precise, unambiguous and demonstrated a measure of reliability. Klales and colleagues' revision of Phenice's method has greatly improved upon the technique by yielding consistency in the scoring of traits, and producing posterior probabilities that measures effectively how certain the decedent actually belongs to its estimated sex. By meeting the guidelines outlined by *Daubert,* Klales et al.'s (2012) method is currently being utilized in modern forensic casework in the U.S. and internationally.

Klales et al.'s (2012) revision of the Phenice (1969) technique for the estimation of sex has since been tested on three samples outside the U.S. Furthermore, a re-calibration equation to "fit" the population in question increased classification accuracy. External samples derived from South Africa (Kenyhercz, 2012), Mexico (Gomez-Valdes et al., *in review*). The classification accuracy for the South African sample reached 90%. A re-calibration of the original regression equation gauged to fit the South African and Mexican sample improved accuracy by 9%, raising the final accuracy to 99% in the South African sample. The Mexican population had also reached an initial accuracy as high as 90%, and as with the South African sample, a re-calibration of the original equation also achieved a classification accuracy of 99%. Given such results, it is reasonable to assume population specific re-calibrations more closely related to the original calibration sample (HTH) would vastly improve the classification accuracy when applied to a modern U.S. sample (Table 4).

Table 4. Classification accuracies in validation studies of the Klales et al. (2012) method.

Validation Study	Classification Accuracy (%)	Recalibration Accuracy (%)
Kenyhercz (2012)	90.6	99.2
Gomez-Valdes et al. (*in review*)	90.0	99.0

In light of the *Daubert* decision and given its consistent classification accuracy on samples outside the U.S., it is imperative that methods such as Klales et al. (2012) be tested, implemented in SOPs (Standing Operating Procedures), and utilized in forensic casework here in the U.S; however, Klales et al. (2012) method had not been tested on a modern American sample by an outside researcher. Active practitioners explain that once Klales et al. (2012) technique has been validated on a U.S. sample by a third party, with success, it can be incorporated into SOP.

2.11 Asymmetry

The departure from identical development of paired traits on both sides of the body is known as bilateral asymmetry (Van Valen, 1962; Zachos et al., 2007). Bilateral asymmetry has been cited extensively in the evolutionary literature as a way to measure developmental stability of an organism (Moller and Swaddle, 1997). Developmental stability refers to the ability of an organism to produce a given phenotype under a range of environmental and genetic conditions (Moller and Swaddle, 1997; DeLeon, 2007). In other words, it is the capacity of an organism to reach its intended developmental form, resisting happenstantial deviations from the genetically encoded developmental pathway during fetal development and beyond into childhood and adolescence. A greater degree of stability during development (e.g., homeostasis) will result in a phenotype that faithfully reflects the genetically encoded phenotype, while bouts of instability, regardless of cause, disrupts homeostasis and thereby has the potential to result in deviations of growth away from the genetically encoded genotype. Such deviations from normative growth and development are either manifested as subtle fluctuations between right and left sides of paired structures or may result in one side consistently being larger or more developed than the opposite side. As such, asymmetries may either reflect the level of environmental stress experienced by an individual during growth and development, or they may reflect a derived adaptation. Previous studies that have focused on bilateral asymmetries suggest that if structures are primarily under genetic control and if there is little to no disruption of homeostasis during growth and development, then equal bilateral expression of the trait generally occurs. However, if a significant amount of environmental stress is encountered during growth and development, fluctuating asymmetrical expression of that trait is commonly found (Trinkaus, 1978).

The genetic basis of developmental stability is highly complex; however, inbreeding has been shown to decrease the ability of an organism to buffer developmental perturbations during ontogenetic growth and development (Wadington, 1940). In addition, the environmental factors that have been identified as contributing to developmental instability are many. At a minimum, these include diet (Kirpichnikov, 1981), disease (Moller and Swaddle, 1997), climate (Skinnes and Buras, 1987), chemical exposure (Yablokov, 1986), biomechanical stress, and even loud noises (Moller and Swaddle, 1997).

Directional asymmetry, as defined by Van Valen (1962), occurs when there is a propensity for a trait to develop to a greater extent on one side more than the other. Consequently, frequency distributions of a trait affected by directional asymmetry will be skewed either positively or negatively depending upon where the right or left side is affected to the greater extent. Directional asymmetry has been attributed to adaptive strategies such as the development of a larger left claw (crushing claw) in snapping shrimp (Bethe et al., 1930) as well as to the relative hypertrophication of the left ventricle in the mammalian heart. In cases of directional asymmetry it is possible to predict which side of a trait will be larger before the organism has reached maturity (Moller and Swaddle, 1997). By contrast, fluctuating or random asymmetries represent random deviations from perfect symmetry and can occur on either right or left sides, often with statistically equal frequencies. Less commonly, frequency distributions of a specific trait can take on either a bimodal (where there is an even distribution in either direction with two distinct peaks) or platykurtic (where the distribution in both directions is highly dispersed without a distinct peak) distribution on a histogram when fluctuating asymmetry occurs in a sample. Because these deviations are random, they are challenging to predict.

While the study of bilateral asymmetry has a long history across many evolutionary and biological fields this thesis focuses on its relevance to the study of skeletal biology. Studies of bilateral asymmetry have been extensively cited as an indicator of handedness (Steele and Mays, 1995; McManus et al., 2010). Many anthropologists and bio-archaeologists concerned with incidences of bilateral asymmetry on human skeletal remains have attributed such differences to habitual behavior and differential mechanical loading. Increased mechanical loading of the dominant side of the body during endochondral bone growth results in greater robusticity (Steele and Mays, 1995). In accordance with Wolff's Law, this process occurs through the modeling and remodeling of bone undergoing habitual stress, especially in areas associated with the muscles and joints employed during systematic activity. While long-bone elements such as the humerus, radius, femur, and tibia tend to experience the brunt of morphological changes due to differential mechanical loading, mechanically related anatomical regions, such as the pelvis and spine are affected as well (Plochocki, 2002; Overbury et al., 2009; Bussey, 2010).

2.12 Asymmetry of the Innominate

The non-metric traits outlined by Phenice (1969) are all located in the pelvic region of the skeleton, an area under strict hormonal control, but are also susceptible to morphological changes as pointed out by Plochocki (2002) who found significant levels of bilateral asymmetry in the human sacrum due to differential mechanical loading. Overbury et al. (2009) also reported asymmetry of age-related markers of the pubic symphysis to be a real and remarkably frequent phenomenon, occurring in over 60% of a sample of 140 modern white males from the Hamann-Todd Osteological Collection. This prevalence of asymmetry may be significant to the entire pubis, particularly, the traits identified by Phenice and forensic anthropological literature in general (Bass, 2005; Byers, 2011; Christensen et al., 2014) as the best indicators of sex.

The ventral arc and medial aspect of the ischio-pubic ramus are the points of origin of several muscles responsible for movements involved in: adduction of the thigh and hip, flexion of the leg at the knee joint, medial rotation and flexion of the hip and lateral rotation of the thigh and hip (Bowden and Bowden, 2003). Throughout the course of everyday life, individuals experience these movements, such as walking, squatting, sitting with legs crossed, horseback riding, skiing, and exercise. The ventral arc, directly lateral to the pubic symphysis, is the point of origin for adductor magnus inferiorly, and adductor brevis superiorly (Todd, 1921; Anderson, 1990). This pair of muscles acts as adductors of the thigh, and to stabilize the pelvis when bearing weight. The ischio-pubic ramus is also a point of origin for the adductor muscles as well as the gracilis muscle and both obturator externus on its outer surface and obturator internus on its internal surface (Bowden and Bowden, 2003). While the subpubic concavity is not a point of origin for muscle attachment, its function is also relevant. A recent study by Bussey (2010) who examined the incidences of pelvic asymmetry among athletic and non-athletic females, found that athletes participating in unilateral sports prior to puberty were at greater risk for developing pelvic asymmetries than those who do not participate in such activities. During puberty, the pelvis undergoes morphological changes in response to hormonal signals. The pelvis in females begins to broaden to accommodate potential vaginal delivery during childbirth. The subpubic concavity widens in females, and remains relatively narrow in males (Phenice, 1969). During this process, individuals who engaged in habitual activities that involve unilateral use of the limbs are at greater risk to developmental disruption. Given the reported high frequencies of asymmetry in the pelvic region, and the functional responsibilities of the ventral arc, subpubic concavity, and ischio-pubic ramus it is reasonable to expect the innominate to be exposed to symmetrical instability due to differential mechanical loading and/or habitual activity.

Lovell's (1989) test of Phenice (1969) was the first to offer a possible explanation for the discrepancy in classification accuracy. Lovell's results demonstrated a decrease in accuracy with an increase in age; however, information regarding sex estimates from a particular side was not mentioned. It is hard to know if the degenerative breakdown described by Lovell (1989) was restricted to the left side of the pubis, representing fluctuating asymmetry or something else such as directional asymmetry due to the preponderance of right-handedness among humans. MacLaughlin and Bruce (1990) however, had focused their observations of the three traits on the right innominate on a sample of 275 individuals of known sex from three European skeletal collections. The authors reported much lower accuracy (59%) than any other evaluation of the Phenice method; all of which report observing traits of the left innominate or the protocol prescribed by Phenice (the left innominate). Due to the lack of investigation into the bilateral expression of these traits it is impossible to discern if asymmetry may have impacted the results reported by MacLaughlin and Bruce (1990). Moreover, a method for measuring those differences in a quantifiable way, such as a scoring system, wasn't readily available. Bilateral asymmetry caused by age progressive degenerative processes could arise through two trajectories. First, through natural breakdown of bone, obscuring the features in question or, through the intensification of divergence acquired during skeletal development over a lifetime (Albert and Greene, 1999). However, in response to Lovell's findings, Sutherland and Suchey (1991) stated that the ventral arc had been easily discernible with up to 93% accuracy in individuals up to the eighth decade of life. While ambiguity over the presence or absence of the three traits had been a common argument, and age dependent variability mentioned, little attention has been paid to the possibility of asymmetrical instability of the traits as a factor in classification error. This is perhaps due to the difficult interpretation of the biomechanical function of these elements and their relationship with the appendicular bones (Plochocki, 2002; Auerbach and Raxter, 2008).

In Phenice's original study and all subsequent validation studies, classification accuracy for women has consistently surpassed that of males with classification of females reaching 100%

in some cases. As mentioned before, sub sequent validations, like the original study have focused primarily on the left pubis. The pubis, like other skeletal elements of the pelvis serve as important points of origins for muscles involved in daily activities; therefore, are also susceptible to the affects of differential loading. Even though the majority of humans are found to be right handed studies have found significant levels of sexual dimorphism in handedness with left handedness being more common in males (Peters et al., 2006). With the propensity of left handedness being more common in males it is reasonable to expect incidences of bilateral asymmetry to affect male classification accuracy based on observations restricted to one side of an individual.

Klales et al. (2012) have identified and illustrated expanded levels of expression of the three traits of Phenice (1969), once assumed to simply be discrete. Most studies interested in the effects of mechanical loading employ metrical analyses to quantify bilateral asymmetry. However, due to the nature of non-metric traits, traditional measurements fail to capture the shapes and curves that may be apparent to the naked eye. Klales et al.'s (2012) ordinally graded expansion offers a unique way in which to quantify the traits and to examine symmetrically stability. The investigation into Klales et al.'s (2012) revision offers a platform with which to ascertain bilateral differences in the pelvis in regard to the most widely used indicators of sex to and examine how they may affect sex estimation.

Chapter 3 MATERIALS AND METHODS

3.1 Skeletal Material

The skeletal sample used in this study is derived from the William Bass Donated Skeletal Collection (WBSC) housed at the University of Tennessee, Knoxville. This sample was selected as it is one of the largest samples available in the United States and is most similar to modern forensic cases in terms of time period currently encountered by forensic anthropologists in the United States. The majority of the individuals within the WBSC were donated by individuals in their will prior to their death, or by the decision of close relatives for the purpose of scientific research. While the majority of the collection is composed of well-documented, donated remains, a small fraction of individuals, with unknown ages, represent forensic cases donated by medical examiner's office. The collection includes individuals born between 1892 and 2011, with most individuals born after 1940. For this research, a stratified random sample of Euroamerican innominates and the innominates of all American Indian, Hispanic, White/Asian, and American blacks of known sex was selected from the WBSC (Table 5) via a list of ID numbers provided to me from the curator of the collection to ensure balance between males and females. This protocol yielded a sample of 204 individuals. The documented age of individuals in this sample range between 23 and 99 years of age with a median age of 62 (Figure 3). The ancestral variation is markedly asymmetric, with individuals of European ancestry comprising 86% of the sample. In an effort to diversify the sample by ancestry, the age of some individuals is unknown because many of the black males and females in the collection were donated by medical examiners that were unable to determine their age-at-death.

Table 5. Sample composition.

Ancestry	Male	Female	(n)
American Indian	-	2	2
Black	15	7	22
Hispanic	2	-	2
White	85	91	176
White/Asian	2	-	2
Total	104	100	204

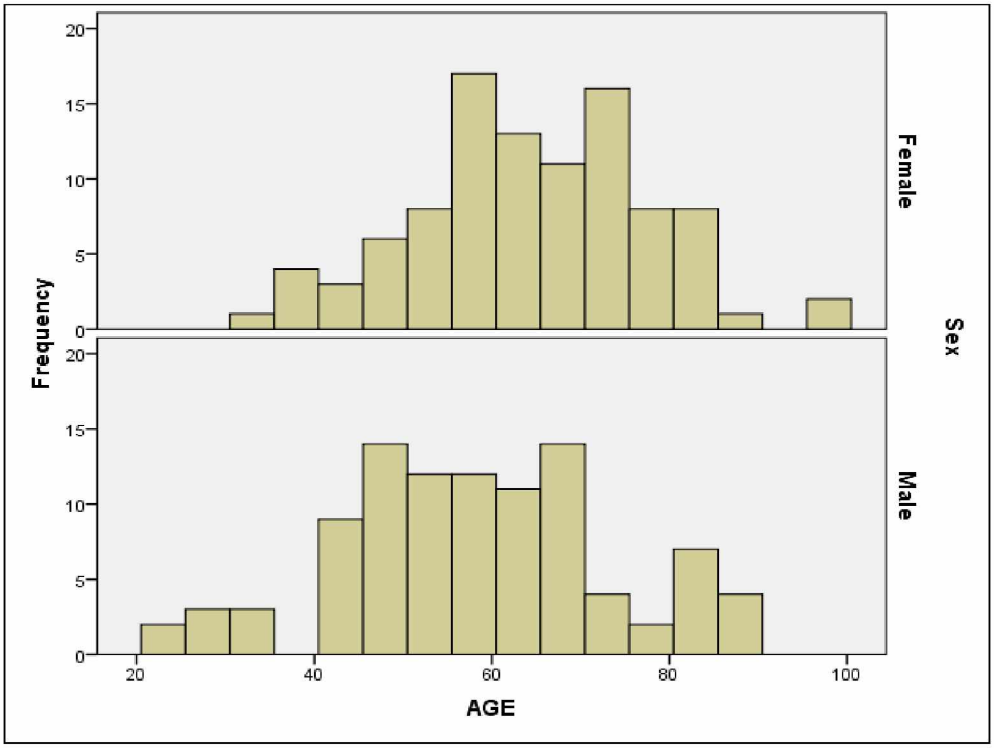

Figure 3. Frequency distribution of the age of the 204 individuals used in the present study.

3.2 Trait Scoring

For the present research, I scored each of the three Phenice (1969) traits using the methodology of Klales et al. (2012), which includes schematic figures, pictures, a website, and descriptions (Figure 2). Samples of left and right innominates were scored with no prior knowledge of the sex of the individuals being scored.

3.3 Trait Distribution

Frequency distributions were calculated to better understand the distribution of the three traits within the sample, and their relationship to biological sex.

3.4 Statistical Methods for Validation

3.4.1 Validity (Classification Accuracy)

Ordinal logistic regression was chosen to assess validity because it does not assume that a variable is normally distributed and that the dependent variable is measured on an ordinal scale.

In addition, Klales et al. (2012) and all subsequent validation studies of the Klales et al. (2012) method employed ordinal logistic regression; therefore, results produced by this study will be directly comparable. Ordinal logistic regression is used to predict the state of a dependent variable given one or more independent variables. This process involves the direct calculation of posterior probabilities for classification. Posterior probability measures the likelihood of a membership for an unknown to belong to an estimated group based on relative distance. Four assumptions are associated with ordinal logistic regression: 1) the dependent variable is measured on an ordinal scale, 2) there are one or more dependent variables that can be either continuous, ordinal, or nominal, 3) dependent variables should be mutually exclusive, and 4) there needs to be a linear relationship between any continuous independent variables and the logit transformation of the dependent variable (Norusis, 2012).

In the original study, Klales et al. (2012) used a leave-one-out-cross-validation (LOOCV) (Lachenbruch and Mickey, 1968) procedure to attain classification accuracy. The aim of LOOCV is to reduce bias of error rate estimation when using re-substitution. In using LOOCV each individual in a reference group is removed from its group one at a time; the predictive parameters are recalculated using the remaining individuals, and that individual is then classified into one of the reference groups (Klales et al., 2012). That individual is then added back into its reference group and the next individual is classified in an identical manner. When all individuals have been classified, the total number of correctly classified individuals equals the expected unbiased variance (Lachenbruch and Mickey, 1968).

Using the logistical regression equation developed by Klales et al. (2012), classification accuracy was tested for 204 left and right innominates. Sex was estimated by inputting the scores into the ordinal logistic equation provided in Klales et al. (2012): 2.726(VA) + 1.214(MA) + 1.073(SPC) – 16.312. Classification accuracy was compared with that of the original study. The original equation calculated posterior probabilities in accordance with Press and Wilson (1978): probability of being female ($p_f = 1/(1 + e^{score})$ and the probability of being male ($p_m = 1 - p_r$). Following sex estimation for all left and right pubes, sex bias was calculated to determine whether there is a significant difference in classification accuracy between males and females.

3.4.2 Reliability

Intra- and inter- observer tests will be conducted to evaluate whether the standardization of trait expression into the five ordinal grades, by Klales et al. (2012) is a reliable method for assessment of the traits initially identified by Phenice (1969). Reliable techniques will render high levels of agreement, while methods with low levels of agreement are not *Daubert* compliant as they cannot be replicated by observers.

Intra- and inter observer agreement is often reported as a Cohen's Kappa statistic (Viera and Garrett, 2005). Kappa is calculated using how much agreement is observed compared to how much agreement would be expected purely by chance (Viera and Garrett, 2005). However, a simple Kappa statistic cannot reveal the level of agreement across several categories for which there is a meaningful difference (Viera and Garrett, 2005). For this reason, a weighted Kappa was used in this study and is in accordance with Klales et al. (2012), thereby making the results comparable. Weighted Kappa measures agreement for meaningful differences between observations that are extended along an ordinal scale. Weighted Kappa assigns greater weights to those observations that are furthest apart on an ordinal scale and lesser weights to those observations that are close or more similar along that scale. Five assumptions must be met when using Cohen's Kappa for rater reliability: 1) the observations must be made on a mutually exclusive nominal or ordinal scale, 2) the raters observe the same phenomenon on the same specimen, 3) the variable must have the same number of possible scores during each scoring bout, 4) the raters or rating sessions are independent, 5) the raters making the observations are the only observers involved in the study (Berry et al., 2014).

3.4.3 Intra-observer Agreement

Intra-observer agreement will be calculated using Cohen's (1968) Weighted Kappa (K). All assumptions for appropriate use of a Weighted Kappa in this study will be met. The author randomly selected and scored the three traits on 25 left innominates. Several days following the initial observations, the same 25 left innominates were rescored by the author to test the level of intra-observer agreement. Levels of agreement for K were determined based on the criteria of Landis and Koch (1977) in which: K = 0.0 no agreement, K = 0.01 to 0.20 represents slight agreement, K = 0.21 to 0.40 represents fair agreement, K = 0.41 to 0.60 represents moderate agreement, K = 0.61 to 0.80 represents substantial agreement, and K = 0.81 to 1 represents near perfect to perfect agreement.

3.4.4 Inter-observer Agreement

The intra-class correlation coefficient (ICC) is currently considered the best measurement of inter rater reliability for ordinal and interval scale data (Landers, 2015). The ICC is a descriptive statistic used to measure the level of agreement between two or more raters making the same observation. What the correlation effectively measures is the proportion of differences in the observations made by the observers caused by either a disparity between the raters themselves or in the object the raters are observing. In this study, a two-way mixed model was selected as this model assumes that the two raters are the only raters involved in the study, that the subjects are random, and that there is no change in the phenomenon being observed between its assessments by the raters. A sample of thirty individuals originally scored by Klales (from the Klales et al. 2012 study) was later scored by the author and compared for inter-observer differences.

3.5 Statistical Methods to Assess Asymmetry

Given that the Phenice (1969) study, the Klales et al. (2012) revision, and subsequent validation studies of Klales et al. methodology (Kenyhercz, 2012; Gomez-Valdes et al., *in review*) primarily focused on the left innominate for evaluation, the right side was tested in this study for overall accuracy, as well as assessed for asymmetrical differences from that of the accompanying left side within the same individual. Asymmetry measures the difference in ordinal scores between the left and right sides of the same individual with regard to the character states outlined by Klales et al. (2012). The technique recognizes five character states for each trait; therefore, several scores or combinations thereof have the ability to produce correct or incorrect classifications when incorporated into statistical classification functions. In this study, the focus is on how well the Klales et al. (2012) method performs if asymmetry is present. However, if present, asymmetrical differences do not necessarily produce incorrect sex classification.

3.5.1 Presence of Asymmetry and Wilcoxon Signed-ranked Test

To establish the presence of asymmetry, all scores for the 204 left innominates were compared to the scores for the corresponding right innominates. The percentage of those pairs found to score asymmetrically within one score difference or more was recorded. A series of Wilcoxon matched-pairs signed-ranks tests were used to evaluate statistical differences in the trait scores between the left and right innominates for each of the three traits. The significance level was set at 0.05. The Wilcoxon signed rank test was selected as it does not assume that variables are normally distributed and as trait scores are matched pairs belonging to the same population, this test is appropriate for the data being analyzed (Bremer and Doerge, 2010).

3.5.2 Type of Asymmetry

Establishing the type of asymmetry is important for constructing potential strategies when utilizing Klales et al.'s (2012) technique to estimate sex, such as systematically selecting for a left or right innominate, or the higher or lower or the two scores, as one tends to yield more accurate sex classifications in the presence of asymmetry, as this study aims to determine. In determining the type of asymmetry present, individuals who were not asymmetrical were not included in the following analysis

To determine whether the asymmetry, if present, is random (e.g., fluctuating) or directional in nature, the preponderance of each trait to either increase or decrease in expression on the right side relative to the left side was recorded. If the trait is found to show no prediction for either right or left sides, then it was adjudged to reflect random or fluctuating asymmetry. If asymmetrical trait scores showed a preponderance to be greater on a specific side in 51% of cases or greater, asymmetry was determined to be directional as it can be predicted in a way that is greater than chance.

In the event of directional asymmetry, directionality of each trait was classified as 51 to 60% (weak directionality), 61 to 70% (moderate directionality), 71% or greater (strong directionality). This scale was arbitrarily created by the author as a means to measure the level on tendency and/or strength of directionality.

3.5.3 Patterns of Asymmetry

Age, sex, and ancestry are three factors that may potentially impact the patterning of asymmetry. To investigate the impact of these factors on asymmetry a rank-based method known as the Scheirer-Ray Hare test, an extension of the Kruskal-Wallis H-test was performed. This test is a non-parametric substitute for the two-way ANOVA with replication used to determine if the interaction between two factors affect a data set in a significant way. Like the Kruskal-Wallis H-test, the Scheirer-Ray-Hare extension assumes that: 1) the samples are drawn from the populations are random, 2) the cases in each group are independent, and 3) the scale of measurement is at least ordinal (Bremer and Doerge, 2010) This examination will determine if sex, age, or ancestral groups are most affected by asymmetrical differences and to what degree.

The first step in the Scheirer-Ray-Hare requires that all factors are ranked. Data for American Indians, Hispanics, and White/Asians were removed from the analysis as there were too few individuals within the sample (two of each) to produce any statistically meaningful results. In addition, individuals of unknown age were also removed from the analysis, thereby reducing the sample size from 204 to 189 individuals for this particular analysis.

Like the Kruskal-Wallis test, the Scheirer-Ray-Hare test requires that the data be distributed symmetrically about an axis. Being that the variables in this study scored along an ordinal scale, the axis is represented by the median score. To find the median score of the left and right sides for each of the three traits, all scores were listed in ascending order and the value at the 50^{th} percentile was identified as the median. The same process was used to identify median age. The median score for both the left and right VA is 3. The median score for both the left and right SPC is 2, while the median score for the left and right MA is 3.

The median scores were used to verify normality in the distribution of cases about the axis of age, as well as for both left and right sides of each of the three traits. The difference between the number of scores (or cases) above the axis and below the axis were divided by the sample size to assess normality, with a difference of only 1.58%, the distribution of left and right VA was identical. The distribution of the right SPC was one percent greater at 9% than the left SPC whose distribution was 8%. The right MA was also slightly greater than the left at 5.82% while the left was 5.29%. The distribution of age was found to be 4.2%. Each of these values is very low and falls close to the expected value of 9.45% due to random chance. Given such results, the variables in this study are decidedly symmetrical, thereby satisfying the requirement for the nonparametric equivalent of normality.

Next, in order to perform the ranked test, each of the three traits and age had to be placed in ranked order. All scores were placed in ascending order and summed up from top to bottom. The scores were replaced with the new value assigned to the number of scores in each of those categories. A similar process was performed to rank age categories except that values for individuals who were identical for age received an average rank for all members of that age.

3.5.4 Interactions and Correlations

Six comparisons were made using the H value, completing the Scheirer-Ray-Hare test for significant interactions. In a manner similar to that of univariate ANOVA, the scores for each of the three traits were analyzed for correlations and interactions between manifestation of each trait (represented by its score), with age, ancestry, and biological sex. These interactions are important as they may reveal relationships between demographic characteristics and the asymmetrical expression of the three traits analyzed in the current research. Such interactions would be especially valuable should asymmetry be found to compromise the Klales et al. (2012) method and provide data to aide in reconciling such problems.

The Scheirer-Ray-Hare test discards the mean squares (MS) and p values produced by ANOVA and shifts F values to H values. This is important as the H statistic is used to determine statistically significant differences using ordinal scale data and is based on the Chi-square distribution. Next, the sum of squares for each factor, the sum of squares for all interactions and the sum of squares for error are calculated. This sum is divided by the total degrees of freedom producing the new adjusted MS that accommodates ordinally graded data. The new MS was used to calculate the H value in Microsoft Excel using the formula: CHISQ.DIST.RT(H,df).

3.6 Comparison of Classification Accuracy Between Left and Right Innominates

Finally, classification accuracy of the 204 right pubes was calculated using Klales and colleagues' ordinal logistic regression in the same manner as the left innominate. Classification accuracy of the left innominate was compared with that of the right to test whether asymmetry compromises the Klales et al. (2012) method for estimation of biological sex. Additionally, classification accuracy combining both the left and corresponding right innominate of the same individual was analyzed.

3.7 Re-Calibration

The validity of sex estimations are affected by several factors with inter- and intra-sample differences among them. Both Kenyhercz (2012) and Gomez-Valdes et al. (*in review*) reported a 9% increase in classification accuracy after re-calibrating the Klales et al. (2012) logistic regression equation to fit their specific samples, respectively. The use of sample specific equations for sex estimation meets the best practice standards published by the Scientific Working Group for Forensic Anthropology (2010). The re-calibration entails finding the best fit for classification accuracy and will produce the re-calibrated ordinal linear regression equation based on the best classification. The Klales et al. ordinal logistic regression equation will be re-calibrated to fit the modern U.S. sample used in the present research to determine if a sample specific equation will classify the whites and blacks sampled in the WBSC with greater accuracy. Like the original equation, the posterior probabilities for the re-calibration will be calculated in accordance with Press and Wilson (1978): probability of being female ($p_f = 1/(1 + e^{score})$ and the probability of being male ($p_m = 1 – p_r$).

The Klales et al. (2012) study did not assess the accuracy rate of each of the individual traits: therefore, the classification accuracy of each trait individually and in combination will be assessed to elucidate how well each trait performs individually for sex estimation. Ordinal logistic regression equations for each trait and trait combination will be produced and such equations will be especially useful in situations where all three traits are not available for analysis. Classification accuracy and sex bias based on these six equations will be evaluated for the entire sample of 204 individuals.

3.8 Software

All statistical calculations were conducted in SPSS 23.0 (SPSS Inc., 2015) and Microsoft Excel version (2007).

Chapter 4 RESULTS

4.1 Trait Distribution

The conformation of the ventral arc (VA) was found to be the least variable trait among males and females; with females having predominately (161/200 = 80.5%) low scores (grades 1-2) and males predominately (151/208 = 72.6%) having high scores (Figure 4). The ischio-pubic ramus (MA) was found to be most variable among males with scores encompassing the entire scale (Figure 5). Nevertheless, scores among males are dominated by intermediate to high scores (grades 4-5) with 180 of the 208 male innominates (86.5%) receiving such scores. The opposite pattern is observed among females, among whom the overwhelming majority (176/200 = 88%) received low scores (grades 1-2). A large majority of males received high scores (182/208 = 87.5%) for the subpubic concavity (SPC), while females tended to receive low scores (192/200 = 96%) with some overlap between males and females in mid level scores (grades 2-3) (Figure 6).Surprisingly, no scores of five were recorded for either sex.

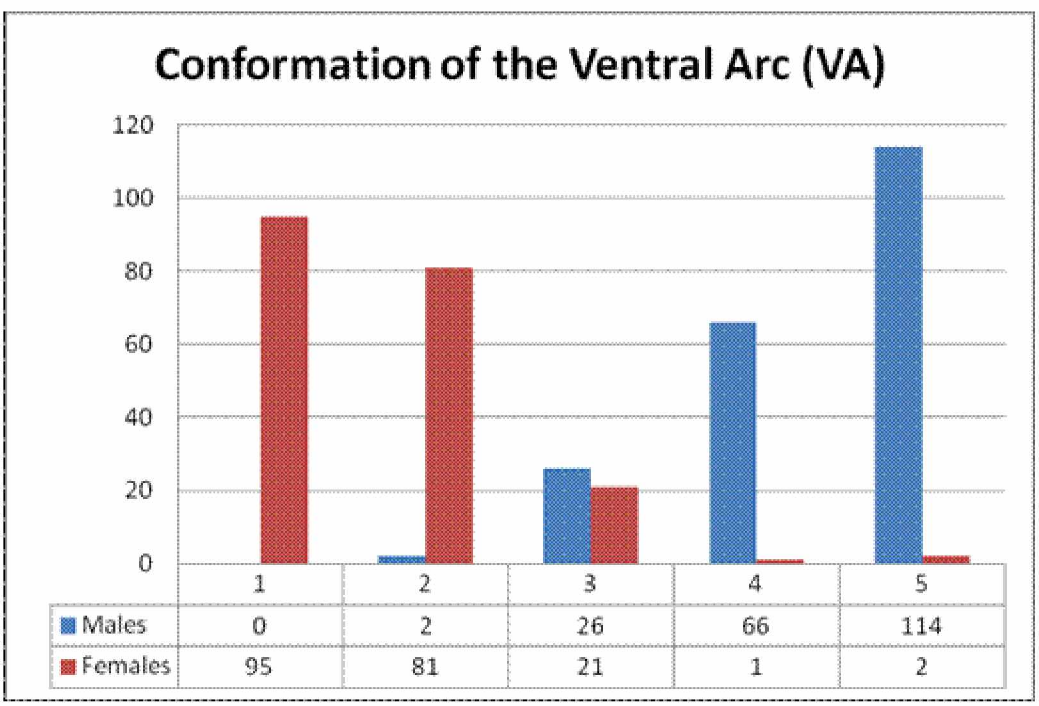

Figure 4. Score frequency distribution by sex for the ventral arc.

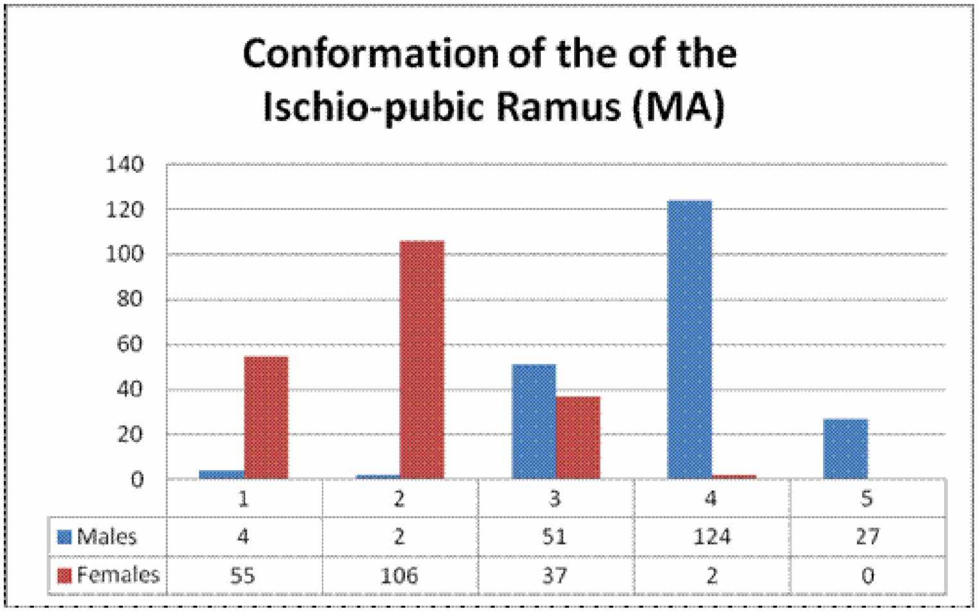

Figure 5. Score frequency distribution by sex for the ischio-pubis ramus.

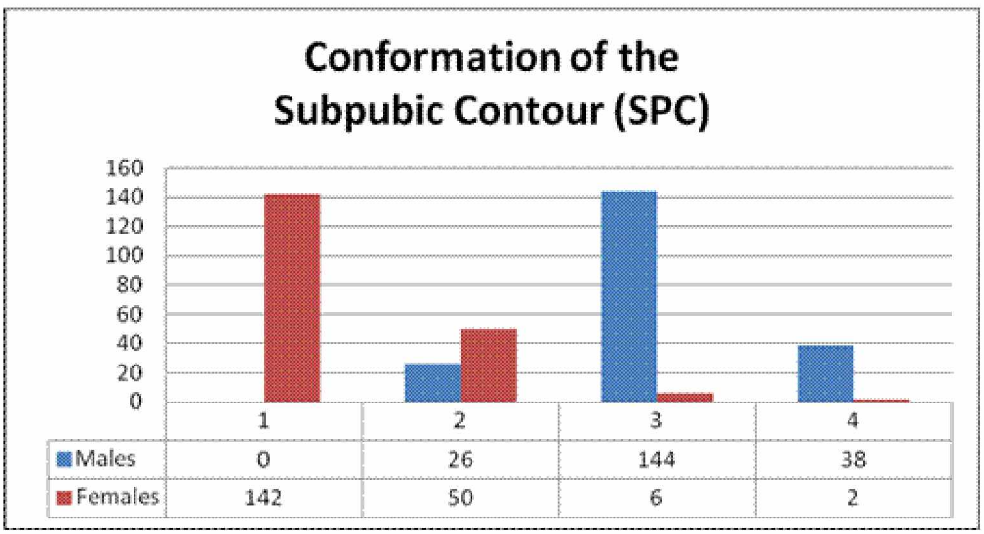

Figure 6. Score frequency distribution by sex for the subpubic concavity.

4.2 Validity (Classification Accuracy)

Klales et al.'s original logistic regression equation yielded high classification accuracy (combined 93.6%) for the innominates from the WBSC considered in the present study. Females were correctly sexed for 99 % (99/100) of individuals scored, while males were correctly sexed for 87.5 % (92/104) of individuals. Such values yield a sex bias of 11.5% (Table 6).

Table 6. Classification accuracies (%) using Klales et al.'s (2012) logistic regression equation.

Biological Sex	*n*	Accuracy	Sex Bias
Males	104	87.5	11.5
Females	100	99.0	11.5
Combined	204	93.6	-

4.3 Reliability

4.3.1 Intra-observer Error

The results obtained with Cohen's weighted Kappa suggest that Klales and colleague's method is highly repeatable and hence reliable (Table 7). Of the original scores, one VA originally scored as a four was scored as a five upon the second observation. Two SPCs that were originally scored as a three were later scored as a four. All MAs were scored identically in observation bouts one and two. Intra-observer agreement for each trait was either perfect or near perfect based on the parameters outlined by Landis and Koch (1977).

Table 7. Intra-observer error results using weighed kappa.

Trait	Weighted Kappa	Significance	*n*	Level of Agreement
VA	0.974	<0.001	25	near perfect
SPC	0.928	<0.001	25	near perfect
MA	1.000	<0.001	25	perfect

4.3.2 Inter-observer Error

Tests of inter-observer error, based on the ICC also rendered high levels of agreement (Table 8). While all the MAs were scored identically by both observers, two SPCs were scored

Table 8. Inter-observer error results using the ICC.

Trait	ICC	Significance	*n*	Level of Agreement
VA	0.995	<0.003	30	high
SPC	0.981	<0.013	30	high
MA	1.000	<0.008	30	high

by the author as a three and were scored by Klales as a four. One VA scored by the author as a 3 was also scored by Klales as a four.

4.4 Presence of Asymmetry and Ranked Pair Signed Test

Asymmetry was based on a difference in score between the left and right innominates of the same individual, was found to be prevalent throughout the sample and for each trait (Table 9). Of the 204 individuals included in the study, one third exhibited asymmetry in one trait or more. The majority (88.2%) of these asymmetrical individuals exhibited asymmetry for a single trait, seven were asymmetrical for two traits (10.2%), and one individual was found to be asymmetric for all three traits (1.4%). With the exception of a single individual, who differed in left and right trait score of the VA by two, all other asymmetrical scores differed by only one score. The VA was found to be the most asymmetric trait (64% of cases), followed by the MA (26%), with the SPC (10%) exhibiting the least asymmetry.

Table 9. Frequencies of asymmetrical traits.

Trait	*n*	Consolidated	(%)
MA	13	20	26
SPC	5	8	10
VA	42	49	64
SPC, MA	1	-	
VA, MA	5	-	
VA, SPC	1	-	
VA, SPC, MA	1	-	
Total	68 (individuals)	77 (traits)	

A series of Wilcoxon matched-pairs signed-ranks test revealed a statistically significant difference in trait scores of the ventral arc between the left and right innominates at the 0.05 significance level (Table 10).

Table 10. Wilcoxon signed-ranks.

Pairs	VA-L – VA-R	SPC-L – SPC-R	MA-L – MA-R
P Value	0.002 < 0.05	0.24 > 0.05	0.14 > 0.05

4.5 Type of Asymmetry

As the left innominate has served as an anchor in this examination, directionality was determined in reference to the left side. Furthermore, as this test aims to recognize predictable patterns, any departure in asymmetry by side from 50% (which is equal to chance alone) is indicative of directional bias. As can be seen in Table 11, asymmetry, when present, asymmetry was more often expressed as higher scores on the right side in relation to the left, but this was not exclusively the case. Of those individuals found to be asymmetrical at the VA, scores increased by at least one score on the right side in 73% of cases (strong directionality). The SPC also favored the right side in 75% of cases (strong directionality) as did, scores for the MA with higher scores on the right side in 65% of cases (moderate directionality). Of the 49 asymmetrical ventral arcs (VA) 13 scored lower on the right side, while 36 scored higher on the right side, a ratio of nearly three to one (2.77:1). In two cases, the sub-pubic contour (SPC) scored lower on the left, while scores increased on the right side in six cases, once again, a ratio of 3:1. Of the 20 asymmetrically scored ischio-pubic rami (MA), seven scored lower on the left, while thirteen scored higher on the right., a ratio of about two to one (1.86:1).

Table 11. Directionality of asymmetry.

Trait	*n*	Scored lower on right	(%)	Scored higher on right	(%)	Directionality
MA	20	7	35	13	65	Moderate
SPC	8	2	25	6	75	Strong
VA	49	13	26	36	73	Strong
Total	77	22	-	55	-	

4.6 Patterns of Asymmetry

4.6.1 Interactions and Correlations

Biological sex was found to be highly correlated with the score of the VA on both left ($H = 121.299$, $p = 0.000$) and right sides ($H = 113.645$, $p = 0.000$). No significant correlation was found for age and ancestry. Expression of the right SPC was highly correlated with biological sex ($H = 237.030$, $p = 0.000$), but unlike the VA, the SPC was also found to be significantly correlated with advancing age at death ($H = 323.320$, $p = 0.000$). No significant correlation was found between the appearance of the right SPC and ancestry. Like the right side, expression of the SPC on the left was also found to be significantly correlated with biological sex (H = 254.092, p = 0.000) and advancing age at death (H = 398.355, p = 0.000), but unlike the right side, the expression of the SPC on the left was also found to be significantly correlated with ancestry (H = .098, p = 0.014). Like expression of the SPC on the left side, manifestation of the MA on the right side was found to be significantly correlated with advancing age at death (H = 370.454, p = 0.000), ancestry (H = 8.587, p = 0.000), and biological sex (H = 221.623, p = 0.000). However, manifestation of the MA on the left side followed the pattern observed for expression of the SPC on the right side. That is, the appearance of the MA on the left side was found to be significantly correlated with advancing age at death (H = 461.732, p = 0.000) and biological sex (H = 300.053, p = 0.000), but not with ancestry.

The Sheirer-Ray-Hare two-way ANOVA extension for ordinally graded data (SRH) revealed no significant interactions between age-at-death, ancestry and sex on the expression of the VA on either the left or right side. By contrast, the SRH identified significant interactions between age-at-death and ancestry (H = 398.355, p = 0.000) as well as age-at-death and biological sex (H = 124.832, and p = 0.000) for the expression of the SPC on the right. Like the expression of the SPC on the right side, SRH also revealed significant interactions between age-at-death and ancestry (H = 89.519, p = 0.000) and between age-at-death and biological sex (H = 167.286, p = 0.000) for the expression of the SPC on the left side. Similar results were obtained with the SRH for expression of the MA, except that the expression of this trait on the right side mirrored expression of the SPC on the left side, while expression of the MA on the left side mirrored the expression of the SPC on the right side. That is, SRH identified significant interactions between age-at-death and ancestry (H = 112.801, p = 0.000) on the right side while significant interactions were identified between age-at-death and ancestry (H = 169.871, p = 0.000) and age-at-death and biological sex (H = 344.630, p = 0.000) on the left side. Details and implications of these interactions will be discussed in the Chapter 5.

4.7 Comparison of Classification Accuracy between Left and Right Innominates

As depicted in Table 12, which compares the classification accuracy attained from the earlier examination of validity of the left innominate with classification accuracy of the right innominate in the current examination, the classification accuracy for the right side improved only slightly for males at 89.4%. Thirteen of 104 males were incorrectly classified as female. Classification accuracy for females actually decreased by 1%, as the innominates of two females were incorrectly classified as male (98/100). Nevertheless, sex bias improved by 2.9%. Such results indicate that even though asymmetrical values had been observed in one-third of the sample, asymmetry does not appear to affect the accuracy of the method when either the left or right side is considered individually.

Table 12. Classification accuracies (%) using Klales et al.'s (2012) logistic regression equation: Lefts and rights considered separately.

Sex	Males	*n*	Females	*n*	Combined	*n*	Sex Bias
Lefts	87.5	104	99	100	93.6	204	11.5
Rights	89.4	104	98	100	93.6	204	8.6

Conversely, when both the left and right innominates of an individual are considered in tandem, and whereby both innominates must estimate sex correctly, the rate of the method's combined accuracy drops by 3.9% to 89.7%, and the level of sex bias increases markedly to 16.3% (Table 13). Five of the 104 males were incorrectly classified as female on both the left and right innominates in tandem (1.5%). Eight of the 104 males (7.7%) were incorrectly classified as females on the left innominate only, while seven of the 104 males (6.7%) were incorrectly sexed as female on the right innominate only. Two of the 100 female innominates (2%) were incorrectly classified as male on the right side, while one female innominate was incorrectly classified as male on the left.

Table 13. Classification accuracies (%) using Klales et al.'s (2012) logistic regression equation: Left and right innominates considered in tandem.

Sex	*n*	Accuracy (%)	Combined	Sex Bias
Males	104	81.7	89.7	16.3
Females	100	98.0	-	-

When those individuals who were affected by asymmetry were removed from consideration, classification accuracy for males increased significantly when both innominates were considered for correct classification (Table 14). With this restricted sample, only four of the 89 symmetrically scored male innominates (4.5%) incorrectly sexed as females, while only one of the 99 symmetrically scored female incorrectly sexed as male.

Table 14. Classification accuracy for left and right innominates of symmetrical individuals only.

Sex	*n*	Accuracy (%)	Combined
Males	69	95	96.4
Females	67	99	
Total	136		

When consideration is limited to asymmetry-affected individuals, the rate of correct sex classification among males decreased significantly when both innominates were considered for correct classification (Table 15). Eight males incorrectly sexed as female on the left side and seven males incorrectly sexed as female on the right side, meaning that 15 of 35 asymmetry-affected males (42.9%) were misclassified by sex. Two females incorrectly sexed as male on the right side.

Table 15. Classification accuracy for left and right innominates of asymmetrical individuals only.

Sex	*n*	Accuracy (%)	Combined
Males	35	43	69.20
Female	33	97	
Total	68		

4.8 Re-calibration

Given that the results of the direct comparison between the left and right innominates individually found classification accuracy to be identical when used in isolation, the scores of the left and right innominates were pooled during the re-calibration to create an equation that is more sensitive to the distribution of trait scores observed for both the left and right innominates; thus creating an equation valid and applicable to either innominate in a forensic or bio-archaeological context.

Klales et al.'s (2012) original ordinal logistic regression equation (2.726 (VA) + 1.073 (SPC) + 1.214 (MA) – 16.312) was re-calibrated in SPSS 23.0 based on the WBSC in order to generate a sample specific equation in an attempt to improve classification accuracy. The resulting equation was applied to the left and right innominates of all 204 individuals and compared with the accuracy of the original equation:

$$2.067\ (VA) + 2.220\ (SPC) + 1.335\ (MA) - 15.396 \qquad [1]$$

This equation produces what is known as a positive weight (or e^{score}). This equation is applied in accordance with Press and Wilson (1978): probability of being female ($p_f = 1/(1 + e^{score})$ and the probability of being male ($p_m = 1 - p_r$). Examination of the coefficient loadings of each trait included in this equation revealed that the most sensitive indicator of sex among members of this sample is the VA (2.067), followed by the SPC (2.220), with the MA (1.335) being the least sensitive of the three. To confirm, the average weighted score by trait was calculated. This was done by multiplying the number of individuals by sex that receive grade one and multiply by one. The same calculation was made for grade two, but multiplied by two. The same calculation was made for all grades and finally averaged by sex. This yields the following weighted average scores by sex: VA (Males = 183.2; Females = 66.8), MA (Males = 158.8; Females 77.2), SPC (Males = 159; Females = 67). Next, the ratios of scores were calculated for each trait by sex: VA: (183.2/66.8 = 2.743), MA (158.8/77.2 = 2.057), SPC (159/67 = 2.373). Finally, the contributions of all three predictor variables were summed: 5.669 (VA) +2.746 (MA) +5.268 (SPC) = 13.683. That sum was next divided by each individual predictor variable to determine the proportionate contribution of that predictor to the response variable: VA = (5.669/13.683)*100 = 41.43%, MA = (2.746/13.683)*100 = 20.07%, SPC = (5.268/13.683)*100 = 38.50%. With these calculations it was determined that the VA provides 7.6% greater explanatory power than SPC (41.43/38.5) = 1.076 and VA provides 106% greater explanatory power than the MA (41.43/20.07) = 2.064. Lastly, SPC provides 91.8% greater explanatory power than the MA (38.5/20.07) = 1.918.

When applied to the sample of 204 individuals from the WBSC the re-calibrated equation improves combined classification accuracy (99.0%) by 5.4% (Table 16) as compared to the overall accuracy of the original equation using all three traits (93.6%). Sex bias decreased significantly to only 1%.

Table 16. Classification accuracies (%) using the re-calibrated logistic regression equation.

Sex	*n*	Accuracy	Sex Bias
Males	104	99.5	1
Females	100	98.5	1
Combined	204	99.0	-

The resulting sample specific equations for estimation of biological sex of a United States sample using individual traits and combinations thereof were created (Table 17). Classification accuracies are based on the left and right innominates of 204 individuals from the WBSC. When

Table 17. Classification accuracy (%) for each individual trait and trait combinations.

Trait	Equation	Males	Females	Combined	Sex Bias
VA	3.146 (VA) -9.390	99.0	88.0	93.6	11.0
SPC	3.615 (SPC) -7.980	87.5	96.0	91.7	8.5
MA	2.995 (MA) -8.623	97.0	80.5	89.0	16.5
MA, VA	1.871 (MA) + 2.624 (VA) -13.275	98.0	94.0	96.0	0.4
MA, SPC	2.061 (MA) + 2.840 (SPC) -12.359	96.2	96.0	96.0	0.2
SPC, VA	2.681 (SPC) + 2.426 (VA) -13.623	99.0	98.5	98.8	0.5

single indicator of sex (traits) are employed, greatest classification accuracy is obtained with the ventral arc (93.6%), however this accuracy is accompanied by a sex bias of 11% in which accuracy is superior for males (99.0%) over females (88.0%). The conformation of the subpubic concavity is the second best indicator of sex (91.7%) with less sex bias (8.5%). In a dramatic reversal of results obtained with the ventral arc, females are classified with greater accuracy

(96.0%) than males (87.5%). Not surprisingly from the ordinal logistic regression equation formula presented above, poorest results were obtained with the conformation of the ischio-pubic ramus (89.0%), which also yielded the most sex bias of all three indicators at 16.5%. Like the ventral arc, classification accuracies were markedly greater among males (97.0%) than among females (80.5%).

Chapter 5 DISCUSSION

Of the studies reviewed on pelvic sex estimation, very few had to contend with the fallibility of their methods in the eyes of the courts, and even fewer meet the criteria outlined by *Daubert* to ensure quality and consistency in the fields of physical and forensic anthropology. This study has analyzed the Klales et al. (2012) method in light of the *Daubert* criteria and the impact of asymmetry on the revised method. The results of this study will be discussed in detail below. This discussion is in three parts. The first will address the results of the validation study in light of the *Daubert* standard, the second examines the impact of re-calibration of the original ordinal logistic regression formula based upon Hamann-Todd Collection (HTC) for use with individuals encompassed by the William Bass Skeletal Collection (WBSC), and the third will address the findings of the impact of asymmetry on the Klales et al. (2012) technique.

5.1 Classification Accuracy

The application of Klales and colleagues' equation to the sample used in the present study produced classification accuracy consistent with that of experienced observers in the original study (93.6%). This technique was consistently more reliable for classifying females which echoed results obtained by Phenice (1969). Based on the observations made by Klales et al. (2012) males were correctly classified in 99% of cases, while females were correctly classified in 92% of cases; thereby yielding a sex bias of 7%. Results of the present study produced a sex bias of 11%. While other factors may be responsible, the difference in sex bias is likely due to sampling error. Not an error representative of fault or incorrectness but one reflecting differences in the ancestral composition of the two samples considered. Perhaps the greatest difference in the composition of the two samples are the proportional representation of blacks relative to whites in the HTC (n = 41 black females, 42 white females, 43 white males, 44 black males) and the inclusion of a few American Indian and Asian Americans in the WBSC. These findings are suggestive of certain sexual and ancestral tendencies where sexual dimorphism is more marked between black males and females than among whites. Even within restricted geographical regions patterns of sexual dimorphism sometimes vary significantly (Walker, 2008). In this case, the differences are reflected in the expression and distribution of the traits of Phenice between males and females. Differences in pelvic morphology between blacks and whites have been reported by Letterman (1941) who found statistically significant disparities of the sciatic notch between blacks and whites with greater variability among white males and females than in black males and females across several dimensions of the ilium. Furthermore, considering that the results of the SRH test identified differences in expression of the three traits between blacks and whites, it is reasonable to attribute the increase in sex bias to a marked skew in the ancestral composition of the WBSC in favor of white males and females. Therefore, it is expected that trait expression would be more consistent with a sample upon which that method is based than to another sample, regardless of whether the latter sample was marked by a similar ancestral composition. The age-at-death profiles of the two samples are similar and hence differences in age are less likely to have contributed to the elevated sex bias in the current research.

Lovell (1989) found a moderate negative correlation between age-at-death and correct classification when applying Phenice's (1969) technique. As-at-death age increased, accuracy decreased. Using a much larger sample, the present research was unable to substantiate Lovell's conclusion. Age in the current sample ranged from 23 to 99 with a median age of 62. While false classifications were low overall, incorrect classifications were proportionately similar for individuals throughout the fourth, fifth, and seventh decades of life, with the greatest percentage of false positives occurring among males during the sixth decade of life. There are too few individuals in the sample who died in the eighth and ninth decade of life to contribute meaningful information.

According to Phenice, the presence of a VA is the female condition. While 54% of the males were assigned a score of 5 for this trait, 46% had a ventral arc of varying degrees and received scores of two, three, or four. Therefore, 46% of males in this sample display the "female" condition of this trait as described by Phenice. According to Phenice, a ridge of bone found on a narrow MA is the female condition. The scale created by Klales and colleagues (2012) however, describes two variations of a narrow MA with either a sharp or rounded ridge, or a narrow to medium ramus with no ridge. While the majority of males were assigned a score of three, four, or five, four males were found to possess narrow ischio-pubic rami, two of these were accompanied with a sharp ridge, while the other two were accompanied by a rounded ridge. Of the five possible variations on the Klales et al. (2012) scale, manifestation of the SPC in the current sample encompassed all but the fifth character state. With regard to the subpubic concavity (SPC), Phenice simply describes a recurve below the symphyseal face in females, while males lack this characteristic. Some of the males in this sample were assigned a score of two 13%, reflecting a slight degree of concavity. Klales and co-worker's expansion of Phenice's original presence or absence of the male or female condition into five grades of expression clearly improves the technique as most of the variations of each trait were observed in this sample and the classification accuracy between this study and that achieved in the original research remained consistent.

Given these results, approximately 50% of VAs and MAs and 13% of SPCs belonging to males are characteristically "female" according to the original standards of Phenice. These numbers are alarming considering the majority rule protocol prescribed by Phenice. The present research concludes that expansion of the scale, paired with the ordinal logistic regression model has greatly improved upon Phenice's technique by departing from the strict dichotomous decision table and thereby providing a superior accounting of the distribution of these traits. However, a more robust sample of various age groups should be examined to understand how these traits change from the onset of skeletal maturity to advanced age. Nevertheless, the results obtained in the current study indicate that age alone does not greatly affect classification.

The *Daubert* standards demand that methods utilized in legal proceedings produce estimated error rates so that their validity may be assessed. Classification accuracy reached 94% with only an error rate of 6% in the present study. Consequently, this method has proven to be highly reliable and greatly accurate. Nevertheless, it must be pointed out that classification accuracy is based on the model as it applies to the specific sample. Classification accuracy in and of itself is only a portion of reality. A model will not perform identically in every set of circumstances. Therefore, it is the opinion of the analysts that is of most value. Considering that the high classification accuracies of the Klales et al. (2012) technique on two external samples are consistent with those produced in this study, and the near perfect results of the intra- and inter-observer tests, the present verification study validates this method as both reliable and accurate.

5.2 Reliability

While tests of intra- and inter-observer error do not affect classification results directly, they do measure the repeatability of a technique, for departures from perfect replication which influence classification accuracy indirectly. An unreliable technique produces results that cannot be replicated by two or more analysts, or by the same observer on separate occasions. Therefore, the ideal technique should be repeatable, clearly understood, and uncompromised by inexperienced analysts.

The present research demonstrates high to near perfect agreement between a single observer (the author) and two other observers of diverse knowledge and experience (the author and Klales). Using a printed copy of the photos and illustrations presented by Klales and colleagues (2012) the author, with moderate training through formal education, scored a subsample of 25 innominates consistently with that of Klales, a professional anthropologist with extensive experience. Although experience levels between the two observers varied, agreement between scores was nevertheless high. These results are a vast improvement from inter-observer error reported by MacLaughlin and Bruce (1990) based on Phenice's (1969) method. The majority of the scores were relatively easy to make by comparing the innominate in proper orientation in accordance with the descriptions and illustrations provided by Klales et al. (2012). However, this research found that, all possible variations of the medial aspect of the ischio-pubic ramus (MA) had not been captured by the ordinal grades devised by Klales et al. (2012). According to Klales et al. (2012), a score of three should be assigned if the "ascending ramus of the MA is narrow dorso-ventrally with no plateau/rounded ridge present below the symphyseal face." A score of four should be assigned when "the ascending ramus is medium width dorso-ventrally, also with no ridge present." Nonetheless, it was my experience that an ascending ramus of medium width dorso-ventrally was observed that also possessed a ridge. This variation did not fall neatly into any of the five character states. In these cases the specimen was assigned a score of four. The width of the ascending ramus may be considered the most important factor, while the presence/ absence of the ridge(s) was of lesser importance. Perhaps, a simple addition to the description of a possible faint ridge on ascending rami of medium width MAs may be added to the protocol published by Klales et al. (2012). This is a minor issue as MA was scored with perfect agreement in both the intra-and inter-observer comparisons.

Scores for the SPC were found to be the most variable between two separate observers and the same individual making the observation at two different points in time. No males were assigned a score of one for this trait and no individuals in the sample were assigned a score of five. These results were unexpected because the absence of an arc (the requisite for assignment to grade five) has been attributed to males historically. This suggests that the slight presence of a concavity of the male SPC, albeit not as pronounced as in females, is far more common than had been realized by Phenice (1969). While 96% of female innominates received scores of either a one or two, a few received scores of three or four; which further substantiates the idea of a classic female SPC as observably convex. As noted in the results section above, during the intra-observer examination, two SPCs initially scored by the author as a three were later scored as a four. Most scoring discrepancies centered on ordinal scores that fell in the mid-range expressions, for in this middle range the variations are often subtle and hence straddle the thresholds that separate them.

It has been argued that visual assessment of the morphological appearance of skeletal elements is not as reliable as metric analyses in which size differences are quantified by traditional measurements. As such, some critics have deemed non-metric analyses as subjective and compromised by the experience (or lack thereof) of the observer (Bruzek, 2002). However, the results obtained in the present study indicate that the ordinal scale created by Klales and colleagues (2012) has standardized the traits of Phenice (1969) in such a way that facilitates consistent results by different observers with varying levels of experience. Therefore, analysts with minimal practice can apply this technique with great accuracy. With the minor adjustment made to include the slight variation of the MA observed by the author, all variants of the trait will be included on the scale and will further refine the parameters of its descriptions and its goal to capture all possible manifestations. *Daubert* stresses that replicable methods are to be used to justify scientific evidence. Therefore, testing the replication of a method and its findings are essential part of verification. The present research concludes that the technique involved in application of the Klales et al. (2012) method is, in fact, replicable.

5.3 The Presence of Asymmetry

Historically, studies that have examined symmetry of non-metric traits have most often used a trait's presence or absence on both sides of the body, cranium or dentition to establish ancestral associations, patterning of kin groups in cemeteries, or genetic links to other contemporary populations. Other studies have focused on traditional metric measurements of long bones to examine the impacts of handedness and mechanical loading. This study involved the examination of non-metric traits and their possible differences in expression due to differential use of the left and right appendages. However, the traits are not measured using traditional metric techniques; instead, the traits of Phenice were measured through visual assessment based on the ordinally graded scale developed by Klales and colleagues (2012). The scores were treated as a measurement along an ordinal scale and used to evaluate asymmetry in this research.

In the current study one-third of all females (33%) and males (34%) were found to express one or more of the three traits asymmetrically. While asymmetrical occurrences had been observed among all three traits a Wilcoxon signed ranked test revealed that statically significant asymmetrical expression was exclusive to the ventral arc (0.05>0.02). As established earlier in this thesis, the VA, and MA are known points of origin for an array of muscles responsible for movements involved in daily activity, while the SPC stretches and elongates during development, especially among females. Therefore, it is not surprising that the present research identified greater prevalence of asymmetry of the VA and MA with very few cases of asymmetry affecting the SPC. This pattern suggests that the habitual movements involving the muscles adductor magnus and brevis as well as the gracilis are manifesting themselves on their points of origin on the pubes of individuals, but only asymmetrically in one-third of the sample. Being that no individuals in the sample were observed to have suffered from deformity due to accidental trauma to the pelvis a few alternative explanations may account for the presence of asymmetry in the sample.

The first involves developmental deformity and/or disease of bone occurring early in life and becoming more pronounced with advanced age. For example, some osteochondrodysplasias (abnormalities of cartilage and/or bone growth), such as scoliosis, are identifiable at birth and can be corrected; however, some are not identifiable until later in life. Scoliosis, a lateral curvature of the spine, often results in uneven posture at the shoulders and/or hips. While extreme cases of such conditions can be surgically corrected, mild cases are often treated with physical therapy or corrective braces. However, years of physical activity with even a mild misalignment becomes readily apparent on skeletal material and can be observed macroscopically in the form of asymmetrical structures. This response is due to the amount of stress, or mechanical loading placed on the bony structures involved in physical activity over a prolonged period. Bone will increase or decrease in mass reflecting the amount of functional pressure forced upon it (Wolff, 1986). This explanation may be possible for some (as other skeletal elements were not assessed during the course of this study) but, certainly not all asymmetrical individuals in the sample.

However, given the rarity of osteochondroplastic disorders, some other underlying cause must be taking place given that asymmetry affects one-third of the sample. Much like long bones that increases in mass when subjected to repetitive stress, smaller bony features react in a similar fashion. This response is caused by the tension when the muscle contracts, which places stress upon the periosteum adjacent to the tendon covering the structures to which they are attached, such as tubercles, crests, and/or ridges. As described in Chapter 1, the ischio-pubic ramus possesses a ridge of bone, the ventral arc, which as its name implies; both the ramus and the arc are points of origin for muscles and are thereby subject to tensile stress, and contributes to the morphology of the innominate. Such tension stimulates bone growth, increasing osteogenesis and the mass of bone to which the tendon for a specific muscle is attached. This could suggest that the "pseudo-arc" (often misidentified as a true arc found on females) and the ridge of the MA observed on some of the pubes of males may be large and visibly apparent due to repetitive muscle contraction, and hence tensile stress placed on its bony origin over time. These structures may also manifest asymmetrically if a dominant appendage is involved in certain movements. However, while such repetitive and often low-level bouts of force stimulate bone growth, it is also the case that rare but excessive muscle pull can lead to cortical recession of bone resulting in resorption rather than deposition.

An overwhelming majority (78%) of the asymmetry observed in the current study occurred for the ventral arc and ischio-pubic ramus. Given that in majority of cases, scores of the arc of the VA and the ridge on the MA increased on the right side and that the ramus of the MA became thicker could suggest repetitive behavior with a dominant right appendage. Previous studies have shown that squatting/flexion of the knee, and sitting with legs crossed for prolonged periods have been linked to occupational activities that lead to macroscopically visible changes to bones in the form of bilateral osteitis, and erosion of the ischial tuberosity, osteoporosis and osteoarthritis accompanied by lesions of the knee and at the condyles of long bones. Repetitive flexion of the knee has been identified in coal miners repetitively engaged in squatting while hewing a low seam of coal, soldiers who jump from a squatting position repetitively and movements involved in martial arts (Capasso et al, 1999). Macroscopic evidence of repetitive sitting with legs crossed has been identified on skeletal material and associated with occupations requiring prolonged sitting such as tailoring (Capasso et al, 1999).

Given these findings it is reasonable to expect that those individuals identified as being asymmetrical at those traits participated in activities that may have been either casual or work related but, above all, habitual. In doing so, the dominant appendage used to stand from a kneeled position, the supporting knee of the crossed leg, the leg used to initiate a sprint or jump, or the kicking leg in some unilateral sports may be more prominent. However, being that so many components of the muscles and bones involved in locomotion and the processes taking place with each movement, it is not clear to as to which of these processes are directly responsible for the asymmetry observed on these particular three traits as they are part of a broader system. What is clear is that the 33% of individuals who are asymmetrical in this sample were distinct from the remaining sample in some way that caused the differential morphology of their pubes. A thorough investigation of their life histories and examination of their entire skeleton is needed to identify the source of these differences definitively.

Both blacks and whites appeared to be affected proportionately by asymmetry suggesting that both ancestral groups were exposed to similar levels of stress. This refers only to the presence of asymmetry, and does not explain the pattern of asymmetry. Intriguingly, asymmetry was not found among individuals who died in their second or third decade of life. This suggests that asymmetrical differences in this region developed near and during the fourth decade of life or later. Further still, the incidence of asymmetry was found to increase with advancing age at death. This is reflected by the fact that 24% of individuals in their forties were asymmetrical; 33% of those in their fifties, 40% of those in their sixties, 41% of those in their seventies, 45%

percent of those in their eighties, and 100% of those in their nineties (albeit the sample size for those in the eighth and ninth decade of life is extremely small), none of which were found to be statistically significant. However, an increase in severity was not observed. Those of unknown age who were asymmetrical were not considered.

5.4 Type of Asymmetry

In an attempt to better understand asymmetry and its implications when present, those who were not asymmetrical were not considered in the identification of the type of asymmetry. Fluctuating asymmetry was ruled out as all three traits were found to display moderate or greater directional asymmetry in favor of the right side. On the Klales et al. (2012) scale, the MA, VA, and SPC progressively transition in form. The VA is first very apparent and well-defined and progressively becomes smaller until ultimately no ridge is observed. This means that the arc of 65% of those found to be asymmetrical at this trait are less apparent on the right side. The MA starts off narrow and widens progressively, ultimately appearing broad and ridgeless. Like the VA, the 75% of the ridges on the MA are becoming less apparent, while at the same time the MA is increasing in width dorso-ventrally on the right side. The SPC is artfully curved ending in a rugged convexity. This convexity however, is less visible on the right pubis in 73% of those found to be asymmetrical for this trait.

As mentioned above, habitual force on sites of muscle attachment are capable of increasing bone mass in accordance with Wolff's Law. It is more likely that muscle-induced morphology is responsible for the directional asymmetry found in this sample than disease or early developmental defects, for no apparent evidence pointed to disease. However, while previous studies often attribute directional asymmetry to biomechanical loading based on Wolff's Law, given the proportion of asymmetrical individuals in the present research, it is difficult to conclusively identify the physiological processes responsible for the asymmetry identified in this sample. A more detailed study and deeper understanding of how these traits change over a lifespan is needed.

5.5 Interactions and Correlations

Results obtained with the Sheirer-Ray-Hare extension of the Kruskal-Wallis H test revealed interesting relationships between the left and right pubes. While expression of the ventral arc on the left and right pubes is highly correlated with each other, as well as with biological sex, no significant interactions were found. This suggests that the VA is the most neutral of the three traits with regard to how it manifests. This is not surprising as this trait is the least variable of the three; however, it was found to be the most asymmetrical. As noted above, this is likely a consequence of the severity and recurrence of tension placed upon the periosteum in this region of the pubis due to the muscles that originate at the VA.

The results indicate significant interactions between age and ancestry and age and sex on the subpubic concavity (SPC). The interaction between age and ancestry shows a greater variability of the SPC among white males of all ages than observed among black males. This could simply be due to the prevalence of white males and very few black males in the sample. While several white males were found to be asymmetrical at this trait, no such incidences occurred among black males. Black and white females appeared to be more similar at this trait with a tendency for a slight age progressive increase in score among white females. One black female and two white females were found to be asymmetrical at this trait. The interaction between age and sex reveals a similar pattern for manifestation of the SPC. Overall, it appears that females generally scored lower at this trait than males regardless of age. By contrast, males of all ages were more variable at this trait. However, it would appear as though age is a viable factor for manifestation in this particular test except that we are aware that females generally score lower than males, even in advanced age. These interactions are not surprising being that the scores of white males were so variable. This information suggests that the clear and apparent convexity of the SPC is more common in youthful females and is more likely to appear less convex in white females but still easily observable with advanced age. Males, on the other hand, do not exhibit any age dependent changes at this trait; nevertheless, white males are more likely to be asymmetrical.

A significant interaction between age and ancestry on the left and right MA reveals that white males and females tend to score higher at this trait than their black counterparts. Furthermore, while several white males and females were found to be asymmetrical for this trait, no black females and only two black males were. Again, white males of all ages appear to be distinctly more variable at this trait relative to black males who there are simply fewer of in the sample. However, the interaction between age and sex on the left MA is noteworthy. The MA is determined to be the most variable trait of the three. Test results reveal that scores for males span the entire scale with some male pubis scores falling into each of the five categories. Generally an increase in score corresponds with an increase in age among males. Scores for women, however, do not span the entire scale at this trait with scores of one, two, and three for mixed age groups with only two females scoring a four. This suggests that the expression of this trait is not as age dependent for females as it is for males, who tend to score higher for this trait as age increases. This suggests that the left MA is becoming thicker dorso-ventrally with an increase in age. However, this relationship does not hold true for the right MA where no such interaction was found. Given this information, and the physiological rules outlined by Wolff's law, it is reasonable to attribute the increase in width of the left MA to the pull and force of the muscles attached to this area of the pubis.

5.6 Comparison of Classification Accuracy between Left and Right Innominates

In comparing the classification accuracy of left innominates with that of right innominates across the entire sample, it became apparent that asymmetry does not affect the accuracy of the Klales et al. (2012) method critically. Though the combined classification accuracies for the left and rights were identical (93.6%), classification of females was slightly higher and sex bias slightly lower when sex estimation was based on assessment of the right innominate. Classification accuracies of males changed quite a bit when both the left and corresponding right innominates were required to yield the correct biological sex to be considered a correct classification. When asymmetrical individuals were removed from the sample classification accuracy was highest (96.4%). On the contrary, classification accuracy plunged to a meager 69% when only asymmetrical individuals were analyzed. While sex estimation remained high for asymmetrical females at 97%, males were correctly classified in only 43% of cases (less than expected by chance alone!) where both left and rights were considered in tandem. However, when classification accuracy of asymmetrical males is restricted to either the left or right pubis, accuracy is as high as 78%. Given this information, it appears that while both males and females appear to be equally affected by asymmetry in one-third of the sample, asymmetry has little effect on the classification accuracy of females using the left, right, or both innominates in combination. However, asymmetry appears to be most detrimental to males when both innominates are used in combination to render correct sex classification.

Overall, 21 of the 204 individuals (10.3%) included in this sample were falsely identified as the incorrect biological sex. While directional asymmetry in favor of the right side is found to be present in asymmetrical scores, such differences, did not introduce bias of sufficient strength to justify a systematic selection of the left or right pubis. No significant differences were found in the frequency with which the left or right pubis renders correct sex classification. Though slightly higher for females using the right pubis, combined classification accuracy was identical for the left and right pubes considering the entire sample. Nonetheless, directionality is relevant to better understand the processes influencing the asymmetry of the pubes in future research. Also worth noting is that asymmetry appears to increase with age with the fourth decade of life serving as a threshold for manifestation. Being that asymmetrical form of these traits is permanent, older individuals are more likely to be asymmetrical at these traits.

While this asymmetry doesn't appear to be detrimental to the use of these traits for sex estimation, it does shed light on how these traits of the pelvis are affected by habitual activity, which hasn't been investigated prior to this study. Given the proportion of symmetrical individuals and the close proximity of scores of those who are asymmetrical, the traits of Phenice appear to be symmetrically stable. One of the most interesting observations of this study was that all the asymmetrical individuals could be correctly classified by using at least one of their pubes. None of the asymmetrical individuals were incorrectly classified on both pubes. This information suggests that the factors influencing the asymmetry present on the pubes differ from the processes responsible for the progression of the three non-metric traits of Phenice documented and scored by the Klales et al. method.

5.7 Re-calibration Accuracy

While the original ordinal logistic regression equation using all three traits performed well on the WBSC, a rather strong sex bias in favor of females remained nevertheless. It has been proposed that a sample specific equation will decrease sex bias and further increase classification. Results of the re-calibration increased classification accuracy from 93.6% to 99% and decreased sex bias from 11% to 1%. This is because the re-calibration measured the parameters and magnitude of trait manifestation in this specific sample and adjusted the posterior probabilities accordingly. This adjustment increased the precision of the ordinal logistic equation for the estimation of sex within the current sample. This increase in classification accuracy is consistent with results obtained by other researchers in samples of modern South Africans (99.2%) and Mexicans (99.2%) (Kenyhercz, 2012; Gomez-Valdes et al., *in review*). Surprisingly, classification accuracy for males was higher for females following the recalibration. This shift in classification accuracy in favor of males is consistent with that seen for the Mexican sample (Gomez-Valdes et al., *in review*), but is opposite to the results obtained from the South African sample (Kenyhercz, 2012) where classification accuracy was higher among females.

Classification accuracy of individual traits and combinations thereof followed much the same pattern with the exception of the SPC. Classification accuracy using the newly developed equations is high. These equations are significant as they are useful in situations in which one or more traits cannot be observed due to erosion and/or fragmentation. Similar to Klales et al. (2012), the ventral arc individually and paired with the other two traits proved most reliable followed by the subpubic contour (SPC) and the ischio-pubic ramus (MA). These results are consistent with those of Phenice (1969) who also found the VA to be the most reliable but contradicts the findings of Gomez-Valdes et al. (*in review*) who found accuracy to be highest for the sub pubic contour in two Mexican samples. However, sex bias varied greatly for each trait. Sex bias was consistently lowest for the SPC (<1%), both individually and when paired with the other two traits.

The results of this recalibration demonstrate the impact of sample-specificity and should be considered in the development and application of techniques for demographic analysis in physical and forensic anthropology. The scientific working group for forensic anthropology lists the use of population- and period-specific standards as best practice when available for sex assessment. Therefore, application of the sample-specific equations developed in the present study are currently the best suited for anthropological use in the when estimating sex of U.S. blacks and whites.

Furthermore, unlike other studies who have re-calibrated the previous equation based on scores restricted to the left innominate, the present research has created an equation(s) encompassing the full range of variability among both the left and right innominates. This technique accounts for asymmetrical incidences found within my sample. While the combined classification accuracy had been identical for both the left and right innominates using the original regression equation developed by Klales et al. (2012) (93.6%), so had the misclassifications of males on both the left and right innominate drastically decreasing classification accuracy to 81% when both innominates were employed in tandem. Additionally, as the current research has uncovered preponderance for whites to score higher on the right innominate for the VA and MA using both the left and right scores has effectively avoided over-fitting or under-fitting the classification function by only focusing on trait manifestation on a single side. Failing to account for the right side creates a biased equation that could be detrimental to individuals with reverse symmetry, for instance dominance on the less frequent side. This would be a mis-step as those individuals would be misclassified more frequently. Therefore, the utility of this re-calibrated equation is not limited to a specific side, which would be a major limitation for forensics and bio-archaeology.

Chapter 6 CONCLUSION

6.1 Hypothesis

This study aimed to accomplish four primary objectives: 1) to test the validity of the Klales et al. (2012) method on a separate sample of inhabitants of the United States from which served as the basis for the Klales et al. method, 2) to test the reliability of the Klales et al. method, 3) to assess the symmetrical stability of the traits of Phenice and determine to what extent asymmetry may compromise classification accuracy when using the Klales et al. method to estimate biological sex and, 4) to obtain optimal classification accuracy by re-calibrating the original equation to fit a specific sample of U.S. inhabitants (the WBSC). Those endeavors were formally outlined in formal hypotheses and investigated in the present study.

The first hypothesis states:

H_0: Klales et al.'s (2012) revised method of Phenice (1969) is a valid technique, such that the rate of correct classification by sex among known-sex individuals in a specific sample of modern inhabitants of the United States is greater than 85%.

H_1: Klales et al.'s (2012) revised method of Phenice (1969) is not a valid technique, such that the rate of correct classification by sex among known-sex individuals in a specific sample of modern inhabitants of the United States is greater than 85%.

With regard to hypothesis one, the null hypothesis is accepted as the Klales and colleague's original equation had produced accuracy rates in excess of 85%. This conclusion is based on results of the combined classification accuracy obtained by Klales (86.2% > 85%) and the author (93.6% > 85%) In both cases, classification accuracies exceed the accepted level of 85% accuracy.

The second hypothesis states:

H_0: Klales et al.'s (2012) revised method of Phenice (1969) is sufficiently reliable, such that error rates in repeatability occur in fewer than 5% of cases.

H_3: Klales et al.'s (2012) revised method of Phenice (1969) is not sufficiently reliable, for error rates in repeatability occur in greater than 5% of cases.

With regard to hypothesis two, the null hypothesis is accepted. The revised method of Phenice (1969) is sufficiently reliable meaning that error in repeatability rates does not occur in excess of 5% of cases. This conclusion is based on the results of the inter-observer test, which showed near perfect to perfect levels of agreement between Klales and the author with no statistically significant differences in scoring in excess of 5% using a two way mixed model of the ICC. Such results suggest the method is capable of producing replicable results by observers possessing various levels of experience.

The third hypothesis states:

H_0: Klales et al.'s (2012) revised method of Phenice (1969) is not compromised by asymmetry between sides, for less than 5% of cases yield opposite sex identifications by morphological features.

H_4: Klales et al.'s (2012) revised method of Phenice (1969) is compromised by asymmetry between sides, for greater than 5% of cases yield opposite sex identifications by morphological features.

With regard to hypothesis three, the null hypothesis is accepted as combined classification accuracy is identical for the left and right innominates at 93% when only a single innominate is used for estimating the biological sex of human skeletal remains. This conclusion is based on the results of the classification accuracies produced for the left innominates and the right innominates using the original ordinal logistic regression equation produced by Klales et al. (2012). As the application of the equation is applied to only a single side in the proposed method by Klales and colleagues (2012) it appears that the classification accuracy for males improved minimally on the right side (left = 87.5%, right = 89.4%), while the accuracy for females remained high (left = 99%, right = 98%).

The fourth hypothesis states:

H_0: A re-calibration of Klales et al.'s (2012) ordinal LR equation to fit a specific sample of modern inhabitants of the United States does not increase the rate of correct classification by sex among known-sex individuals.

H_2: A re-calibration of Klales et al.'s (2012) ordinal LR equation to fit a specific sample of modern inhabitants of the United States increases the rate of correct classification by sex among known-sex individuals.

With regard to hypothesis four the null hypothesis is rejected and the alternate hypothesis accepted as the re-calibrated ordinal logistic regression equation increases classification accuracy. This conclusion is based on the combined results of the original classification test (93.6%) and the re-calibrated classification test (99.0%). Classification accuracy using all three traits had increased by 5.4%. However, due to the specificity of this equation based on visual observations on a sample of individuals from the WBSC, the re-calibrated equations may not be as accurate for samples or populations of people in other regions of the world. It is the responsibility of the analysts to develop and/or refine existing methods based on firmly established theories as has been done in the present research.

6.2 Implications

Implications of this study suggest that asymmetrical expression the traits of Phenice are potentially problematic for correct classification of males, but only when both the left and right innominates are considered in tandem. However, the potential for correct classification is equal when either the left of right innominate is considered in isolation. By contrast, correct classification of females is not compromised by asymmetrical trait expression. The present research also revealed an age-progressive correlation in asymmetrical incidence. This is perhaps because older individuals have had more time to participate in the habitual activities responsible for the asymmetrical differences observed on the innominate in general and the pubis specifically. The SRH test also revealed differences in the expression of the traits of Phenice between blacks and whites. Generally, whites tend to receive higher scores than blacks. These observations must be considered with great caution as there were few black individuals of known sex in the sample. Furthermore, differences in pubic morphology between blacks and whites have been investigated throughout the anthropological literature and such results support the argument that blacks and whites differ somewhat in pelvic morphology, including the pubis. However, the differences due to genetic ancestry were slight and do not warrant the development of separate sets of standards for blacks and whites as Klales and colleagues (2012) standard does not appear to be greatly compromised by these differences as claimed in the original article.

6.3 Closing Statement

The results presented here strongly support the suppositions made by previous authors (Kenyhercz, 2012; Klales et al., 2012; Gomez-Valdes et al., *in review*) that the newly revised method is a promising development for use in estimating the biological sex from the skeletal remains of unknown individuals. The results of the present research conclude that the Klales et al. (2012) method is valid and capable of producing reliable results. It is also possible to fit the original regression equation to a specific sample, thereby increasing its effectiveness beyond the original sample from which the method was developed. Furthermore, while asymmetrical differences were revealed during the course of this study, those differences do not compromise the current technique. Many aspects of Klales and colleague's revision sought to meet the evidentiary standards of *Daubert* and have effectively accomplished those goals by producing a standardized scale from which observations can be measured against, producing posterior probabilities of which error can also be effectively measured, and has been validated by external researchers on several samples. The current research is in support of the newly proposed method and the effort to launch Phenice (1969) into the new phase of forensic anthropology.

REFERENCES

Acsadi G, Nemeskeri J. 1970. History of human life span and mortality. Budapest: Akademiai Kiado.

Albert AM, Greene DL. 1999. Bilateral asymmetry in skeletal growth and maturation as an indicator of environmental stress. Am J Anthropol 110:341-349.

Anderson BE. 1990. Ventral arch of the os pubis: anatomical and developmental considerations. Am J Phys Anthropol 83:449-458.

Auerbach BM, Raxter MH. 2008. Patterns of clavicular bilateral asymmetry in relation to the humerus: Variation among humans. Hum Evol 54:663-674.

Bass WM. 2005. Human osteology: A laboratory and field manual, 5th ed. Springfield: Special Publication No. 2 of the Missouri Archaeological Society.

Berry KJ, Johnston JE, Mielke PW. 2014. A chronicle of permutation statistical methods. Switzerland: Springer International Publishing Press.

Berry RJ, Berry AC. 1967. Epigenetic variation in the human cranium. J. Anat 101:361-379.

Bethe AB, Bergmann GV, Embden G, Ellinger A. 1930. Handbunch der normalen und pathologischen physiologie. Berlin: Springer-Verlag.

Bookstein F. 1991. Morphometric tools for landmark data. New York: Cambridge University Press.

Bowden B, Bowden J. 2003. An illustrated atlas of the skeletal Muscles. Clin Anat 16: 185-186.

Bremer M, Doerge RW. 2010. Statistics at the bench. New York: Cold Spring Harbor Laboratory Press.

Bruzek J. 2002. A method for visual determination of sex, using the human hip bone. Am J Anthropol 117:157-168.

Buikstra JE, Ubelaker DH. 1994. Standards for data collection from human skeletal remains. Fayetteville: Arkansas Archaeological Survey.

Bussey MD. 2010. Does the demand for asymmetric functional lower body postures in lateral sports relate to structural asymmetry of the pelvis? J Sci Med Sport 13:360-364.

Byers SN. 2011. Introduction to forensic anthropology, 3rd ed. Upper Saddle River: Pearson.

Bytheway JA, Ross AH. 2010. A geometric morphometric approach to sex determination of the human adult os coxa. J Forensic Sci 55:859-864.

Capasso, L, Kennedy KAR, Wilczak CA. 1999. Atlas of occupational markers on human remains. Italy: Edigrafital S.p.A.

Christensen AM, Crowder CM. 2009. Evidentiary standards for forensic anthropology. J Forensic Sci 54:1211–1216.

Christensen AM, Passalacqua NV, Bartelink EJ. 2014. Forensic anthropology: Current methods and practice. Cambridge: Elsevier Academic Press.

Cohen J. 1968. Weighted kappa: Nominal scale agreement with provisions for scaled disagreement or partial credit. Phychol Bull 70:213-220.

Corruccini RS. 1974. An examination of the meaning of cranial discrete traits for human skeletal biological studies. Am J Phys Anthropol 40:425-446.

Daubert v. Merrell Dow Pharmaceuticals. 1993. US Supreme Court 509.U.S.579, 113S.Ct.2786, 125L. Ed.2d 469.

Dee PM. 1981. The preauricular sulcus. Radiology 140:354.

Delcarmen RV. 2010. Criminal procedure: Law and Practice. Bellmont: Wadsworth Publishing.

DeLeon VB. 2007. Fluctuating asymmetry and stress in a medival Nubian population. Am J Phys Anthropol 132:520-534.

Dirkmaat DC, Cabo LL, Ousley SD, Symes SA. 2008. New perspectives in forensic anthropology. Am J Phys Anthropol Suppl 47:33–52.

Drew R. 2013. A review of the ischium-pubis index: Accuracy, reliability, and common errors. Human Biol. 85:579-596.

Dwight T. 1905. The size of the articular surfaces of the long bones as characteristic of sex: An anthropological study. American Journal of Anatomy. 4:19-31.

Gomez-Valdes J., Menendez G, Garcia-Barzola L, Sanchez-Mejorada G, Karam C, Baraybar J, Klales A. (In review) Calibration of Klales' method of sexing the human innominate in contemporary Mexican population. Am J Phys Anthropol.

Gonzalez PN, Bernal V, Perez, SI. 2009. Geometric morphometric approach to sex estimation of human pelvis. Forensic Sci Int 189:68-74.

Grant JC, Boilean. 1965. Grant's method of anatomy, 7th ed. Maryland: Williams and Wilkins.

Gruenberg H. 1952. Genetic studies on the skeleton of mouse IV. Genetics 51:95-114.

Haeckel E. 1879. The evolution of man: A popular exposition of the principal points of human ontogeny and phylogeny, 3rd ed. New York: Appleton.

Hefner JT. 2009. Cranial nonmetric variation and estimating ancestry. J Forensic Sci 54:985-995.

Houghton P. 1974. The relationship of the pre-auricular groove of the ilium to pregnancy. Am J Phys Anthropol 41:381-390.

Hunt DR, Albanese J. 2005. History and demographic composition of the Robert J. Terry anatomical collection. Am J Phys Anthropol 127:406-417.

Jantz RL, Ousley SD. 2005. FORDISC 3 Computerised forensic discriminant functions, Version 3.0. The University of Tennessee, Knoxville.

Kelley MA. 1978. Phenice's visual technique for the os pubis: A critique. Am J Phys Anthropol 48:121-122.

Kenyhercz MK. 2012. Sex estimation using pubic bone morphologyin a modern South African sample: A test of the Klales et al. method. Am J Phys Anthropol S54: 179-180 (abstract).

Kirpichnikov VS. 1981. Genetic bases of fish selection. Berlin: Springer-Verlag.

Klales AR. Ousley SD. Vollner JM. 2012. A revised method of sexing the human innominate using Phenice's nonmetric traits and statistical methods. Am J Phys Anthropol 149:104-114.

Klales AR, Vollner JM, Ousley SD. 2009. A new metric procedure for the estimation of sex and ancestry from the human innominate. Proc Am Acad Forensic Sci 15:311 (abstract).

Krogman WM, 1939. A guide to the identification of human skeletal material. FBI Law Enforcement Bulletin 8(8): 1-29.

Lachenbruch PA, Mickey MR. 1968. Estimation of error rates in discriminant analysis. Technometrics 10:1-11.

Landers R. 2015. Computing intraclass correlations (ICC) as estimates of interrater reliability in SPSS. The Winnower 2:e143518.81744.

Landis JR, Koch GG. 1977. The measurement of observer agreement for categorical data. Biometrics 33:159-174.

Letterman GS. 1941. The greater sciatic notch in American Whites and Negroes. Am J Phys Anthropol 28:99-116.

Listi G, Bassett E. 2006. Test of an alternative method for determining sex from the os coxae: Applications for modern Americans. J Forensic Sci 51:248-252.

Lovell NC. 1989. Test of Phenice's technique for determining sex from the os pubis. Am J Phys Anthropol 28:105-109.

MacLaughlin SM, Bruce MF. 1990. The accuracy of sex identification in European skeletal remains using Phenice characters. J Forensic Sci 35:1384-1392.

McBride DG, Dietz MJ, Vennemeyer MT, Meadors SA, Benfer RA, Furbee NL. 2001. Bootstrap methods for sex determination from the os coxae using the ID3 algorithm. J Forensic Sci 46:427-431.

McKern TW, Stewart TD. 1957. Skeletal age changes in young American males. Natik, MA.: Quartermaster Research and Development Command [Technical Report] EP-45.

McManus IC, Moore J, Freegard M, Rawles R. 2010. Science in the making right hand, left hand III: Estimating historical rates of left-handedness. Laterally 15:186-208.

Melnick RL. 2005. A Daubert motion: a legal strategy to exclude essential scientific evidence in toxic tort litigation. Am J Pub Health 95: 30-35.

Moller PM, Swaddle JP. 1997. Asymmetry, developmental stability, and evolution. Oxford: Oxford University Press.

National Academy of Sciences. 2009. Strengthening forensic sciences in the United States: A path forward. Washington, DC: National Academic Press.

Norusis, MJ. 2012. IBM SPSS Statistics 19 Guide to Data Analysis. Upper Saddle River: Prentice Hall.

Novak L, Schultz JJ, McIntyre M. 2012. Determining sex of the posterior ilium from the Robert J. Terry and William Bass collection. J Forensic Sci 57:1155-1160.

Overbury RS, Cabo LL, Dirkmaat DC, Symes SA. 2009. Asymmetry of the os pubis: implications for the Suchey_Brooks method. Am J Phys Anthropol 138:261-268.

Peters M., Reimers S., Manning JT. 2006. Hand preference for the writing and associations with selected demographic and behavioral variables in 255,100 subjects: The BBC internet study. Brain Cogn 62:177-189.

Phenice TW. 1969. A newly developed visual method of sexing the os pubis. Am J Phys Anthropol 30:297–301.

Plochocki JH. 2002. Directional bilateral asymmetry in human sacral morphology. Int J Osteoarchaeol 2:349-355.

Press SJ, Wilson S. 1978. Choosing between logistic regression and discriminant analysis. J Am Stat Assoc 73:699-705.

Rogers T, Saunders S. 1994. Accuracy of sex determination using morphological traits of the human pelvis. J Forensic Sci 39:1047-1056.

Sachdeva K, Singla RK, Kalsey G. 2014. Role of ischio-pubic index in sex identification from innominate bones in North Indian population. Int J Anat Res 2:515-520.

Saunders SR. 1989. Nonmetric skeletal variation. In: Iscan MY, Kennedy KAR, editors. Reconstruction of life from the skeleton. New York: Wiley-liss, Inc. p 95-108.

Schmitt A, Cunha E, Pinheiro J. 2006. Forensic anthropology and medicine: Complementary sciences from recovery to cause of death. Totowa: Humana Press.

Schultz, AH. 1930. The skeleton of the trunk and limbs of higher primates. Human Biol 2:165-200.

Scientific Working Group for Forensic Anthropology: sex assessment. 2010. Available at: http://swaganth.startlogic.com/SexRev0.pdf.

Seidler, H. 1980. Sex-diagnosis of isolated os coxae by discriminant functions. J Hum Evol 9:597–600.

Sellke T, Bayarri MJ, Berger JO. 2001. Calibration of P values for testing precise null hypotheses. Am Stat 5: 165-200.

Singh S, Potturi BR. 1978. Greater sciatic notch in sex determination. J Anat 125:619-624.

Skinnes, H, Buras T., 1987. Developmental stability in wheat to differences in the temperature during seed set and seed development. Acta Agric Scand. 37:281-289.

SPSS for Windows, Rel. 14.0. 2005. Chicago: SPSS.

Steadman D, Adams BJ, Konigsberg LW. 2006. Statistical basis for positive identification in forensic anthropology. Am J Phys Anthropol 131:15-26.

Steele J, Mays SA. 1995. Handedness and directional asymmetry in the long bones of the human upper limb. Int J Osteoarchaeol 5:39-49.

Stewart, TD. 1954. Sex determination of the skeleton by guess and by measurement. Am J Phys Anthropol 12:385–392.

Steyn M, Pretorius E, Hutten L. 2004. Geometric morphometric analysis of the greater sciatic notch in South Africans. HOMO - J Comp Hum Biol 54:197-206.

Sutherland LD, Suchey JM. 1991. Use of the ventral arc in pubic sex determination. J Forensic Sci 36: 501-511

Todd TW. 1921. Age changes in the pubic bone: III. The pubis of the white female. Am J Phys Anthropol 4:26-40.

Trinkaus E. 1978. Bilateral asymmetry of human skeletal non-metric traits. Am J Phys Anthropol 49:315-318.

Trotter M, Gleser C. 1952. Estimation of stature from long bones of American whites and Negroes. Am J Phys Anthropol 10:463-514.

Ubelaker DH, Volk CG. 2002. A test of the Phenice method for the estimation of sex. J Forensic Sci 47:19-24.

Van Valen L. 1962. A study of fluctuating asymmetry. Evolution 16:125-142.

Viera AJ, Garrett JM. 2005. Understanding interobserver agreement: The kappa statistic. Fam Med 37:360-363.

Wadington, CH. 1940. Organisers and genes. Cambridge: Cambridge University Press.

Walker PL. 2005. Greater sciatic notch morphology: Sex, age, and population differences. Am J Phys Anthropol. 127: 385-391.

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

Washburn SL. 1948. Sex differences in the pubic bone. Am J Phys Anthropol 6:199-207.

Wolff J. 1986. The law of bone remodeling. Berlin: Springer-Verlag.

Yablokov AV. 1986. Population biology. Progress and problems of studies on natural populations. Moscow: Mir.

Zachos FE., Hartl GB., Suchentrunk F. 2007. Fluctuating asymmetry and genetic variability in roe deer (Capreolus capreolus): A test of the developmental stability hypothesis in mammals using natural molecular markers. Heredity 98:392-400.

Appendix

ID Numbers [a]	VA [b]	SPC [c]	MA [d]	PPF [e]	PPM [f]	Pos. Weight [g]	Age [h]	Race [i]	Est. Sex [j]	Act. Sex [k]
001L	5	3	3	0.015	0.984	4.179	23	BLACK	MALE	MALE
001R	5	3	3	0.015	0.984	4.179	23	BLACK	MALE	MALE
002L	2	1	1	0.999	0.000	-8.573	66	BLACK	FEMALE	FEMALE
002R	2	1	1	0.999	0.000	-8.573	66	BLACK	FEMALE	FEMALE
003L	1	1	3	0.999	0.000	-8.871	59	BLACK	FEMALE	FEMALE
003R	1	1	3	0.999	0.000	-8.871	59	BLACK	FEMALE	FEMALE
004L	2	1	2	0.999	0.000	-7.359	63	WHITE	FEMALE	FEMALE
004R	1	1	2	0.999	0.000	-10.085	63	WHITE	FEMALE	FEMALE
005L	1	1	2	0.999	0.000	-10.085	69	WHITE	FEMALE	FEMALE
005R	1	1	2	0.999	0.000	-10.085	69	WHITE	FEMALE	FEMALE
006L	2	1	3	0.997	0.002	-6.145	80	WHITE	FEMALE	FEMALE
006R	3	1	3	0.968	0.031	-3.419	80	WHITE	FEMALE	FEMALE
007L	1	1	2	0.999	0.000	-10.085	59	WHITE	FEMALE	FEMALE
007R	2	2	2	0.998	0.001	-6.286	59	WHITE	FEMALE	FEMALE
008L	2	2	3	0.993	0.006	-5.072	UNK	BLACK	FEMALE	FEMALE
008R	2	2	3	0.993	0.006	-5.072	UNK	BLACK	FEMALE	FEMALE
009L	1	1	1	0.999	0.000	-11.299	53	WHITE	FEMALE	FEMALE
009R	1	1	1	0.999	0.000	-11.299	53	WHITE	FEMALE	FEMALE
010L	3	1	2	0.990	0.009	-4.633	76	WHITE	FEMALE	FEMALE
010R	3	1	2	0.990	0.009	-4.633	76	WHITE	FEMALE	FEMALE
011L	5	4	3	0.005	0.994	5.252	57	WHITE	MALE	MALE
011R	4	4	3	0.074	0.925	2.526	57	WHITE	MALE	MALE
012L	1	2	2	0.999	0.000	-9.012	52	WHITE	FEMALE	FEMALE
012R	1	2	2	0.999	0.000	-9.012	52	WHITE	FEMALE	FEMALE
013L	4	3	4	0.064	0.935	2.667	43	BLACK	MALE	MALE
013R	4	3	4	0.064	0.935	2.667	43	BLACK	MALE	MALE
014L	4	3	4	0.064	0.935	2.667	60	BLACK	MALE	MALE
014R	4	3	4	0.064	0.935	2.667	60	BLACK	MALE	MALE
015L	5	3	4	0.004	0.995	5.393	UNK	BLACK	MALE	MALE
015R	5	3	4	0.004	0.995	5.393	UKN	BLACK	MALE	MALE
016L	3	3	4	0.514	0.485	-0.059	58	WHITE	FEMALE	MALE
016R	3	3	4	0.514	0.485	-0.059	58	WHITE	FEMALE	MALE
017L	4	3	4	0.064	0.935	2.667	66	WHITE	MALE	MALE
017R	4	3	4	0.064	0.935	2.667	66	WHITE	MALE	MALE
018L	5	2	4	0.013	0.980	4.32	82	WHITE	MALE	MALE
018R	5	2	4	0.013	0.980	4.32	82	WHITE	MALE	MALE
019L	1	1	2	0.999	0.000	-10.085	53	WHITE	FEMALE	FEMALE
019R	1	1	3	0.999	0.000	-8.871	53	WHITE	FEMALE	FEMALE

[a]Identification number; [b]score of ventral arc; [c]score for sub pubic contour; [d]score for ischio pubic ramus; [e]probability of being male; [f]probability of being female; [g]positive weight; [h]age of individual at death; [i]race; [j]estimated sex based on probability; [k]actual documented sex

ID Numbers [a]	VA[b]	SPC[c]	MA[d]	PPF[e]	PPM[f]	Pos. Weight [g]	Age [h]	Race [i]	Est. Sex [j]	Act. Sex [k]
020L	1	1	2	0.999	0.000	-10.085	72	WHITE	FEMALE	FEMALE
020R	2	1	2	0.999	0.000	-7.359	72	WHITE	FEMALE	FEMALE
021L	1	1	2	0.999	0.000	-10.085	69	WHITE	FEMALE	FEMALE
021R	1	2	2	0.999	0.000	-10.085	69	WHITE	FEMALE	FEMALE
022L	2	1	2	0.999	0.000	-7.359	66	WHITE	FEMALE	FEMALE
022R	2	1	2	0.999	0.000	-7.359	66	WHITE	FEMALE	FEMALE
023L	5	3	4	0.004	0.995	5.393	50	WHITE	MALE	MALE
023R	5	3	4	0.004	0.995	5.393	50	WHITE	MALE	MALE
024L	1	1	2	0.999	0.000	10.085	58	WHITE	FEMALE	FEMALE
024R	1	1	1	0.999	0.000	-11.299	58	WHITE	FEMALE	FEMALE
025L	5	3	4	0.004	0.995	5.393	48	WHITE	MALE	MALE
025R	5	3	5	0.001	0.998	6.607	48	WHITE	MALE	MALE
026L	1	1	1	0.999	0.000	-11.299	52	WHITE	FEMALE	FEMALE
026R	2	1	2	0.999	0.000	-7.359	52	WHITE	FEMALE	FEMALE
027L	4	3	4	0.064	0.935	2.667	UNK	BLACK	MALE	MALE
027R	4	3	4	0.064	0.935	2.667	UNK	BLACK	MALE	MALE
028L	4	3	4	0.064	0.935	2.667	46	BLACK	MALE	MALE
028R	4	3	4	0.064	0.935	2.667	46	BLACK	MALE	MALE
029L	2	1	1	0.999	0.000	-8.573	55	WHITE	FEMALE	FEMALE
029R	2	1	2	0.000	0.000	-7.359	55	WHITE	FEMALE	FEMALE
030L	3	3	3	0.781	0.218	-1.273	62	WHITE	FEMALE	MALE
030R	4	3	3	0.064	0.935	2.667	62	WHITE	MALE	MALE
031L	2	1	1	0.999	0.000	-8.573	62	WHITE	FEMALE	FEMALE
031R	1	1	2	0.999	0.000	-10.085	62	WHITE	FEMALE	FEMALE
032L	2	1	2	0.999	0.000	-7.359	60	WHITE	FEMALE	FEMALE
032R	3	1	2	0.990	0.009	-4.633	60	WHITE	FEMALE	FEMALE
033L	3	3	3	0.781	0.218	-1.273	UNK	BLACK	FEMALE	MALE
033R	3	3	3	0.781	0.218	-1.273	UNK	BLACK	FEMALE	MALE
034L	4	2	5	0.056	0.943	2.808	49	BLACK	MALE	MALE
034R	5	2	5	0.003	0.996	5.534	49	BLACK	MALE	MALE
035L	4	4	4	0.023	0.976	3.74	55	BLACK	MALE	MALE
035R	5	4	5	0.000	0.999	7.68	55	BLACK	MALE	MALE
036L	4	3	4	0.064	0.935	2.667	27	BLACK	MALE	MALE
036R	4	3	4	0.064	0.935	2.667	27	BLACK	MALE	MALE
037L	4	3	3	0.189	0.810	1.453	54	BLACK	MALE	MALE
037R	4	3	4	0.064	0.935	2.667	54	BLACK	MALE	MALE
038L	4	3	4	0.064	0.935	2.667	47	BLACK	MALE	MALE
038R	5	3	4	0.000	0.995	5.393	47	BLACK	MALE	MALE
039L	1	1	1	0.999	0.000	-11.299	73	WHITE	FEMALE	FEMALE
039R	2	1	2	0.999	0.000	-7.359	73	WHITE	FEMALE	FEMALE

[a]Identification number; [b]score of ventral arc; [c]score for sub pubic contour; [d]score for ischio pubic ramus; [e]probability of being male; [f]probability of being female; [g]positive weight; [h]age of individual at death; [i]race; [j]estimated sex based on probability; [k]actual documented sex

ID Numbers [a]	VA [b]	SPC [c]	MA [d]	PPF [e]	PPM [f]	Pos. Weight [g]	Age [h]	Race [i]	Est. Sex [j]	Act. Sex [k]
040L	3	3	3	0.781	0.218	-1.273	UNK	BLACK	FEMALE	MALE
040R	4	3	3	0.189	0.810	1.453	UNK	BLACK	MALE	MALE
041L	2	1	3	0.997	0.002	-6.145	43	WHITE	FEMALE	FEMALE
041R	1	1	2	0.999	0.000	-10.085	43	WHITE	FEMALE	FEMALE
042L	3	2	2	0.972	0.027	-3.56	68	WHITE	FEMALE	FEMALE
042R	3	2	2	0.972	0.027	-3.56	68	WHITE	FEMALE	FEMALE
043L	4	2	4	0.168	0.831	1.594	62	WHITE	MALE	MALE
043R	4	2	4	0.168	0.831	1.594	62	WHITE	MALE	MALE
044L	5	4	4	0.000	0.999	7.539	64	WHITE	MALE	MALE
044R	5	4	4	0.000	0.999	7.539	64	WHITE	MALE	MALE
045L	1	1	2	0.999	0.000	-10.085	81	WHITE	FEMALE	FEMALE
045R	1	1	2	0.999	0.000	-10.085	81	WHITE	FEMALE	FEMALE
046L	2	1	2	0.999	0.000	-7.359	75	WHITE	FEMALE	FEMALE
046R	2	1	2	0.999	0.000	-7.359	75	WHITE	FEMALE	FEMALE
047L	2	1	2	0.999	0.000	-7.359	58	WHITE	FEMALE	FEMALE
047R	2	1	2	0.999	0.000	-7.358	58	WHITE	FEMALE	FEMALE
048L	3	1	3	0.969	0.031	-3.419	85	WHITE	FEMALE	FEMALE
048R	3	1	3	0.969	0.031	-3.419	85	WHITE	FEMALE	FEMALE
049L	3	2	3	0.912	0.406	-2.346	71	WHITE	FEMALE	FEMALE
049R	4	2	3	0.087	0.593	0.38	71	WHITE	FEMALE	FEMALE
050L	1	1	1	0.999	0.000	-11.299	60	AM IND	FEMALE	FEMALE
050R	1	1	1	0.999	0.000	-11.299	60	AM IND	FEMALE	FEMALE
051L	1	2	2	0.999	0.000	-9.012	64	WHITE	FEMALE	FEMALE
051R	1	2	2	0.999	0.000	-9.012	64	WHITE	FEMALE	FEMALE
052L	5	4	4	0.001	0.998	6.466	71	WHITE	MALE	MALE
052R	5	4	5	0.000	0.999	7.68	71	WHITE	MALE	MALE
053L	4	4	4	0.023	0.976	3.74	33	WHITE	MALE	MALE
053R	4	4	4	0.023	0.976	3.74	33	WHITE	MALE	MALE
054L	4	3	4	0.064	0.935	2.667	30	WHT/ASN	MALE	MALE
054R	4	3	4	0.064	0.935	2.667	30	WHT/ASN	MALE	MALE
055L	4	3	4	0.064	0.935	2.667	41	WHITE	MALE	MALE
055R	4	3	4	0.064	0.935	2.667	41	WHITE	MALE	MALE
056L	2	1	2	0.999	0.000	-7.359	73	BLACK	FEMALE	FEMALE
056R	2	2	2	0.998	0.001	-6.286	73	BLACK	FEMALE	FEMALE
057L	2	2	2	0.998	0.001	-6.286	71	BLACK	FEMALE	FEMALE
057R	2	2	2	0.998	0.001	-6.286	71	BLACK	FEMALE	FEMALE
058L	1	1	3	0.999	0.000	-8.871	81	WHITE	FEMALE	FEMALE
058R	2	1	3	0.997	0.002	-6.145	81	WHITE	FEMALE	FEMALE
059L	5	3	4	0.004	0.995	5.393	70	WHITE	MALE	MALE

[a]Identification number; [b]score of ventral arc; [c]score for sub pubic contour; [d]score for ischio pubic ramus; [e]probability of being male; [f]probability of being female; [g]positive weight; [h]age of individual at death; [i]race; [j]estimated sex based on probability; [k]actual documented sex

ID Numbers [a]	VA[b]	SPC[c]	MA[d]	PPF[e]	PPM[f]	Pos. Weight [g]	Age [h]	Race [i]	Est. Sex [j]	Act. Sex [k]
059R	5	3	4	0.004	0.995	5.393	70	WHITE	MALE	MALE
060L	4	3	4	0.064	0.935	2.667	45	WHITE	MALE	MALE
060R	4	3	4	0.064	0.935	2.667	45	WHITE	MALE	MALE
061L	1	2	2	0.999	0.000	-9.012	86	WHITE	FEMALE	FEMALE
061R	2	2	2	0.998	0.001	-6.286	86	WHITE	FEMALE	FEMALE
062L	1	1	1	0.999	0.000	-11.299	60	WHITE	FEMALE	FEMALE
062R	2	1	1	0.999	0.000	-8.573	60	WHITE	FEMALE	FEMALE
063L	2	2	2	0.998	0.000	-6.286	57	WHITE	FEMALE	FEMALE
063R	2	2	2	0.998	0.000	-6.286	57	WHITE	FEMALE	FEMALE
064L	1	1	2	0.999	0.000	-10.085	31	WHITE	FEMALE	FEMALE
064R	1	1	2	0.999	0.000	-10.085	31	WHITE	FEMALE	FEMALE
065L	3	3	4	0.514	0.485	-0.059	89	WHITE	FEMALE	MALE
065R	4	3	4	0.064	0.935	2.667	89	WHITE	MALE	MALE
066L	4	4	5	0.007	0.992	4.954	61	WHITE	MALE	MALE
066R	5	4	5	0.000	0.999	7.68	61	WHITE	MALE	MALE
067L	4	3	4	0.064	0.935	2.667	54	WHITE	MALE	MALE
067R	5	3	4	0.004	0.995	5.393	54	WHITE	MALE	MALE
068L	4	3	3	0.189	0.810	1.453	67	WHITE	MALE	MALE
068R	3	3	3	0.781	0.218	-1.273	67	WHITE	FEMALE	MALE
069L	4	3	4	0.023	0.976	3.74	58	WHT/ASN	MALE	MALE
069R	4	3	4	0.023	0.976	3.74	58	WHT/ASN	MALE	MALE
070L	5	3	4	0.004	0.995	5.393	88	WHITE	MALE	MALE
070R	5	4	4	0.001	0.998	6.466	88	WHITE	MALE	MALE
071L	4	3	5	0.020	0.979	3.881	83	WHITE	MALE	MALE
071R	5	3	5	0.001	0.998	6.607	83	WHITE	MALE	MALE
072L	4	3	4	0.064	0.935	2.667	63	WHITE	MALE	MALE
072R	4	3	4	0.064	0.935	2.667	63	WHITE	MALE	MALE
073L	4	3	3	0.180	0.810	1.453	83	WHITE	MALE	MALE
073R	4	3	3	0.180	0.810	1.453	83	WHITE	MALE	MALE
074L	5	3	4	0.004	0.995	5.393	79	WHITE	MALE	MALE
074R	5	3	4	0.004	0.995	5.393	79	WHITE	MALE	MALE
075L	2	1	3	0.997	0.002	-6.145	99	BLACK	FEMALE	FEMALE
075R	3	1	3	0.968	0.031	-3.419	99	BLACK	FEMALE	FEMALE
076L	2	2	2	0.998	0.001	-6.286	67	WHITE	FEMALE	FEMALE
076R	3	2	2	0.972	0.027	-3.56	67	WHITE	FEMALE	FEMALE
077L	1	1	2	0.999	0.000	-10.085	79	AM IND	FEMALE	FEMALE
077R	1	1	2	0.999	0.000	-10.085	79	AM IND	FEMALE	FEMALE
078L	1	1	3	0.999	0.000	-8.871	78	WHITE	FEMALE	FEMALE
078R	2	1	3	0.997	0.002	-6.145	78	WHITE	FEMALE	FEMALE

[a]Identification number; [b]score of ventral arc; [c]score for sub pubic contour; [d]score for ischio pubic ramus; [e]probability of being male; [f]probability of being female; [g]positive weight; [h]age of individual at death; [i]race; [j]estimated sex based on probability; [k]actual documented sex

ID Numbers [a]	VA [b]	SPC [c]	MA [d]	PPF [e]	PPM [f]	Pos. Weight [g]	Age [h]	Race [i]	Est. Sex [j]	Act. Sex [k]
079L	5	3	4	0.004	0.995	5.393	77	WHITE	MALE	MALE
079R	5	3	5	0.001	0.998	6.607	77	WHITE	MALE	MALE
080L	1	1	3	0.999	0.000	-8.871	UNK	BLACK	FEMALE	FEMALE
080R	1	1	3	0.999	0.000	-8.871	UNK	BLACK	FEMALE	FEMALE
081L	4	4	4	0.189	0.810	1.453	55	WHITE	MALE	MALE
081R	3	3	3	0.781	0.218	-1.273	55	WHITE	FEMALE	MALE
082L	4	2	4	0.168	0.831	1.594	71	WHITE	MALE	MALE
082R	4	2	4	0.168	0.831	1.594	71	WHITE	MALE	MALE
083L	4	3	3	0.064	0.935	2.667	81	WHITE	MALE	MALE
083R	4	3	5	0.020	0.979	3.881	81	WHITE	MALE	MALE
084L	1	1	2	0.999	0.000	-10.085	39	WHITE	FEMALE	FEMALE
084R	1	1	2	0.999	0.000	-10.085	39	WHITE	FEMALE	FEMALE
085L	5	3	5	0.001	0.998	6.607	50	WHITE	MALE	MALE
085R	5	3	5	0.001	0.998	6.607	50	WHITE	MALE	MALE
086L	5	3	4	0.004	0.995	5.393	44	WHITE	MALE	MALE
086R	5	3	4	0.004	0.995	5.393	44	WHITE	MALE	MALE
087L	2	1	2	0.999	0.000	-7.359	64	WHITE	FEMALE	FEMALE
087R	2	1	2	0.999	0.000	-7.359	64	WHITE	FEMALE	FEMALE
088L	1	1	1	0.999	0.000	-11.299	73	WHITE	FEMALE	FEMALE
088R	1	1	1	0.999	0.000	-11.299	73	WHITE	FEMALE	FEMALE
089L	1	1	2	0.999	0.000	-10.085	50	WHITE	FEMALE	FEMALE
089R	1	1	2	0.999	0.000	-10.085	50	WHITE	FEMALE	FEMALE
090L	2	2	4	0.979	0.020	-3.858	82	WHITE	FEMALE	FEMALE
090R	2	2	4	0.979	0.020	-3.858	82	WHITE	FEMALE	FEMALE
091L	1	1	3	0.999	0.000	-8.871	60	WHITE	FEMALE	FEMALE
091R	1	1	3	0.999	0.000	-8.871	60	WHITE	FEMALE	FEMALE
092L	2	2	2	0.998	0.001	-6.286	49	WHITE	FEMALE	FEMALE
092R	2	2	2	0.998	0.001	-6.286	49	WHITE	FEMALE	FEMALE
093L	1	1	1	0.999	0.000	-11.299	60	WHITE	FEMALE	FEMALE
093R	1	1	1	0.999	0.000	-11.299	60	WHITE	FEMALE	FEMALE
094L	2	1	2	0.999	0.000	-7.359	64	WHITE	FEMALE	FEMALE
094R	2	1	2	0.999	0.000	-7.359	64	WHITE	FEMALE	FEMALE
095L	2	1	2	0.999	0.000	-7.359	75	WHITE	FEMALE	FEMALE
095R	2	1	2	0.999	0.000	-7.359	75	WHITE	FEMALE	FEMALE
096L	5	3	4	0.004	0.995	5.393	42	WHITE	MALE	MALE
096R	5	3	4	0.004	0.995	5.393	42	WHITE	MALE	MALE
097L	1	1	1	0.999	0.000	-11.299	74	WHITE	FEMALE	FEMALE
097R	1	1	1	0.999	0.000	-11.299	74	WHITE	FEMALE	FEMALE
098L	5	3	4	0.004	0.995	5.393	54	WHITE	MALE	MALE

[a]Identification number; [b]score of ventral arc; [c]score for sub pubic contour; [d]score for ischio pubic ramus; [e]probability of being male; [f]probability of being female; [g]positive weight; [h]age of individual at death; [i]race; [j]estimated sex based on probability; [k]actual documented sex

ID Numbers [a]	VA[b]	SPC[c]	MA[d]	PPF[e]	PPM[f]	Pos. Weight[g]	Age[h]	Race[i]	Est. Sex[j]	Act. Sex[k]
098R	5	3	4	0.004	0.995	5.393	54	WHITE	MALE	MALE
099L	5	4	5	0.000	0.999	7.68	61	WHITE	MALE	MALE
099R	5	4	5	0.000	0.999	7.68	61	WHITE	MALE	MALE
100L	2	1	3	0.981	0.018	-3.999	97	WHITE	FEMALE	FEMALE
100R	3	1	3	0.968	0.317	-3.419	97	WHITE	FEMALE	FEMALE
101L	5	3	4	0.004	0.995	5.393	55	WHITE	MALE	MALE
101R	5	3	3	0.015	0.984	4.179	55	WHITE	MALE	MALE
102L	5	3	4	0.005	0.994	5.252	70	WHITE	MALE	MALE
102R	5	4	4	0.001	0.998	6.466	70	WHITE	MALE	MALE
103L	5	3	4	0.004	0.995	5.393	81	WHITE	MALE	MALE
103R	5	3	4	0.004	0.995	5.393	81	WHITE	MALE	MALE
104L	4	2	5	0.056	0.943	2.808	54	WHITE	MALE	MALE
104R	4	2	5	0.056	0.943	2.808	54	WHITE	MALE	MALE
105L	5	3	4	0.004	0.995	5.393	55	WHITE	MALE	MALE
105R	5	3	4	0.004	0.995	5.393	55	WHITE	MALE	MALE
106L	3	3	4	0.514	0.485	-0.059	60	WHITE	MALE	MALE
106R	4	3	4	0.064	0.935	2.667	60	WHITE	MALE	MALE
107L	5	2	3	0.042	0.957	3.106	53	WHITE	MALE	MALE
107R	5	2	3	0.042	0.957	3.106	53	WHITE	MALE	MALE
108L	5	3	4	0.004	0.995	5.393	58	WHITE	MALE	MALE
108R	5	3	4	0.004	0.995	5.393	58	WHITE	MALE	MALE
109L	5	3	4	0.004	0.995	5.393	69	WHITE	MALE	MALE
109R	5	3	4	0.004	0.995	5.393	69	WHITE	MALE	MALE
110L	2	1	1	0.999	0.000	-8.573	66	WHITE	FEMALE	FEMALE
110R	2	1	1	0.999	0.000	-8.573	66	WHITE	FEMALE	FEMALE
111L	1	1	2	0.999	0.000	-10.085	66	WHITE	FEMALE	FEMALE
111R	2	1	2	0.999	0.000	-7.359	66	WHITE	FEMALE	FEMALE
112L	1	1	2	0.999	0.000	-10.085	60	WHITE	FEMALE	FEMALE
112R	1	1	2	0.999	0.000	-10.085	60	WHITE	FEMALE	FEMALE
113L	2	1	2	0.999	0.000	-7.359	61	WHITE	FEMALE	FEMALE
113R	2	1	2	0.999	0.000	-7.359	61	WHITE	FEMALE	FEMALE
114L	2	1	2	0.999	0.000	-7.395	57	WHITE	FEMALE	FEMALE
114R	2	1	2	0.999	0.000	-7.395	57	WHITE	FEMALE	FEMALE
115L	1	2	3	0.999	0.000	-7.798	57	WHITE	FEMALE	FEMALE
115R	2	2	3	0.993	0.006	-5.072	57	WHITE	FEMALE	FEMALE
116L	1	1	2	0.999	0.000	-10.085	36	WHITE	FEMALE	FEMALE
116R	1	1	2	0.999	0.000	-10.085	36	WHITE	FEMALE	FEMALE
117L	1	2	1	0.999	0.000	-10.226	83	WHITE	FEMALE	FEMALE
117R	2	2	1	0.999	0.000	-7.5	83	WHITE	FEMALE	FEMALE

[a]Identification number; [b]score of ventral arc; [c]score for sub pubic contour; [d]score for ischio pubic ramus; [e]probability of being male; [f]probability of being female; [g]positive weight; [h]age of individual at death; [i]race; [j]estimated sex based on probability; [k]actual documented sex

ID Numbers [a]	VA [b]	SPC [c]	MA [d]	PPF [e]	PPM [f]	Pos. Weight [g]	Age [h]	Race [i]	Est. Sex [j]	Act. Sex [k]
118L	5	3	4	0.004	0.995	5.393	89	WHITE	MALE	MALE
118R	5	3	4	0.004	0.995	5.393	89	WHITE	MALE	MALE
119L	5	3	4	0.004	0.995	5.393	88	WHITE	MALE	MALE
119R	5	3	5	0.001	0.998	6.607	88	WHITE	MALE	MALE
120L	5	4	4	0.001	0.998	6.466	58	WHITE	MALE	MALE
120R	5	4	4	0.001	0.998	6.466	58	WHITE	MALE	MALE
121L	4	3	3	0.064	0.935	2.667	46	WHITE	MALE	MALE
121R	3	3	3	0.781	0.218	-1.273	46	WHITE	FEMALE	MALE
122L	3	3	3	0.781	0.218	-1.273	46	WHITE	FEMALE	MALE
122R	3	3	3	0.781	0.218	-1.273	46	WHITE	FEMALE	MALE
123L	5	3	4	0.004	0.995	5.393	70	WHITE	MALE	MALE
123R	5	3	4	0.004	0.995	5.393	70	WHITE	MALE	MALE
124L	5	2	5	0.042	0.957	3.106	65	WHITE	MALE	MALE
124R	5	2	5	0.042	0.957	3.106	65	WHITE	MALE	MALE
125L	5	3	4	0.004	0.995	5.393	82	WHITE	MALE	MALE
125R	5	3	4	0.004	0.995	5.393	82	WHITE	MALE	MALE
126L	5	3	4	0.004	0.995	5.393	46	WHITE	MALE	MALE
126R	5	3	4	0.004	0.995	5.393	46	WHITE	MALE	MALE
127L	5	3	4	0.004	0.995	5.393	49	WHITE	MALE	MALE
127R	5	3	4	0.004	0.995	5.393	49	WHITE	MALE	MALE
128L	5	3	4	0.004	0.995	5.393	67	WHITE	MALE	MALE
128R	5	3	4	0.004	0.995	5.393	67	WHITE	MALE	MALE
129L	5	3	5	0.001	0.998	6.607	82	WHITE	MALE	MALE
129R	5	3	5	0.001	0.998	6.607	82	WHITE	MALE	MALE
130L	5	3	4	0.004	0.995	5.393	59	WHITE	MALE	MALE
130R	5	3	4	0.004	0.995	5.393	59	WHITE	MALE	MALE
131L	4	4	5	0.007	0.992	4.954	60	WHITE	MALE	MALE
131R	5	4	5	0.000	0.999	7.68	60	WHITE	MALE	MALE
132L	3	3	4	0.514	0.485	-0.059	67	WHITE	FEMALE	MALE
132R	3	4	3	0.064	0.935	2.667	67	WHITE	MALE	MALE
133L	5	3	3	0.015	0.098	4.179	56	WHITE	MALE	MALE
133R	5	3	3	0.015	0.984	4.179	56	WHITE	MALE	MALE
134L	5	2	3	0.042	0.957	3.106	70	WHITE	MALE	MALE
134R	5	2	3	0.042	0.957	3.106	70	WHITE	MALE	MALE
135L	5	3	4	0.004	0.995	5.393	65	WHITE	MALE	MALE
135R	5	4	4	0.001	0.998	6.466	65	WHITE	MALE	MALE
136L	5	4	5	0.000	0.999	7.68	59	WHITE	MALE	MALE
136R	5	4	4	0.001	0.998	6.466	59	WHITE	MALE	MALE
137L	5	2	4	0.013	0.986	4.32	32	WHITE	MALE	MALE

[a]Identification number; [b]score of ventral arc; [c]score for sub pubic contour; [d]score for ischio pubic ramus; [e]probability of being male; [f]probability of being female; [g]positive weight; [h]age of individual at death; [i]race; [j]estimated sex based on probability; [k]actual documented sex

ID Numbers [a]	VA [b]	SPC [c]	MA [d]	PPF [e]	PPM [f]	Pos. Weight [g]	Age [h]	Race [i]	Est. Sex [j]	Act. Sex [k]
137R	5	2	4	0.013	0.986	4.32	32	WHITE	MALE	MALE
138L	4	3	5	0.020	0.979	3.881	66	WHITE	MALE	MALE
138R	4	3	5	0.020	0.979	3.881	66	WHITE	MALE	MALE
139L	5	3	4	0.004	0.995	5.393	55	WHITE	MALE	MALE
139R	4	3	4	0.064	0.935	2.667	55	WHITE	MALE	MALE
140L	5	3	4	0.004	0.995	5.393	58	WHITE	MALE	MALE
140R	5	3	4	0.004	0.995	5.393	58	WHITE	MALE	MALE
141L	5	3	3	0.015	0.984	4.179	42	WHITE	MALE	MALE
141R	5	3	3	0.015	0.984	4.179	42	WHITE	MALE	MALE
142L	5	2	4	0.013	0.986	4.32	43	WHITE	MALE	MALE
142R	5	2	4	0.013	0.986	4.32	43	WHITE	MALE	MALE
143L	2	1	2	0.999	0.000	-7.359	56	WHITE	FEMALE	FEMALE
143R	2	1	2	0.999	0.000	-7.359	56	WHITE	FEMALE	FEMALE
144L	5	4	3	0.005	0.994	5.252	61	WHITE	MALE	FEMALE
144R	5	4	3	0.005	0.994	5.252	61	WHITE	MALE	FEMALE
145L	2	1	2	0.999	0.000	-7.359	49	WHITE	FEMALE	FEMALE
145R	2	1	2	0.999	0.000	-7.359	49	WHITE	FEMALE	FEMALE
146L	2	2	3	0.993	0.006	-5.072	60	WHITE	FEMALE	FEMALE
146R	2	2	3	0.993	0.006	-5.072	60	WHITE	FEMALE	FEMALE
147L	1	2	1	0.999	0.000	-10.226	62	WHITE	FEMALE	FEMALE
147R	1	1	1	0.999	0.000	-11.299	62	WHITE	FEMALE	FEMALE
148L	1	1	2	0.999	0.000	-10.085	51	WHITE	FEMALE	FEMALE
148R	1	1	2	0.999	0.000	-10.085	51	WHITE	FEMALE	FEMALE
149L	2	1	1	0.999	0.000	-7.5	72	WHITE	FEMALE	FEMALE
149R	2	1	1	0.999	0.000	-7.5	72	WHITE	FEMALE	FEMALE
150L	1	1	1	0.999	0.000	-11.299	41	WHITE	FEMALE	FEMALE
150R	1	1	1	0.999	0.000	-11.299	41	WHITE	FEMALE	FEMALE
151L	2	1	1	0.999	0.000	-8.573	65	WHITE	FEMALE	FEMALE
151R	2	1	1	0.999	0.000	-8.573	65	WHITE	FEMALE	FEMALE
152L	1	2	2	0.999	0.000	-9.012	73	WHITE	FEMALE	FEMALE
152R	1	2	2	0.999	0.000	-9.012	73	WHITE	FEMALE	FEMALE
153L	2	2	1	0.999	0.000	-7.5	85	WHITE	FEMALE	FEMALE
153R	2	2	1	0.999	0.000	-7.5	85	WHITE	FEMALE	FEMALE
154L	1	1	2	0.999	0.000	-10.085	73	WHITE	FEMALE	FEMALE
154R	1	1	2	0.999	0.000	-10.085	73	WHITE	FEMALE	FEMALE
155L	2	3	1	0.998	0.001	-6.427	77	WHITE	FEMALE	FEMALE
155R	1	3	1	0.999	0.000	-9.153	77	WHITE	FEMALE	FEMALE
156L	2	2	2	0.998	0.001	-6.286	62	WHITE	FEMALE	FEMALE
156R	3	2	2	0.972	0.027	-3.56	62	WHITE	FEMALE	FEMALE

[a]Identification number; [b]score of ventral arc; [c]score for sub pubic contour; [d]score for ischio pubic ramus; [e]probability of being male; [f]probability of being female; [g]positive weight; [h]age of individual at death; [i]race; [j]estimated sex based on probability; [k]actual documented sex

ID Numbers [a]	VA [b]	SPC [c]	MA [d]	PPF [e]	PPM [f]	Pos. Weight [g]	Age [h]	Race [i]	Est. Sex [j]	Act. Sex [k]
157L	1	1	1	0.999	0.000	-11.299	78	WHITE	FEMALE	FEMALE
157R	1	1	1	0.999	0.000	-11.299	78	WHITE	FEMALE	FEMALE
158L	3	2	1	0.991	0.008	-4.774	50	WHITE	FEMALE	FEMALE
158R	3	2	1	0.991	0.008	-4.774	50	WHITE	FEMALE	FEMALE
159L	1	1	2	0.999	0.000	-10.085	62	WHITE	FEMALE	FEMALE
159R	1	1	2	0.999	0.000	-10.085	62	WHITE	FEMALE	FEMALE
160L	1	1	1	0.999	0.000	-11.299	50	WHITE	FEMALE	FEMALE
160R	1	1	1	0.999	0.000	-11.299	50	WHITE	FEMALE	FEMALE
161L	3	4	4	0.266	0.733	1.014	70	WHITE	MALE	MALE
161R	3	4	3	0.549	0.450	-0.2	70	WHITE	FEMALE	MALE
162L	3	3	4	0.485	0.514	-0.059	49	WHITE	MALE	MALE
162R	3	3	4	0.485	0.514	-0.059	49	WHITE	MALE	MALE
163L	3	4	1	0.932	0.067	-2.628	68	WHITE	MALE	MALE
163R	4	4	1	0.475	0.524	0.098	68	WHITE	MALE	MALE
164L	5	3	4	0.004	0.995	5.393	69	WHITE	MALE	MALE
164R	5	3	4	0.004	0.995	5.393	69	WHITE	MALE	MALE
165L	4	4	3	0.074	0.925	2.526	43	WHITE	MALE	MALE
165R	4	4	3	0.074	0.925	2.526	43	WHITE	MALE	MALE
166L	3	3	4	0.524	0.485	-0.059	65	WHITE	FEMALE	MALE
166R	4	3	4	0.064	0.935	2.667	65	WHITE	MALE	MALE
167L	5	3	4	0.004	0.995	5.393	UNK	BLACK	MALE	MALE
167R	5	3	4	0.004	0.995	5.393	UNK	BLACK	MALE	MALE
168L	5	3	3	0.015	0.984	4.179	UNK	BLACK	MALE	MALE
168R	5	3	3	0.015	0.984	4.179	UNK	BLACK	MALE	MALE
169L	2	1	2	0.999	0.000	-7.359	71	WHITE	FEMALE	FEMALE
169R	1	1	2	0.999	4.169	-10.085	71	WHITE	FEMALE	FEMALE
170L	5	3	4	0.001	0.998	6.607	50	WHITE	MALE	MALE
170R	5	3	4	0.001	0.998	6.607	50	WHITE	MALE	MALE
171L	2	1	2	0.999	0.000	-7.359	38	WHITE	FEMALE	FEMALE
171R	2	1	2	0.999	0.000	-7.359	38	WHITE	FEMALE	FEMALE
172L	3	1	3	0.968	0.031	-3.419	62	WHITE	FEMALE	FEMALE
172R	3	1	3	0.968	0.031	-3.419	62	WHITE	FEMALE	FEMALE
173L	1	1	1	0.999	0.000	-11.299	62	WHITE	FEMALE	FEMALE
173R	1	1	1	0.999	0.000	-11.299	62	WHITE	FEMALE	FEMALE
174L	1	2	2	0.999	0.000	-9.012	76	WHITE	FEMALE	FEMALE
174R	2	1	2	0.999	0.000	-7.359	76	WHITE	FEMALE	FEMALE
175L	2	2	2	0.998	0.001	-6.286	39	WHITE	FEMALE	FEMALE
175R	2	2	2	0.998	0.001	-6.286	39	WHITE	FEMALE	FEMALE
176L	1	1	2	0.999	0.000	-10.085	67	WHITE	FEMALE	FEMALE

[a]Identification number; [b]score of ventral arc; [c]score for sub pubic contour; [d]score for ischio pubic ramus; [e]probability of being male; [f]probability of being female; [g]positive weight; [h]age of individual at death; [i]race; [j]estimated sex based on probability; [k]actual documented sex

ID Numbers [a]	VA [b]	SPC [c]	MA [d]	PPF [e]	PPM [f]	Pos. Weight [g]	Age [h]	Race [i]	Est. Sex [j]	Act. Sex [k]
176R	1	1	2	0.999	0.000	-10.085	67	WHITE	FEMALE	FEMALE
177L	1	1	2	0.999	0.000	-9.012	82	WHITE	FEMALE	FEMALE
177R	2	2	2	0.998	0.001	-6.286	82	WHITE	FEMALE	FEMALE
178L	5	3	3	0.015	0.984	4.179	73	WHITE	MALE	MALE
178R	5	3	3	0.015	0.984	4.179	73	WHITE	MALE	MALE
179L	4	2	1	0.085	0.114	-2.048	73	WHITE	FEMALE	MALE
179R	5	2	1	0.336	0.663	0.678	73	WHITE	MALE	MALE
180L	3	4	4	0.266	0.733	1.014	54	WHITE	MALE	MALE
180R	2	3	4	0.941	0.058	-2.785	54	WHITE	FEMALE	MALE
181L	3	4	4	0.266	0.733	1.014	UNK	HISP	MALE	MALE
181R	3	4	4	0.266	0.733	1.014	UNK	HISP	MALE	MALE
182L	1	1	1	0.999	0.000	-11.299	71	WHITE	FEMALE	FEMALE
182R	1	1	1	0.999	0.000	-11.299	71	WHITE	FEMALE	FEMALE
183L	5	2	3	0.042	0.570	3.106	29	WHITE	MALE	MALE
183R	5	2	3	0.042	0.570	3.106	29	WHITE	MALE	MALE
184L	4	3	3	0.189	0.810	1.453	33	WHITE	MALE	MALE
184R	4	3	3	0.189	0.810	1.453	33	WHITE	MALE	MALE
185L	1	2	2	0.999	0.000	-9.012	67	WHITE	FEMALE	FEMALE
185R	2	2	2	0.998	0.000	-6.286	67	WHITE	FEMALE	FEMALE
186L	2	1	1	0.999	0.000	-8.573	82	WHITE	FEMALE	FEMALE
186R	1	1	1	0.999	0.000	-11.299	82	WHITE	FEMALE	FEMALE
187L	3	1	2	0.997	0.002	-5.847	80	WHITE	FEMALE	FEMALE
187R	3	1	2	0.997	0.002	-5.847	80	WHITE	FEMALE	FEMALE
188L	4	2	4	0.168	0.831	1.594	63	WHITE	MALE	MALE
188R	3	2	4	0.756	0.243	-1.132	63	WHITE	FEMALE	MALE
189L	5	3	3	0.015	0.984	4.179	55	WHITE	MALE	MALE
189R	5	3	3	0.015	0.984	4.179	55	WHITE	MALE	MALE
190L	1	2	1	0.999	0.000	-10.226	73	WHITE	FEMALE	FEMALE
190R	1	2	1	0.999	0.000	-10.226	73	WHITE	FEMALE	FEMALE
191L	2	3	2	0.994	0.000	-5.213	67	WHITE	FEMALE	FEMALE
191R	2	3	3	0.981	0.018	-3.999	67	WHITE	FEMALE	FEMALE
192L	4	4	3	0.074	0.924	2.526	24	HISP	MALE	MALE
192R	4	4	3	0.074	0.924	2.526	24	HISP	MALE	MALE
193L	2	1	1	0.999	0.000	-8.573	58	WHITE	FEMALE	FEMALE
193R	2	1	1	0.999	0.000	-8.573	58	WHITE	FEMALE	FEMALE
194L	1	3	3	0.998	0.001	-6.725	51	WHITE	FEMALE	FEMALE
194R	1	3	3	0.998	0.001	-6.725	51	WHITE	FEMALE	FEMALE
195L	5	3	3	0.015	0.984	4.179	62	WHITE	MALE	MALE
195R	5	3	4	0.004	0.995	5.393	62	WHITE	MALE	MALE

[a]Identification number; [b]score of ventral arc; [c]score for sub pubic contour; [d]score for ischio pubic ramus; [e]probability of being male; [f]probability of being female; [g]positive weight; [h]age of individual at death; [i]race; [j]estimated sex based on probability; [k]actual documented sex

ID Numbers [a]	VA [b]	SPC [c]	MA [d]	PPF [e]	PPM [f]	Pos. Weight [g]	Age [h]	Race [i]	Est. Sex [j]	Act. Sex [k]
196L	3	4	4	0.266	0.730	1.014	47	WHITE	MALE	MALE
196R	2	4	4	0.847	0.152	-1.712	47	WHITE	FEMALE	MALE
197L	4	3	3	0.020	0.979	3.881	66	WHITE	MALE	MALE
197R	4	3	3	0.020	0.979	3.881	66	WHITE	MALE	MALE
198L	1	1	1	0.999	0.000	-11.299	59	WHITE	FEMALE	FEMALE
198R	1	1	1	0.999	0.000	-11.299	59	WHITE	FEMALE	FEMALE
199L	1	1	1	0.999	0.000	-11.299	52	WHITE	FEMALE	FEMALE
199R	1	1	1	0.999	0.000	-11.299	52	WHITE	FEMALE	FEMALE
200L	1	2	2	0.999	0.000	-9.012	50	WHITE	FEMALE	FEMALE
200R	1	2	2	0.999	0.000	-9.012	50	WHITE	FEMALE	FEMALE
201L	5	3	3	0.015	0.984	4.179	45	WHITE	MALE	MALE
201R	5	3	3	0.015	0.984	4.179	45	WHITE	MALE	MALE
202L	3	1	2	0.990	0.000	-4.633	42	WHITE	FEMALE	FEMALE
202R	3	1	2	0.990	0.000	-4.633	42	WHITE	FEMALE	FEMALE
203L	1	1	3	0.999	0.000	-8.871	73	WHITE	FEMALE	FEMALE
203R	2	1	3	0.997	0.002	-6.145	73	WHITE	FEMALE	FEMALE
204L	5	3	2	0.049	0.950	2.965	46	WHITE	MALE	MALE
204R	5	3	2	0.049	0.950	2.965	46	WHITE	MALE	MALE

[a] Identification number; [b] score of ventral arc; [c] score for sub pubic contour; [d] score for ischio pubic ramus; [e] probability of being male; [f] probability of being female; [g] positive weight; [h] age of individual at death; [i] race; [j] estimated sex based on probability; [k] actual documented sex

Copyright Permissions

Licensee: Sandra J Call

License Date: Jan 21, 2016

License Number: 3793631444660

Publication: American Journal of Physical Anthropology

Title: A revised method of sexing the human innominate using Phenice\'s nonmetric traits and statistical methods

Type of Use: Dissertation/Thesis

Licensee: Sandra J Call

License Date: Jan 21, 2016

License Number: 3793640234122

Publication: American Journal of Physical Anthropology

Title: A newly developed visual method of sexing the os pubis

Type of Use: Dissertation/Thesis