USING SHAPE ANALYSIS AND HUMAN VARIATION TO BETTER PREDICT SEX IN THE HUMAN COXAL BONE

by

Heather Robertson

B.A., Simon Fraser University, 2004

M.A., University of Victoria, 2013

A DISERTATION SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Anthropology)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

April 2020

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the dissertation entitled:

Using shape analysis and population inclusivity to better predict sex in the human coxal bone

submitted by Heather Robertson in partial fulfillment of the requirements for

the degree of Doctor of Philosophy

in Anthropology

Examining Committee:

Dr. Darlene Weston, Associate Professor, Anthropology, UBC
Supervisor

Dr. David Pokotylo, Professor, Anthropology, UBC
Supervisory Committee Member

Dr. Hugo Cardoso, Associate Professor, Archaeology, SFU
Supervisory Committee Member

Dr. Camilla Speller, Assistant Professor, Anthropology, UBC
University Examiner

Dr. Siddharth Vora, Assistant Professor, Dentistry, UBC
University Examiner

Abstract

Metric methods of sex estimation are often less powerful than visual methods because linear measurements represent too may isometric measures of body size and lack sufficient allometric measures of body form (size and shape). This study uses geometric morphometrics to identify 17 landmarks that most effectively represent sex-based shape in right and left coxal bones (n = 394, f = 191, m = 203), these are: the anterior superior iliac spine; posterior superior iliac spine; posterior inferior iliac spine; iliac crest; apex of the auricular surface; greater sciatic notch; ischial spine; superior, inferior and distal points on ischial tuberosity; superior, inferior and midpoint on the symphyseal face; arcuate eminence; ischiopubic ramus; and posterosuperior and anterosuperior points on the acetabular rim. The first and second principal components (PCs) correctly predicted sex in 98.5% of cases; better than previous studies on whole coxal bone sex-based shape.

Linear measurements from Langley et al. (2016) that correspond with the 17-landmarks were used to generate a reliable discriminant function (DF) equation and logistic regression model (LRM) for sex estimation. The DF equation correctly predicted sex 99.7% of the time in cross-validation, the LRM correctly predicted sex in all individuals. Both equations accounted for allometric size, isometric size, and fluctuating asymmetry to help discern sex from other variants of shape. When tested on an independent population (n = 120; f = 60/60, m = 60/60), the DF equation correctly predicted sex with 99.2% accuracy (f = 191/191, 100%, m = 202/203, 99.7%), and the LRM correctly predicted sex in all test specimens.

Measurements and landmarks were further tested for use in fragmented coxal bones. The most successful DFs and LRMs accurately predicted sex between 98.7 – 99.2% for

measurements representing coxal bones completeness between 50-25%. DF and LRM equations representing coxal bones no less than 25% complete predicted sex with similar accuracies (DF = 99.0%; LRM = 99.2%) and correctly assigned 100% of the test population. These equations excelled at sex estimation because the measurements account for variations in sex, size (allometry and isometry) and fluctuating asymmetry. These DF and LRM equations are recommended for forensic applications.

Lay Summary

Interpreting sex-based shape in the coxal bone (hip bone) not only involves uncovering measurements that best lead to correct estimates of sex in the human skeleton, but also involves including other sources of bone shape variation like differences in body size and asymmetry. True sex differences can be exposed when these variables of variation are accounted for in the method. Estimating sex with population inclusivity in mind is important in forensic and biological anthropological contexts where population affinity is unknown and population-specific sex estimation methods are consequently rendered moot. This study also discovers that landmarks and measurements representing true sex differences are also maintained in equations designed for fragmented coxal bones. These methods are reliable 98.7– 100% of the time and repeatable 98.3 – 100% of the time.

Preface

This dissertation reflects original thought, research, and work by the author, H, I. Robertson.

A version of Chapter 2 has been published [Robertson, H. I., Pokotylo, D. L., & Weston, D. A. (2019). Testing landmark redundancy for sex-based shape analysis of the adult human os coxa. *American Journal of Physical Anthropology*, *169*, 689-703.] Editors from the American Journal of Physical Anthropology assisted with the research design and manuscript edits of Chapter 2. Pokotylo, D. L., and Weston, D. A. assisted with research design and manuscript edits. I formulated the research concept, wrote all major areas of the manuscript, conducted all research, and executed all statistical analysis.

Recommendations from *Forensic Science International* editors were incorporated into the statistical tests of Chapter 3. Weston, D. A., Pokotylo, D. L, and Cardoso, H. assisted with research design and manuscript edits. I formulated the research concept, wrote all major areas of the manuscript, conducted all research, and executed the statistical analysis.

Cardoso, H., Weston, D. A. and Pokotylo, D. L, assisted with research design and manuscript edits of Chapter 4. I formulated the research concept, wrote all major areas of the manuscript, conducted all research, and executed the statistical analysis.

Table of Contents

List of Tables

List of Figures

List of Symbols

Chapter 2

† Hispanic (n = 4), Native American (n = 3), mixed White and Native American (n = 2), Polynesian (n = 1), mixed White Polynesian (n = 1), Filipino (n = 1), Japanese (n = 1)

* Variance

+ Cumulative Variance

× Total Female Variance

§ Total Male Variance

Chapter 3

- no data

Chapter 4

* Do not use if the acetabulum has excessive osteophytosis.

+ 1 standard deviation from the centroid

List of Abbreviations

ANOVA: Analysis of Variance

Anth.Var.: Anthroposcopic Variable

Cum.%: Cumulative Variance presented as a percentage

DF: Discriminant Function

DFA: Discriminant Function Analysis

Fem.Var.: Female Variance

GM: Geometric Morphometrics

Lbs: Pounds

LR: Logistic Regression

LRM: Logistic Regression Model

Male.Var.: Male Variance

MANOVA: Multivariate Analysis of Variance

MANCOVA: Multivariate Analysis of Covariance

PCA: Principal Component Analysis

PC: Principal Component

SD: Standard Deviation

Std. Error: Standard Error

Test Pop.: Test Population

Var. %: Variance as a percentage

Yrs.: Years

Acknowledgements

This project would not have been possible without the supervision and mentorship of Drs Darlene Weston and David Pokotylo Department of Anthropology, University of British Columbia, and Dr. Hugo Cardoso, Department of Archaeology, Simon Fraser University. Thanks also to Dr. Dawne Steadman, Dr. Lee Meadows-Jantz, Jieun Kim, and Haley Horbaly, Forensic Anthropology Centre, University of Tennessee, Knoxville, for access to the William M. Bass Donated Collection. Dr. Barbara Winter and Dr. Deborah Merrett, Department of Archaeology, Simon Fraser University, and Dr. Helen Kurki, Department of Anthropology, University of Victoria, for the loan of their NextEngine laser scanners and peripheral tools. Special thanks to Dr. Michael Richards, Department of Archaeology, Simon Fraser University, and Dr. R.G. Matson for their support.

I would also like to thank the UBC undergraduate students who volunteered their time to assist with the inter/intraobserver error studies: Claire Allen, Alara Aydin, Paneet Bahia, Jade Beaupré, Meghan Brown, Ellen Dipple, Laura Fisher, Tessa Grogan, Caroline Hanna, Bella Hogan, Bethany Khalil, Ester Lo, Curt Logan, Hayley McLeod, Liam O'Reilly, Polina Orlov, Rebecca Patterson, Suparada Philips, Andrew Raitt, Indri Smith, Sarah Russ, Kate Sensenig, Mine Sher, Jennifer Sugar, Danika Vincent, Kathryn Vit, JK Vo, Ashley Wand, Teagan West, and Karly Zouppas,

Dedication

To my daughter

Chapter 1: **Introduction**

In biological anthropology, sex estimation from the os coxa (hip bone) is accomplished in one of two ways: 1) by interpreting morphological characteristics, where non-metric sex traits are assessed for "maleness" and "femaleness" based on the strength or weakness of their expression, and is the standard in biological anthropology or 2) by metric methods, using discriminant functions or algebraic equations (Albanese, 2003; Arun et al., 2012; Baumgarten & Ousley, 2015; Buikstra & Ubelaker, 1994; Brůžek, 2002; Klales et al., 2012; Phenice, 1969). Non-metric trait assessment is used most often to identify sex in archaeological skeletons of unknown population affinity and has one of the highest percentages of accurate sex prediction when interpretations are made by well-trained biological anthropologists (Buikstra & Ubelaker, 1994; Brůžek, 2002; Klales et al., 2012; Phenice, 1969). Osteometric methods, on the other hand, are more reliable across multiple observers, which makes them more desirable in forensic analyses of skeletal remains despite having slightly lower predictive power than non-metric traits (Albanese, 2003; Arun et al., 2012; Baumgarten & Ousley, 2015; Brůžek, 2002). The problem with metric methods, however, is they largely depend on population specific measurements, which makes them incomparable to each other and difficult to standardize (Biwasaka et al., 2012; MacLaughlin & Bruce, 1986; Listi, 2010). Osteometric methods of sex estimation are also more inclined to extract size variables from the coxal bone that are typically conflated with sex due to sex-based body size and adolescent growth differences (Coleman, 1969; Kurki, 2011, 2013; Tague, 2000). Shape analysis, on the other hand, is a highly accurate technique that is gaining popularity in the field of biological anthropology but requires further exploration before

sex-based shape analysis can be applied to skeletons of both known and unknown population affinity and before landmark standardization can take effect.

Previous coxal bone shape analyses between geographically distinct human populations often explain within- or between-sex differences in the study sample as population variation (Anastasiou & Chamberlain, 2013; Betti et al., 2013; Bilfeld et al., 2012; Bytheway & Ross, 2010; Gómez-Valdés et al., 2012; González et al., 2009; Pretorius et al., 2006; Steyn et al., 2004; Velemínská et al., 2013). However, when stature, body mass, and body breadth measurements are included in these analyses, using traditional Euclidean or linear osteometrics, these variables correlate with coxal bone shape and not with geographic location (Albanese et al., 2008; Fischer & Mitteroecker, 2015; Hierneaux, 1985; Kurki, 2011, 2013; Papaloucas et al., 2008; Ridgeway et al., 2008, Ridgeway et al., 2011). Stature, body mass, and body breadth variables correlate with sex in osteometric analyses of shape differences while geographic origin suggests a weak relationship between coxal bone shape and geographic origin. The relationship between coxal bone shape and geograpic origin is weaker than the relationshp between coxal bone shape and sex.

To investigate relationships between stature, body mass, and sex-based shape in the coxal bone, this study used geometric morphometric and multivariate statistical analyses to find true sex differences in the os coxa. This dissertation defines true sex differences as differences in the coxal bone that relate to obstetric function. Essentially, the goal of this research was to locate bodies that are biomechanically equipped to give birth regardless of population affiliation or geographic origin. This research had three main objectives: 1) to identify sex-based shapes in the coxal bone that best represent individuals who are biomechanically equipped to give birth and those who are not; 2) to test the landmarks representing true sex differences in the coxal bone

against linear measurements to develop a reliable and repeatable discriminant function equation to estimate sex; and 3) apply the above mentioned landmark and measurement data to develop discriminant function equations that are reliable and repeatable for sex estimation in fragmented coxal bones. This research dismantles the assumptions that population affinity contributes to sex differences and proposes an alternative population inclusive method to interpreting sex-based shape. The methods presented in this dissertation strategically captured multiple sources of skeletal variation to improved sex estimation methods in skeletal remains.

1.1 Background and Theory

1.1.1 Sex Estimation Methods

Sex estimation can be performed in many ways on different parts of the skeleton. This study is interested in sex estimation methods employed by the most sexually dimorphic bone in the human skeleton, the coxal bone. The numerous methods that estimate sex from the coxal bones fall into three general categories: visual (analysis of non-metric traits), osteometric, and geometric morphometric (shape analysis). Each method has an application that is important in biological anthropology, but they all tend to reproduce the same theoretical assumptions that limit their functionality; thus, they must be understood within the framework of population affiliation.

The visual method, in which non-metric sex traits are identified, compared, and interpreted as Male and Female, is a highly accurate method that can be applied universally to any skeletal collection, and currently is the standard by which undocumented archaeological

material is identified (Buikstra & Ubelaker, 1994; Walker, 2005; Wescott, 2015). However, the visual method has been criticized for being inaccurate between observers and difficult to master because it requires a high level of experience in recognizing traits within the spectrum of human variation (Brůžek, 2002). On the other hand, the osteometric method of sex estimation is easier to duplicate between observers because it is formulaic (Brůžek, 2002). Osteometric methods are also preferred over visual methods in forensic applications of sex estimation because the results can be reproduced accurately, consistently, and objectively (Brůžek & Murail, 2006). The downside to osteometric sex assessments is they are generally not as accurate as visual methods. The accuracy of osteometric sex equations tends to decrease when applied to skeletal populations that differ geographically or temporally (Patriquin et al., 2003; Walker, 2005).

To improve the accuracy of osteometric sex equations, scholars have developed or suggested the development of population- and temporal-specific sex equations (MacLaughlin & Bruce, 1986; Patriquin et al, 2003; Small et al., 2012; Vercellotti et al., 2011; Walker, 2005). This has proved problematic since population affiliation in forensic and biological anthropological contexts is often unknown, thereby negating the practical usefulness of population-specific sex equations. Other authors have suggested that population specificity does not impact the statistical significance of sex determination in large sample sizes (Steyn & Patriquin, 2009) and that variation observed between populations is a consequence of stature or body mass variation rather than ancestral affiliation (Albanese et al., 2008). When studies account for differences in stature or body mass, osteometric methods tend to be more accurate and lose their correlation with population affiliation or geographic origin, making these osteometric methods widely applicable among different skeletal populations (Albanese et al, 2008; Cabo et al., 2012; Kurki, 2011; Ridgeway et al., 2008; Ridgeway et al, 2011).

Sex differences in the human skeleton can also be investigated via geometric morphometrics. Geometric morphometrics is the analysis of object shape and size represented by landmark coordinate data (Bookstein, 1991; Sholts et al., 2011; Zelditch et al, 2012). Multivariate statistical methods are used to detect patterns in object shape and size that contribute to similarities or differences between groups. Geometric morphometrics has been used to accurately estimate sex in both whole and partial coxal bones (Anastasiou & Chamberlain, 2013; Betti et al, 2013; Bytheway & Ross, 2010; Gómez-Valdés et al., 2012; González, 2009; Pretorius et al., 2006; Wilson et al., 2015). Landmarks that represent shape can be selected to capture morphological variation of sex traits, such as the greater sciatic notch and sub-pubic contour, as well as variations in coxal bone length and thickness (Betti et al., 2013; Bilfeld et al., 2012; Bytheway & Ross, 2010; Djorojevic et al, 2014; Gómez-Valdés et al., 2012; González et al., 2009; Pretorius et al., 2006; Steyn et al., 2004; Velemínská, et al 2013). Geometric morphometrics has the flexibility of incorporating the strengths of both visual and osteometric methods by capturing important morphological data that contribute to highly accurate sex estimations with a high level of objectivity like osteometric measurements (Bookstein, 1978; Corner et al., 1992; Zelditch et al., 2012).

1.1.2 Biological Anthropology Theory

Despite the methodological diversity in which sex can be estimated in the human skeleton, a common theory unites them - skeletal sex can be interpreted in the same way that biological sex is interpreted: as static and insulated from external influences such as gendered norms and adaptation (see critiques in Geller, 2005, 2008; Meskell, 2000, 2007; Nordbladh &

Yates, 1990). There has been increasing sentiment among gender theorists and gender archaeology theorists regarding the true complexity of biological sex and how its use in biological anthropology has been oversimplified. Biological sex is the culmination of genital morphology, chromosomal karyotype, and hormonal makeup (Butler, 1993; Doyle & Paludi, 1991; Fausto-Sterling, 2000; Geller, 2005; Gilchrist, 1999; Hollimon, 2011; Knudson & Stojanowski, 2008; Meckel, 2000; Sofaer, 2013; Stone & Walrath, 2006). Skeletal sex, on the other hand, is defined in this dissertation as the manifestation of coxal bone growth as directed by hormones, nutrition, and indirectly influenced by genitals and chromosomes.

Traditionally, biological anthropology assumes the direct correlation between skeletal sex and biological sex. This is not surprising since many of the chromosomal and hormonal stimulants that influence the development of secondary sex characteristics and body forms that signal sexually maturity are the same stimulants that trigger changes in the human skeleton to accommodate the sexually mature body form (Doyle & Paludi, 1991; Hollimon, 2011). Gender scholars challenge the idea of a binary biological sex by highlighting the complexity of chromosomal karyotype (XO, XXY, and XYY) and hormonal makeup (androgen insensitive males, gonadal dysgenesis) that produce biological sex (Butler, 1993; Doyle & Paludi, 1991; Fausto-Sterling, 2000; Geller, 2005, 2008; Nordbladh & Yates, 1990; Thomas, 2007). Gender scholars have also highlighted the social construction of sex as science's attempt to distance it from gender (Butler, 1993; Doyle & Paludi, 1991; Fausto-Sterling, 2000; Geller, 2005, 2008). The process of ascribing sex at birth based on genital phenotype is an act of construction, particularly when it excludes other genital forms (Butler, 1993; Fausto-Sterling, 2000). In Western culture, non-typical genital forms are pathologized and "corrected" based on a constructed medicalized wisdom of what is "normal" (as opposed to typical) and what is not

(Butler, 1990, 1993; Fausto-Sterling, 2000; Sitek et al., 2012). The heteronormative construction of sex that gender scholars such as Judith Butler, and Ann Fausto-Sterling, have illuminated is a post-modern feminist critique of the scientific assumptions regarding human biological sex.

Some biological anthropologists such as Pamela Geller, Jarl Nordbladh and Tim Yates have adopted these post-modern feminist critiques in their perception of the importance of sex in biological anthropology, although in a slightly different way. Sex estimation is often one of the first steps in an osteological analysis of archaeological human remains, along with determining age and ancestral affiliation (Agarwal, 2012; Buikstra & Ubelaker, 1994). However, in paleopathological or mortuary studies, identifying sex patterns in the assemblage prior to pathological analyses can bias or presuppose patterns of disease that might not be socially relevant (Agarwal, 2012; Stone & Walrath, 2006). Gender archaeology studies have already deemphasized the importance of sex estimation as the first step in an osteological profile in exchange for increasing the focus on understanding changes in the life course and identity among individuals in a society over time (Agarwal, 2012; Armelagos, 2003; Gilchrist, 1999, 2004; Joyce, 2005; Meskell, 2000). Although the importance of sex estimation as a mainstay of osteological description is declining in archaeology in favour of a more gendered approach to the interpretation of human organization, there is by no means a push to eliminate it altogether from the osteological tool kit. On the contrary, authors like Joanna Sofaer (2013) call for improvements to the predictive power of sex estimation so it can make interpretations of life courses more accurate (Meskell, 2000; Sofaer, 2013).

Because sex estimation is largely based on the interpretation of sex-traits that develop during adolescence, one could say that biological anthropologists are examining the side effect of biological sex that is being expressed in the skeleton, and are not examining biological sex

directly, as you would in molecular methods of sex determination that detect the chromosomal make up of individuals, or in *in vivo* studies that document primary sex characteristics (Coleman, 1969). As hormones change with age, so too does skeletal morphology, making skeletal sex more fluid within an individual over his or her lifetime (Sitek et al., 2012). This idea that skeletal sex is more fluid than biological sex reflects the theories upheld by gender scholars. For example, among very young individuals, sex determination might not always be possible, but the likelihood of discerning sex increases with age along with sexual maturity (Wilson et al., 2015). As individuals reach adulthood, their skeletal sex differences become more apparent, however, a sex bias exists among very young adults and old adults (Walker, 1995, 2005; Weiss, 1971). Very young Male and Female adults (< 20 years old) tend to have more Female like skeletal sex morphologies, and as an individual reaches the age of 40 years, more Male-like morphologies develop (Walker, 1995). The gradual masculinization of sex traits with increasing age could explain why there is often a 12% sex bias in favour of males when using bioarchaeological methods of sex estimation (Weiss, 1971). It is this age-based change in skeletal sex that suggests that it is fundamentally different from biological sex and should receive renewed attention in bioarchaeological theory and method.

The way in which gender theory incorporates the fluidity and changeability of skeletal sex as it is manifested in the skeleton as we age, opens the doors for incorporating other theoretical similarities; specifically, that skeletal sex is constructed through methodological categorization. Given that the skeleton undergoes continuous change during life, impacting the interpretation of skeletal sex, a deeper scrutinization of the methods is warranted in order to improve the practice of sex estimation. One way to deepen the scrutinization of biological anthropological methods of sex estimation is to identify the assumptions made in the

interpretation of sex that could be confounding the predictive power of the sex estimation methods being used. The term *skeletal sex* will be used to highlight the argument outlined above and typical categories of sex such as *Female* and *Male* will be capitalized to identify them as constructed categories.

1.1.3 Confounding Factors of Sex Estimation

When biological anthropologists investigate skeletal sex in the coxal bone using either visual, osteometric, or geometric morphometric methods, they are often comparing sexual dimorphism as a function of stature and body mass differences between Males and Females, as well as sexually distinctive shapes and non-metric traits (Betti, 2014; Cabo et al., 2012; Clark, 2014; Kurki, 2013; MacLaughlin & Bruce, 1986; Ruff, 2002; Vercellotti et al., 2011). Size sexual dimorphism is the physiological consequence of Males having a longer adolescent growth period and a greater growth velocity compared to Females (Cabo et al., 2012; Hiernaux, 1985). As a result, allometric body proportions (body breadth) between Males and Females are differentiated and these differences are reflected in the coxal bone, with shorter individuals having shorter coxal bone heights (Kurki, 2011; Fischer & Mitteroecker, 2015). However, stature is also affected by nutritional variation, which can hinder adolescent growth (Vercellotti et al., 2011; Zakrzewski, 2003). Within groups of shorter and taller individuals, the magnitude of sexual dimorphism between Males and Females is consistent, meaning if geometric size differences can be controlled for, all that would remain would be allometric differences related to development (Kurki, 2011).

Body mass does not exhibit sexual dimorphic differences among populations from varying latitudes, but it does have a significant correlation with sex (Ruff, 2002). Body mass, as represented in biological anthropology, is the skeletal mass of an individual derived from the cross-sectional geometric size of long bones and cortical bone thickness (Cabo et al, 2012). Males and Females differ in body mass in much the same way as they differ in stature, with differing rates of skeletal growth and development leading to thicker, heavier bones among men and lighter, thinner bones among women. Like stature, nutritional variances can negatively impact an individual's skeletal body mass; however, unlike stature, physical activity can maximize the skeletal body mass potential in an individual, while inactivity can minimize it (Cabo et al, 2012). Weight, self-reported by participants in the Body Donation Program, was observed as body mass.

Population differences, derived from climate adaptation, have inspired researchers to develop population-specific sex estimation techniques, particularly in osteometric methods of sex estimation (Macaluso, 2010; MacLaughlin & Bruce, 1986; Listi, 2010; Rosenberg, 2002; Steyn & Patriquin, 2009). Population variation is the basis for creating ancestral profiles that categorize individuals, however, studies on human population variation in the skeleton have identified more variation within population groups than between groups (Mays, 2010; Walker, 2000; White & Folkens, 2005; Zakrzewski, 2011). Some authors suggest that when examining a large population (around 600 individuals), within-group variation ceases to display statistically significant patterns (Steyn & Patriquin, 2009). When osteometric analyses of sex factor in such variables as stature, body mass, or body breadth, within-sex and between-sex differences correlate with these variables rather than with population afiliation (Albanese et al., 2008;

Fischer & Mitteroecker, 2015; Hierneau, 1985; Kurki, 2011, 2013; Papaloucas et al., 2008; Ridgeway et al., 2008, Ridgeway et al., 2011).

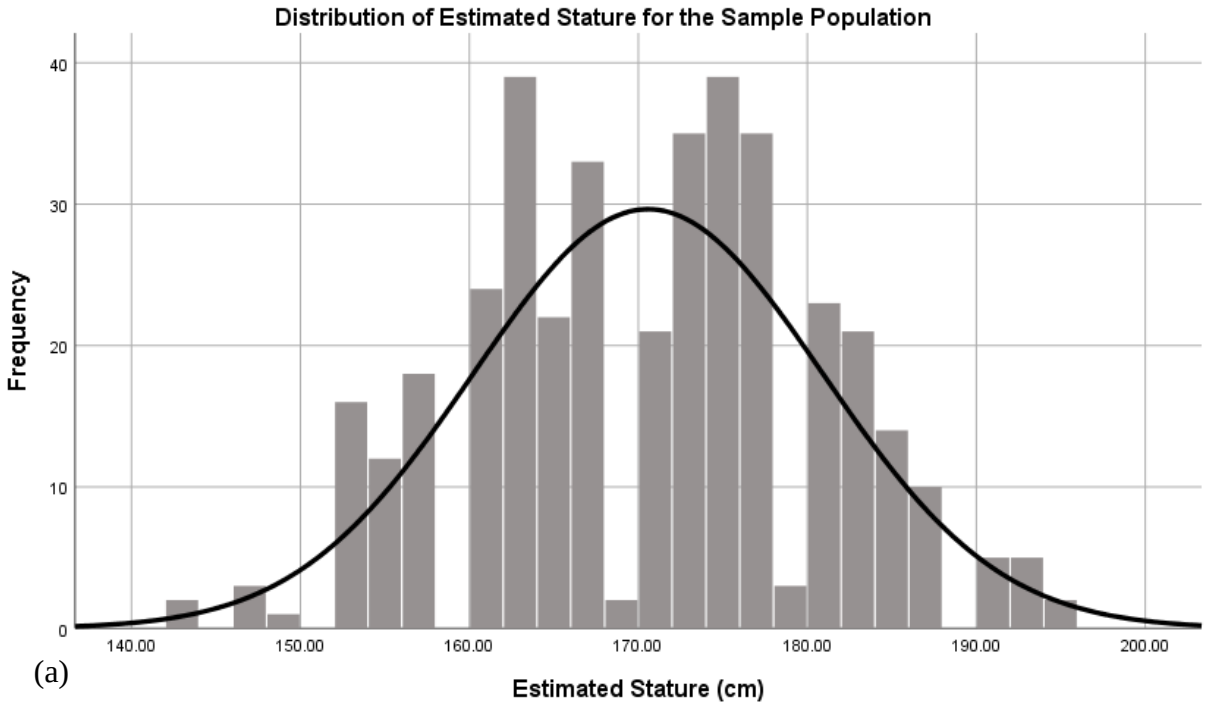
In osteometric analyses allometric size differences of body size and body mass are frequently considered when trying to improve the predictive capability of sex estimation methods (Betti, 2014; Kurki, 2013). However, it is often because osteometric measurements also capture isometric, or non-sex related variables associated with stature, body breadth, and body mass measurements that the predictive power obtained from these measurements are not as strong as visual methods. Allometric size differences have been attributed to variations in sexual dimorphism, variations between geographical populations, nutritional variation, and hormonal variation (Betti et al, 2013; Kurki, 2011). Using osteometric methods of sex estimation, it is difficult to determine if allometry or isometry is contributing to size-based sex difference. In order to improve osteometric methods of sex estimation, these non-sex related elements accompanying linear measurements must be accounted for so true sex differences can be isolated and incorporated into future methodologies. Geometric morphometrics has the added advantage of analysing sex-based shape differences independently from geometric size that aids in disassociating linear measurements with isometric size (Klingenberg, 2016). There is a need for a standardized method of sex estimation where non-sex related size components can be removed, leaving only sex-related size differences rooted in the biological function of childbirth. This type of method would provide biological anthropologists with a more focused interpretation of sex-based differences and would eliminate the need to develop population or temporally specific methods of sex estimation, thereby making a geometric method of sex estimation more universally applicable.

1.2 **Methodology**

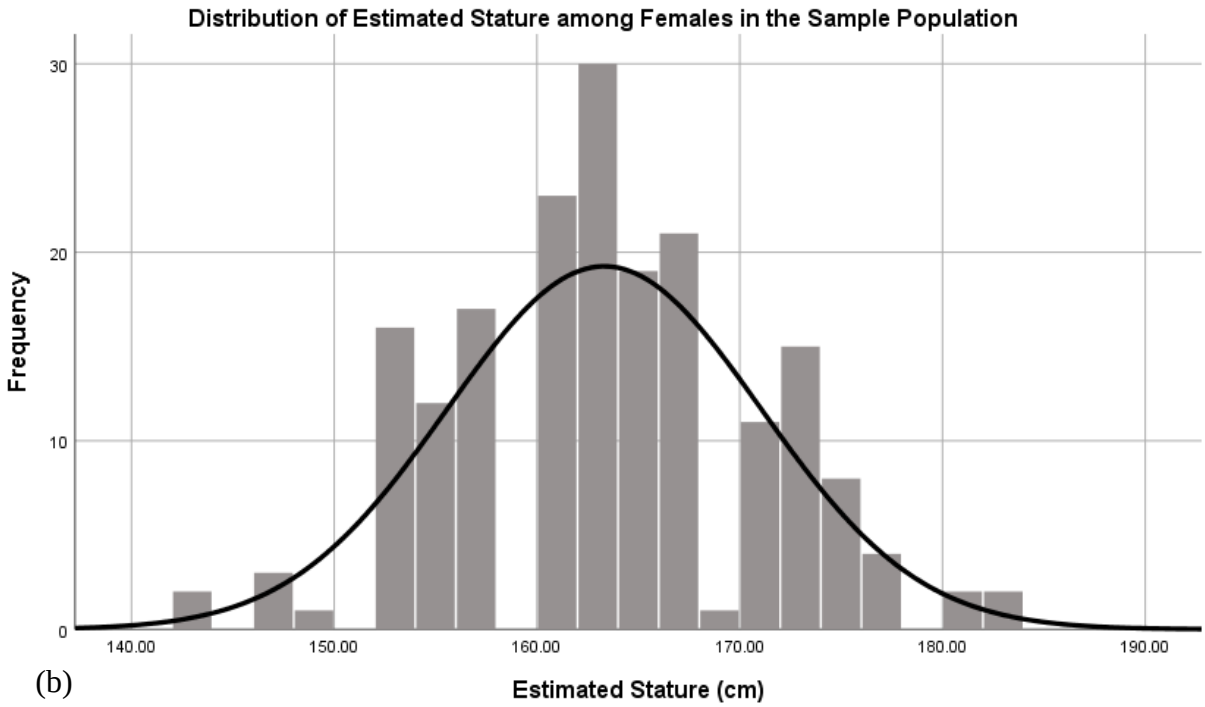
1.2.1 Specimens

This research was conducted on a population of 394 individuals of known sex, age, and ancestry from the William Bass Donated Skeletal Collection (Bass Collection), at the University of Tennessee, Knoxville TN (Males n = 203, Females n = 191). Living applicants to the Body Donation Program (BDP) provided age, sex, and ancestry information *pre mortem.* Skeletons were procured *post mortem* and catalogued into the Bass Collection. References to ancestry are capitalized in this dissertation to identify them as discrete categories. The study population is primarily made up of individuals categorized as White (n = 369). Non-white individuals included Black (n = 12), Hispanic (n = 4), Native American (n = 3), mixed White and Native American (n = 2), Polynesian (n = 1), mixed White and Polynesian (n = 1), Filipino (n = 1), and Japanese (n = 1). Although the ethnic origin of the sample population is clearly biased toward individuals classified as "White" this study is most interested in how sex measurements relate to body form variables. Ancestry is included in this study as a comparison to correlations with body form and to demonstrate where and how population variation becomes an important variable in analyses of coxal bone shape and sex-based measurements. Body size or stature ranges are more varied in the study sample (142.24cm – 195.58cm) as is weight (80lbs – 500lbs), suggesting methods derived from this study can be applied to any skeletal population within these ranges. Mean stature for the sample population is 170.57cm (SD = 10.36cm) and is normally distributed (Figure 1.1a, skewness = -0.060; kurtosis = -0.497). Average Female stature is 163.31cm (SD = 7.75cm) and is also normally distributed (Figure 1.1b, skewness = 0.049; kurtosis = -0.125,). Average Male stature is roughly 14cm taller than Female stature at 177.43cm, but both have

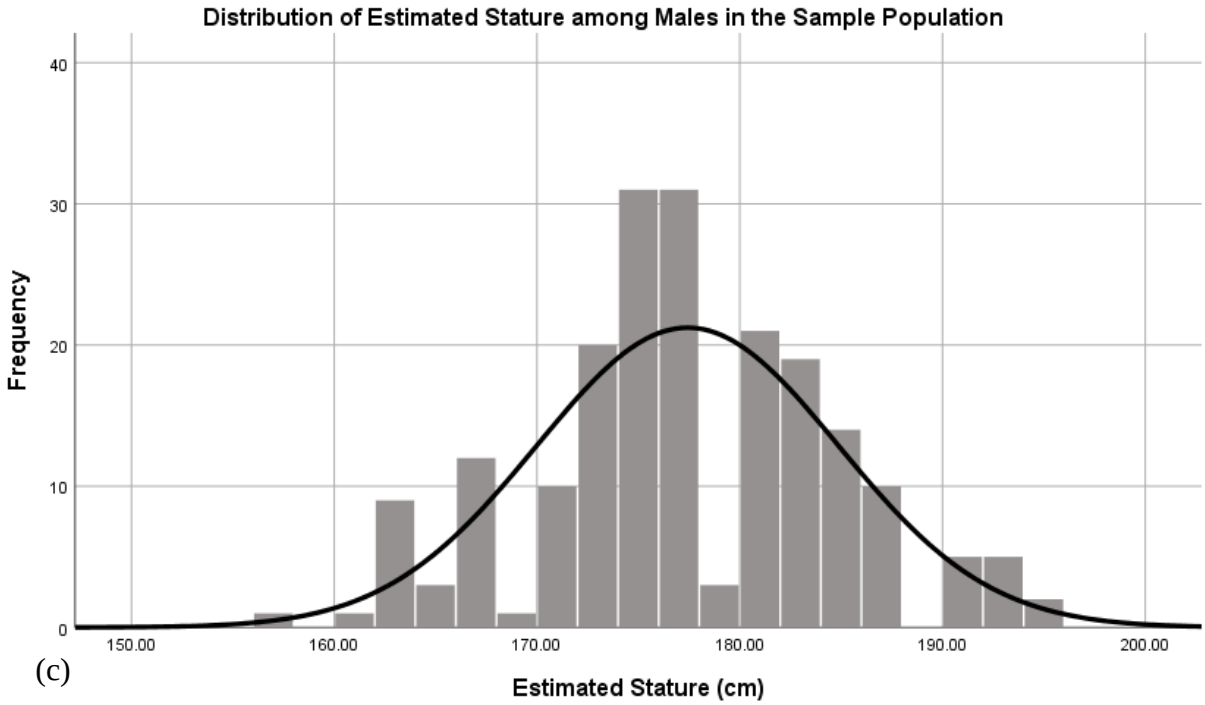
similar specimen distributions (Male SD = 7.44cm; skewness = -0.041; kurtosis = -0.032, Figures 1.1c).



(a)



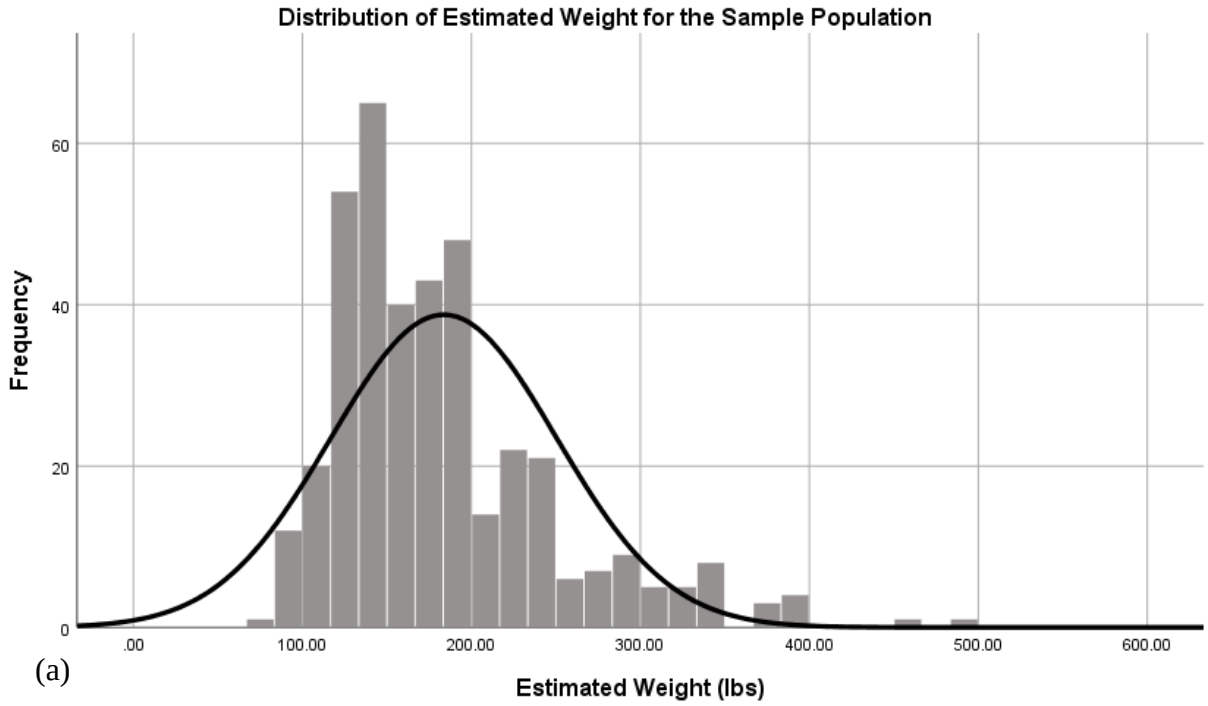
(b)



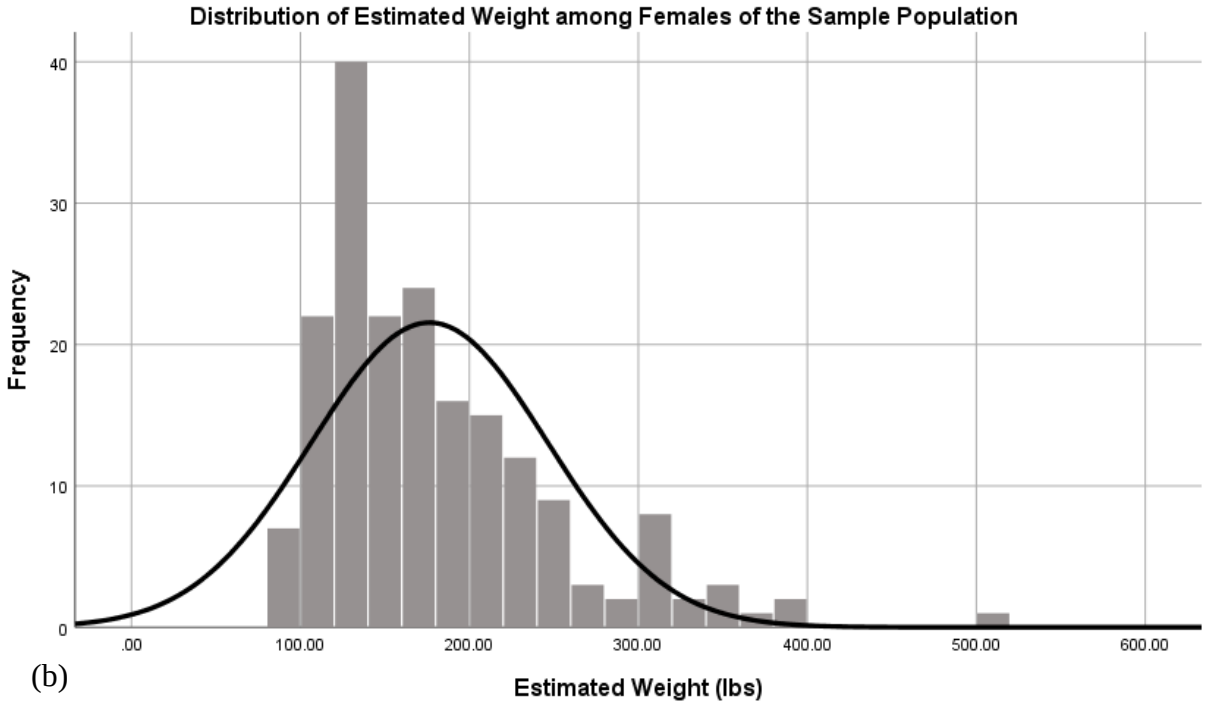
(c)

Figure 1.1 Distribution of Estimated Stature (cm) among the sample population (a), grouped by Females (b) and Males (c).

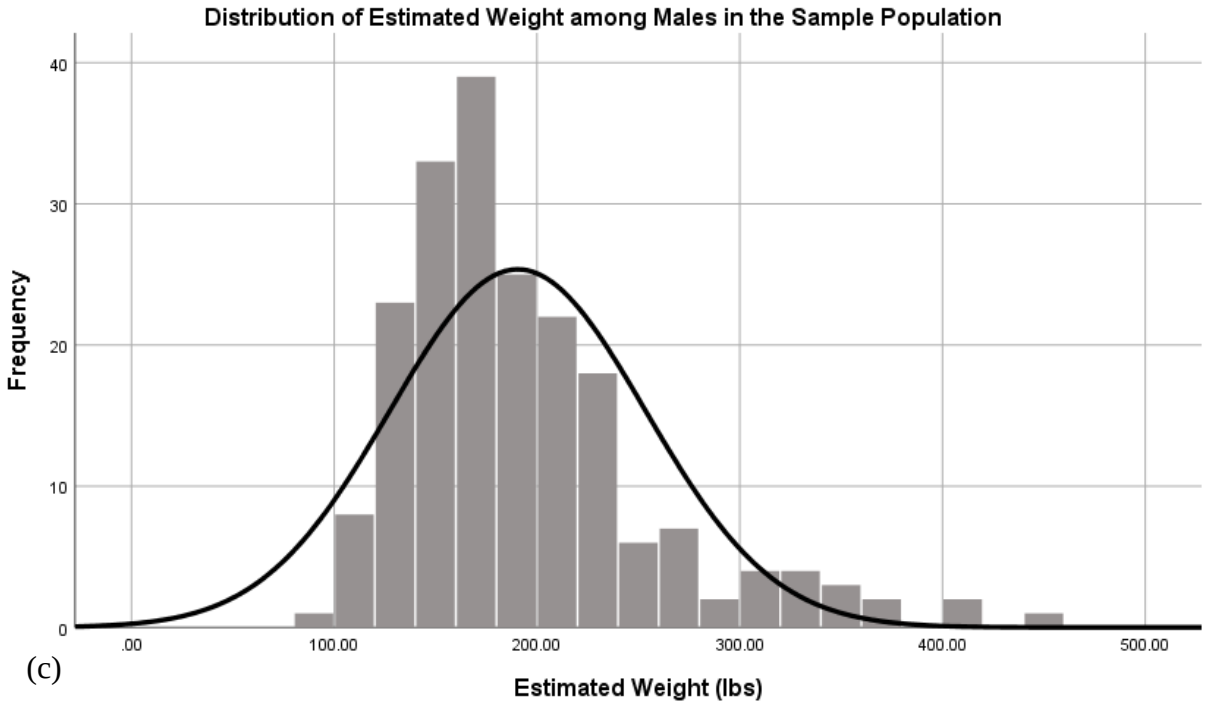
Estimated weight (Figure 1.2a) among the sample population averaged 183.64lbs (SD = 66.72lbs) and is negatively skewed (1.402) and leptokurtic (kurtosis = 2.408) due to a few heavy individuals (>400lbs) increasing the weight range. Estimated Female weight (Figure 1.2b) averaged 176.40lbs, roughly 13lbs lighter than the sample population average, and possesses a slightly larger weight distribution than the total sample population (SD = 69.96lbs). Female Estimated Weight is more negatively skewed (1.462) and leptokurtic (kurtosis = 2.564) than the total sample population. Male Estimated Weight (Figure 1.2c) averaged roughly 7lbs heavier than the sample population average (mean = 190.48lbs) and possesses a slightly narrower range in weight values (SD = 62.92lbs). Estimated Male weight is also negatively skewed and leptokurtic (skewness = 1.456; kurtosis = 2.527).



(a)



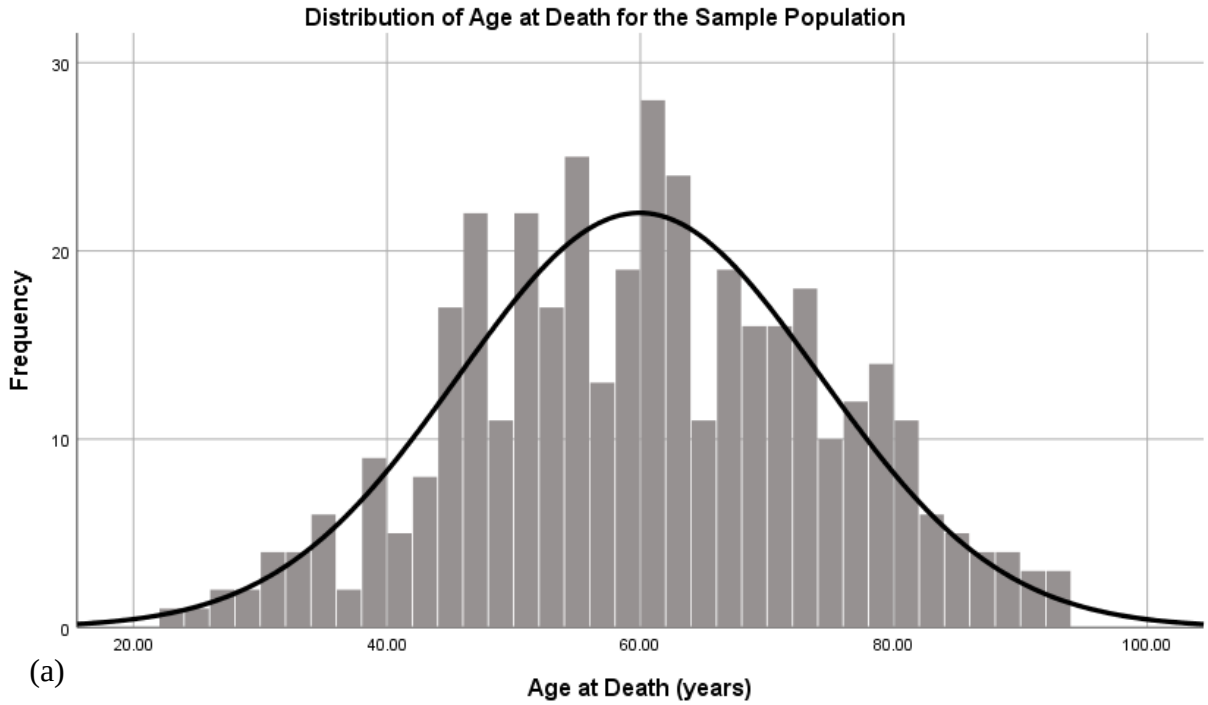
(b)



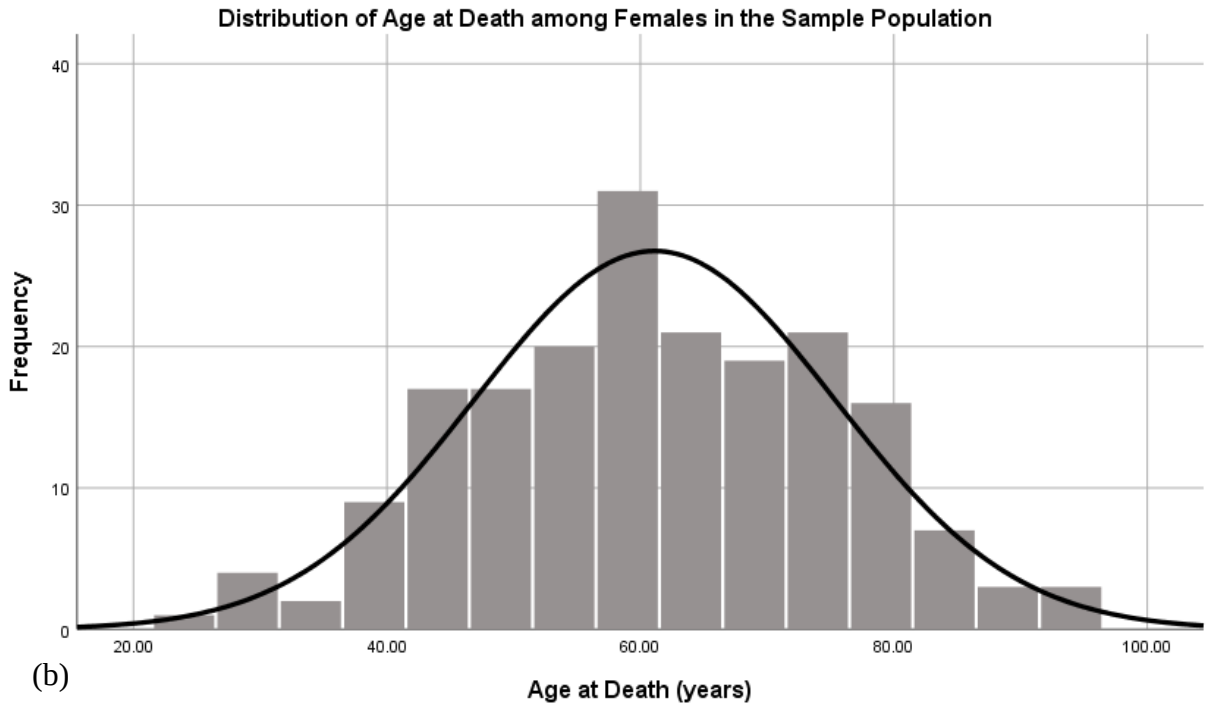
(c)

Figure 1.2 Distribution of Estimated Weight (lbs.) among the sample population (a), grouped by Females (b) and Males (c).

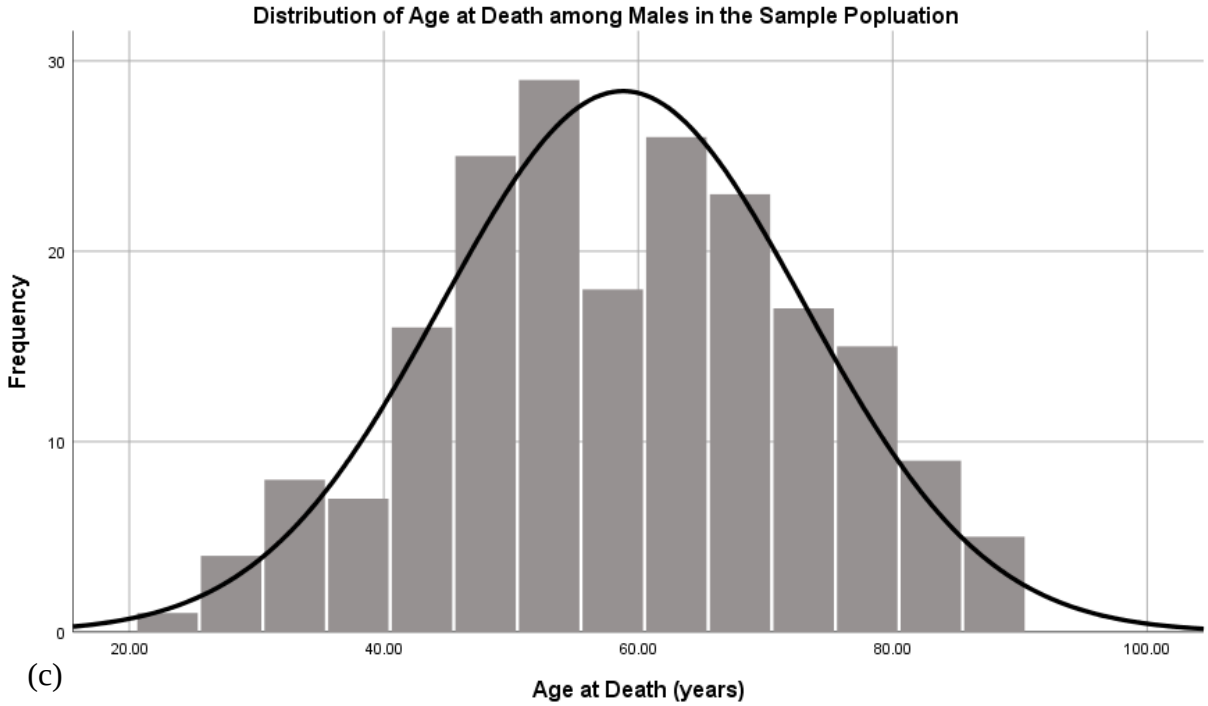
Age at death among the total sample population (Figure 1.3a) ranged between 23-93 years (mean = 59.93 years; SD = 14.27 years) and is normally distributed (skewness = -0.031; kurtosis -0.454). Female age at death (Figure 1.3b) averaged 61.12 years (SD = 14.24 years) roughly 1 year older than the total sample population, and is also normally distributed (skewness = -0.020; kurtosis = -0.409), while Male age at death (Figure 1.3c) averaged roughly a year younger than the total sample population at 58.80 years (SD = 14.25) and shares a similar distribution pattern as the total sample population (skewness = -0.040), however, Male age at death is lightly more leptokurtic (kurtosis = -0.496) than the female and total distributions.



(a)



(b)



(c)

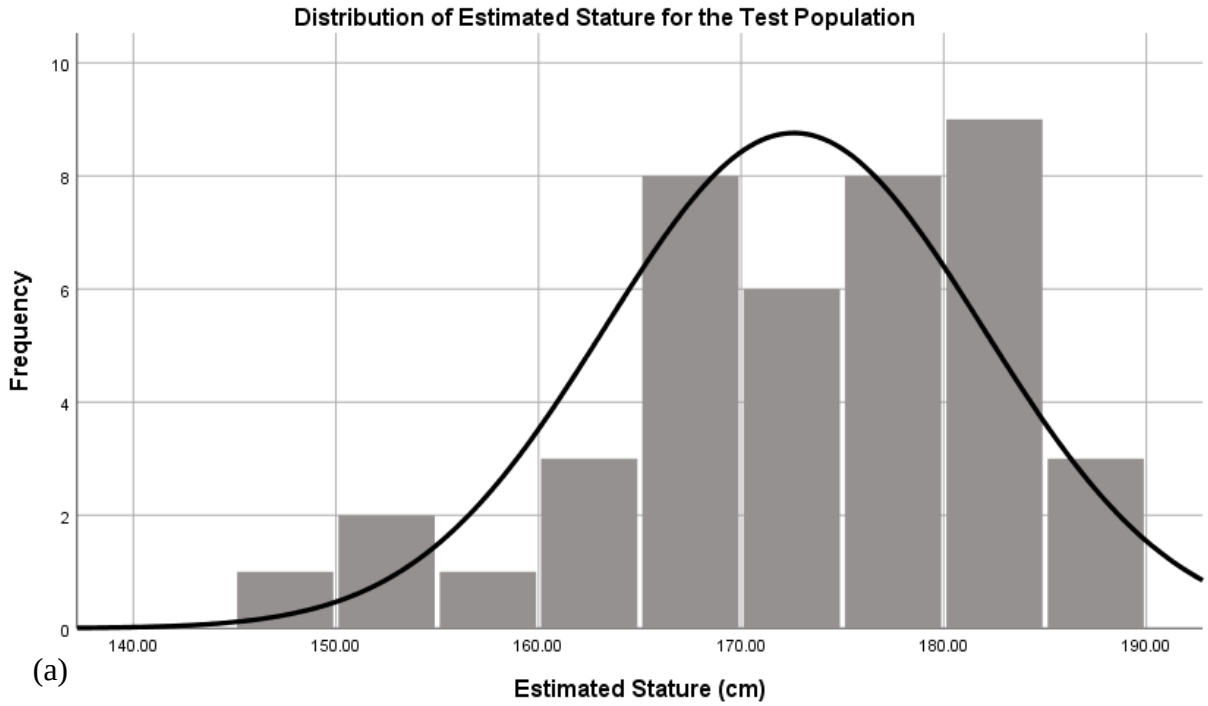
Figure 1.3 Distribution of Age at Death in years among the sample population (a), grouped by Females (b) and Males (c).

Studies of sexual dimorphism incorporating individuals aged between 15-19 years have demonstrated higher rates of misclassification of Males as Females (González et al., 2009; Walker, 1995). Age-related changes to the coxal bone will be minimized in this study by selecting individuals that do not exhibit extensive acetabular osteophitosis. Individuals who have undergone *pre mortem* hip replacement surgery were not included in the study poplulation. Individuals included in the study population were born during the late 19th to mid 20th centuries and could have experienced varied nutritional profiles related to secular food trends. Observable nutrition-related pathologies were documented from skeletal individuals when available.

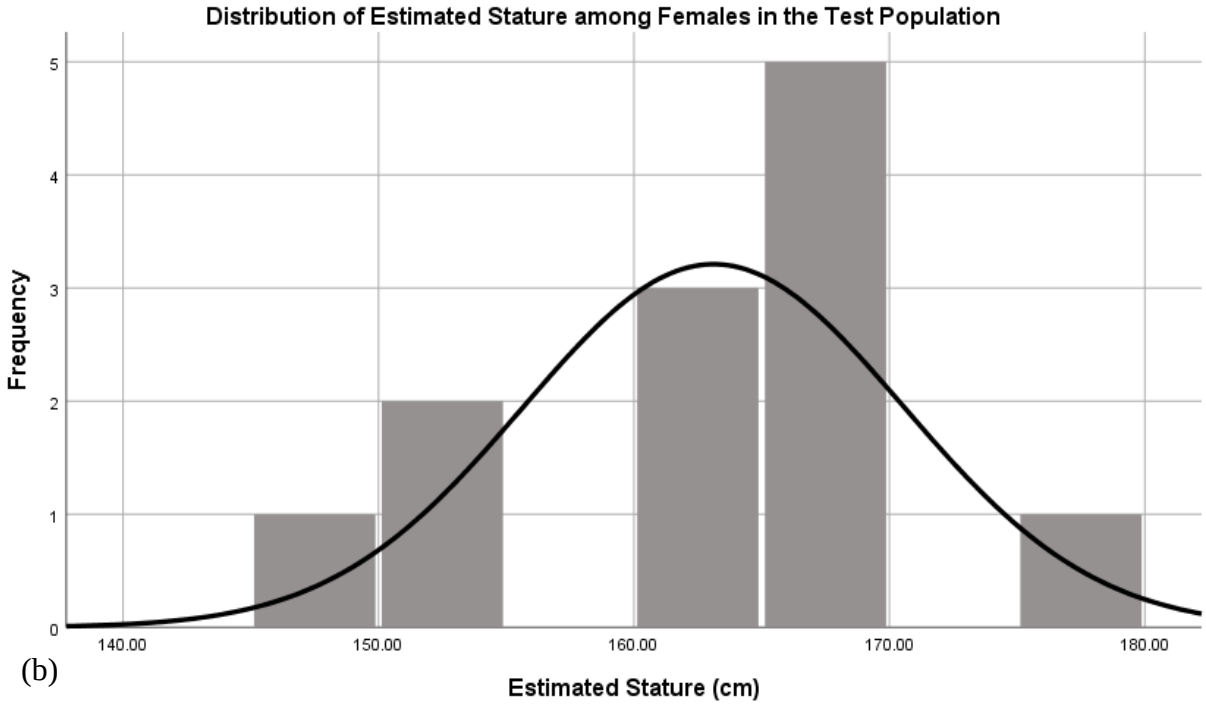
The shape data obtained from individuals in this study provided the basis for generating predictive models to determine sex in archaeological and forensic skeletal remains. The accuracy of these predictive models was tested 120 individuals, also obtained from the Bass Collection but independent from the sample population (Males n = 60, Females n = 60), varying in body size (stature 149.86cm – 187.96cm; weight 103lbs – 315lbs), and age at death (25-96 years old). The test population compised largely of individuals classified as White (n = 99; f = 57; m = 42), 14 individuals classified as Black (f = 3, m = 11), 4 males classified as "hispanic", and 3 males classified as "American Indian". These last two categories were combined into a collective "american" category (n = 7).

Average Estimated Stature among the test population is 172.60cm (SD = 9.34cm) and comparable with the sample population, however unlike the sample population the distribution of stature variables among the test population is slightly positively skewed (Figure 1.4a, skewness = -0.502; kurtosis = -0.305). Female stature (Figure 1.4b) averages roughly 10cm shorter (mean = 163.12cm), is less varied (SD = 7.46cm) and more normally distributed (skewness = -0.046, kurtosis 0.558) than the test population as a whole, while the distribution of Male stature is only

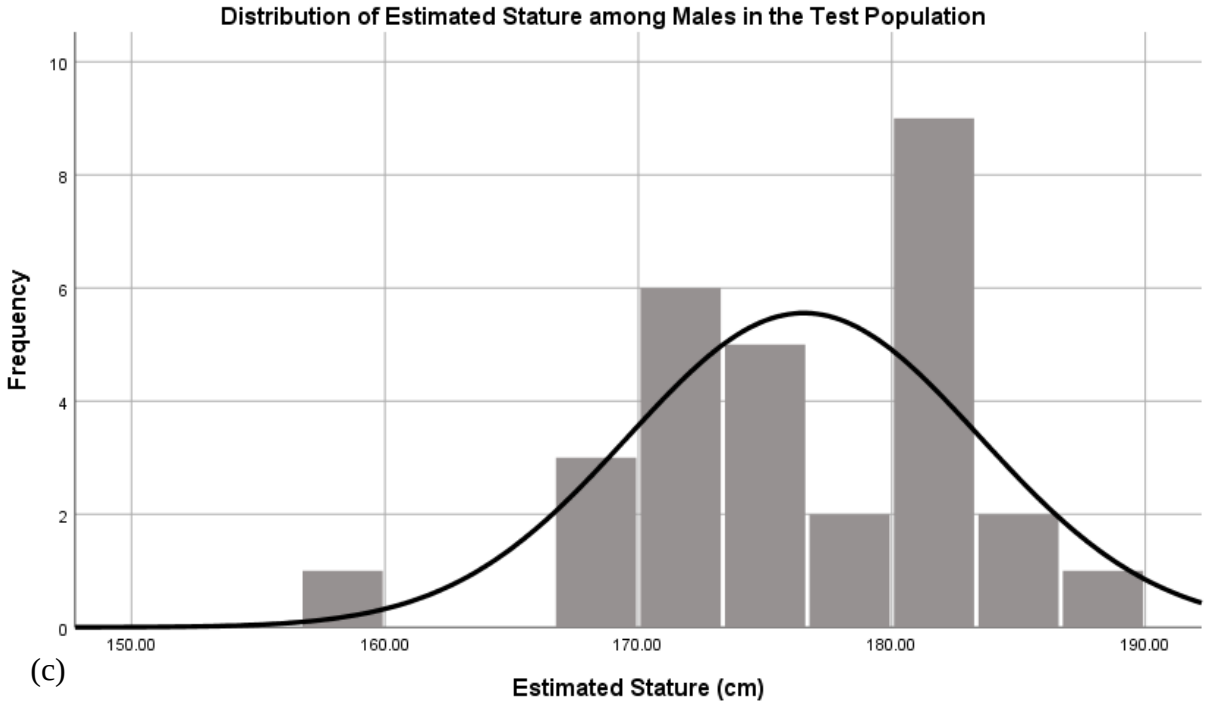
4cm taller than the group average (mean = 176.53cm, SD = 6.94cm) and slightly more skewed (skewness = -0.615, kurtosis = 0.389 than the test population as a whole (Figure 1.4c).



(a)



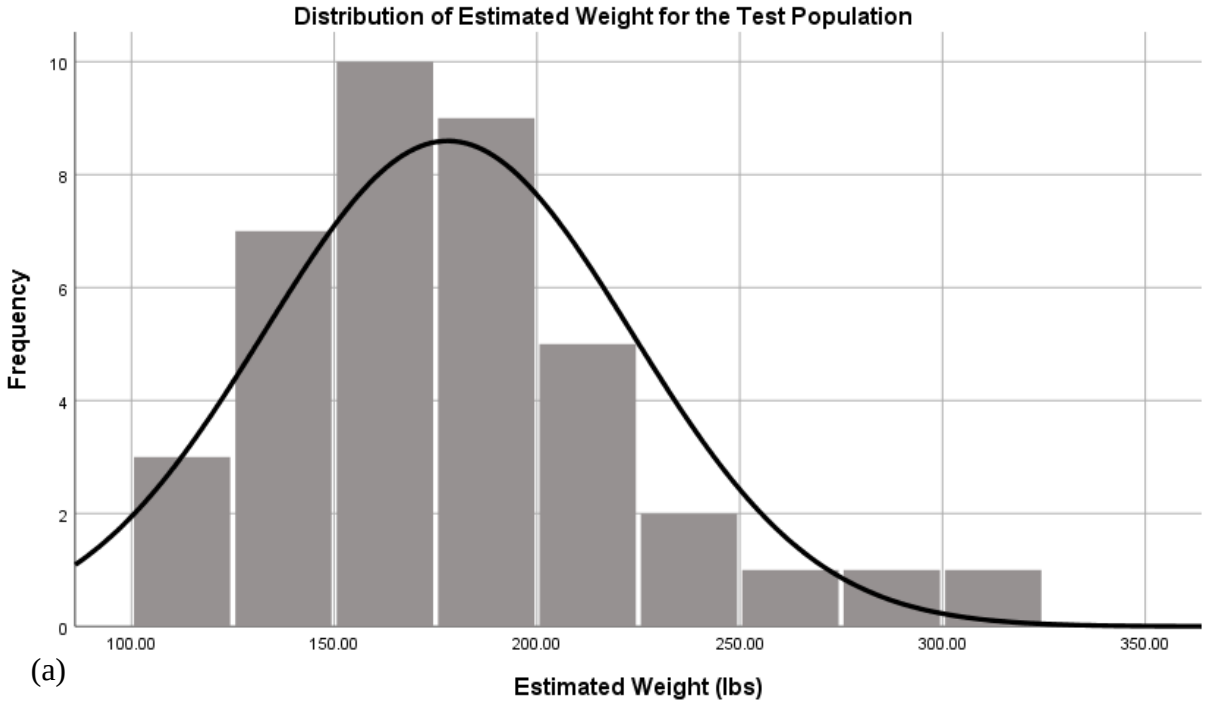
(b)



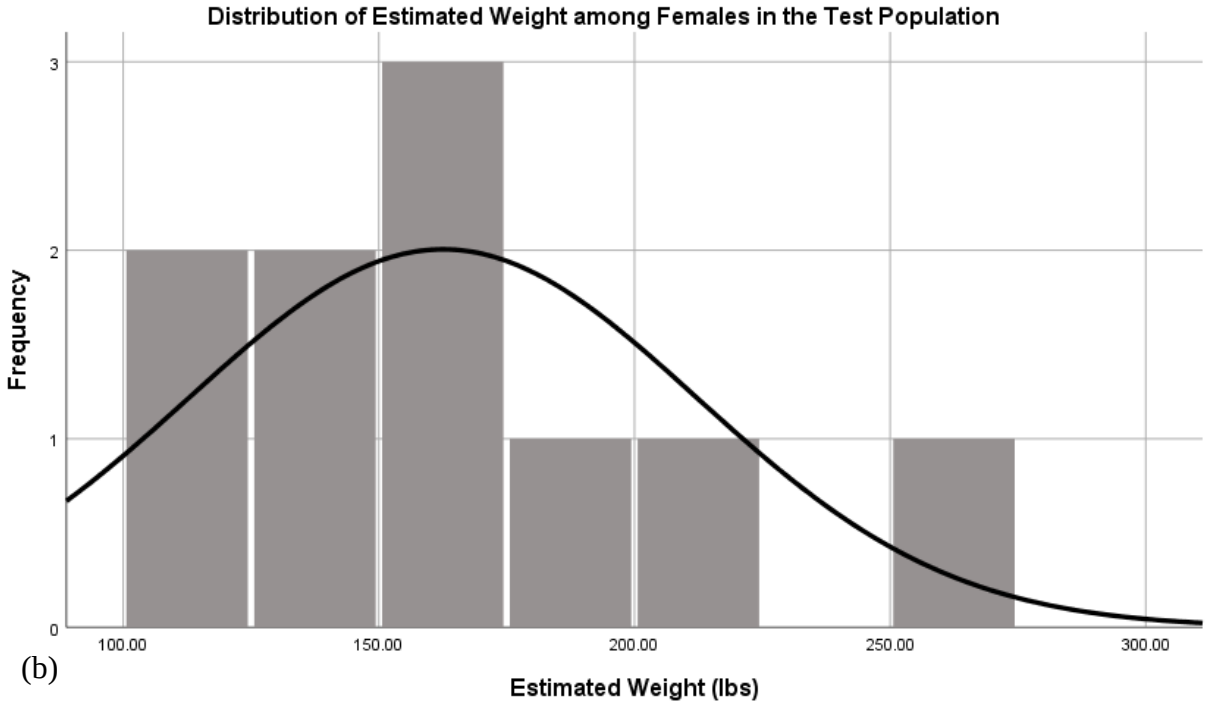
(c)

Figure 1.4 Distribution of Estimated Stature (cm) among the test population (a), grouped by Females (b) and Males (c).

Average Estimated Weight is roughly 14lbs lighter in the test population than in the sample population (mean = 178.02lbs; SD = 45.24lbs). Estimated Weight (Figure 1.5a) is slightly less negatively skewed (0.944) and less leptokurtic (kurtosis = 1.372) among the test population than the sample population. The distribution of Female weight in the test population (mean = 162.53lbs, SD = 49.74lbs) is slightly less skewed (0.769) and more mesokurtic (0.283) than the total test population (Figure 1.5b). Male weight is driving the skewness and leptokurtic distribution of the test population as a whole (mean = 183.36lbs, SD = 43.21lbs, skewness = 1.283, kurtosis 2.28, Figure 1.5c).



(a)

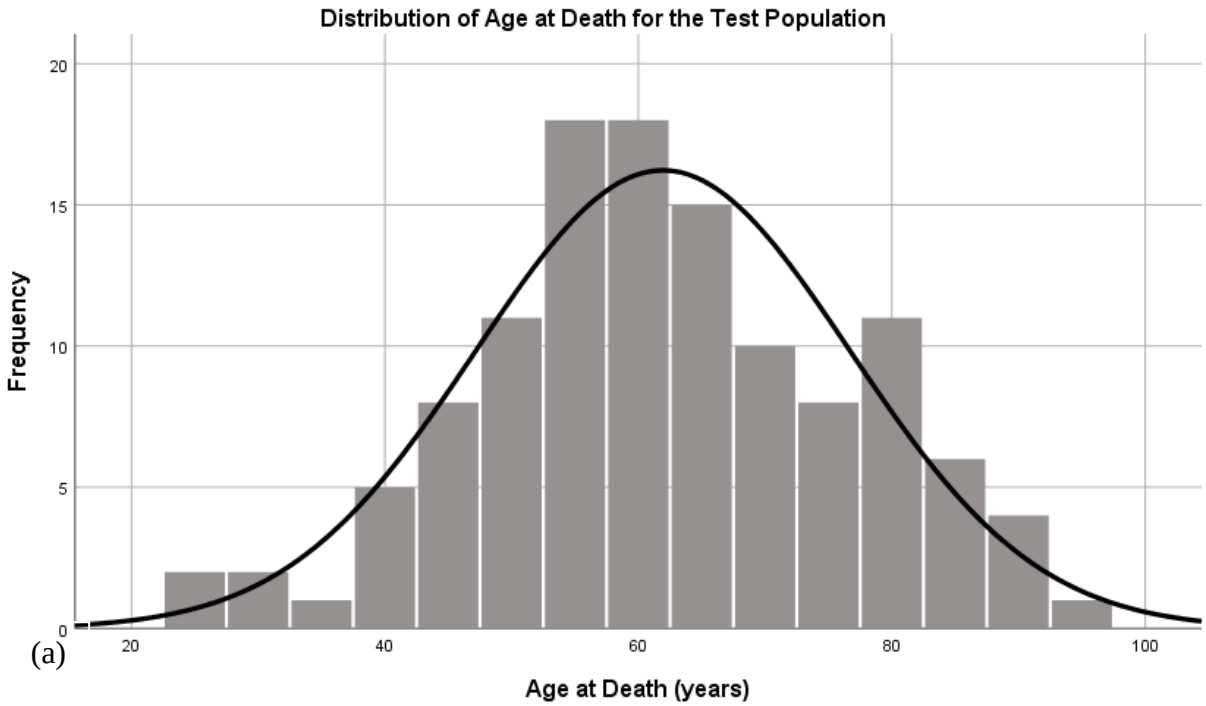


(b)

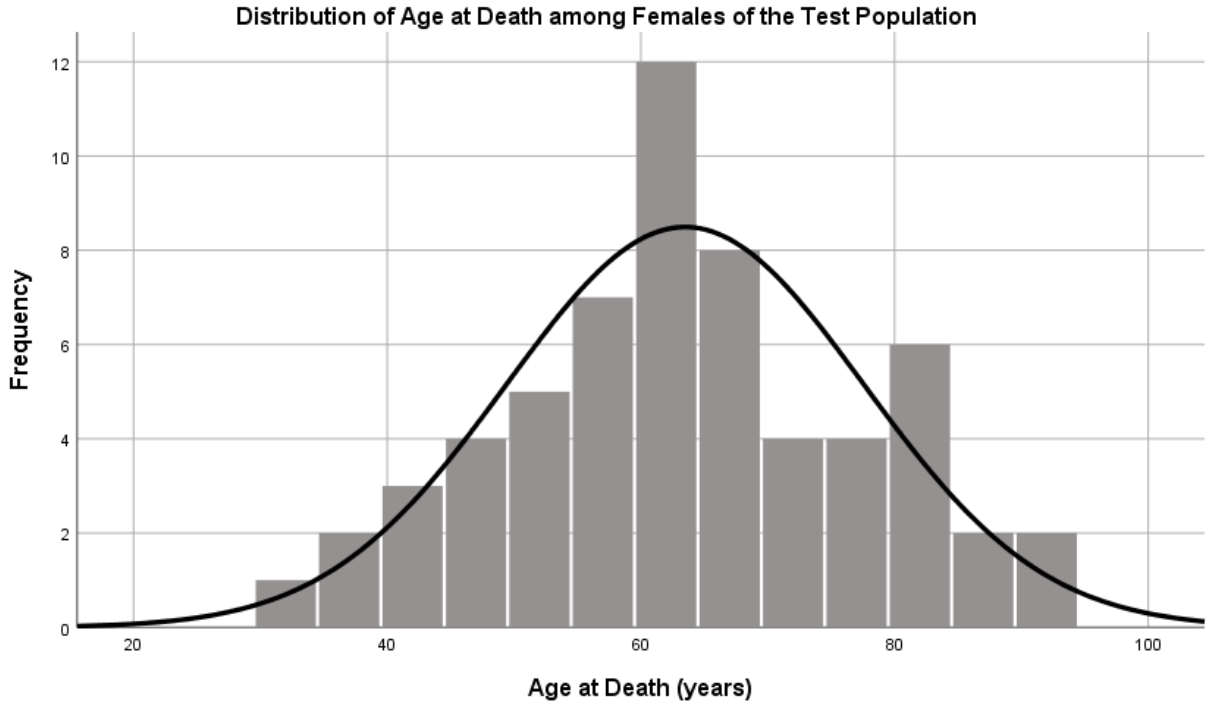
(c)

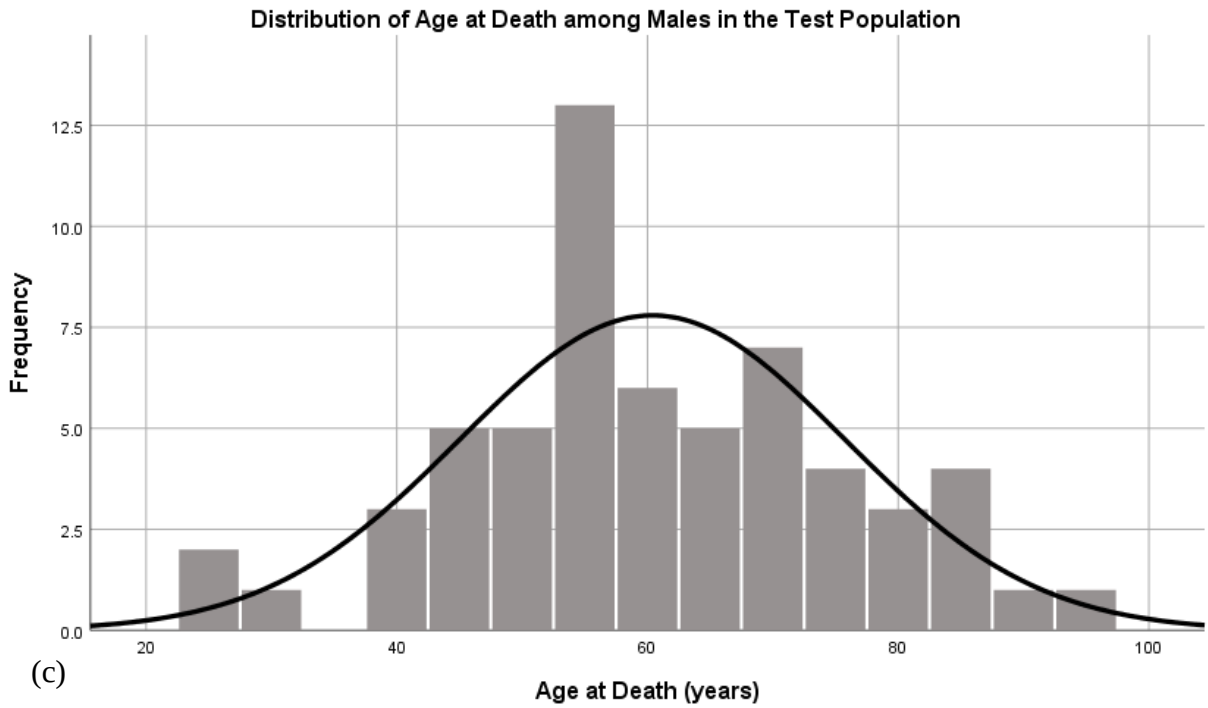
Figure 1.5 Distribution of Estimated Weight (lbs.) among the test population (a), grouped by Females (b) and Males (c).

Mean age at death among the test population was roughly two years older than among the sample population (mean = 61.93 years, SD = 14.75 years, skewness = -0.005, kurtosis = -0.319, Figure 1.6a), this pattern is consistent among Females (mean = 63.48 years, SD = 14.09 years, skewness = -0.004, kurtosis -0.421) and Males (mean = 60.38 years, SD = 15.35 years, skewness = 0.043, kurtosis = -0.211) in the test population (Figures 1.6b and c respectively).



(a)





(c)

Figure 1.6 Distribution of Age at Death in years among the test population (a), grouped by Females (b) and Males (c).

1.2.2 Study Variables

This analysis used both left and right coxal bones to account for individual fluctuating asymmetry. Although fluctuating asymmetry does not appear to influence the result of sex estimation in greater sciatic notch shape analysis specifically (Biwasaka et al., 2012), a comparison between left and right coxal bones is necessary in order to account for as many contributors to human variation as possible. As Cabo et al. (2012) note, it is important to check the linear relationship between new variables of sex (in this case variables of shape) and other variables of human variation, such as age, stature, and body mass, to determine the influences acting on the new shape variable. Fluctuating asymmetry can be added to this list of *other variables* to determine if it is confounding true sex differences.

1.2.3 Landmarks and 3D Models

Geometric morphometric data can be captured through a manual landmark selection process or a semi-automated semi-landmark process (Anastasiou & Chamberlain, 2013; Arnqvist & Martensson, 1998; Baab et al., 2012; Bigoni et al., 2010; Bookstein, 1978, 1991; Bytheway & Ross, 2010; Franklin et al., 2010; González et al., 2017; Gómez-Valdés et al., 2012; Hallgrimsson et al., 2008; Humphries et al., 1981; Mahato, 2010; Mitteroecker & Gunz, 2009; Pretorius et al., 2006; San-Millán et al., 2017; Sholts et al., 2011; Steyn et al., 2004; Velemínská et al., 2013; Zelditch et al., 2012). Points, or landmarks, are strategically positioned on an object to capture shape variation. These points are reproduced exactly on all subsequent specimens and registered as three-dimensional coordinates in tangent space. Positional differences among the

same landmark points in this tangent space provide the special information needed to measure shape differences. Landmarks and semi-landmarks highlight the unique features contributing to the shape of an object and highlight features of biological importance. Bookstein (1991) describes three types of landmarks used in geometric morphometrics: Type I landmarks are discrete juxtapositions, defined in terms of intersecting structures such as between sutures or the branching points of a tree. Type II landmarks are "maxima of curvature or other local morphogenetic processes" (Bookstein, 1991: 64). This category includes points that experience push or pull forces such as the tips of muscle attachments on bone, or ranges of a curve such as the "'corner' of the jaw and the 'corners' of the orbital rim" (Bookstein, 1991:65). Type II landmarks are defined in terms of local features, but they are not surrounded by structures as in type I. In contrast, Type III landmarks are points at extreme ends. Two other landmark types that are important but not as well-defined as the previous three are "constructed" and "fuzzy" landmarks. Constructed landmarks are a combination of traditional landmarks and geometric information while "fuzzy" landmarks are larger than single point landmarks but are still recognized as areas of biological significance (Bytheway & Ross, 2010). In addition to being biologically relevant to the research question, a landmark should be selected according to how easy it is to locate on the object by inexperienced users and how repeatable the expression of the landmark is on all the specimens being studied (Bytheway & Ross, 2010).

Landmarks that record maximum shape differences can omit subtle morphological variation that lies between the points (Bookstein, 1991). This problem can be rectified by applying semi-landmarks through a thin-plate spline (TPS). TPS allows semi-landmarks to capture the shape of curves between well-defined anatomical landmarks on three-dimensional surfaces (Bytheway & Ross, 2010; Mitteroecker & Gunz, 2009). Deformation grids also operate

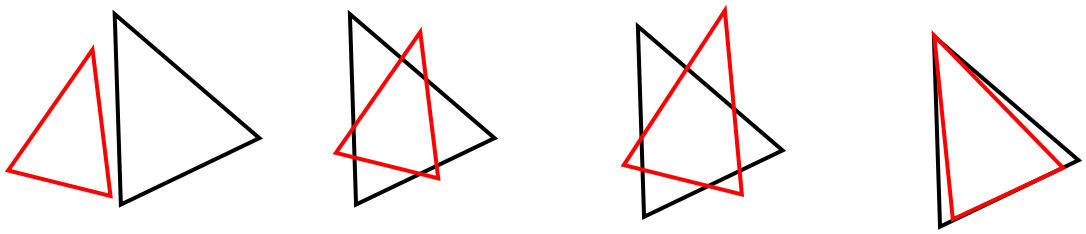
on the TPS and visually display the shape differences of objects. The shape differences located on the landmarks are displayed on the grid and morphed or “deformed” to indicate in which direction shape difference occurs (Mitteroecker & Gunz, 2009). Mitteroecker and Gunz, (2009: 240) suggest that deformation grids are not as effective for visualizing three-dimensional shape differences, but a sequence of warped surfaces can be used to describe shape differences in three dimensions.

This analysis used landmarks that have demonstrated biological significance in previous studies (Bilfeld et al., 2012; Bytheway & Ross, 2010; González et al., 2009) including landmarks that represent the anterior and posterior regions of the coxal bone, the auricular surface (Anastasiou & Chamberlain, 2013), the greater sciatic notch (Velemínská et al., 2013), and landmarks used in research previously undertaken (Robertson, 2013). There are two methods of obtaining landmark data for geometric morphometric analysis; one uses a 3D digitizer that captures landmark coordinates directly from the surface of the object being studied, while another way is to collect 3D surface scans of each specimen and postition the landmarks on these 3D computer models. Both methods have their own strengths and weaknesses. The 3D digitizer has been shown to reduce the measurement error of Type I and Type II landmarks while a 3D surface scanner more accurately measures Type III landmarks (Sholts et al., 2011). However, it is more difficult to duplicate specimens position when using a 3D digitizer compared to a 3D surface scanner (Arnqvist & Martensson, 1998; von Cramon-Taubadel et al., 2007). Hallgrimsson et al. (2008) recommend using a laser scanner to obtain 3D computer models of specimens because it allows for more flexibility when sharing data. Different landmarks can be applied to the same 3D model without having to verify compatibility between models or between scanning equipment. Computer models also provide a means to change the location of landmarks

when new sex-based shape stuides emerge. In this study, the NextEngine surface scanner was used to generate 3D models of coxal bone specimens. A more detailed description of the scanning procedure is provided in Chapter 2.

1.2.4 Geometric Morphometric Methods

The most widely used method of shape analysis is the Generalized Procrustes Analysis (Bookstein, 1991; Mitteroecker & Gunz, 2009; Zelditch et al., 2012). This method calculates the best fit among shapes using the principle of least squares. To calculate the best fit among all shapes, the Generalized Procrustes Analysis applies the Procrustes superimposition, in which the landmark configurations are superimposed onto one another and centred to a common centroid (Figure 1b), scaled to a unit centroid size (Figure 1c), and rotated to a common position (Figure 1d).



(a) (b) (c) (d)

Figure 1.7 Illustration of Procrustes Superimposition. Image (a) raw landmark configurations data. Images (b) centred at centroid. Images (c) centred and scaled. Images (d) centred, scaled and rotated.

To find the centre of the landmark configuration, after Procrustes superimposition, the centroid position coordinate (the average coordinates among all the landmark configurations) is

subtracted from the corresponding coordinates of each landmark configuration (Zelditch et al., 2012). This step is called translation because the centroid position has been translated from individual centres to a common centroid of all landmark configurations (Zelditch et al., 2012). The next step is to scale the newly translated landmark configurations either to a common size, a mean centroid size, or to an optimized least-squares estimate of scale (Mitteroecker & Gunz 2009). The centroid size is calculated as "the square root of the summed squared distances of each landmark from the centroid of the landmark configuration" (Zelditch et al., 2012:457). The sum of the squared distances between corresponding landmarks used to generate the unit centroid size is referred to as a squared Procrustes distance (Bookstein, 1991; Goodall, 1991; Mitteroecker & Gunz, 2009; von Cramon-Taubadel et al., 2007; Zelditch et al., 2012). The final step is to rotate the landmark configurations around their centroids, so all the corresponding landmarks are in their closest proximity, which minimizes the Procrustes distance between the coordinates (Mitteroecker & Gunz, 2009).

When the Procrustes superimposition calculates the best fit among the shapes, there is a risk that one or more landmarks, and their corresponding Procrustes coordinates, could express large shape differences. Large displacement can occur at one or more landmark location due to best fit among most remaining landmarks (Zelditch et al., 2012). This displacement of the landmark and corresponding Procrustes coordinate is referred to as the Pinocchio effect, and can cause a disproportionate amount of localized shape difference (von Cramon-Taubadel et al., 2007).

1.3 **Organization**

This dissertation is organized as three distinct papers focusing on three distinct areas of research. Chapter 2 has previously been published in the *American Journal of Physical Anthropology* and outlines a method for validating landmarks that best represent sex-based shape differences (Robertson et al., 2019). Numerous studies have been conducted to compare sex differences between Male and Female coxal bones using shape analysis with varying degrees of success and little to no transferability between methods. This study illuminates a comprehensive method for validating landmarks used to represent a specific research question, in this case, landmarks that represent true sex differences in the coxal bone and eliminates landmarks that obscure sex differences.

Chapter 3 used the sex-based landmarks identified in Chapter 2 to generate a discriminant function equation and logistic regression model that best predict sex in the whole coxal bone. Debates regarding which of the two methods is more appropriate to use in biological anthropology are ongoing (Albanese, 2003, 2008), consequently both equations were applied to evaluate their usefulness and applicability to the study population. The success of the sex prediction in Chapter 2 (98.5%) lies in use of landmarks that represent isometric and allometric size as well as fluctuating asymmetry. By examining sex-based shape differences from within the array of variation inherent in coxal bone form, sex predictions are improved and applicable within that array of variation. In other words, both the discriminant function (DF) equation and the logistic regression model (LRM) are applicable to both right and left coxal bones and to skeletons of various geographical origins. Consequently, these equations can be applied to either the right or left coxal bone and are population inclusive. The DF equation accurately identified

specimens in 99.7% of cases with an option for an intermediate sex classification. The LRM correctly identified all specimens with one less measurement than the DF equation and no option for intermediate sex. Both equations were verified against a test population (n = 120) varying in stature and body mass, again the LRM correctly identified all specimens and the DF equation correctly identified 99.2% of specimens. Landmarks associated with linear measurements were tested for intra- and inter-observation error. Fluctuating asymmetry was used as a benchmark for acceptable inter- and intra-observer error in this study.

Chapter 4 investigates the applicability of converting the population inclusive linear measurements presented in Chapter 3 to accommodate fragmented coxal bones. DF and LRM were also used in this chapter. The DF equation from Chapter 3 was repeated in Chapter 4 to represent coxal bones greater than 75% complete and to compare accuracy levels between new LRM and discriminant functions. Two more completeness categories represented coxal bones between 50-25% complete and approximately 25% complete. As in the previous chapter, the discriminant function equations generated to predict sex in fragmented coxal bones were subjected to a repeatability test on an independent population. Chapter 4 was not limited to linear measurements presented in Chapter 3 but included linear measurements associated with the other landmarks of sex-based shape introduced in Chapter 2.

The conclusion of this manuscript explores two arenas; the methodology of validating landmarks in future shape analyses in the skeleton, and decoupling ideas of biological vs skeletal sex.

Chapter 2: **Testing landmark redundancy for sex-based shape analysis**

2.1 **Synopsis**

Objectives: To test the individual effectiveness of common landmarks used in sex estimation of whole adult coxal bones in sex-based shape analysis and propose a method to determine how many principal components of sex-based shape to include for discriminant function analysis.

Methods: Three-dimensional models (NextEngine desktop laser scanner) of left and right os coxae from 394 individuals (William Bass Donated Skeletal Collection, Forensic Anthropology Centre, University of Tennessee, Knoxville, TN) were subjected to shape analysis using thirty-two landmarks (Landmark 3.6, Institute for Data Analysis and Visualization). Each landmark was individually removed and subjected to a new Principal Component Analysis (PCA) to identify the effect omitting a landmark has on PC1/PC2 ordination. Landmarks that poorly discriminated sex-based shape were considered redundant for analysis on sex estimation.

Results: This study identified 17 landmarks that represent sex-based shape of right and left coxal bones most effectively, these are: the anterior superior iliac spine; posterior superior iliac spine; posterior inferior iliac spine; iliac crest; apex of the auricular surface; greater sciatic notch; ischial spine; superior, inferior and distal points on ischial tuberosity; superior, inferior and midpoint on the symphyseal face; arcuate eminence; ischiopubic ramus; and posterosuperior and anterosuperior points on the acetabular rim. The first and second PCs of the 17-landmark configuration correctly predicted sex in 98.5% of cases; better than a 32-landmark configuration (96%) and better than previous landmark studies on whole coxal bone sex-based shape.

Conclusions: These 17 landmarks represent more meaningful data for sex-based shape analysis in PC1 and 2 and concentrate meaningful sex-based shape data to the first five PCs that make up over 50% of the total shape variance.

2.2 Introduction

In shape analysis, such as geometric morphometrics, individual landmarks adhere to the following criteria (Bookstein, 1991; Webster & Sheets, 2010): 1) be reproduced in the same relative position on every object of study; 2) summarize object morphology; 3) be replicated with high accuracy; and 4) compare similar shapes. The problem lies in a lack of consensus regarding how densely or thinly landmarks should summarize object morphology and the number of landmarks needed to address a research question. This is an obvious problem when comparing studies on whole coxal bone sexual dimorphism. Three studies have prepared the way for sex-based shape studies in complete human os coxa; Bytheway and Ross (2010) used 36 landmarks and predicted sex with 98% probability using 23 principal components of shape, Bilfeld et al. (2012) used 15 landmarks to predict sex with 87.62% probability reporting 13 principal components of shape, and González et al. (2017) used 28 landmarks to predict sex within 90.5% probability in females and 80% probability in males reporting 18 principal components of shape. However, the landmarks used in these analyses differ in number, predictive probability, and are non-specific to their research questions.

Landmarks in the whole coxal bone have not been systematically evaluated to test their individual effectiveness in sex-based shape analysis. Shape analysis is directed by the landmarks selected and the morphology of the objects of study. However, selected landmarks must also

coincide with the research question and be consistent between studies to ensure comparability of results between different study populations. Roth (1993) outlines three criteria for landmark selection based on previous work by Bookstein (1991). First, landmarks should be repeatable on all study specimens; this ensures data consistency. Adding to Bookstein and Roth's concepts of landmark repeatability is that landmarks should also be repeatable between studies to ensure method consistency and comparability of data. Naturally, different research questions might require a more detailed look at specific bone features, but in studies of sex-based shape of the whole human coxal bone consistent landmarks generate directly comparable data on sexual dimorphism between a variety of human populations.

The second criteria Roth (1993) outlines for selecting landmarks is they should represent features of biological importance. As stated previously, bone features considered to be biologically important will no doubt differ depending on the research question. In studies of sex-based shape in the os coxa, areas of biological importance are varied: pubic bone and greater sciatic notch have reported accuracy between 88-95% (Phenice, 1969; Singh & Potturi, 1978; Walker, 2005); The ilium, auricular surface, acetabulum, obturator foramen, and ischium are less accurate at predicting sex using osteometric analysis (80-84%) and are considered of secondary importance for estimating sexual dimorphism (Bierry et al., 2010; Brůžek, 2002; Macaluso, 2010; Rogers & Saunders, 1994). Sex-based shape analysis of these characteristics can predict sex-based shape between 87-98% (Anastasiou & Chamberlain, 2013; Bilfeld et al., 2012; Bytheway & Ross, 2010; González et al., 2017; Gómez-Valdés et al., 2012). Determining which areas of the coxal bone are of greater or lesser importance to sex-based shape analysis is a crucial step towards generating consistent and comparable data for future analyses.

A possible method to determine which landmarks could be made consistent for whole coxal bone sex-based shape analysis lie with Roth's (1993) third criteria for landmark selection: landmarks on key features; in this case, features of sexual dimorphism should generate meaningful data. In other words, it is important to investigate how to best represent sexually dimorphic bone features using landmarks considered to be meaningful for sex-based shape. According to Bookstein (1991), it is important to allow some landmark redundancy when representing shape, however, the level of landmark redundancy is not clear by either Bookstein or Roth or at what point redundant landmarks stop generating meaningful shape data. Studies using semi landmarks on the greater sciatic notch, ischiopubic ramus, and acetabulum isolated landmarks that consistently performed better at identifying meaningful sex-based shape data (Gómez-Valdés et al., 2012; González et al., 2009; San-Millán et al., 2017; Velemínská et al., 2013), which suggests some landmarks generate more meaningful shape data than others on sexually dimorphic features of both primary and secondary importance.

Shape analysis of cranial morphology is more prolific than coxal bone shape analysis because there are available standardized cranial landmarks that allow between-study shape comparisons on an ever-growing dataset of human, hominin, and non-human primate crania (Baab et al., 2012; Bigoni et al., 2010; Bookstein, 1991; Bruner, 2004; Franklin et al., 2010; González et al., 2011; Kimmerle et al., 2008; Manzi et al., 2000; Rosas & Bastir, 2002; Viðarsdóttir et al., 2002). The lack of standardized landmarks in the os coxa renders the wealth of data being generated on sex-based shape differences moot, as the results are not directly comparable between studies involving different populations or across time. This study seeks to determine which landmarks commonly used in sex-based shape analyses satisfy Roth's (1993) third criterion of landmark selection, that of generating meaningful shape data. It is central to

understand the degree to which coxal bone forms are related to allometric size differences and which are important for sex estimation. It is also vital to understand the degree to which sexual dimorphic forms should be represented by landmarks before overly redundant allometric data obscures the goal of the analysis. This study takes a qualitative and quantitative approach to interpreting meaningful sex-based shape data through comparisons of principal components (PCs). Additionally, this study evaluates how principal component analysis (PCA) of sex-based shape data are interpreted and reported for future studies.

2.3 **Materials and Methods**

2.3.1 Specimens

The specimens used in this study are from the William Bass Donated Skeletal Collection, Forensic Anthropology Centre, University of Tennessee, Knoxville, TN. The individuals (n = 394) died within the last 35 years and were between 23-93 years old at their time of death (Table 2.1). The individuals sampled in this study represent the demographics of the Bass collection (93% White, 3% Black, and 3% other), but do not reflect current North American or global population demographics. This analysis incorporated both left and right os coxae to capture variation in bilateral asymmetry, or individual variation, among the samples. In total, 788 coxal bones were analyzed (Male n = 406; Female n = 382).

	Age at death	White	Black	Other†	Total
Males	23-90 (mean = 58.8)	189	8	6	203
Females	24-93 (mean = 61.1)	180	4	7	191
Total		**369**	**12**	**13**	**394**

† = Hispanic (n = 4), Native American (n = 3), mixed White and Native American (n = 2), Polynesian (n = 1), mixed White Polynesian (n = 1), Filipino (n = 1), Japanese (n = 1)

Table 2.1 Demographics of study population.

2.3.2 Image Capture and Shape Analysis

The NextEngine desktop three-dimensional surface scanner (NextEngine, Inc.) was used to scan the specimens. Geometric point resolution was set to capture 310 points per cm^2 for objects in a wide field of view with an accuracy of up to 0.038cm (resolution published as 2.0k points per square inch with an accuracy of 0.015-inch). Each specimen required two 380° rotations at eight scans (10 minutes) per rotation to capture all levels of pelvic topography. The first rotation consisted of the hip bone positioned in anatomic position, parallel to the vertical arm of the PartGripper (Fig. 2.1a). The ischiopubic ramus was centred on the PartGripper platform and the ilium oriented so the anterior superior and posterior superior iliac tuberosities were within the borders of the turntable base. The second rotation consisted of the hip bone laying parallel to the base of the NextEngine turntable. The iliac fossa lay on the PartGripper platform parallel to the turntable base, the pubic bone pointing downward and the ischium pointing upwards (Fig. 2.1b). This orientation captured the superior surface of the iliac crest, inferior surface of the ischiopubic ramus, and the inside of the acetabulum that were missed in the first rotation. The second rotation required that the ischiopubic ramus and the apex of the ilium were inside the turntable boundaries or their surfaces would be missed by the scanner.

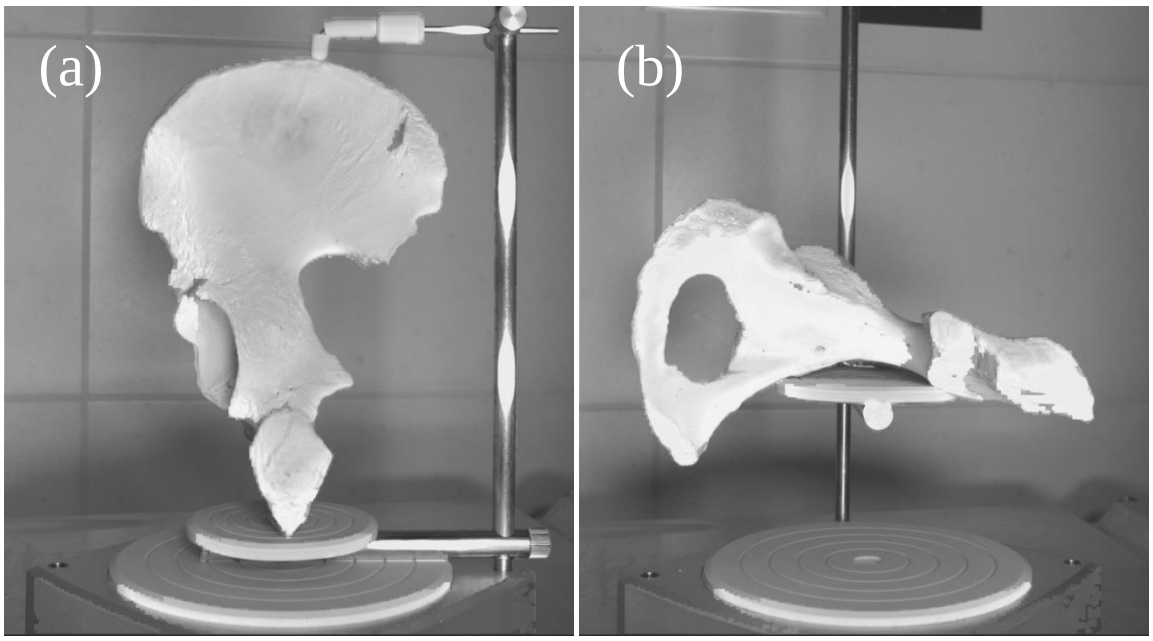


Figure 2.1 Position of the coxal bone to the arm of the NextEngine PartGripper; (a) parallel (b) perpendicular.

The scans were fused into a single three-dimensional image using ScanStudio software (ScanStudio HD 1.2.0, NextEngine, Inc). The fuse settings included no hole filling, to avoid artificially created coxal bone surfaces and terminus points, and a resolution ratio of 0.9 (default setting) to maintain the same mesh triangle size (0.0225 inches) as the original scans. Although a digitizer more accurately represents landmarks at sutures (type I landmarks) and at a maximum point on a curve (type II landmarks), such as the greater sciatic notch or the apex of the auricular surface, a 3D scanner is more accurate at representing type III landmarks (Sholts et al., 2011), extreme end points that are the furthest away from each other, such as between the maximum arch of the iliac crest and the most inferior point on the ischial tuberosity. Constructed landmarks, landmarks that are equidistant between two type III landmarks, are also more easily determined on 3D models using a measurement tool. This study used 3D models, as opposed to

two-dimensional photographs because coxal bone growth and development are three-dimensional. The associated measurements between end point landmarks will be useful to develop sex estimation methods.

The three-dimensional models were imported into the *Landmark* 3.6 software (Institute for Data Analysis and Visualization, 2007) as ply files for landmark placement using a semi-automated process. The process consisted of selecting a reference specimen and manually placing all landmarks on the object surface. Both right and left coxal bones used the same reference specimen for landmark semi-automation (Figure 2.2). For the remaining specimens, the first six landmarks in the sequence required manual placement to orient the semi-automation of the remaining landmarks. Each semi-automated landmark required manual refinement to ensure correct placement (Institute for Data Analysis and Visualization, 2007).

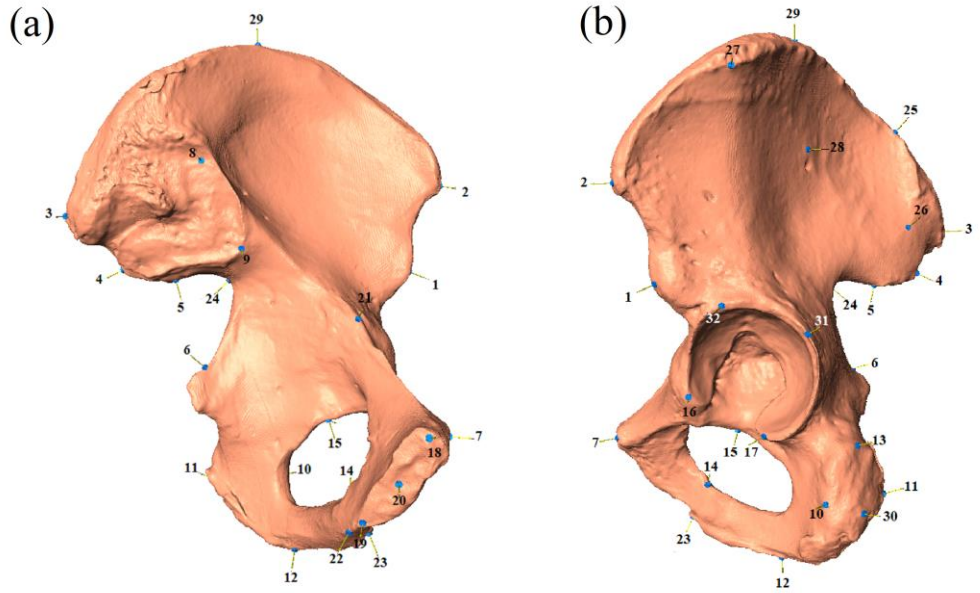


Figure 2.2 Landmark positions of reference specimen as seen in Landmark 3.6 Software; (a) ventral view (b) dorsal view.

2.3.3 Landmarks

Thirty-two landmarks (Table 2.2), common in a variety of morphometric studies and os coxal measurements (Bilfeld et al., 2012; Bytheway & Ross, 2010; Brůžek, 2002; González et al., 2017; Langley et al., 2016), were examined in this study. Landmarks 1-4, 6, 7, 9, 12, 14-22, 24, 26-28, 30-32 are derived from Bytheway and Ross's (2010) study. No significant sex-based shape differences were found in the acetabulum in Bytheway and Ross's study, so the number of landmarks in that area was reduced to points used in osteometric measurements of the acetabulum (Langley et al., 2016,). Landmarks 8 and 29 were included from González et al.,'s (2017) study. Landmark 29 is constructed from landmark 30 (most posterior point on the ischial tuberosity) to the point on the iliac crest that represents the maximum height of the os coxa. Landmarks from non-standard anthroposcopic sex traits included Brůžek's (2002) composite arch (8) and the phallic ridge (23). Landmark 5 was used as a terminus for greater sciatic notch shape and it, along with landmarks 10, 11, 13, and 25, are specific to this study.

Landmark Number	Landmark Description
1	Anterior inferior iliac spine
2	Anterior superior iliac spine
3	Posterior superior iliac spine
4	Posterior inferior iliac spine
5	Posterior inferior point of the preauricular sulcus
6	Ischial spine
7	Pubic tubercle
8	Superior point of the auricular surface
9	Apex of the auricular surface
10	Most lateral point on the ischial tuberosity
11	Most medial point on the ischial tuberosity

12	Most inferior point on the ischial tuberosity
13	Most superior point of the ischial tuberosity
14	The most inferior medial point on the obturator foramen rim
15	The most superior lateral point on the obturator foramen rim
16	The most anterior inferior point of the lunate surface
17	The most posterior inferior point of the lunate surface
18	The most superior point on the symphyseal face
19	The most inferior point on the symphyseal face
20	Midpoint of the pubic symphysis
21	Arcuate (iliopubic) eminence
22	Narrowest point on the ischiopubic ramus inferior to the pubic symphysis
23	Most lateral point on the phallic ridge on the ischiopubic ramus
24	Maximum arch of the greater sciatic notch
25	Point of intersection of the posterior gluteal line to the iliac crest
26	Midpoint on the posterior gluteal line
27	Iliac tubercle
28	Midpoint on the anterior gluteal line
29	Most superior point on the iliac crest
30	Most posterior point on the ischial tuberosity
31	Most posterior superior point on the acetabular rim
32	Most anterior superior point on the acetabular rim

Table 2.2 Landmark description.

2.3.4 Digitizing and Measurement Error

Digitization of the specimens required two NextEngine laser scanners to operate simultaneously. One machine was dedicated to scanning left os coxae while the other scanned only right os coxa specimens in the entire study. A precision study between scanners was required to measure differences in digitization. The precision study used 12 specimens selected at random (Males n = 7, Females n = 5). Left and right coxal bones were digitized once using both scanners. A MANOVA (IBM SPSS Statistics v.25.0, SPSS Inc. 2017) examined Procrustes

coordinates of each landmark as the dependent variables, scanner and side as independent variables. The results did not find significant digitizing error (F (46, 1) = 10.51, p = 0.241, Wilks' Λ = 0.002) between the two scanners or bilateral asymmetry (F (46, 1) = 2.064, p = 0.510, Wilks' Λ = 0.010) between the right and left os coxae.

2.3.5 Statistical Analysis

Principal Component Analysis (PCA) was used to assess each of the 32 landmarks for meaningful sex-related data. Systematic individual removal of each landmark followed by a new PCA revealed the impact each landmark made to the spread and separation of Male and Female individuals on a scatter plot comparing PC1 and 2. Left and right coxal bones were evaluated together. Landmarks To the author's knowledge, this method of evaluating landmark effectiveness had not been used before. However, because PCA results are directed by correlated variables, it seemed a logical method to evaluate landmark effectiveness within this data set. Landmark assessment was visually assessed for spread and separation of individuals on a PC1/PC2 scatter plot. Landmarks that when removed improved the overall separation between Males and Females on the new PC1/PC2 scatter plot, from the original 32-landmark PC1/PC2 scatterplot, were eliminated from the 32-landmark configuration. The PCs from the remaining landmark configuration were subjected to discriminant function analysis (DFA) and cross-validation to compare sex-based shape prediction of the new configuration from the original 32-landmark configuration.

It is important, when confronted with many PCs of shape, to reduce the dimensionality of shape data to a manageable level. This study will determine which PCs contain not only shape

data that is biologically important to interpreting sex but also shape data that is statistically meaningful. There are three ways to determine PCs of biological importance (Zelditch et al., 2012); the first is to consider all PCs that account for at least 80% of shape variance, which would require the evaluation of over 20 PCs, may of which would contain shape variance of less than 3% and is not likely to garner new information independent of PCs containing greater amounts of shape variance. The second method is to interpret the scree plot of PC variance - or the cumulated percentage of variance by PC - and limit interpretation of biological importance to those PCs to the left of a clear drop of in shape variance, called the inflection point. This method would not be useful, however, if there is no clear inflection point in the scree plot. A third method for evaluating biological importance is to consider all PCs containing over 5% of shape variance. This third method was deemed the most useful to this study as it provided a more manageable number of PCs to interpret that the first method of determining biological importance and a clearer cut off point than the second method.

Once PC of biological importance have been ascertained, the next step is to determine which of these PCs to use in statistical tests. Not all PCs of biological importance are statistically relevant to answer the specific question of sex in the coxal bone. Coxal bone sex-based morphology is conflated with body form (size and shape). The broken stick model is one way to determine which PCs of biological importance contain statistically meaningful data to answer this study's specific question of sex difference (Jackson, 1993). This final step is vital when there is no discernable inflection point on the scree plot between PCs. The broken stick model divides the total variance (eigenvalue sum from each PC) by the eigenvalue for each PC. If the value from the broken stick model exceeds the variance for a given PC, that PC is considered not statistically meaningful. Values from the broken stick model that do not exceed the variance for a

given PC can be retained for further statistical testing. Jackson (1993) reports that the broken stick model can underestimate the number of statistically meaningful PCs, which is why this study will report PCs of biologically importance along side PCs that meet the criteria of the broken stick model.

PCA and shape analysis were conducted using *MorphoJ* software, which uses a full Procrustes fit based on principle axes and a covariance matrix (Klingenberg, 2011). DFA was conducted using SPSS (IBM SPSS v25). Centroid size was subjected to a Spearman's rho to test its association with PCs and sex.

2.4 **Results**

In a Procrustes ANOVA (MANCOVA) using *MorphoJ* (Klingenberg, 2011), sex-based shape differences were not significantly impacted by bilateral asymmetry (F(df=89) = 0.58, p (*parim.*) = 0.999; Pillai's trace = 0.12, p (*parim.*) = 0.43), nor by centroid size (F(df=1) = 47558.60, p *(parim)* = 0.0029) based on a randomized test of 10,000 permutation at an alpha level of 0.001 (default *MorphoJ* alpha level).

2.4.1 **Principal Component Analysis**

The original 32-landmark configuration generated 89 principal components (PCs) and is presented in Table 2.3. The first two PCs meet the criteria of the broken stick model for inclusion in subsequent statistical tests (Jackson, 1993) and the first 4 PCs contain biologically meaningful shape data. The results of the broken stick model will only be included for the first 45 PCs. The

first 38 PCs capture 90% of the total variance and 50% of the total variance is described in the first seven PCs.

PC	Eigenvalues	Var.%*	Cum.% +	Broken Stick Model	PC	Eigenvalues	Var.% *	Cum.% +
1.	**0.00105132**	**14.554**	**14.554**	**6.870762**	46.	0.00002445	0.338	93.513
2.	**0.00090448**	**12.522**	**27.076**	**7.986213**	47.	0.00002383	0.330	93.843
3.	**0.00058142**	**8.049**	**35.125**	12.42367	48.	0.00002369	0.328	94.171
4.	**0.00040613**	**5.622**	**40.748**	17.78586	49.	0.00002218	0.307	94.478
5.	0.00034339	4.754	45.502	21.03547	50.	0.00002135	0.296	94.773
6.	0.00029731	4.116	49.617	24.29575	51.	0.00002018	0.279	95.053
7.	0.00022899	3.170	52.788	31.54448	52.	0.00001982	0.274	95.327
8.	0.00021557	2.984	55.772	33.50823	53.	0.00001910	0.264	95.592
9.	0.00019030	2.635	58.407	37.95780	54.	0.00001873	0.259	95.851
10.	0.00018027	2.496	60.902	40.06973	55.	0.00001738	0.241	96.092
11.	0.00016507	2.285	63.188	43.75944	56.	0.00001720	0.238	96.330
12.	0.00014810	2.050	65.238	48.77360	57.	0.00001596	0.221	96.551
13.	0.00013568	1.878	67.116	53.23828	58.	0.00001552	0.215	96.765
14.	0.00012696	1.758	68.874	56.89485	59.	0.00001458	0.202	96.967
15.	0.00011737	1.625	70.499	61.54358	60.	0.00001418	0.196	97.164
16.	0.00010840	1.501	71.999	66.63625	61.	0.00001364	0.189	97.352
17.	0.00009897	1.370	73.370	72.98545	62.	0.00001262	0.175	97.527
18.	0.00009276	1.284	74.654	77.87160	63.	0.00001225	0.170	97.697
19.	0.00009181	1.271	75.925	78.67738	64.	0.00001150	0.159	97.856
20.	0.00007968	1.103	77.028	90.65474	65.	0.00001133	0.157	98.013
21.	0.00007701	1.066	78.094	93.79782	66.	0.00001066	0.148	98.160
22.	0.00007022	0.972	79.066	102.86770	67.	0.00001047	0.145	98.305
23.	0.00006822	0.944	80.011	105.88347	68.	0.00000993	0.137	98.443
24.	0.00006654	0.921	80.932	108.55681	69.	0.00000970	0.134	98.577
25.	0.00006368	0.882	81.813	113.43232	70.	0.00000905	0.125	98.702
26.	0.00005895	0.816	82.629	122.53384	71.	0.00000858	0.119	98.821
27.	0.00005678	0.786	83.415	127.21680	72.	0.00000848	0.117	98.938
28.	0.00005429	0.752	84.167	133.05157	73.	0.00000785	0.109	99.047
29.	0.00005151	0.713	84.880	140.23238	74.	0.00000741	0.103	99.150
30.	0.00004897	0.678	85.558	147.50602	75.	0.00000734	0.102	99.251

31.	0.00004818	0.667	86.225	149.92466	76.	0.00000693	0.096	99.347
32.	0.00004561	0.631	86.856	158.37251	77.	0.00000647	0.090	99.437
33.	0.00004517	0.625	87.482	159.91521	78.	0.00000609	0.084	99.521
34.	0.00004365	0.604	88.086	165.48385	79.	0.00000526	0.073	99.594
35.	0.00004160	0.576	88.662	173.63870	80.	0.00000500	0.069	99.663
36.	0.00003994	0.553	89.215	180.85553	81.	0.00000466	0.064	99.728
37.	0.00003782	0.524	89.739	190.99339	82.	0.00000390	0.054	99.782
38.	0.00003601	0.499	90.237	200.59345	83.	0.00000348	0.048	99.830
39.	0.00003353	0.464	90.701	215.43006	84.	0.00000316	0.044	99.873
40.	0.00003234	0.448	91.149	223.35714	85.	0.00000290	0.040	99.914
41.	0.00003121	0.432	91.581	231.44409	86.	0.00000228	0.032	99.945
42.	0.00003028	0.419	92.000	238.55251	87.	0.00000185	0.026	99.971
43.	0.00002898	0.401	92.401	249.25362	88.	0.00000165	0.023	99.994
44.	0.00002841	0.393	92.795	254.25449	89.	0.00000045	0.006	100.000
45.	0.00002744	0.380	93.175	263.24235			Total variance:	0.00722337

* Variance
\+ Cumulative Variance

Table 2.3 PCA of the L32 configuration.

Plotting PCs 1 and 2 against each other (Figure 2.3) reveals a separation of individuals by sex along PC1 with overlapping confidence ellipses between -0.02 and 0.02. The corresponding wireframe graphs depicting shape changes along PC1 illustrates expected Female coxal bone shape with a wide greater sciatic notch, short and wide ilium, long pubic bone, short symphyseal face, and smaller ischium. Mean coxal shape between Males and Females in the sample population is represented by light grey lines and is the same configuration for both PC1 and PC2 means at the 0.0 point. The negative end of PC1 represents expected Male coxal bone shape being taller and narrower overall, with a narrower greater sciatic notch, shorter pubic bone, and longer symphyseal face. Other notable sex-based shape changes are a wider Male and narrower

Female acetabulum, a more anterior Female and posterior Male auricular surface, and a smaller Female and larger Male ischium.

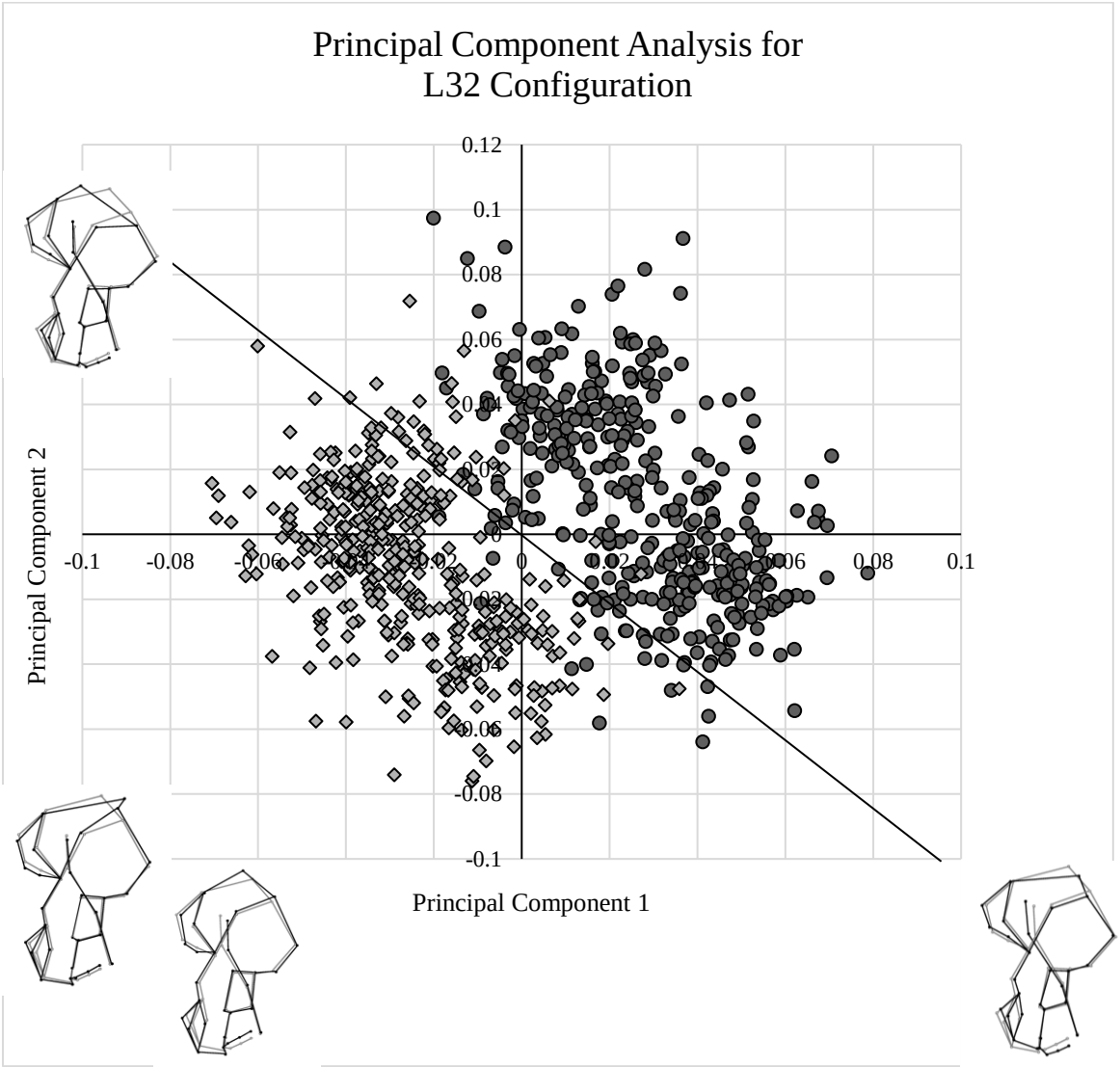


Figure 2.3 Scatterplot of PC1 against PC2 for the 32-landmark configuration. Dark circles are Female and light diamonds are Male. The corresponding wireframe graphs for PC1 and PC2; dark lines represent coxal bone shape at the respective -0.1 or 0.1 scale, light lines represent the average coxal bone shape (at scale 0.0). Linear equation *y* = -1*x*.

The second PC also illustrates expected coxal bone sex-based shape differences at different magnitudes compared to PC1. In the greater sciatic notch, PC2 captures more extreme Female and Male shapes compared to PC1, whereas sex-based pubis length and ischium size are less variable along PC2. A striking difference along PC2 is the position of landmark 29, representing the maximum arch of the iliac crest, compared to its position on the PC1 axis. Along the PC1 axis, landmark 29 is more anteriorly positioned at the positive end of the scale and more posteriorly positioned at the negative end. This pattern is reversed along PC2, a more posterior landmark represents the positive end of the axis, and landmark 29 is more anterior at the negative end of PC2.

Each landmark was then systematically removed from the 32-landmark configuration and a new PC scatter plot was generated to observe the effect each eliminated landmark had on the covariance matrix through visualization of the new distribution of individuals. If a landmark, when removed, increased the spread and separation of individuals on the PC1/PC2 scatterplot, that landmark was flagged as a poor contributor to sex-based shape analysis in this population. If the distribution of data narrowed and generated more overlap between Males and Females when a landmark was removed, that landmark was considered important for sex-based shape analysis. If there was no change to the distribution of individuals when a landmark was removed, that landmark could either be retained or eliminated based on whether it strengthened the predictive power of the landmark configuration and its usefulness for cross-comparison with other coxal bone studies. Accompanying the qualitative description of specimen movement in ordination during systematic removal of landmarks, presented in Table 2.4 are the associated values for Male variance, Female variance, and total variance pooled between Male and Female individuals.

Landmark Removed	Description Distribution	Total Fem.Var.×	Total Male.Var.§	Total Variance	Diff. from Original
0	**Original Configuration**	**0.00676**	**0.00607**	**0.00722**	0.0
1	No change in data distribution	0.00660	0.00602	0.00713	0.00009
2	Plot is inverse of original on PC2	0.00660	0.00591	0.00710	0.00012
3	Wider spread of data on PC1	0.00682	0.00609	0.00730	-0.00008
4	Less discrimination between sex-based shape	0.00671	0.00605	0.00713	0.00009
5	Less discrimination between sex-based shape	0.00668	0.00600	0.00712	0.00010
6	Less discrimination between sex-based shape	0.00671	0.00601	0.00715	0.00007
7	Wider spread of data on PC1	0.00677	0.00614	0.00730	-0.00008
8	More discrimination between sex-based shape	0.00682	0.00606	0.00721	0.00001
9	Fewer Females present in overlap	0.00659	0.00600	0.00704	0.00018
10	More discrimination between sex-based shape	0.00676	0.00606	0.00723	-0.00001
11	Wider spread of data on PC1	0.00677	0.00604	0.00720	0.00002
12	Less discrimination between sex-based shape	0.00684	0.00610	0.00726	-0.00004
13	No change in data distribution	0.00670	0.00600	0.00715	0.00007
14	No change in data distribution	0.00686	0.00614	0.00733	-0.00011
15	No change in data distribution	0.00675	0.00605	0.00721	0.00001
16	No change in data distribution	0.00673	0.00601	0.00718	0.00004
17	No change in data distribution	0.00673	0.00601	0.00719	0.00003
18	Less discrimination between sex-based shape	0.00695	0.00619	0.00736	-0.00014
19	Less discrimination between sex-based shape	0.00696	0.00622	0.00736	-0.00014
20	Less discrimination between sex-based shape	0.00698	0.00623	0.00738	-0.00016
21	Less discrimination between sex-based shape	0.00663	0.00594	0.00741	-0.00019
22	Wider spread of data on PC1	0.00692	0.00623	0.00741	-0.00019

23	Less discrimination between sex-based shape	0.00690	0.00620	0.00737	-0.00015
24	Less discrimination between sex-based shape	0.00671	0.00603	0.00717	0.00005
25	Wider spread of data on PC1	0.00696	0.00619	0.00744	-0.00022
26	No change in data distribution	0.00680	0.00610	0.00728	-0.00006
27	More discrimination between sex-based shape	0.00653	0.00578	0.00698	0.00024
28	No change in data distribution	0.00676	0.00610	0.00726	-0.00004
29	More condensed spread of data on PC2	0.00609	0.00579	0.00676	0.00046
30	No change in data distribution	0.00683	0.00610	0.00726	-0.00004
31	No change in data distribution	0.00667	0.00595	0.00712	0.00010
32	No change in data distribution	0.00671	0.00601	0.00717	0.00005

× Total Female Variance
§ Total Male Variance

Table 2.4 Description of changes in PCA distributions between PC1 and 2, results of changes in variance between Females and Males, and overall variance with systematic removal of each individual landmark. Variance values from the original L32 configuration.

When removed, landmarks 1, 13-17, 26, 28, 30-32 do not change the distribution of specimens on either the first or second PC axes; landmarks 2 and 29 change the distribution of shape data on PC2; and landmarks 4-6, 12, 18-21, 23, 24 discriminate poorly between Male and Female specimens, which suggests they are important for maintaining sex-based shape distinction. Landmarks that dispersed specimens along PC1 when removed (3, 7, 11, 22, and 25) contribute statistically significant changes to patterns of total Female variance (mean variance = -0.0001; $p = 0.450$), total Male variance (mean variance = > -0.0001; $p = 0.419$), and total pooled variance (mean variance = -0.0001; $p = 0.182$) compared to landmarks that did not change the distributions of specimens. When compared to the variance of the original 32-landmark configuration, these landmarks showed statistical significance (mean variance = 0.0001; p =

0.182) over landmarks that did not change the distribution of specimens. Therefore, these landmarks could be good candidates for removal. Landmarks that generated more discrimination between sex-based shape among specimens when removed (8, 10, and 27) show statistically significant patterns of total Male variance (mean variance = 0.0001; p = 0.241) and pooled variance (mean variance = 0.0001; p = 0.333) compared to landmarks that show less discrimination between Males and Females. There was also a significant difference between landmarks that show less sex discrimination and those that show more sex discrimination (mean variance = - 0.0001; p = 0.333) when compared to the variance of the 32-landmark formation and should also be considered for removal.

Common landmarks used in linear and geometric morphometric measurements are largely maintained for compatibility with other studies. These include: landmarks 2 and 3 measuring ilium breadth; 8, 9, 6, and 24 representing the composite arch; 16, 17, 31, 32 representing transverse and vertical acetabular diameters; landmarks 18 to 16 measuring minimum pubis length; landmarks 18 to 31 measuring maximum pubis length; 19 and 30 measuring maximum ischiopubic ramus length; and 12 and 29 measuring maximum coxal bone height (Arsuaga & Carretero, 1994; Brůžek, 2002; Buikstra & Ubelaker, 1994; Langley et al., 2016). Landmarks that did not alter the distribution of individuals on the PC1/PC2 scatter plot could be removed to improve the percentage of variance captured by PCA; these include landmarks 14-17 and 28. Other landmarks to eliminate are 25, to increase the complexity of sex-based shape data, and 27, to increase discrimination between Male and Female shapes, on PC1.

The landmarks considered redundant for sex-based shape analysis are: the anterior inferior iliac spine, the posterior inferior point of the preauricular sulcus, the pubic tubercle, the superior point of the auricular surface, the most lateral and medial points on the ischial

tuberosity, the most inferior and superior points on the obturator foramen rim, the most anterior and posterior points on the lunate surface, the most lateral point on the phallic ridge on the ischiopubic ramus, the iliac tubercle, and all of the landmarks representing the gluteal muscles (25, 26, and 28). The remainders form a configuration of 17 landmarks that uses many common points between studies of sex-based shape (Table 2.5), osteometric measurements (Langley et al., 2016), and that display the best possible distribution of Male and Female individuals along both PC1 and PC2 in a PCA (Figure 2.4).

Num.	Description	Source
1	The apex of the anterior superior iliac spine	Bilfeld et al., 2012; Bytheway & Ross, 2010; González et al., 2017
2	The apex of the posterior superior iliac spine	Bilfeld et al., 2012; Bytheway & Ross, 2010; González et al., 2017
3	The apex of the posterior inferior iliac spine	Bilfeld et al., 2012; Bytheway & Ross, 2010; González et al., 2017
4	The apex of the ischial spine	Bilfeld et al., 2012; Bytheway & Ross, 2010; González et al., 2017
5	Apex of the auricular surface	Bytheway & Ross, 2010; González et al., 2017
6	Most inferior point on the ischial tuberosity	Bilfeld et al., 2012; Bytheway & Ross, 2010; González et al., 2017
7	Most superior point on the ischial tuberosity	Bytheway & Ross, 2010
8	Most superior point on the symphyseal face	Bilfeld et al., 2012; Bytheway & Ross, 2010; González et al., 2017
9	Most inferior point on the symphyseal face	Bilfeld et al., 2012; Bytheway & Ross, 2010; González et al., 2017
10	Midpoint of the pubic symphyseal face (constructed landmark calculated between landmarks 8 and 9)	Bytheway & Ross, 2010; González et al., 2017
11	The apex of the arcuate (iliopubic) eminence	Bytheway & Ross, 2010; González et al., 2017
12	Ischiopubic ramus at the narrowest point inferior to the pubic symphysis	Brůžek, 2002

13	Point of maximum curvature in the greater sciatic notch	Bilfeld et al., 2012; Bytheway & Ross, 2010; González et al., 2017
14	Maximum arc of the iliac crest	Bilfeld et al., 2012; González et al., 2017; Langley et al., 2016
15	Most distal point on the ischial tuberosity	Bytheway & Ross, 2010; Langley et al., 2016
16	Most posterior point on the superior portion of the acetabular rim	Bilfeld et al., 2012; Langley et al., 2016
17	Most anterior point on the superior portion of the acetabular rim	Bytheway & Ross, 2010; Langley et al., 2016

Table 2.5 Description of 17 landmarks that best distinguish sex-based shape based on systematic removal and their appearance in previous studies.

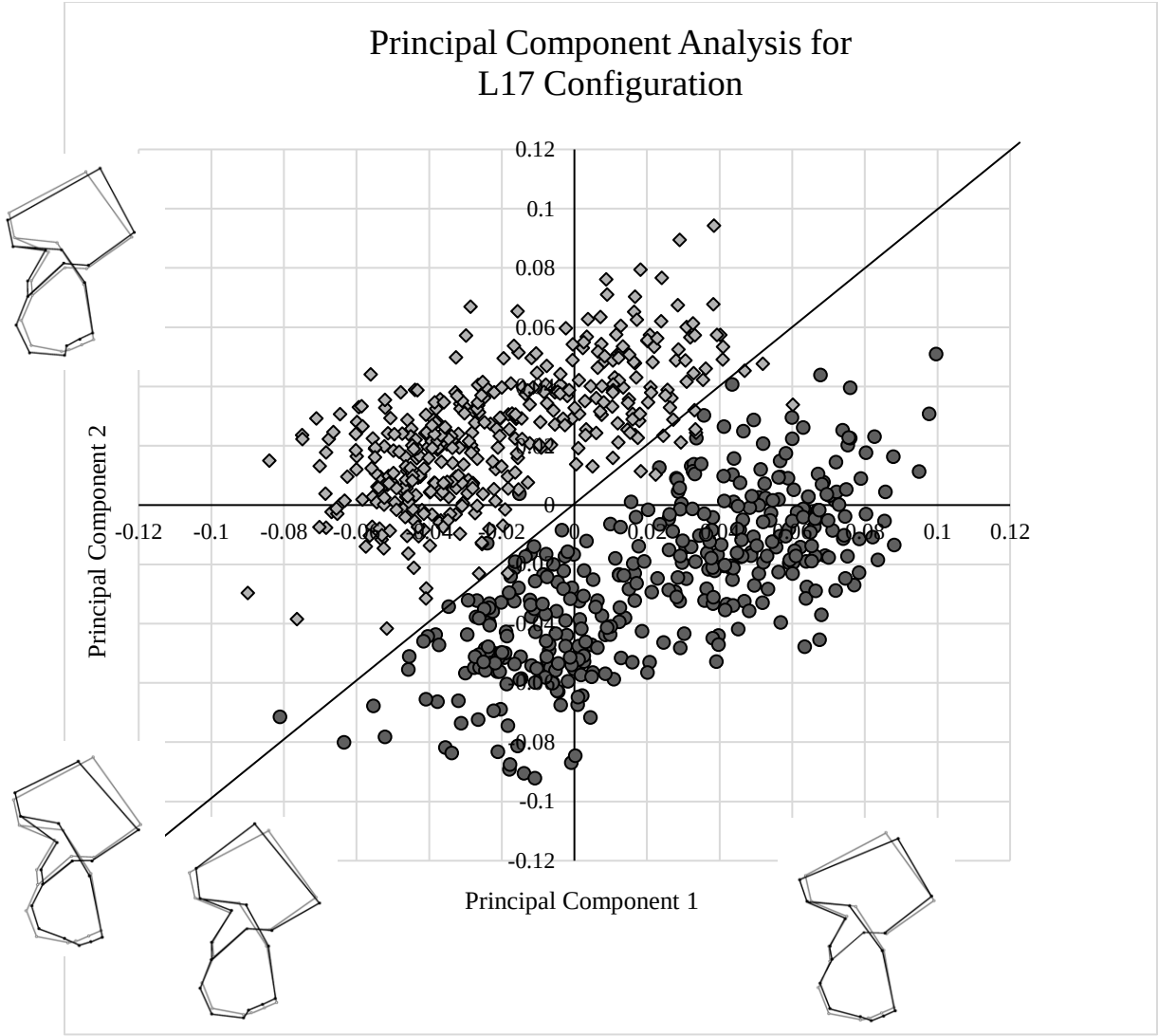


Figure 2.4 Scatterplot of PC1 against PC2 for the 17-landmark configuration. Dark circles are Female and light diamonds are Male. The corresponding wireframe graphs for PC1 and PC2; dark lines represent coxal bone shape at the respective -0.1 or 0.1 scale, light lines represent the average coxal bone shape (at scale 0.0). Linear equation *y* = 1*x*.

Like the 32-landmark configuration, the first two PCs of the remaining 17 landmarks met the broken stick model criteria for inclusion in further statistical tests and make up 37% of shape variance. However, 5 PCs contain biologically meaningful shape data in the 17-landmark configuration, one more than the 32-landmark configuration. There are fewer PCs (n=44) because there are fewer landmarks. Twenty-two PC report the results of the broken stick model,

21 PCs describe 90% of the variance, and 50% of the variance is described in the first 4 PCs (Table 2.6).

PC	Eigenvalues	Var.%*	Cum.%+	Broken Stick Model	PC	Eigenvalues	Var.%*	Cum.%+
1.	**0.00153911**	**21.004**	**21.004**	**4.761044**	23.	0.00006489	0.886	92.186
2.	**0.00119944**	**16.368**	**37.372**	**6.109326**	24.	0.00005909	0.806	92.993
3.	**0.00052757**	**7.200**	**44.572**	13.88966	25.	0.00005801	0.792	93.785
4.	**0.00050616**	**6.907**	**51.479**	14.47718	26.	0.00005349	0.730	94.515
5.	**0.00037270**	**5.086**	**56.565**	19.66131	27.	0.00004779	0.652	95.167
6.	0.00032359	4.416	60.981	22.64523	28.	0.00004363	0.595	95.762
7.	0.00029779	4.064	65.045	24.60717	29.	0.00004226	0.577	96.339
8.	0.00024768	3.380	68.425	29.58563	30.	0.00003701	0.505	96.844
9.	0.00021853	2.982	71.407	33.53210	31.	0.00003454	0.471	97.315
10.	0.00018730	2.556	73.963	39.12317	32.	0.00002779	0.379	97.694
11.	0.00015705	2.143	76.107	46.65883	33.	0.00002461	0.336	98.030
12.	0.00014339	1.957	78.063	51.10377	34.	0.00002391	0.326	98.357
13.	0.00013221	1.804	79.868	55.42523	35.	0.00002214	0.302	98.659
14.	0.00012596	1.719	81.586	58.17537	36.	0.00001914	0.261	98.920
15.	0.00011647	1.589	83.176	62.91551	37.	0.00001843	0.251	99.171
16.	0.00010994	1.500	84.676	66.65245	38.	0.00001485	0.203	99.374
17.	0.00009871	1.347	86.023	74.23534	39.	0.00001285	0.175	99.549
18.	0.00008855	1.208	87.232	82.75291	40.	0.00001154	0.158	99.707
19.	0.00008162	1.114	88.346	89.77910	41.	0.00000927	0.126	99.833
20.	0.00007517	1.026	89.371	97.48260	42.	0.00000744	0.102	99.935
21.	0.00007205	0.983	90.355	101.7040	43.	0.00000385	0.053	99.987
22.	0.00006934	0.946	91.301	105.6788	44.	0.00000093	0.013	100.00
							Total variance:	0.00732777

* Variance
\+ Cumulative Variance

Table 2.6 Principal Component Analysis of L17 configuration.

Comparing the Procrustes coordinate data between the two landmark configurations, Goodall's *F*-ratio of sex-based shape variation from unexplained variation indicates a greater deviation of *a priori* sex groups from mean Procrustes distances (shape data) in the 17-landmark configuration (F(df = 44) = 276.06, p = <0.0001) compared to the 32-landmark configuration (F(df = 89) = 178.57, p = <0.0001). The 17-landmark configuration had a smaller but significant Mahalanobis distance value (D = 6.379, p <0.0001) suggesting there is less sample variation than the 32-landmark configuration (D = 7.042, p <0.0001), where the distance between the landmark points and the distribution of points is slightly greater. The Mahalanobis distance (represented as D in *MorphoJ*) is scaled relative to sample variation and is inverse from the variance-covariance matrix (Marcus, 1993). Sex was not significant for bilateral asymmetry (F(df = 44) = 0.51, p *(parim)* = 0.9971; Pillai's trace = 0.04, p *(parim)* = 0.9174) in a Procrustes ANOVA (MANCOVA).

2.4.2 Discriminant Function Analysis

Discriminant function analysis (DFA) predicted known sex of both the 32 and 17-landmark configurations based on PCs 1 and 2 (Table 2.7). Figure 2.5 illustrates the average Male and Female coxal bone shape among the sample collection for the 32-landmark configuration. Sex was correctly predicted in 96.3% of cases in a cross-validated DFA of the 32-landmark configuration, and in 98.5% of cases for the 17-landmark configuration. Figure 2.6 illustrates the average Male and Female shape for the 17-landmark configuration. Group prediction improved with the 17-landmark configuration due to more Females correctly classified compared to the 32-landmark configuration.

L32 configuration				
	Sex	**Predicted Group Membership**		
		Female	**Male**	**Total**
Original	Female	97.6% (n = 373)	2.4% (n = 9)	100.0% (n = 382)
	Male	2.2% (n = 9)	97.8% (n = 397)	100.0% (n = 406)
Cross-validated	Female	97.6% (n = 373)	2.4% (n = 9)	100.0% (n = 382)
	Male	2.5% (n = 10)	97.5% (n = 396)	100.0% (n = 406)

L17 configuration				
	Sex	**Predicted Group Membership**		
		Female	**Male**	**Total**
Original	Female	99.2% (n = 379)	0.8% (n = 3)	100.0% (n = 382)
	Male	2.2% (n = 9)	97.8% (n = 397)	100.0% (n = 406)
Cross-validated	Female	99.2% (n = 379)	0.8% (n = 3)	100.0% (n = 382)
	Male	2.2% (n = 9)	97.8% (n = 397)	100.0% (n = 406)

Table 2.7 Accuracy of *a priori* sex predictions for L32 and L17 configurations.

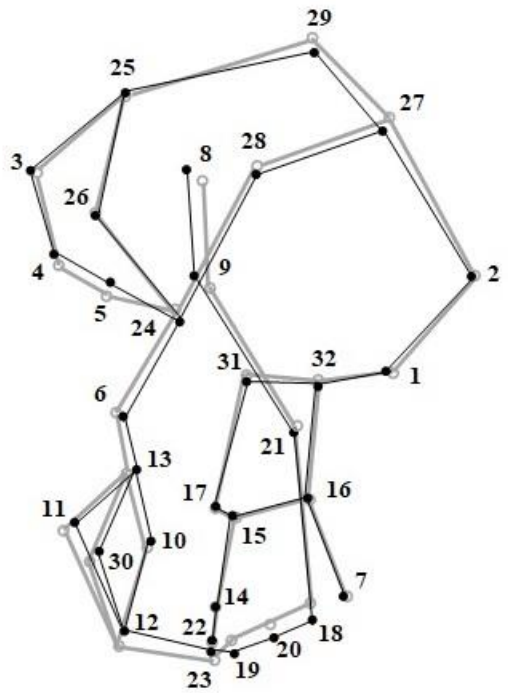


Figure 2.4 Sex-based shape as wireframe graph of L32 configuration. Typical Female in black and typical Male in grey.

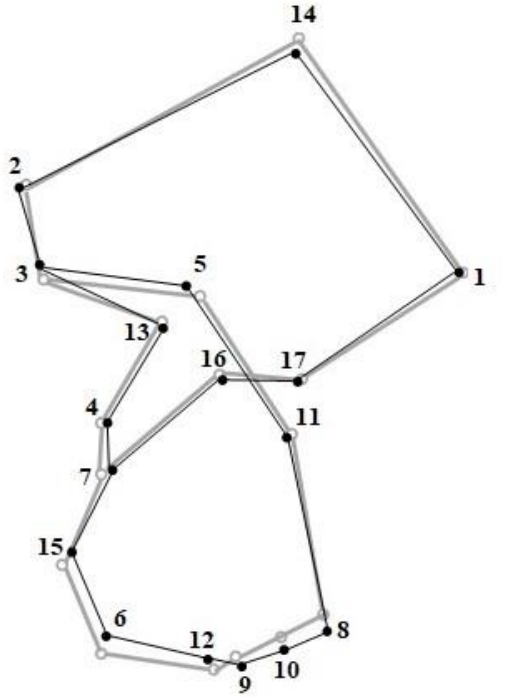


Figure 2.5 Sex-based shape as wireframe graph of L17 configuration. Typical Female in black and typical Male in grey.

2.5 **Discussion**

This study evaluated type II and type III landmarks on digital models, similar to studies conducted by Bilfeld et al. (2012), Bytheway and Ross (2010), and González et al. (2017). Bytheway and Ross identified the least effective areas in the coxal bone for sex-based shape analysis as the acetabulum – represented by eight landmarks – and obturator foramen – represented by five landmarks. The current study identified 12 landmarks, considered redundant in the 32-landmark dataset, present in Bytheway and Ross's landmark configuration: the anterior inferior iliac spine, the pubic tubercle, lateral and medial points on the ischial tuberosity, inferior and superior points on the obturator foramen, anterior and posterior points on the lunate surface, the iliac tubercle, and three points representing the gluteal lines. Although Bytheway and Ross's landmark configuration correctly predicted sex 98% of the time they used the first 23 PC of shape, the current study predicted sex 98.5% of the time using the first two PCs of shape. This suggests that Bytheway and Ross's landmarks are capturing allometric shape differences as well as sex-based shape differences and more PCs of shape are required to distinguish sex-based shape in a meaningful way.

Zelditch et al. (2012) and Jackson (1993) outline two important rules for determining which PCs represent shape data of biological importance and which PCs contain statistically meaningful data respectively. The first four PCs in the current study's 32-landmark configuration represented PCs of biological importance, however, biological importance was captured in the first five PCs of the 17-landmark configuration. Fewer landmaks seem to suggest biologicaly importat shape data must be captured by more PCs. This pattern is consistent when measuring the amount of biologically meaningful coxal bone shape data captured by other studies. Bilfeld et al.(2012)'s 15 landmarks captured biologcally meaningful shape in the first six PCs. However,

González et al.(2017) also captured biologically meaningful shape data in the first six PCs with 28 landmarks. Biologically meaningful data is captured not only by the number of landmarks, but by the quality of the sex-based shape data represented by those landmarks. The ilium is the common feature contributing to shape variance between the Bilfeld and González studies. Both studies represent ilium shape with landmarks at the anterior inferior iliac spine – a redundant landmark in the current study – the anterior superior iliac spine, the most lateral/salient point on the iliac crest – the iliac tubercle, also considered a redundant landmark in the current study – the superior point of the iliac crest the posterior superior iliac spine, the posterior inferior iliac spine, and the widest point of the greater sciatic notch. When the anterior inferior iliac spine and the point on the iliac tubercle were omitted from the 32-landmark configuration of this study, meaningful sex-based shape data was represented in the first five PCs of the 17-landmark configuration.

However, when more landmarks are used to represent features of coxal bone shape, particularly in the ilium, the number of PCs that represent meaningful shape data decreases slightly. In the current study's 32-landmark configuration, 13 landmarks are in the ilium, not including the acetabulum; as a result, the number of PCs representing meaningful shape data were reduced to four. As a comparison, González et al. (2017) used 14 landmarks to represent the ilium, however, 7 PCs represent meaningful shape data. Roth (1993) explains how redundancy in landmarks might be desirable if the landmark points represent not only a fixed point of reference, but also an axis of orientation and a plane of reference, which the landmarks on the semi-circular coxal bone do. Having landmarks that represent the medial and lateral surfaces of the ilium, such as on the auricular surface and gluteal lines, as well as the anterior and posterior extremal points of the iliac spines and superior aspect of the iliac crest, will help

discern patterns of object shape. However, that shape may be meaningful to sex-based shape differences or allometric shape differences. The use of landmarks on the surface of the ilium as well as outlining the shape of the ilium (32-landmark configuration) discriminated between Males and Females less accurately than the 17-landmark configuration because the previous configuration captured greater allometric shape differences, whereas landmarks outlining the shape of the ilium captured greater sex-based shape differences and yielded greater sex-predicting power.

The 4 PCs representing shape data of biologial importance in the 32-landmark configuration contribute less than 50% of the total shape variance while the 5 PCs of biological importance in the 17-landmark configuration represent more than 50% of the total shape variance, as did the 6 landmarks of biological importance from both Bilfeld et al. (2012) and González et al. (2017). The landmarks in the 17-landmark configuration are more effective than the 32-landmark configuration for identifying shapes of biological importance since these fewer landmarks identify shapes of biological importance in greater than half the total variance. By evaluating how much shape variance is being taken up by PCs of biological importance, we can get a sense of whether the landmarks themselves are meaningful. This is important methodologically to get a sense of where sex-based shape variance may reside.

An alternative to including PC's of biological importance to DFA would be to simply use the broken stick model to determine which PCs are statistically meaningful, which in this study were the first two PCs of both landmark configurations (Jackson, 1993; Jolliffe, 2002). In the 32-landmark configuration, where statistically meaningful shape data contributed less than half the total shape variance, DFA on PC1 and 2 predicted sex in only 96.3% of cases, a 1.3% decrease over sex prediction using the five PCs of biological importance. In the 17-landmark

configuration the overall predictive power did not change between PCs of biological importance and predictions based on PC1 and 2 (statistically meaningful components of shape). However, there was a slight decrease in the predictive power among Males (98.3% to 97.8%) and an increase in the predictive power among Females (98.8% to 99.2%) using the first two PCs compared to five, respectively. Focusing only on the first two PCs in a DFA would provide a more directly comparable dataset and reflect the predictability of ordination so long as they meet criteria for inclusion using the broken stick model. The 17-landmark configuration continues to have greater predictive power than the 32-landmark configuration and greater predictive power than previous studies using landmarks to identify sex-based coxal bone shape regardless of which PCs are included in DFA. Landmark choice is more likely to be the deciding factor for improving sex prediction using whole coxal bone shape. These 17 landmarks would be the first step in understanding modularity in human coxal bone. By isolating the landmarks of sex-based shape, remaining landmarks - repressing the gluteal lines, the acetabulum, and the obturator foramen for example - are likely to represent non-specific differences in allometry (Bierry et al., 2010; Ridgeway et al., 2008; Schulter-Ellis et al., 1983; Zeng et al., 2012).

Only 6% of the population in this study are non-white, which could suggest a lack of variation in coxal bone shape. However, according to studies in coxal bone (Betti, 2014) and pelvic shape dimorphism (Fischer & Mitteroecker, 2015; Kurki, 2011, 2013) between human population samples representing climate and latitude, variation of within-sex and between-sex shape is predicated on a complex conjunction of body form (stature and body breadth). Mean stature among the small bodied of Kurki's (2013) samples fall within this current study's range of 142.24cm - 195.58cm (self-reported before death), representing variation in body form. Most of the misclassified individuals in this study are White, one individual classified as an "American

Male" fell close to the Female side of the negative end of the PC1/PC2 scatter plot and one "Asian Male" fell close to the Female side of the positive end of the PC1/PC2 scatter plot (see Appendix A) suggesting ancestry does not play a large role in the misclassification of individuals.

Stature values in the current study are significant between sexes (Male mean = 177.43cm; Female mean = 163.31cm; p = <0.001) but sex is not significant for coxal bone centroid size, which suggests that either centroid size is representing differences in body form that are not sex-based or the Procrustes ANOVA lacks the power to detect patterns between centroid size and residual sex-based shape. Because centroid size captures isometric size variants, it is more likely that non sex-based variables of coxal bone form (i.e. in coxal bone height and breadth) are being represented by centroid size, since sex-based variables of coxal bone height and breadth were captured by PC1. This would suggest further that centroid size could be used as a variable of human variation. In a Spearman's rho centroid size is statistically significant in the second PC of the 32-landmark configuration (PC1 r = -0.030, p = 0.403; PC2 r = -0.097, p = 0.007), for both PCs in the 17-landmark configuration, (PC1 r = 0.125, p = <0.001; PC2 r = -0.101, p = 0.004), and for sex (r = -0.111, p = 0.028). When centroid size is included as a variable in a DFA, it increases the predictive power of PC1 and PC2 by 0.2% in both landmark configurations. This means that isolated shape variants of sex, captured by PC1 and PC2, as well as variants of isometric variation in the form centroid size, and allometric variation in the form of right and left coxal bone asymmetry, can predict sex with up to 98.7% accuracy. More work should be done to relate centroid size to osteometric correlates before generating a sex estimation method.

2.6 **Conclusion**

The landmarks that best represent sex-based shape in the coxal bone are: the four iliac spines, ischial spine, apex of the auricular surface, ischial tuberosity superior and inferior, distal ischial tuberosity, symphyseal face height, center of the symphyseal face, arcuate eminence, ischiopubic ramus, iliac crest, acetabular rim posterior superior and anterior superior. Sex-based shape variation was not significantly influenced by bilateral asymmetry or allometry. The configuration of these 17 landmarks not only more accurately predicts sex in the os coxa, but they more accurately predict sex using PCs 1 and 2. The first two PCs that met the criteria for statistically meaningful data possessed the same total predictive capabilities as the 5 PCs representing biologically important shape data that contributes over 50% of the total variance. Sex predictions based on whole coxal bone shape is improved slightly by including allometric variables captured by centroid size. While the 32-landmark configuration captured more variation in object shape, too many landmarks, or too much variation in object shape, limit predictive power.

We strongly encourage applying the proposed 17 landmarks to future sexual dimorphism studies involving collections of both documented and undocumented human skeletal material and allometric variation to test the repeatability of this landmark configuration.

Chapter 3: **Investigating isometric and allometric effects of coxal bone measurements to develop a variation-inclusive sex estimation method**

3.1 **Synopsis**

Objectives: To identify linear measurements from Langley et al. (2016) that correspond with the 17-landmark configuration identified in Chapter 2 (Robertson et al., 2019), and use those measurements to generate reliable and repeatable sex estimation equations. Linear measurements were also evaluated to determine which represented centroid, isometric, and allometric size.

Methods: Parametric correlations between linear measurements and demographic variables identified which linear measurements represent important demographic variables. A component matrix identified linear measurements associated with principal components 1 and 2. An allometric equation using both Log Centroid Size and Log Stature identified measurements that corresponded with isometric and allometric size. A discriminant function (DF) and logistic regression (LR) equation were generated and tested on a subsample of the W.H. Bass Skeletal Collection (n = 120; f = 60, m = 60) for repeatability.

Results: The maximum pubic length (XPL), a measurement from the anterior superior iliac spine to the symphysion (ASISS), and the minimum distance between the apical border of the auricular surface to the symphysion (WAS) were measurements that significantly correlated with ancestry. WAS significantly correlated with side. ASISS and the maximum distance between the posterior superior iliac spine to the symphysion (PSISS) were not correlated with sex. Maximum iliac breadth (XIB) and XPL correlated with age at death. Maximum coxal bone height (XCH)

represented isometric size when the allometric coefficient was Log Centroid Size. XCH, XIB, ASISS, and PSISS represented isometric size when the allometric equation used Log stature. **Conclusions:** Measurements representing isometric size, ancestry, age, fluctuating asymmetry, and sex used to generate a DF equation for sex estimation were XCH, XIB, XPL, ISL, and WAS. The equation correctly predicted sex in 99.7% of cases (f = 191/191, 100%, m = 202/203, 99.5%) and demonstrated repeatability in 99.2% of cases (f = 59/60, 98.3%, m 60/60, = 100%). An equation using LR predicted sex in 100% of the study and test populations using XCH, XPL, ISL, and WAS.

3.2 **Introduction**

Sex estimation using the coxal bone has been a subject of interest for decades, yielding a myriad of methods (Albanese, 2003; Baumgarten & Ousley, 2015; Bierry et al., 2010; Brůžek, 2002; Gómez-Valdés et al., 2012; Karakas et al., 2013; Klales et al., 2012; Listi & Bassett, 2006; Milne, 1990; Naňka et al., 2007; Passalacqua et al., 2013; Phenice, 1969; Rogers & Saunders, 1994; Schulter-Ellis et al., 1983; Singh & Potturi, 1978; Sutherland & Myers Suchey, 1991; Vacca & Di Vella, 2012; Walker, 2005; Washburn, 1948). Geometric morphometrics (GM) has been used in previous studies to isolate sex-based shape with a high degree of accuracy but without much emphasis on developing a sex estimation method from these results (Anastasiou & Chamberlain, 2013; Betti et al., 2013; Bilfeld et al., 2012; Biwasaka et al., 2012; Bytheway & Ross, 2010; González et al., 2009; Gómez-Valdés, et al., 2012). The accuracy of these previous GM methods has been matched or bested by the 3D sex-based shape analysis outlined in Chapter 2 (Robertson et al., 2019).

Chapter 2 identified 17-landmarks that predicted sex in the adult coxal bone in 98.5% of cases (Robertson et al.,2019). Chapter 2 also observed that centroid size improved sex estimation prediction in discriminant function analysis (DFA) by 0.2%. Centroid size is calculated by taking the square root of the sum of the square distances of each landmark from the object centroid (Klingenberg, 2016). Centroid size is isometric size, or geometric size, that is uncorrelated with shape (Bookstein, 1991). The alternate function of size is allometric, which is size covariant with object shape and isolated morphology (Bookstein, 1987; Marcus et al., 1993; Zelditch et al., 2012). In Chapter 2, allometric size was captured in PCs 1 and 2 not only due to principal component analysis that reduces the number of random variables in a dataset to isolate allometric components of size to principal components and isometric components of size to centroid size (Fischer & Mitteroecker, 2015), but also because the PCA dataset included shape variants from both right and left coxal bones. Fluctuating asymmetry is also a component of allometric size. However, since centroid size improved sex prediction in the previous chapter, isometric as well as allometric components of size should both be incorporated when creating a sex estimation method in this study.

The landmarks in Chapter 2 correspond with various coxal bone measurement points outlined in Langley et al.'s (2016) *Data Collection Procedures for Forensic Skeletal Material* (Robertson et al., 2019). This study used Langley et al.'s corresponding measurements to generate a sex estimation method. Because sex estimation is a routine method in bioarchaeological and forensic analyses, it is essential that the method be universally applicable and as accurate as possible. This study will attempt to improve sex estimation using linear measurements that correspond with allometric size differences captured by PCs 1 and 2, and isometric size differences captured by centroid size in Chapter 2.

Previous studies aimed to address issues of bone form differences (size and shape) by generating population-specific models (İşcan, 1983; MacLaughlin & Bruce, 1986; Patriquin et al., 2003). However, other scholars recognize that sex estimation results are also impacted by within-group insults to body size, such as consequences of poor health or nutrition (Betti, 2017; Bytheway & Ross, 2010; Clark, 2014; Fischer & Mitteroecker, 2017; González et al., 2017; Guégan et al., 2000; Rosenberg, 2002; Ruff, 2002; Velemínská, et al., 2013). Albanese (2003) developed a variation-inclusive sex estimation method for the coxal bone with an accuracy of 98%, similar to the acuracy of the 17-landmark configuration in Chapter 2, however, Ablanese reports a 2% measurement error could lead to an incorrect sex classificationin using logistic regression. Using the landmarks from Chapter 2 a sex estimation method could be developed with a predictive power similar to Albanese's model, while simultaneously addressing the limitation of Albanese's pubic bone landmark at the junction in the acetabulum where the ilium, ischium, and pubis join. This point on the lunate surface of the acetabulum is not always easy to isolate and contributes to measurement error that conflates sex estimation (Albanese 2003). It would be interesting to see if an equation using Langley et al.'s (2016) measurement of maximum pubis and ischial length from the rim of the acetabulum, also used by Schulter-Ellis et al., (1983 and 1985), could improve Albanese's measurement error and therefore improve sex prediction. Such a method could improve sex estimation prediction and benefit forensic and biological anthropological analyses where population affiliation is an unknown variable.

Deconstructing isometry from allometry in coxal bone measurements can be difficult since factors that affect human variation, such as proportional changes due to growth (isometric variation) and non-proportional changes due to fluctuating asymmetry and climatic adaptation (allometric variation) are not distinguishable by sex-independent linear measurements (Betti,

2014; Fischer & Mitteroecker, 2017; Gustafsson & Lindenfors, 2004; Kurki, 2017; Pfeiffer, 2012; Plavcan, 2011; Vercellotti et al., 2011). Centroid size, which represents isometric size, improved sex prediction in Chapter 2, and this study aims to determine which linear measurements also reflect isometric and allometric variables to generate a highly accurate sex estimation method from those landmarks. The goals of this research were to: 1) isolate isometric and allometric size within Langley et al.'s (2016) osteometric measurements that correspond to landmarks from Chapter 2, 2) attempt to unravel the interconnectedness of isometry and allometry in sex estimation to provide much-needed insight into a variation-inclusive model for sex estimation, and 3) generate a powerful sex predictive function based on results from the first and second goals.

3.3 **Materials and Methods**

3.3.1 Specimens

The study sampled 394 osteological specimens from the William Bass Donated Skeletal Collection (Bass Collection), Forensic Anthropology Center (FAC), University of Tennessee, Knoxville, TN. The FAC receive donated cadavers through the Forensic Anthropology Body Donation Program (BDP). Living applicants to the BDP self-report sex, place of birth, ancestral affiliation, stature, and weight. Executors for the deceased provide age at death. Ancestry is categorized by the FAC as "White", "Black", "American Indian", "Hispanic", "Polynesian", "East Asian", or "Asian". This study combined categories of American Indian and Hispanic as "American", while Polynesian, East Asian, and Asian were reclassified as "Asian". These two

categories were combined and labelled “American/Asian” to increase group sample size for statistical tests. The distribution of individuals within this study population was overwhelmingly White (93.7%), with 3% of individuals classified as Black (n = 12,) and 3.3% classified as American/Asian (n = 13), however, it is proportional to the demographics of the Bass Collection (White = 93%; Black = 4%; Multiple[1], American Indian, Asian and Hispanic = 3% [Forensic Anthropology Center, n.d.]). For certain statistical tests it was necessary to further condense the ancestral categories into a collective “Non-White” category (n = 25). Although flawed, this categorization strategy offered a statistical comparison to previous studies given the low numbers of Non-White individuals available in the Bass Collection (Betti et al., 2013; Brown, 2015; Djorojevic et al., 2014; Gómez-Valdés et al., 2011; Kurki, 2011; Patriquin et al., 2003; Rosenberg, 2002; Steyn & Patriquin, 2009; Tague, 2000; Velemínská et al., 2013).

Table 3.1 outlines the demographic information of the sample population by sex and by population origin. Maximum Femoral Length (FML) and Femoral Head Diameter (FHD) were used as alternative measures of body form to compare with reported Stature and Weight. FML and FHD were measured according to Langley et al. (2016). FML was measured using an osteometric board from the most proximal to the most distal point on the femur, FHD was obtained using a digital sliding caliper measured in anterior/posterior (AP) orientation. Osteometric measurements were not taken on femora if medical devices or implants were present on the bone; for example, measurements were not taken from individuals with knee replacement

[1] “Multiple” refers to reported mixed ancestry. One individual in our sample population was reported as “White/ American Indian” and was incorporated into the “American/Asian” category.

implants or from broken femora repaired by surgical screws. Reported Stature and Reported Weight was not available for all participants.

Sex	Ancestry	Anth.Var.	N	Min.	Max.	Mean	SD.
		Descriptive Statistics of Demographic Data					
Female	American	Age at Death	4	60	79	65.50	9.04
		Stature (cm)	4	152.40	172.72	164.46	8.64
		Weight (lbs.)	4	98.00	212.00	135.00	52.11
		FHD (mm)	4	42.00	45.50	43.25	1.55
		FML (mm)	4	424.50	460.50	438.50	16.29
	Asian	Age at Death	3	46	80	68.33	19.35
		Stature (cm)	3	147.32	168.91	156.21	11.29
		Weight (lbs.)	3	110.00	275.00	168.33	92.51
		FHD (mm)	3	37.00	46.00	41.67	4.51
		FML (mm)	3	389.50	474.00	418.00	48.50
	Black	Age at Death	4	24	73	47.00	20.77
		Stature (cm)	2[2]	167.64	167.64	167.64	<0.001[3]
		Weight (lbs.)	4	104.00	374.00	203.00	124.15
		FHD (mm)	4	38.00	41.00	40.25	1.50
		FML (mm)	4	409.00	456.00	439.75	21.50
	White	Age at Death	180	29	93	61.22	14.02
		Stature (cm)	178	142.24	182.88	163.36	7.70
		Weight (lbs.)	178	80.00	500.00	176.86	68.84
		FHD (mm)	179	37	49.50	42.38	2.19
		FML (mm)	178	376.50	492.50	436.03	21.72
	Total	**Age at Death**	**191**	**24**	**93**	**61.12**	**14.24**
		Stature (cm)	**187**	**142.24**	**182.88**	**163.31**	**7.75**
		Weight (lbs.)	**189**	**80.00**	**500.00**	**176.40**	**69.96**
		FHD (mm)	**190**	**37.00**	**49.50**	**42.34**	**2.22**
		FML (mm)	**189**	**376.50**	**492.50**	**435.87**	**22.05**
Male	American	Age at Death	5	29	64	40.80	14.10

[2] Reported stature only available for 2 out of 4 individuals in this category
[3] The two individuals shared the same data; verified with original FAC data.

	Stature (cm)	2	162.56	190.50	176.53	19.76
	Weight (lbs.)	5	135.00	224.00	166.60	33.69
	FHD (mm)	5	41.00	51.50	46.10	3.83
	FML (mm)	5	411.00	517.50	446.30	46.83
Asian	Age at Death	1	72	72	72.00	-
	Stature (cm)	1	185.42	185.42	185.42	-
	Weight (lbs.)	1	320.00	320.00	320.00	-
	FHD (mm)	1	53.00	53.00	53.00	-
	FML (mm)	1	483.00	483.00	483.00	-
Black	Age at Death	8	23	79	51.88	17.11
	Stature (cm)	8	162.56	190.50	179.55	9.28
	Weight (lbs.)	8	105.00	250.00	194.37	45.07
	FHD (mm)	8	44.00	51.00	47.81	2.14
	FML (mm)	8	457.50	512.00	495.38	18.57
White	Age at Death	189	26	90	59.50	13.82
	Stature (cm)	187	157.48	195.58	177.30	7.26
	Weight (lbs.)	186	84.00	459.00	190.26	63.64
	FHD (mm)	189	42.50	57.00	48.31	2.40
	FML (mm)	189	414.50	534.50	471.86	22.09
Total	**Age at Death**	**203**	**23**	**90**	**58.80**	**14.25**
	Stature (cm)	**198**	**157.48**	**195.58**	**177.43**	**7.44**
	Weight (lbs.)	**200**	**84.00**	**459.00**	**190.48**	**62.92**
	FHD (mm)	**203**	**41.00**	**57.00**	**48.26**	**2.46**
	FML (mm)	**203**	**411.00**	**534.50**	**472.21**	**23.41**

Table 3.1 Descriptive statistics for demographic data by self-reported sex and ancestry.

Osteometric measurements of the coxal bone were obtained using a digital spreading caliper at corresponding landmarks from Chapter 2 (Figure 3.1). Osteometric measurements are defined in Table 3.2 and follow descriptions from Langley et al. (2016).

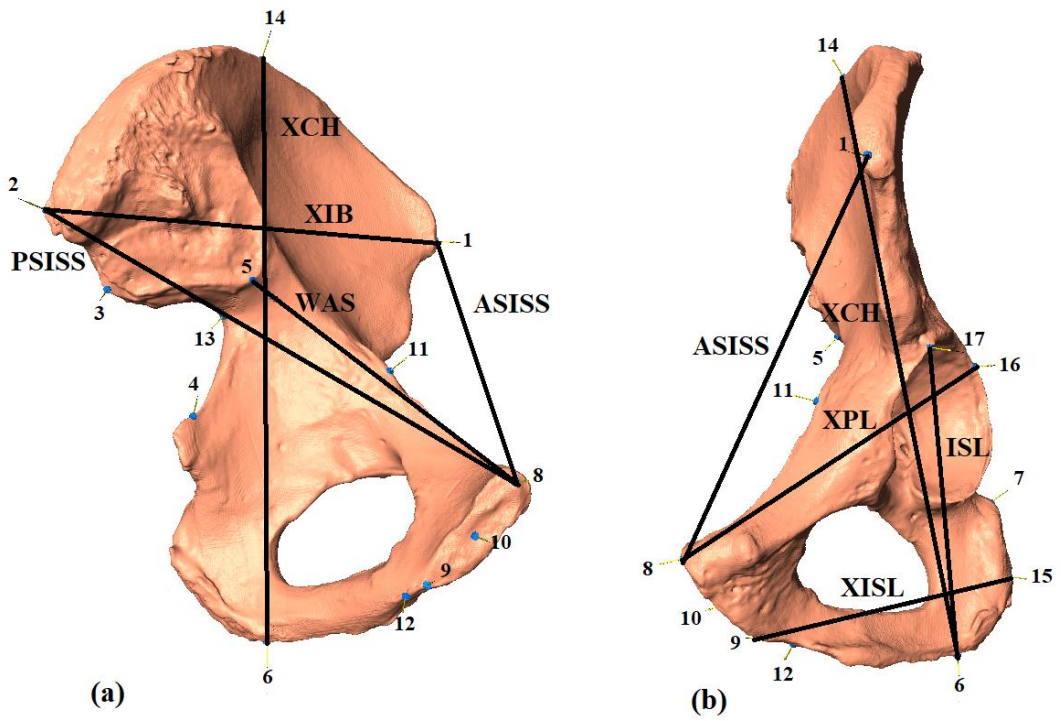


Figure 3.1 Illustration of coxal bone measurements used in this study. Left coxal bone; (a) anterior view (b) lateral view.

Variable	Code	Description
Maximum Coxal Height	XCH	Distance from the most superior point on the iliac crest to the most inferior point on the ischial tuberosity; distance between landmarks 6 and 14.
Maximum Iliac Breadth	XIB	Distance from the anterior superior iliac spine to the posterior superior iliac spine; distance between landmarks 1 and 2.
Maximum Pubic length	XPL	Distance between the most superior point on the symphyseal face to the farthest point on the acetabular rim; distance between landmarks 8 and 16.
Ischial Length	ISL	Distance from the point on the acetabular rim where the iliac blade meets the acetabulum to the most inferior point on the ischial tuberosity; distance between landmarks 6 and 17.

Maximum Ischiopubic Ramus Length	XIRL	Distance from the most inferior point on the symphyseal face to the most distant point on the ischial tuberosity; distance between landmarks 8 and 15.
Anterior Superior Iliac Spine to Symphysion	ASISS	Distance from the apex of the anterior superior iliac spine to the most superior point on the symphyseal face; distance between landmarks 1 and 8.
Maximum Posterior Superior Iliac Spine to Symphysion	PSISS	Distance from the posterior border of the posterior superior iliac spine to the most superior point on the symphyseal face; distance between landmarks 2 and 8.
Minimum Apical Border to Symphysion	WAS	Minimum distance from the most superior point on the symphyseal face to the anterior apex of the auricular surface; distance between landmarks 5 and 8.

Table 3.2 List of coxal bone measurement variables (from Langley et al., 2016). All measurements are in mm.

3.3.2 Statistical Methods

Data were compiled into two datasets in order to answer questions concerning variation in coxal bone form. Data was first evaluated between left and right coxal bones to identify patterns of fluctuating asymmetry that could affect interpretations of sex-based shape, population, or age-related changes to coxal bone shape. The second dataset contains centroid size, principal components of shape, and averages by side of all linear coxal bone and femoral measurements, for overall correlation (Pearson's correlation; IBM SPSS Statistics v.25.0, SPSS Inc. 2017) and means testing (ANOVA, Independent *t*-test; IBM SPSS Statistics v.25.0, SPSS Inc. 2017) between sex and among ancestral groups.

To determine which linear measurements and demographic variables are correlated with centroid size, namely which variables are uncorrelated with shape (Bookstein, 1991), centroid size was subjected to a Spearman's correlation with sex, side, and ancestry (White, Black, and American/Asian), age at death along with all averaged coxal bone and femoral linear

measurements were subjected to a Pearson's correlation. To isolate patterns of coxal bone shape in the study population, sex and ancestry were correlated with demographic variables (age at death, stature, weight, and femoral measurements) and coxal bone measurements. Principal components of shape (n = 44) from Chapter 2's evaluation of 17 landmarks were also correlated with sex, ancestry, age at death, stature, weight, and femoral measurements to isolate patterns of association between coxal bone shape and demographic variables.

Linear coxal bone measurements were log-transformed and applied to an allometric equation ($Y = bX^k$) to determine the allometric coefficients of each measurement as they relate to Log Centroid Size and Log Stature. There is no formal recommendation for selecting the measure of size, so this study compared results from both Log Centroid Size and Log Stature (Table 3.7). Centroid size is a measure of isometric size and this study expects the log transformed variable to discriminate measures of isometric size well (Bookstein, 1991). Stature is not strictly isometric or allometric in nature, which suggests the log transformed variable would be less sensitive at discriminating between these relationships (Arcini et al., 2014; Guégan et al., 2000; Kurki, 2013; Vercellotti et al., 2011; Zakrzewski, 2003). Similar results between these two measures of size would be a strong indicator of the measurement's underlying isometric or allometric relationship (Fischer & Mitteroecker, 2015).

The allometric coefficients determine which measurements scale allometrically and which scale isometrically with two measures of size (i.e. centroid size or stature). Isometric size is represented by allometric coefficient $k = 1$ and allometric size is represented by a value less than or greater than 1. Allometric coefficient b indicates the magnitude of difference between the measures of size (X) and the linear measurement (Y), which is important for comparing differences between Male and Female measurements and sex differences from the population

average. Measurements were also correlated with stature and centroid size to determine the nature of the relationships between isometry, allometry and coxal bone form.

A Levene's test was performed to assess the homogeneity of variance for linear coxal bone measurements to warrant discriminant function analysis (DFA). DFA was used to develop a sex estimation equation because it is more flexible when determining sectioning points and provides space between sectioning points for intermediate sex classifications. Linear measurements were also applied to logistic regression (LR) analysis as a robust test of dichotomous categories.

3.3.3 Observer Error

Inter-observer error was calculated between four observers on the right and left coxal bones from two specimens to verify the reliability of linear measurements and associated landmarks using the criteria outlined in Corner et al. (1992). Results are presented in Appendix B. Corner et al. recommend using a coefficient of variation [(standard deviation/mean distance) *100] of less than 3% for acceptable measurement error. This study elected to lower the level for acceptable error to 2.5% to capture measurement error that contributes to sex missclassification. A precision study of two landmark placement events by a single observer evaluated intra-observer error. The Procrustees coordinates of each landmark were averaged between 12 specimens grouped by sex (Males = 7; Females = 5) and side. Procrustes coordinate average and standard deviation for each landmark were subjected to a Student's *t*-test to evaluate significant differences between landmark events.

3.4 **Results**

3.4.1 Observer Error

Landmark coordinates were largely consistent between landmark events in the intra-observer landmark precision study (Appendix C). There was a stignificant difference in the anterior superior point on the acetabular rim in the medial/lateral direction (landmark 17x, p = 0.015) among right coxal bones of Males. A landmark coordinate that came close to significance between landmark placement events was the most inferior point of the ischial tuberosity in the anterior/posterior direction (landmark 6z, p = 0.073) among left coxal bones of Male individuals. Measurement error between four observers (inter-observer error) occurred in the maximum ischail length (ISL) measurement in both right and left coxal bones.

3.4.2 Centroid Size, Fluctuating Asymmetry, and Age at Death

In a Spearman's correlation, Centroid Size did not correlate with side (r = 0.032, p = 0.371) or ancestry (r = -0.047, p = 0.184) but correlated only with sex (r = -0.422, p = <0.001). In a Pearson's correlation Centroid size correlated significantly with all linear measurements, age at death did not correlate with Centroid size (r = 0.058, p = 0.105). Significant patterns of fluctuating asymmetry (Table 3.3) were found in three linear measurements, minimum apical boarder to the symphysion (WAS), femoral head diameter (FHD), and maximum femoral length (FML). Age at death was significantly correlated with Stature (r = -0.234, p = <0.001), Weight (r = -0.224, p = <0.001), maximum iliac breadth (XIB: r = 0.106, p = 0.035), maximum pubic length (XPL r = 0.121, p = 0.017), and FML (r = -0.103, p = 0.041). FHD correlated with both

Stature (r = 0.774; p = <0.001) and Weight (r = 0.164; p = 0.001) as did FML (Stature: r = 0.885, p = <0.001; Weight: r = 0.263, p = <0.001).

	Descriptive Statistics				Levene's test		*t*-Test	
Variable	**Side**	**Mean**	**SD**	**Std. Error**	***F***	**Sig**	***t***	**Sig**
XCH	Left	211.17	14.83	0.747				
	Right	210.99	14.33	0.722	0.279	0.598	0.170	0.865
XIB	Left	157.28	8.99	0.453				
	Right	158.51	8.94	0.450	0.074	0.786	-1.920	0.055
XPL	Left	118.53	5.98	0.301				
	Right	118.82	6.05	0.305	0.002	0.960	-0.660	0.509
ISL	Left	108.81	8.31	0.419				
	Right	109.38	8.34	0.420	0.077	0.781	-0.965	0.335
XIRL	Left	96.22	5.52	0.278				
	Right	96.07	5.69	0.287	0.005	0.943	0.394	0.694
ASISS	Left	140.54	10.09	0.508				
	Right	140.58	11.38	0.573	0.000	0.997	-0.054	0.957
PSISS	Left	174.64	10.39	0.523				
	Right	175.01	11.78	0.594	0.412	0.521	-0.462	0.644
WAS	Left	119.09	8.07	0.407				
	Right	121.04	8.50	0.428	0.690	0.406	**-3.288**	**0.001**
FHD	Left	45.34	3.72	0.188				
	Right	45.48	3.81	0.193	0.232	0.630	**-2.378**	**0.018**
FML	Left	455.27	29.52	1.52				
	Right	454.42	29.41	1.52	0.162	0.688	**3.570**	**<0.001**

Table 3.3 Descriptive statistics and independent t-test of measurement variables grouped by side (n = 394, a = 0.05, df= 786). Measurements of statistical significance in bold font.

3.4.3 Sex *t*-test for Demographic and Measurement Variables

Significant relationships existed (Table 3.4) between sex and all measurements averaged by side except the distance between the anterior superior iliac spine to the symphysion (ASISS) and the maximum distance from the posterior superior iliac spine to the symphysion (PSISS). Stature (t = 25.815, df = 768, p = <0.001) and Weight (t = 3.250, df = 770, p = 0.001) were also significant between sex categories. Sex did not correlate with side (r = 0.000, p = 1.000), age at death (r = -0.037, p = 0.304), or ancestry (r = -0.031, p = 0.378).

		Levene's Test					*t*-Test	
Variable	**Sex**	**Mean**	**SD**	**Std. Error**	***F***	**Sig**	***t***	**Sig**
XCH	Male	220.96	10.78	0.757				
	Female	200.58	9.77	0.707				
	Total	**211.08**	**14.49**	**0.730**	0.599	0.439	**19.608**	**<0.001**
XIB	Male	160.45	8.79	0.617				
	Female	155.19	8.13	0.589				
	Total	**157.90**	**8.87**	**0.447**	1.335	0.249	**6.152**	**<0.001**
XPL	Male	119.52	5.89	0.413				
	Female	117.78	5.90	0.427				
	Total	**118.68**	**5.95**	**0.300**	0.011	0.918	**2.930**	**0.004**
ISL	Male	115.51	5.17	0.363				
	Female	102.28	4.67	0.338				
	Total	**109.10**	**8.25**	**0.416**	1.456	0.228	**26.567**	**<0.001**
XIRL	Male	94.96	5.22	0.366				
	Female	97.41	5.36	0.387				
	Total	**96.15**	**5.42**	**0.273**	0.079	0.778	**-4.594**	**<0.001**
ASISS	Male	140.95	11.04	0.775				
	Female	140.15	9.46	0.685				
	Total	**140.56**	**10.30**	**0.519**	1.387	0.240	0.768	0.443
PSISS	Male	175.34	11.39	0.799				
	Female	174.27	9.76	0.706				
	Total	**174.82**	**10.63**	**0.536**	1.079	0.299	1.000	0.318
WAS	Male	117.77	7.46	0.524				
	Female	122.49	7.77	0.562				
	Total	**120.07**	**7.96**	**0.401**	0.624	0.430	**-6.127**	**<0.001**

FHD	Male	48.26	2.46	0.173				
	Female	42.34	2.21	0.161				
	Total	**45.40**	**3.78**	**0.191**	1.131	0.288	**24.983**	**<0.001**
FML	Male	472.21	23.41	1.643				
	Female	435.87	22.05	1.604				
	Total	**454.69**	**29.11**	**1.470**	0.311	0.577	**15.790**	**<0.001**

Table 3.4 Descriptive statistics and independent t-test (a=0.05, df=392) of measurement variables averaged between left and right coxal bones and grouped by Male (n=203) and Female (n=191). Measurements of statistical significance in bold font.

Similar results were observed for log-transformed linear measurements of the coxal bone. The minimum distance from the apical border to the symphysion (WAS) was also significantly different in fluctuating asymmetry for the transformed variables. Log-transformed XIB had a closer relationship with fluctuating asymmetry (t = -1.944, df = 786, p = 0.052), but again, is not statistically significant. There were no significant patterns between asymmetry and centroid or log-centroid size.

3.4.4 Ancestry *t*-test for Demographic and Measurement Variables

Age at death was statistically significant when compared across ancestry groups, however the variances between the two groups were not equal (t = 2.539, df = 53.169, p = 0.0014). Individuals classified as Black (n = 12, Σ age = 50.25 years) were on average 10 years younger than individuals classified as White (n = 369, Σ age = 60.38 years; p = 0.048). The American/Asian category (n = 13, Σ age = 57.15 years) did not display significant patterns with Age at death when compared to individuals classified as White (p = 1.000) or Black (p = 0.675). There were no significant patterns between stature or weight across ancestry groups.

In a Bonferroni *post hoc* test among White, Black, and American/Asian groups and all linear measurements (Table 3.5) XPL, ASISS, WAS, and FML displayed significant differences between groups. The main differences within XPL and WAS measurements were between individuals classified as White and Black. The individuals classified as Black were distinct from either the White or American/Asian groups for ASISS and FML.

Variable	(I)Ancestry	(J) Ancestry	Mean Difference (I-J)	Std. Error	Sig.
XCH	White	Black	6.699	4.244	0.346
		American/Asian	3.005	4.083	1.000
	Black	White	-6.699	4.244	0.346
		American/Asian	-3.693	5.793	1.000
	American/Asian	White	-3.005	4.083	1.000
		Black	3.693	5.793	1.000
XIB	White	Black	4.281	2.598	0.300
		American/Asian	1.605	2.500	1.000
	Black	White	-4.281	2.598	0.300
		American/Asian	-2.677	3.546	1.000
	American/Asian	White	-1.605	2.500	1.000
		Black	2.677	3.546	1.000
XPL	White	**Black**	**5.236**	**1.728**	**0.008**
		American/Asian	2.031	1.662	0.668
	Black	**White**	**-5.236**	**1.728**	**0.008**
		American/Asian	-3.205	2.358	0.525
	American/Asian	White	-2.031	1.662	0.668
		Black	3.205	2.358	0.525
ISL	White	Black	1.477	2.424	1.000
		American/Asian	2.133	2.332	1.000
	Black	White	-1.477	2.424	1.000
		American/Asian	0.656	3.308	1.000
	American/Asian	White	-2.133	2.332	1.000
		Black	-0.656	3.308	1.000
XIRL	White	Black	3.619	1.583	0.068

		American/Asian	0.156	1.523	1.000
	Black	White	-3.619	1.583	0.068
		American/Asian	-3.462	2.107	0.330
	American/Asian	White	-0.156	1.523	1.000
		Black	3.462	2.107	0.330
ASISS	White	**Black**	**14.307**	**2.941**	**<0.001**
		American/Asian	1.662	2.829	1.000
	Black	**White**	**-14.307**	**2.941**	**<0.001**
		American/Asian	**-12.645**	**4.013**	**0.005**
	American/Asian	White	-1.662	2.829	1.000
		Black	**12.645**	**4.013**	**0.005**
PSISS	White	Black	3.602	3.121	0.747
		American/Asian	0.604	3.002	1.000
	Black	White	-3.602	3.121	0.747
		American/Asian	-2.998	4.259	1.000
	American/Asian	White	-0.604	3.002	1.000
		Black	2.998	4.259	1.000
WAS	White	**Black**	**8.802**	**2.296**	**<0.001**
		American/Asian	2.487	2.209	0.782
	Black	**White**	**-8.802**	**2.296**	**<0.001**
		American/Asian	-6.314	3.133	0.134
	American/Asian	White	-2.487	2.208	0.782
		Black	6.314	3.133	0.134
FHD	White	Black	0.135	1.110	1.000
		American/Asian	0.696	1.068	1.000
	Black	White	-0.135	1.110	1.000
		American/Asian	0.561	1.515	1.000
	American/Asian	White	-0.696	1.068	1.000
		Black	-0.561	1.515	1.000
FHL	White	**Black**	**-22.354**	**8.450**	**0.025**
		American/Asian	14.287	8.129	0.239
	Black	**White**	**22.354**	**8.450**	**0.025**
		American/Asian	**36.641**	**11.531**	**0.005**
	American/Asian	White	-14.287	8.129	0.239
		Black	**-36.641**	**11.531**	**0.005**

Table 3.5 Bonferroni *post hoc* comparison of measurement variables and expanded categories of ancestry. Measurements of statistical significance in bold font.

3.4.5 Correlations with Principal Components

Of the 44 principal components (PC) representing coxal bone shape differences, only those that Chapter 2 identified as containing biologically important data were examined (PCs 1 to 5). In an independent *t*-test of PCs 1-5, averaged between the right and left coxal bones, PC3 was statistically significant with fluctuating asymmetry (t = 2.341, df = 786, p = 0.019, see Appendix D). Averaged PCs of biological importance that demonstrate statistically significant patterns between Males and Females included PC1 (t = -15.831, df = 392, p = <0.0001) and PC2 (t = 22.423, df = 392, p = <0.0001). Statistically significant patterns with ancestry (White and Non-white) are observable in PC4 (t = 3.470, df = 392, p = 0.001) and PC5 (t = -3.647, df = 392, p = <0.0001). A scatter plot of PC4 against PC5 in Appendix E displays the distribution of individuals according to ancestral affiliation. In a Bonferroni post hoc test, individuals categorised as Black differed significantly from the other two groups in PC4 (White*Black mean difference = 0.030, SE = 0.0061, p = <0.0001; American/Asian*Black mean difference = 0.029, SE = 0.0084, p = 0.002). Differences between White and Black groups differentiated shape in PC5 (mean difference = -0.017, SE = 0.0051, p = 0.002).

Principal components that are significantly correlated with stature are PC1 (r = -0.406, p = <0.0001), PC2 (r = 0.520, p = <0.0001), and PC3 (r = 0.187, p = <0.0001). Stature explains 16% of the variance in PC1 and 27% of the variation in PC2. Estimated weight explains a small amount of the variance in PC2 (r = -0.134, p = 0.008). Age at death contributed a small amount of the variance in PC3 (r = -0.137, p = 0.006), but was greater than side for the same component of shape (PC3: r = -0.083, p = 0.019). Sex contributed 39% of the variance to PC1 (r = 0.625, p

= <0.0001) and 56% of the variance in PC2 (r = -0.750, p = <0.0001). Correlations between PCs and averaged linear measurements are presented in Appendix F.

Several PCs of shape contain multiple variants and by examining the component matrix of each PC representing allometry, direct comparisons can be made between the three-dimensional landmark orientation and the linear measurement that would correspond to those landmarks (Table 3.6). Landmarks at the iliac crest, posterior superior and inferior iliac spines, apex of the auricular surface, greater sciatic notch curvature, inferior and distal points on the ischial tuberosity, iliac spine, and the anterior superior point on the acetabular rim correspond largely with shape changes in PC1. The associated linear measurements to PC1 are XCH, XIB, ISL, XIRL, PSISS, and WAS, respectively. As PC1 corresponds with coxal bone change due to sex and stature differences, it makes sense that measurements reflecting overall coxal bone size would be represented. Landmarks representing PC2 include all three points on the pubic symphysis, the posterior inferior iliac spine, greater sciatic notch curve, ischial spine, arcuate eminence, anterior and posterior superior points on the acetabulum, distal and superior points on the ischial tuberosity, apex of the auricular surface, and the ischiopubic ramus. Landmarks representing PC2 detected sex related changes in the pubic region, greater sciatic notch, and the ischium (XPL, ISL, XIRL, ASISS, PSISS, WAS).

	PC1	PC2	PC3	PC4	PC5	Landmarks	Corresponding Measurements
Log 14y	-0.900	0.183	-0.106	0.019	-0.032	Maximum arc of iliac crest	XCH
Log 14z	0.677	-0.622	0.023	-0.035	-0.136	Maximum arc of iliac crest	XCH
Log 14x	0.643	0.334	0.156	-0.126	-0.181	Maximum arc of iliac crest	XCH
Log 2z	-0.639	0.119	0.427	-0.002	-0.099	Posterior superior iliac spine	XIB/PSISS
Log 3y	0.602	0.520	-0.046	0.003	-0.202	**Posterior inferior iliac spine**	
Log 5z	-0.598	-0.342	-0.021	0.222	0.013	Apex of auricular surface	WAS
Log 5y	0.557	0.272	-0.013	0.076	-0.092	Apex of auricular surface	WAS
Log 6x	-0.501	-0.597	0.270	-0.091	-0.204	Inferior point on ischial tuberosity	XCH/ISL
Log 15x	-0.479	-0.370	0.389	-0.096	-0.263	Distal point on ischial tuberosity	XIRL
Log 3z	-0.448	0.115	0.306	0.296	0.011	**Posterior inferior iliac spine**	
Log 6y	-0.359	-0.331	0.143	-0.049	0.267	Inferior point on ischial tuberosity	XCH/ISL
Log 13x	0.347	0.308	-0.175	-0.027	0.052	**Maximum curvature of GSN**	
Log 17z	-0.252	-0.212	-0.228	0.204	0.023	Most anterior point on superior portion of acetabular rim	ISL
Log 9y	0.301	0.657	-0.377	0.201	-0.159	Inferior point on symphyseal face	XIRL
Log 10x	0.445	0.640	0.216	0.087	0.407	**Constructed landmark on symphyseal face**	
Log 3x	-0.018	-0.632	-0.274	0.282	-0.055	**Posterior inferior iliac spine**	
Log 8x	0.408	0.630	0.178	0.076	0.398	Superior point on symphyseal face	XPL/ASISS/PSISS/WAS
Log 11y	-0.224	-0.606	0.137	0.106	0.034	**Arcuate eminence**	
Log 4y	-0.265	-0.572	0.136	-0.050	0.153	**Ischial spine**	
Log 10y	0.288	0.566	-0.408	0.325	-0.279	**Constructed landmark on symphyseal face**	
Log 16x	-0.034	0.546	0.208	0.021	0.018	Posterior superior point on acetabular rim	XPL
Log 9x	0.395	0.527	0.254	0.107	0.304	Inferior point on symphyseal face	XIRL
Log 15z	0.221	0.511	0.342	0.154	-0.345	Distal point on ischial tuberosity	XIRL

Log 5x	-0.122	-0.505	-0.196	0.110	-0.232	Apex of auricular surface	WAS
Log 12x	-0.454	-0.482	0.218	-0.011	0.117	**Ischiopubic ramus**	
Log 4z	0.117	0.477	-0.035	0.181	0.247	**Ischial spine**	
Log 7z	0.231	0.452	-0.017	0.127	0.034	**Superior point on ischial tuberosity**	
Log 13y	0.145	-0.432	0.194	0.047	-0.139	**Maximum curvature of GSN**	
Log 17x	-0.076	0.427	0.128	0.060	0.079	Most anterior point on superior portion of acetabular rim	ISL
Log 13z	-0.247	0.392	-0.057	0.177	0.319	**Maximum curvature of GSN**	
Log 12z	-0.003	-0.355	-0.058	-0.285	-0.173	**Ischiopubic ramus**	
Log 11x	0.023	0.209	-0.053	-0.025	0.059	**Arcuate eminence**	
Log 17y	0.259	0.152	0.507	0.078	0.073	Most anterior point on superior portion of acetabular rim	ISL
Log 2x	0.316	-0.410	-0.485	0.444	0.272	Posterior superior iliac spine	XIB/PSISS
Log 16y	0.196	0.061	0.466	-0.090	0.141	Posterior superior point on acetabular rim	XPL
Log 7x	-0.322	-0.055	0.450	0.015	-0.163	**Superior point on ischial tuberosity**	
Log 6z	0.103	0.184	0.426	-0.053	-0.375	Inferior point on ischial tuberosity	XCH/ISL
Log 1z	-0.344	-0.279	-0.414	-0.073	0.134	Anterior superior iliac spine	XIB/ASISS
Log 12y	-0.285	-0.163	-0.287	0.129	0.007	**Ischiopubic ramus**	
Log 4x	0.002	-0.028	0.200	-0.099	-0.105	**Ischial spine**	
Log 10z	0.332	0.260	-0.281	-0.621	0.205	**Constructed landmark on symphyseal face**	
Log 1y	0.306	-0.431	0.112	-0.602	0.211	Anterior superior iliac spine	XIB/ASISS
Log 1x	-0.333	0.267	-0.519	-0.589	-0.212	Anterior superior iliac spine	XIB/ASISS
Log 8z	0.154	0.093	-0.348	-0.565	0.239	Superior point on symphyseal face	XPL/ASISS/PSISS/WAS
Log 9z	0.381	0.327	-0.131	-0.439	0.045	Inferior point on symphyseal face	XIRL
Log 8y	0.180	0.309	-0.385	0.430	-0.310	Superior point on symphyseal face	XPL/ASISS/PSISS/WAS
Log 16z	0.043	0.095	-0.168	0.322	0.086	Posterior superior point on acetabular rim	XPL

Log 2y	0.302	-0.021	-0.141	0.176	-0.643	Posterior superior iliac spine	XIB/PSISS
Log 15y	-0.194	-0.194	0.021	-0.133	0.613	Distal point on ischial tuberosity	XIRL
Log 7y	-0.073	-0.127	0.153	-0.141	0.453	**Superior point on ischial tuberosity**	
Log 11z	-0.182	-0.135	-0.208	0.074	0.242	**Arcuate eminence**	

Table 3.6 Component matrix of 17-landmark configuration, description and linear measurements associated with landmark. Landmark with no associated linear measurement highlighted in bold.

3.4.6 Allometric Coefficients

An allometric equation, $Y = bX^k$, was applied to log-transformations of each measurement variable (Y) against a "measure of size" (X) to identify allometric and isometric relationships from the coxal bone measurements (Zelditch et al., 2012). Log XCH was the only measurement with a demonstrable relationship with isometry for both measures of size. When the measure of size was Log Stature, Log XCH was joined by Log XIB, Log ASISS, and Log PSISS as isometric measurements. Both Log XCH and Log XIB had significant patterns with sex, however, Logs ASISS and PSISS were not significantly patterned with sex. Log Centroid Size identified greater magnitudes of isometric size difference between Males and Females, representing sexual dimorphism in coxal bone height. Log Stature identified sexual dimorphism in iliac breadth. Measurements that differentiated Male allometric size from Female allometric size when the measure of size was Log Centroid Size were: Log XCH, Log ISL, Log XIRL, and Log WAS. Male differences were only slightly larger than Females for Log ASISS and Log PSISS for the same measure of size. When the measure of size was Log Stature, the magnitude of difference was greater between sex groups and involved different linear measurements. Males were larger than Females for Log XIB, Log XPL, Log XIRL, Log ASISS, and Log WAS and were only slightly larger than Females in Log XCH. Females were larger than Males in Log PSISS and only slightly larger than Males in Log ISL.

Allometric Coefficients by Log Centroid Size						
	Average		**Males**		**Females**	
Log Centroid Size (X)	2.39	X-Y	2.40	X-Y	2.39	X-Y
Y	***k***	***b***	***k***	***b***	***k***	***b***
Log XCH	**0.97**	**0.07**	**0.98**	**0.06**	**0.96**	**0.09**
Log XIB	0.92	0.19	0.92	0.19	0.92	0.19
Log XPL	0.87	0.32	0.87	0.32	0.87	0.32
Log ISL	0.85	0.35	0.86	0.34	0.84	0.38
Log XIRL	0.83	0.41	0.82	0.42	0.83	0.40
Log ASISS	0.90	0.24	0.89	0.25	0.90	0.24
Log PSISS	0.94	0.15	0.93	0.16	0.94	0.15
Log WAS	0.87	0.31	0.86	0.33	0.87	0.30

Allometric Coefficients by Log Stature						
	Average		**Males**		**Females**	
Log Stature (X)	2.23	X-Y	2.25	X-Y	2.21	X-Y
Y	***k***	***b***	***k***	***b***	***k***	***b***
Log XCH	**1.04**	**-0.09**	**1.04**	**-0.10**	**1.04**	**-0.09**
Log XIB	**0.99**	**0.03**	**0.98**	**0.04**	**0.99**	**0.02**
Log XPL	0.93	0.16	0.92	0.17	0.94	0.14
Log ISL	0.91	0.19	0.92	0.19	0.91	0.20
Log XIRL	0.89	0.25	0.88	0.27	0.90	0.22
Log ASISS	**0.96**	**0.08**	**0.96**	**0.10**	**0.97**	**0.07**
Log PSISS	**1.00**	**-0.01**	**1.00**	**0.01**	**1.01**	**-0.03**
Log WAS	0.93	0.15	0.92	0.18	0.94	0.13

Table 3.7 Allometric coefficients by Log Centroid Size (scaled to 0.5) and Log Stature. Allometric coefficients representing isometric size (k = 1) in bold.

It is reasonable that Log Stature would identify XCH and XIB as measures of isometric size as both measurements have been identified as variants of size in previous studies (Auerbach et al., 2018; Betti, 2014; Kurki, 2007; Pinhasi et al., 2005). Log Stature also identified ASISS and PSISS as isometric measurements.

Untransformed centroid size, stature, and body weight were correlated with isometric and allometric measurements to determine the strength of these relationships with measures of size. FHD and FML are included to compare stature and centroid size correlations with non-coxal bone elements that represent body size. These results are presented in Table 3.8. All coxal bone and femoral measurement variables have a statistically significant relationship with stature and centroid size but the strength of their relationships very. Weight is uncorrelated with only one measurement variable, WAS. Centroid size and stature have strong statistically significant correlations (r = greater than 0.7) with FHD and FML as expected for osteometric measurements of body size. Centroid size strongly correlates with stature and is weakly correlated with weight (r value is less than 0.4). Weight is weakly correlated with all variables, but the greatest correlation value is with XIB. Centroid size and stature have strong relationships with XCH, and ISL. Centroid size is moderately correlated (r value less than 0.7 and greater than 0.4) with XIRL, ASISS, and WAS. Stature is moderately correlated with XIB, XPL, ASISS, and PSISS, and weakly correlated with XIRL and WAS.

Variable		**Stature (cm)**	**Weight (lbs.)**	**Centroid Size**
Centroid Size	r	**0.750**	**0.330**	**-**
	p	**< 0.001**	**< 0.001**	**-**
XCH	r	**0.833**	**0.305**	**0.897**
	p	**< 0.001**	**< 0.001**	**< 0.001**
XIB	r	**0.596**	**0.370**	**0.822**
	p	**< 0.001**	**< 0.001**	**< 0.001**
XPL	r	**0.537**	**0.299**	**0.850**
	p	**< 0.001**	**< 0.001**	**< 0.001**
ISL	r	**0.811**	**0.282**	**0.780**
	p	**< 0.001**	**< 0.001**	**< 0.001**
XIRL	r	**0.162**	**0.175**	**0.523**
	p	**0.001**	**0.001**	**< 0.001**
ASISS	r	**0.403**	**0.210**	**0.691**

	p	**< 0.001**	**< 0.001**	**< 0.001**
PSISS	r	**0.462**	**0.277**	**0.808**
	p	**< 0.001**	**< 0.001**	**< 0.001**
WAS	r	**0.115**	0.052	**0.557**
	p	**0.024**	0.306	**< 0.001**
FHD	r	**0.770**	**0.158**	**0.713**
	p	**< 0.001**	**< 0.001**	**< 0.001**
FML	r	**0.885**	**0.259**	**0.744**
	p	**< 0.001**	**< 0.001**	**< 0.001**

- no data

Table 3.8 Pearson Correlations of Centroid Size, Estimated Stature, and Estimated Weight with coxal bone and femoral measurement variables. Significant correlations are in bold.

In an exploratory correlation between linear measurements and demographic variables, age at death was significantly correlated with XIB (r = 0.106, *p* = 0.035), and XPL (r = 0.121, *p* = 0.017).

3.4.7 Sex Estimation Equation

The following measurements were then applied to a stepwise DFA to determine the factor that best predicted sex. The measurements that predicted sex with the highest level of accuracy (99.7%) in a cross-validation discriminant function analysis were maximum coxal bone height (XCH), maximum iliac breadth (XIB), maximum pubic length (XPL), ischial length (ISL) and the minimum distance from the apex to symphysion (WAS). The discriminant function equation is as follows:

$$DF = XCH\ (0.082) - XIB\ (0.030) - XPL\ (0.167) + ISL\ (0.233) - WAS\ (0.073) - 9.464$$

Males were correctly predicted 202 out of 203 individuals (99.7%), while Females were all correctly predicted (191/191 = 100%). The respective Male and Female means using this equation are 2.252 (SD = 1.072) and -2.394 (SD = 0.918). Positive values predict Males while negative values predict Females. One White Male was misclassified in the DFA due to a wide ilium and ischium, and shorter overall XCH, which are characteristically Female traits. This individual was of average height (172 cm) but below average weight (128 lbs) for Males. The individual's DF score, averaged values between right and left coxal bones, was -0.304, which suggests Female. Applying separate right and left measurements to the equation above, the left coxal bone scored -0.164, which brings it closer to the Male side, whereas the right coxal bone scored -0.444, much closer to the Female side.

Logistic Regression (LR) was performed on the same set of coxal bone variables to predict sex. The logistic regression model explained 100% (Nagelkerke R^2) of the variance and was statistically significant (χ^2(df = 4) = 545.83, p <0 .0001). Values greater than 0.5 are Male, values less than 0.5 are Female.

$$LR = XCH\ (11.925) – WAS\ (11.589) – XPL\ (39.871) + ISL\ (50.359) – 1825.677$$

When these two equations were applied to an independent test population of 120 individuals (Males = 60, Females = 60) also from the William Bass Skeletal Collection, DFA correctly identified 119 individuals in cross-validation (99.2%). One Female was misidentified. Females in the test population were largely homogeneous (White n = 57, Black n = 3), though less so among Males (White n = 42; Black n = 11, American n = 7). However, there is body size

diversity within both Male (157.48cm – 187.96cm) and Female (149.86cm – 177.80cm) groups among the few individuals with reported stature. LR correctly identified all specimens in the study sample and in the test population. The logistic regression equation predicted sex better without the XIB measurement.

3.5 **Discussion**

The discriminant function analysis eliminated the distance from the anterior superior iliac spine to the symphysion (ASISS), the maximum distance from the posterior superior iliac spine to the symphysion (PSISS) and maximum ischiopubic ramus length (XIRL) as contributors to a highly accurate sex prediction equation. Although Lance et al. (2000) suggests that omitting variants from a discriminate function analysis (DFA) could lessen the predictive function of the equation, ASISS and PSISS did not have a significant relationship with sex, which explains why these measurements would not be included in the discriminant function. However, measurement XIRL had a significant relationship with sex but only a moderate relationship with centroid size, and weak relationships with stature and weight. Sex and allometry are well represented by other measurements of the pubis and ischium in this study, particularly maximum pubic length (XPL) and ischial length (ISL). Measurement XPL does double duty as a variable of ancestry as well as sex. Landmarks that represent XIRL could be redundant since they are less discriminating than independent measures of the pubic bone and ischium by XPL and ISL respectively in this study.

Logistic Regression (LR) also omitted ASISS, PSISS and XIRL as well as XIB. XIB was a measure of isometric size as were ASISS, PSISS, and maximum coxal bone height (XCH), if stature represented allometric coefficients. XCH is the only measurement of isometric size

present in the LR equation and the allometric coefficients by Log Centroid Size identified only XCH as a variable of isometric size. This suggests that Centroid Size, rather than Log Stature, should be used in an allometric equation to identify variables of isometric size.

The question when determining accurate methods of sex estimation is whether to lump human populations together to create an inclusive sex estimation or whether to generate population-specific models. The problem with generating population-specific models is that methods become moot when the population of origin is unknown. However, the problem with generating a universal sex estimation model is determining which linear measurements best represent true sex-based shape differences when the human coxal bone varies by side, age, and population. Based on the analysis presented in this study, the inclusion of variables that represent isometric size, such as XCH and XIB, fluctuating asymmetry represented by WAS, age at death represented by XPL, and population variation represented by both WAS and XPL, generate highly accurate equations because of their inclusivity. Other successful population inclusive methods of sex estimation involve measuring the femoral head or superior pubis ramus length and acetabular ischium length (Albanese, 2003; Albanese et al., 2008). However, these methods, reached an accuracy of 98% and 89.4% respectively, lower than the DF and LR accuracies presented in this study.

Albanese (2003) incorporated variables of body size into his sex evaluation study using femoral head diameter (FHD) and epicondylar breadth, while coxal bone measurements included XCH and XIB. These four measurements correctly predicted sex in 98.5% of the test sample. The two equations presented in this study accounted for an additional source of human variation omitted by Albanese: fluctuating asymmetry, represented by the minimum distance from the apical border to the symphysion (WAS). Age-related coxal bone changes resided in the XPL

measurement. WAS and XPL also represented differences between individuals classified as White and Black, which makes both the DF and LR equations in this study applicable as population inclusive models. XCH and ISL represented sex differences and differences in body size (Stature, Weight, and Centroid Size).

It is well documented that body size variation can affect the accuracy of sex predictions particularly when those body size variations are more readily identified as human stature variation than between-population variation (Biwasaka, et al., 2012; Clark, 2014; Dixit et al., 2007; Đurić et al., 2005; İşcan, 1983; MacLaughlin & Bruce, 1986; Patriquin et al., 2005; Ridgeway et al., 2011; Vacca & Di Vella, 2012; Vercellotti et al., 2011). It is not surprising that individual populations with different body size ranges would yield different discriminant function equations, thus prompting the need for population-specific discriminant functions. However, when populations are pooled together there does not appear to be glaring differences in body size variation to interfere with sex estimation (Albanese et al., 2008; Macaluso, 2010; Steyn & Patriquin, 2009).

Steyn and Patriquin (2009) tested a pooled population of Greeks, Black South Africans, and White South Africans that yielded an average sex accuracy of roughly 94% (cross-validated results) among all populations using seven coxal bone measurements. The accuracy of sex predictions was similar among Greek and White South African populations, while Black South Africans performed 0.5% below this average. The population of Black South Africans were reported by Patriquin et al. (2002) to have come from unclaimed remains and could likely have previously lived at a lower socioeconomic status than the White South African and the Greek populations. Steyn and Patriquin conclude that coxal bone size, body size, and sample size were contributing factors to the predictive power of their discriminant functions in pooled populations.

The authors acknowledge that their population of Black South Africans would likely have represented lower socioeconomic status compared to their White South African population. Coxal bone size and body size could therefore be impacted by the adverse effects of growing up in lower socioeconomic conditions with reduced health and nutrition (Vercellotti et al., 2011; Zakrzewski, 2003), particularly if factors of individual variation, in the form of fluctuating asymmetry, were not accounted for in their equation. We suggest the inclusion of certain measurements, such as the greater sciatic notch breadth and acetabular diameter, and the exclusion of right/left averaged values limited Steyn and Patriquin's sex predictions.

The greater prediction accuracy of sex this study generated is likely because this study included body size and fluctuating asymmetry as variables of isometry and allometry. Body size is considered of greater importance to variations in dimensions of the pelvic complex, and of the coxal bone specifically, rather than population origin (Betti, 2014; Kurki, 2013). Stature variation exists within populations as well as between populations, which explains why there were no significant patterns between ancestry and stature. Stature differences can occur within populations due to differences in nutrition likely due to social stratification (Arcini et al., 2014; Fischer & Mitteroecker, 2017; Guégan et al., 2000; Vercellotti et al., 2011; Zakrzewski, 2003). However, socioeconomic data was not obtained in this study to support this hypothesis. Age at death has been used as a marker for lower socieconomic status in previous studies (İşcan, 1983). Black Females were the youngest members of this study's test population as well as Black Males, but Black Males were also among the oldest members of the study population. Statistical differences of XPL and WAS between White and Black categories could be related to socioeconomic differences between those two groups influencing coxal bone growth in these areas, but not enough information is available in this study to make such a claim. ASISS was

statistically different between all three ancestral categories as a marker of human variation and is best omitted for greater confidence in sex prediction. It would be interesting to see if this measurement could be used to identify ancestral groups, which would prove useful in a forensic context.

This study used the averaged measurements between the right and left coxal bones in the DFA rather than values from the left side only, to account for individual variation in fluctuating asymmetry. Fluctuating asymmetry was statistically significant with the WAS measurement, an important measurement in this sex estimation equation, which pressed the need for taking average coxal bone measurements into consideration in DFA. In Kurki's (2017) study of directional asymmetry, the right side is significantly larger than the left in the anterior portion of the pelvic inlet, whereas the posterior portion of the pelvic inlet is larger on the left side. Kurki also demonstrated that directional asymmetry in the pelvis occurred in roughly 90% of Females and 89% of Males in her study, which was made up of diverse human populations. By incorporating fluctuating asymmetry as a variable, sex predictions are improved over osteometric functions that exclude this variable, as demonstrated in the formula by Baumgarten and Ousley (2015).

This study arrived at a discriminant function equation that is very similar to that of Baumgarten and Ousley (2015), using the same measurements from Langley et al. (2016) but giving the measurements different weights. Predictive powers were also comparable; Baumgarten and Ousley correctly predicted sex in 97.5% of cases compared to 99.7% of cases in this study. Baumgarten and Ousley's study used 82 individuals classified as Black (Males, n = 23, Females, n = 59) and 118 as White (Males, n = 71, Females, n = 41) from the Hamann-Todd collection, with the sectioning point for their DF equation indicating negative values for Males

and positive values for Females. Baumgarten and Ousley used left coxal bones to generate their formula. When the Baumgarten and Ousley formula was applied to left coxal bones of the current study, 12 cases were misclassified (Females n = 2; Males n = 10), including the previously misclassified White Male mentioned above, resulting in an accuracy of 97%. It is encouraging to know that similar functions can be generated from different collections and from different population demographics. It suggests these methodologies are reproducible despite coming from different reference specimens. However, the Bass collection (used in this study) and Hamann-Todd collection (used in Baumgarten and Ousley's study) might be homogeneous in terms of the socioeconomic status of the individuals in their respective collections, which as previously mentioned, affects overall stature.

The DF and LR equations presented in this study along with Baumgarten and Ousley's (2015) DF equation not only conform to the 17-landmark configuration identified in Chapter 2 but were also used as part of Brůžek's (1996) analysis of sexual stability across time and demonstrated temporal stability in a Green's *t*-test. The pubic bone was shown to fluctuate in length with time; Brůžek's Epipaleolithic samples demonstrated longer pubic bones among Males compared to Epipaleolithic Females, which could lead to misclassification of older archaeological specimens using the DF equation we are proposing. Patriquin et al. (2002) found that pubic bone length fluctuated between populations; people in the White category possessed longer pubic bones, and shorter pubic bones largely occurred among the Black category, which were consistent with Washburn's (1948) findings. However, the length of the pubic bone must not be unique to populations if pubic lengths can also vary over time within similar populations (Brůžek, 1996; MacLaughlin & Bruce, 1986).

Inter-observer measurement error suggests that sex missclassification could occur when ischial length is not a maximum length. When measuring Specimen 2, two observers under-measured (< 108 mm) the length of ISL on the left side, and one observer under-measured ISL on the right side. The under-measurements by these observers likely resulted in a Female missclassification in these cases. However, the ISL measurements of observer 4 was greater than the ISL measurement of observer 3, and yet observer 4 classified this specimen as intermediate. Despite this anomaly, under-measurement of the maximum ischial length is likely to result in sex misclassification. If Corner et al.'s (1992) benchmark of 3% was used to evaluate the measurment error, ISL values from the right coxal bone would be accepted despite sex misclassification, but not the ISL values from the left coxal bone.

Having a sex estimation method that accounts for body size, as it relates to sexual dimorphism in unhealthy populations or monomorphism in healthy populations, might not be as confounding as previously thought. Body size and dimorphism ratios are affected by changes in nutrition and socioeconomic status, which can be reflected in the growth and development of the os coxa, particularly if nutritional deficiencies occur during the adolescent growth phase (Cameron, 1991; Clark, 2014; Vercellotti et al., 2011; Zakrzewski, 2003). A single equation to account for human variation in all its manifestations helps to improve sex estimation rather than inhibit it. Although our method was unsuccessful at classifying one individual out of 394 who was a Male of shorter than average stature, our method improved overall sex estimation, particularly in Males, in a population with diverse body sizes. More work is recommended to evaluate the effectiveness of this formula among a variety of populations of smaller stature.

3.6 **Conclusion**

Two equations presented in this study are accurate 99.7% and 100% of the time. DF equation can be applied when it is important to include individuals of indeterminate sex. LR equation can be applied when more dichotomous results are desired. The success of this method lies in its conception, devising equations from variables representing individual variation (fluctuating asymmetry), population variation (stature, age, and morphology), isometric size variation (XCH, XIB), and allometric size variation (XPL, ISL). These measurements are based on landmarks that are known to best predict sex and have been evaluated to enable highly accurate sex predictions, rather than conflate sex with erroneous data. Therefore, the sex estimation equations are variation inclusive.

Chapter 4: **Variation-inclusive discriminant function analysis and logistic regression modeling for sex estimation in partial coxal bones**

4.1 **Synopsis**

Objectives: This analysis tests the reliability and repeatability of discriminant function (DF) and logistic regression (LR) equations designed to represent fragmented coxal bone based on the equations for whole coxal bone presented in Chapter 3. This analysis also explores the applicability of other Langley et al. (2016) measurements associated with landmarks eliminated from Chapter 2's final configuration that could be used to estimate sex in fragmented coxal bone at a reduced accuracy.

Methods: Discriminant functions were designed (n = 369; f = 191, m = 203, W.H. Bass Skeletal Collection) to predict sex on hypothetical scenarios of bone completeness by removing measurements to represent bone completeness greater than 75%, 50-25%, and at least 25%. Sex predictions were tested for repeatability on an independent subsample from the same collection (n = 120; m = 60, f = 60), ranging in age from 25-96 years and in body size from 149.86cm – 187.96cm. Measurements between left and right sides were averaged and applied in the DFA and test study. An inter-observer error study measured the accuracy of the DF equations when used by inexperienced observers.

Results: Function 1 from Chapter 3 (at least 75% complete), Function 2 (50-25% complete), and Function 3 (at least 25% complete) are considered reliable (99-99.7% accurate) and repeatable (99.2-100%) DF equations for sex estimation. Function 7 and Function 8 are less reliable

(98.7%) and less repeatable (96.7%; 98.3% respectively) but acceptable equations for coxal bones between 50-25% complete.

Conclusions: Functions 1 and 3 are recommended for forensic applications. These equations excel at estimating sex because the landmarks are those that best predict sex, and linear measurements account for sex variation, population variation, and fluctuating asymmetry.

4.2 Introduction

Sex estimation based on true sex-based shape differences presented in Chapter 2 is illuminating the possibility of improving sex estimation in both biological and forensic anthropology (Robertson et al., 2019). The discriminant function (DF) equation for sex estimation presented in Chapter 3 based on the landmarks that best predict sex from chapter 2, amounted to 99.7% accuracy in cross-validation discriminant function analyses (DFA) and 99.2% accuracy when applied to an independent test population. The measurements used in the DF equation outlined in Chapter 3 include Maximum Coxal Height (XCH), Maximum Iliac Breadth (XIB), Maximum Pubis Length (XPL), Ischial Length (ISL), and Minimum Apical Border to Symphysion (WAS). These measurements account for allometry, fluctuating asymmetry, and diversity of body size that represent individual and human, or population, variation. It would be beneficial to the field of biological and forensic anthropology to test these measurements further to see if their accuracy stands up when applied to fragmented os coxae.

As there are many different methods of estimating sex in fragmented coxal bones, osteometric methods of sex estimation are not always comparable between studies, particularly when comparing measurements of the ischiopubic region (Albanese, 2003; Patriquin et al., 2005;

Washburn, 1948). Washburn (1948) measured the length of both the ischium and pubis from a single point in the acetabulum. Albanese (2003) modified the location of this single measurement point to improve inter-observer error. Patriquin et al. (2005) also modified the landmark location to the superior border of the acetabulum to capture maximum pubic and ischial length. All three studies used different landmarks at the acetabulum to obtain their measurements, resulting in incomparable results and lower sex prediction values. In response, recommendations have been made for further studies devising population or temporal specific methods (Biwasaka et al., 2012; Luo, 1995; MacLaughlin & Bruce, 1986; Patriquin et al., 2005; Rosenberg, 2002; Washburn, 1948; Zeng et al., 2012) rather than interpreting the limited accuracy as incomplete data on human variation. These population-specific methods are often not appropriate for identifying sex of specimens with unknown population affinity, and temporally specific osteometric methods of sex estimation might not be applicable if the population experienced environmental or cultural changes over time that influenced coxal bone morphology.

Robertson et al. (2019), Kurki (2007, 2011, 2013) and Betti (2017) have demonstrated that human variation in body size influences sex-based shape of the pelvis and coxal bones. By incorporating elements of human variation, like body size and stature, with elements of individual variation, like fluctuating asymmetry, methods of sex estimation in fragmented coxal bone material could increase in accuracy, as Chapter 3 proposed for whole coxal bones. This study will simulate fragmented coxal bone material by eliminating certain measurements at locations commonly absent in fragmented coxal bones and determine if replacement measurements from Langley et al., (2016) that correspond with some of 17 landmarks from Chapter 2 could maintain high accuracies in sex prediction.

4.3 **Method**

4.3.1 Specimens

This study used measurement data from 394 documented skeletons (Females = 191, Males = 203) from the William M. Bass Donated Skeletal Collection (Bass Collection), Forensic Anthropology Center (FAC), University of Tennessee, Knoxville, TN. Skeletons were from individuals who enrolled in the Forensic Anthropology Body Donation Program (BDP) prior to death and registered their date of birth, ancestry, stature, and weight (Table 4.1). The study population ranged in age from 23 – 93 years. Reported stature ranged between 142 – 196cm and reported weigh between 80 – 500lbs. Self-reported ancestral affiliation included categories of "White" or "Black". Chapter 3 condensed more descriptive BDP categories into "Asian" or "American", see Chapter 3 for a full description of ancestral affiliation.

The results were tested on a sub-population (n = 120; Females = 60, Males = 60) of the Bass Collection. The test population ranged in age from 25 – 96 years. Reported stature (150 cm – 188 cm) was available for less than half of the test population (n = 40) as was reported weight (n = 38; 103 lbs – 315 lbs). However, these ranges fell within the range of variation from the study population (Table 4.2), which suggests the test population does not contain known outliers. Ancestral variation was limited among Females in the test population (White n = 57, Black n = 3), but less so among Males (White n = 42; Black n = 11, American n = 7). The test population did not include individuals that represent an "Asian" category, however body size diversity was present with Male stature ranging from 157.48 cm – 187.96 cm and Female stature ranging from 149.86 cm – 177.80 cm, which Chapter 3 and Steyn and Patriquin (2009) suggest is a more valuable qualifier than population affinity.

Sex	Ancestry	Anth.Var.	N	Min.	Max.	Mean	SD	Skewness
Female	American	Age at Death (yrs.)	4	60	79	65.50	9.04	1.951
		Stature (cm)	4	152.40	172.72	164.46	8.64	-1.199
		Weight (lbs.)	4	98.00	212.00	135.00	52.11	1.822
	Asian	Age at Death (yrs.)	3	46	80	68.33	19.35	-1.727
		Stature (cm)	3	147.32	168.91	156.21	11.29	1.346
		Weight (lbs.)	3	110.00	275.00	168.33	92.51	1.709
	Black	Age at Death (yrs.)	4	24	73	47.00	20.77	0.374
		Stature (cm)	2	167.64	167.64	167.64	<0.001[1]	-
		Weight (lbs.)	4	104.00	374.00	203.00	124.15	1.198
	White	Age at Death (yrs.)	180	29	93	61.22	14.02	0.051
		Stature (cm)	178	142.24	182.88	163.36	7.70	0.102
		Weight (lbs.)	178	80.00	500.00	176.86	68.84	1.492
	Total	**Age at Death (yrs.)**	**191**	**24**	**93**	**61.12**	**14.24**	**-0.020**
		Stature (cm)	**187**	**142.24**	**182.88**	**163.31**	**7.75**	**0.049**
		Weight (lbs.)	**189**	**80.00**	**500.00**	**176.40**	**69.96**	**1.462**
Male	American	Age at Death (yrs.)	5	29	64	40.80	14.10	1.501
		Stature (cm)	2	162.56	190.50	176.53	19.76	-
		Weight (lbs.)	5	135.00	224.00	166.60	33.69	1.689
	Asian	Age at Death (yrs.)	1	72	72	72.00	-	-
		Stature (cm)	1	185.42	185.42	185.42	-	-
		Weight (lbs.)	1	320.00	320.00	320.00	-	-
	Black	Age at Death (yrs.)	8	23	79	51.88	17.11	0.141
		Stature (cm)	8	162.56	190.50	179.55	9.28	-0.737
		Weight (lbs.)	8	105.00	250.00	194.37	45.07	-1.008
	White	Age at Death (yrs.)	189	26	90	59.50	13.82	0.005
		Stature (cm)	187	157.48	195.58	177.30	7.26	0.001
		Weight (lbs.)	186	84.00	459.00	190.26	63.64	1.491
	Total	**Age at Death (yrs.)**	**203**	**23**	**90**	**58.80**	**14.25**	**-0.040**

[1] The two individuals shared the same data; verified with original FAC data.

	Stature (cm)	**198**	**157.48**	**195.58**	**177.43**	**7.44**	**-0.041**
	Weight (lbs.)	**200**	**84.00**	**459.00**	**190.48**	**62.92**	**1.456**

Table 4.1 Descriptive statistics of self-reported sex and ancestry in study population.

Sex	Ancestry	Anth.Var.	N	Min	Max	Mean	SD	Skewness
Female	Black	Age at death (yrs.)	3	39	59	46.00	11.269	1.717
		Reported stature (cm)	0	-	-	-	-	-
		Reported weight (lbs.)	0	-	-	-	-	-
	White	Age at death (yrs.)	57	32	92	64.40	13.689	-0.014
		Reported stature (cm)	12	149.86	177.80	163.12	7.456	-0.046
		Reported weight (lbs.)	10	103.00	260.80	162.53	49.740	0.769
Male	American	Age at death (yrs.)	7	27	58	44.86	10.303	-0.569
		Reported stature (cm)	1	157.48	-	-	-	-
		Reported weight (lbs.)	1	225.00	-	-	-	-
	Black	Age at death (yrs.)	11	25	71	54.27	14.136	-0.699
		Reported stature (cm)	2	167.64	182.88	175.26	10.776	-
		Reported weight (lbs.)	2	19000	205.00	197.50	10.607	-
	White	Age at death (yrs.)	42	31	96	64.57	14.343	0.118
		Reported stature (cm)	26	167.64	187.96	177.360	5.825	-0.028
		Reported weight (lbs.)	26	120.00	315.00	180.673	44.657	1.469

Table 4.2 Demographics of test population.

4.3.2 Measurements

Five linear measurements (Figures 4.1) are derived from Langley et al. (2016:77-78) described in Table 4.3. Following the methods outlined in Chapter 3, measurements from the study population were taken from both right and left coxal bones, averaged, and used to generate

the first four functions of the study (Table 4.4). The accuracy of the DF equation generated by canonical discriminant function coefficients, was verified by applying the equations (Table 4.5) to a test population.

The measurements from both the study and test populations were taken by a trained osteologist and it would be useful to understand the inter-observer error of these equations when untrained observers obtain measurements. An inter-observer error study calculated the reliability of linear measurements between four observers on 30 specimens (see Appendix D).

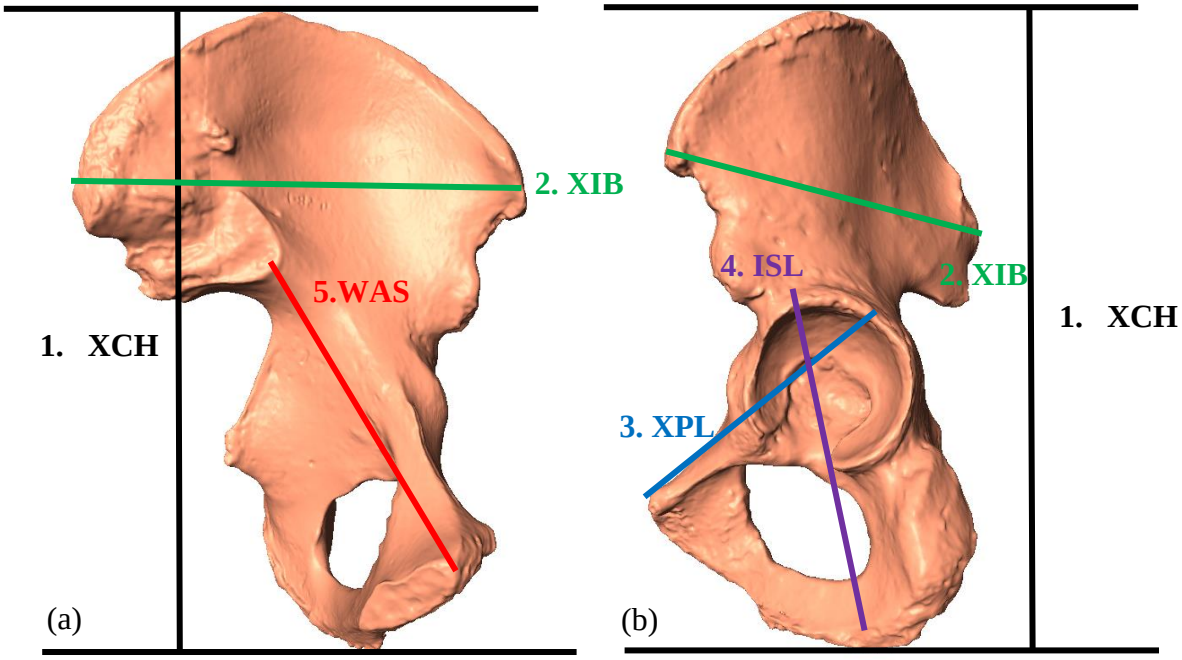


Figure 4.1 Location of linear measurements for Function 1; (a) anterior view b) lateral view

	Measurement	Description
1	*Maximum Coxal Height* (XCH)	The maximum distance from the most superior point on the iliac crest to the most inferior point on the ischial tuberosity.
2	*Maximum Iliac Breadth* (XIB)	The maximum distance from the anterior superior iliac spine to the posterior superior iliac spine.
3	*Maximum Pubis Length* * (XPL)	The maximum distance between the most superior point on the symphyseal face to the farthest point on the superioposterior rim of the acetabulum.
4	*Ischial Length* * (ISL)	The maximum distance from the point on the acetabular rim where the iliac blade meets the acetabulum (most anteromedial point) to the most mediodistal point on the epiphysis of the ischial tuberosity.
5	*Minimum Apical Border to Symphysion* (WAS)	the minimum distance from the anterior apex (anterior border) of the auricular surface to the most superior point on the pubic symphyseal face.

* Do not use if the acetabulum has excessive osteophytosis.

Table 4.3 Description of linear measurements from Langley et al. (2016)

Descriptive statistics for each measurement, and for both right and left coxal bones, are presented in Table 4.4. Right and left measurements were averaged in the study population before being applied to discriminant function analysis (DFA) and logistic regression modeling (LRM) (IBM SPSS Statistics v. 25, SPSS Inc. 2017). Coxal bone completeness, represented by omitting measurements from Chapter 3's initial five measurements, were categorized as approximately 75%, 50-25%, and 25% complete. The initial five linear measurements (Function 1) represent a coxal bone that is approximately 75% complete. Omitting two measurements (Function 2) represents 50-25% completeness and three omitted measurements (Function 3) represents approximately 25% completeness. Measurements omitted from Chapter 3 were also tested in DFA and LRM to determine if they could be useful as measurements in this study.

Functions 4 – 8 use measurements in the ilium not associated with the landmarks in Chapter 2, such as *Maximum Ischiopubic Ramus Length* (XIRL), "the distance from the most inferior point on the symphyseal face to the most distant point on the ischial tuberosity" (Langley et al., 2016:78), *Minimum Iliac Breadth* (WIB), "the minimum distance measured from the area below the anterior inferior iliac spine to the most inward curvature of the greater sciatic notch" (Langley et al., 2016:77), and a new measurement, the *Acetabular Diameter* (AcetD). Here acetabular diameter is defined as the distance between landmarks 16 and 17 that demarcate end points of XPL and ISL (Robertson et al., 2019) rather than maximum acetabular diameter (Dixit et al., 2007; Mahakkanukrauh et al., 2017; Quartrehomme et al., 2017; Schulter-Ellis et al., 1983).

This study also investigated the validity of sectioning points. The strategy for establishing sectioning points for DFA was as follows: 1) Male and Female ranges cannot overlap; 2) the range between Male and Female sectioning points must allow space for a classification of undetermined sex; and 3) sectioning points will be determined by either the group centroid, 1 SD from the group mean (a less conservative sectioning point), or 2 SD from the group mean (a more conservative sectioning point). SPSS recommended a sectioning point of 0.5 for LRM, values greater than 0.5 are considered Male and values less than 0.5 are considered Female.

Sample Population											
Sex		**XCH_L**	**XCH_R**	**XIB_L**	**XIB_R**	**XPL_L**	**XPL_R**	**ISL_L**	**ISL_R**	**WAS_L**	**WAS_R**
Female (n = 191)	Min	174.66	177.95	132.78	135.23	102.21	103.01	89.39	88.97	98.11	100.64
	Max	237.69	233.74	178.51	183.58	133.12	135.41	122.02	113.58	145.87	171.24
	Mean	200.49	200.74	154.52	155.85	117.59	117.96	102.09	102.46	121.2	123.56
	Std. Dev.	9.75	9.95	8.20	8.28	5.91	6.01	4.97	4.60	7.96	8.56
	Skewness	0.389	0.341	0.265	0.342	0.131	0.092	0.198	-0.123	0.009	0.921
Male (n = 203)	Min	198.57	199.03	141.01	141.56	105.64	106.10	104.36	105.40	96.96	100.49
	Max	266.01	259.40	188.02	190.00	137.69	138.01	133.55	138.86	139.64	142.91
	Mean	221.27	220.64	159.88	161.01	119.42	119.62	115.13	115.89	116.90	118.67
	Std. Dev.	11.25	10.66	8.94	8.84	5.92	5.99	5.33	5.26	7.57	7.75
	Skewness	0.674	0.455	0.462	0.455	0.203	0.189	0.329	0.401	0.106	0.141
Test Population											
Sex		**XCH_L**	**XCH_R**	**XIB_L**	**XIB_R**	**XPL_L**	**XPL_R**	**ISL_L**	**ISL_R**	**WAS_L**	**WAS_R**
Female (n = 60)	Min	179.19	180.36	125.29	137.49	103.38	104.36	92.56	92.89	92.65	94.35
	Max	210.01	208.34	171.17	171.49	127.33	129.48	110.02	112.15	137.70	138.10
	Mean	197.44	197.42	151.52	153.33	115.82	116.35	100.38	101.30	119.98	121.78
	Std. Dev.	7.02	6.59	7.14	7.11	5.31	5.45	3.76	3.94	7.18	6.89
	Skewness	-0.416	-0.433	-0.551	-0.049	-0.068	-0.004	0.221	0.266	0.098	0.016
Male (n = 60)	Min	195.59	197.50	141.58	142.03	107.91	106.37	102.34	104.13	92.65	94.35
	Max	240.56	240.89	174.61	178.05	128.98	129.54	127.05	128.72	136.69	137.94
	Mean	218.44	218.28	157.09	158.44	117.96	118.24	114.48	115.49	115.39	116.60
	Std. Dev.	9.80	9.62	7.58	7.63	5.55	5.63	5.76	5.26	7.32	7.30
	Skewness	0.076	0.044	0.103	0.011	0.132	-0.099	0.185	0.182	-0.118	0.034

Table 4.4 Descriptive statistics for the sample and the test population comparing values from right and left coxal bone dimensions for Females and Males. All measurements are in mm.

4.4 **Results**

4.4.1 Repeatability by Inexperienced Observers

Functions 1-3 performed the best in DFA prediction and when tested on an independent population, it would be interesting to see how well these equations perform when used by inexperienced observers. The three functions were applied to the same coxal bone by four observers for 30 specimens and four of the 30 specimens were consistently misidentified by at least three of the four observers. In total, 15 misidentifications were made resulting in an overall inter-observer accuracy of 87.5% (105/120). Function 3 misidentified more Males as Female (n = 8) compared to Males as Intermediate (n = 3), which is not surprising given the narrowness of its sectioning point. Overall accuracy of Function 3 is 66.7% (20/30). Function 2 misclassified more Males as Intermediate (n = 8) and only one Male as Female, resulting in an overall accuracy of 70% (21/30). Function 1 correctly classified sex the most often, 23 times out of 30 (76.7%), with 5 Males being misclassified as Intermediate, and 2 Males misclassified as Female.

4.4.2 Discriminant Function Analysis

The following is a descriptive account of the Canonical DF coefficients generated by the sample population and applied to the test population as DF equations using the unstandardized coefficients. Group membership resulting from DF equation results were evaluated either as 2 SD from the Male or Female group centroids (Functions 1, 2, 3, 7 and 8), or 1 SD from the group centroid (Functions 4, 5 and 6). Sectioning points used in each function are found in Table 4.5. The accuracy of each DF equation on the test population is listed in Table 4.6 and compared to

the predicted accuracy of DFA on the study population. Total accuracy results are listed for each function as well as accuracy by sex.

Function	Variables	Standardized coefficient	Unstandardized coefficient	Centroids (SD)	Sectioning point
1	XCH	0.842	0.082	F: -2.39 (0.918)	F: < -0.554
	XIB	-0.251	-0.030	M: 2.25 (1.072)	M: > 0.106
	XPL	-0.982	-0.167		
	ISL	1.150	0.233		
	WAS	-0.556	-0.073		
			(Constant) -9.464		
2	XPL	-0.970	-0.164	F: -2.21 (0.910)	F: < -0.39
	ISL	1.619	0.328	M: 2.08 (1.078)	M: > -0.08
	WAS	-0.391	-0.051		
			(Constant) -10.102		
3	XPL	-1.271	-0.216	F: -2.12 (0.928)	F: < -0.264
	ISL	1.672	0.339	M: 1.99 (1.064)	M: > -0.138
			(Constant) -11.402		
4	XCH	1.451	0.141	F: -1.54 (0.941)	F: < -0.599^{+}
	XIB	-0.402	-0.047	M: 1.45 (1.053)	M: > 0.397^{+}
	XIRL	-0.798	-0.151		
			(Constant) -7.710		
5	XCH	1.082	0.105	F: -1.44 (0.910)	F: < -0.530^{+}
	XIB	-0.994	-0.117	M: 1.36 (1.078)	M: > 0.282^{+}
	WIB	0.602	0.157		
	AcetD	0.148	0.051		
			(Constant) -15.329		
6	XCH	1.110	0.108	F: -1.43 (0.930)	F: < -0.500^{+}
	XIB	-0.998	-1.118	M: 1.35 (1.062)	M: > 0.288^{+}
	WIB	0.655	0.171		
			(Constant) -15.422		
7	XPL	-0.792	-0.135	F: -2.16 (0.941)	F: < -0.278
	WPL	-0.281	-0.061	M: 2.04 (1.052)	M: > -0.064
	ISL	1.564	0.317		
	XIRL	-0.204	-0.039		
			(Constant) -10.531		

8	XCH	0.687	0.067	F: -2.06 (0.932)	F: < -0.196
	WAS	-0.809	-1.06	M: 1.93 (1.060)	M: > -0.190
	ASISS	-0.444	-0.043		
	ISL	0.809	0.164		
			(Constant) -13.129		

+ 1 standard deviation from the centroid

Table 4.5 Canonical discriminant function coefficients.

4.4.3 Function 1

Function 1 represents the DF from Chapter 3 using all five measurements and is appropriate for coxal bones that are at least 75% complete or greater, or where damage to coxal bones does not impact the measurement points for iliac breadth. In the initial DFA one White Male was misclassified as Female with an overall accuracy of 99.7%. When this function was applied to the test population, the sectioning point was set to 2 SD from the group centroid; values greater than 0.106 were considered Male and values less than -0.554 were considered Female. One Black Female (age = 40) fell between sectioning points resulting in a classification of indeterminate sex. The resulting accuracy for this function on the test population was 99.2% (Females = 98.3%, Males = 100%). When the sectioning point for this function was arbitrarily set to Females < 0.0 and Males > 0.1, this formula was 100.0% successful at identifying Females and Males.

4.4.4 Function 2

This function represents coxal bones between 50-25% complete or if damage to the iliac spines, iliac crest, and ischiopubic ramus prevents measurements XCH and XIB from being taken (Figure 4.2). DFA predicted sex in function 2 at 99.2% accuracy in the sample population; three Females were misclassified (Females = 98.4%, Males = 100%). When function 2 was applied to the test population, using 2 SD from the centroid as the sectioning criterion, one White Male (age = 63) was misclassified resulting in the same level of accuracy as the sample population (99.2%; Females = 100%, Males = 99.2%). Values above -0.08 are considered Male and values below -0.39 are Female.

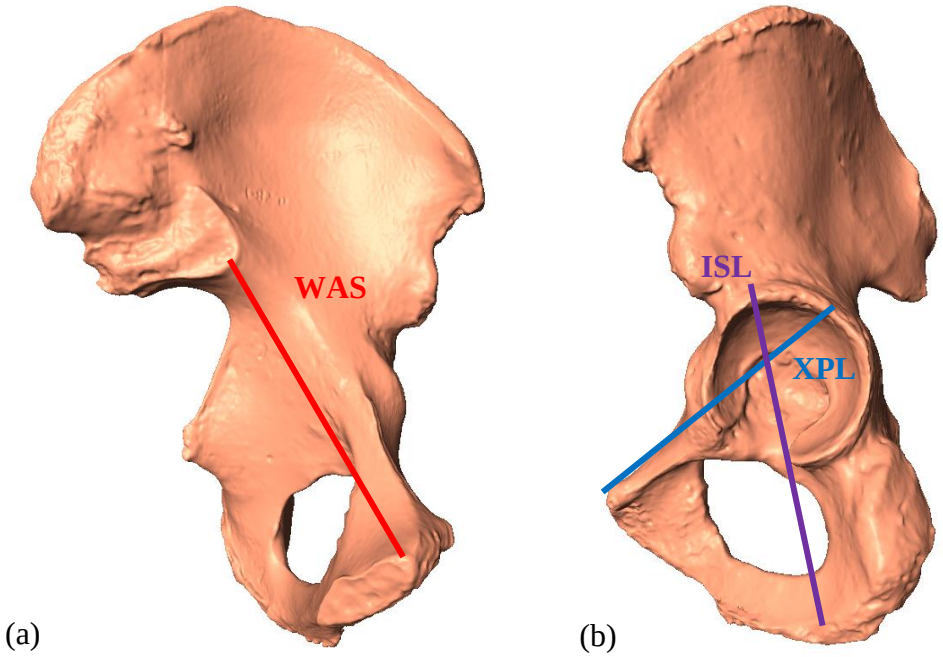


Figure 4.2 Location of linear measurements for Function 2; (a) anterior view, (b) lateral view

4.4.5 Function 3

Function 3 estimates sex using maximum lengths of the pubis (XPL) and ischium (ISL). Superior portions of the symphyseal face, distal ischial tuberosity, and the superior and posterior borders of the acetabulum must be present, representing at least 25% coxal bone completeness. DFA correctly predicted sex with 99% accuracy and two Males (99%) and two Females (99%) were misclassified. When this function was applied to the test population, sex was correctly predicted in all cases using the same sectioning criterion as functions 1 and 2, 2 SD from the centroid; values > -0.138 are Male and values < -0.264 are Female. If a less conservative sectioning strategy was applied, namely if Female values are less than 0.0 and Male values are greater than 0.0 (Females < 0.0 > Males), accuracy is maintained at 100%.

The success of this function is not surprising as it involves maximum pubic and maximum ischial lengths that have long been associated with accurate sex estimation techniques (Albanese, 2003; Brůžek, 2002; Budinoff & Tague, 1990; Kurki, 2007; Patriquin et al., 2005; Schulter-Ellis et al., 1983, Vacca & Di Vella, 2012; Washburn, 1948). The difference with this method is that maximum ischial and pubis lengths were measured from the acetabular ridge the furthest away from the terminus points of either the ischium or the pubis, simlar to measurements taken by Schulter-Ellis et al. (1983).

4.4.6 Function 4

An alternate scenario were a coxal bone could be considered 50-25% complete would include a more intact ilium where total coxal bone length (XCH) and ilium breadth (XIB) measurements could be obtained, and an ischiopubic ramus intact at the inferior boarder of the

pubic symphysis maximum and the ischial tuberosity to obtain the measurement for ischiopubic ramus length (XIRL) illustrated in Figure 4.3. The accuracy of this DFA is among the lowest in the study population at 92.9% (Female accuracy 179/191 = 93.7%; Male 187/203 = 92.1%). One SD from the centroid was selected as opposed to 2 SD as in previous functions because 2 SD created an overlap of Male and Female values, which invalidates the first criteria for establishing sectioning points (see Section 4.3.2). When this function was applied to the test population using a sectioning criterion of 1 SD from the centroid, the predicted accuracy approached that of the study population at 90.0% (Female accuracy 53/60 = 88.3%, Male accuracy 55/60 = 91.7%). Sectioning point values above 0.397 are considered Male and values below -0.599 are considered Female. Two Males (Black = 1, American = 1,) and two White Females were misclassified. Five Females (White = 4, Black = 1) and three White Males fell between sectioning points.

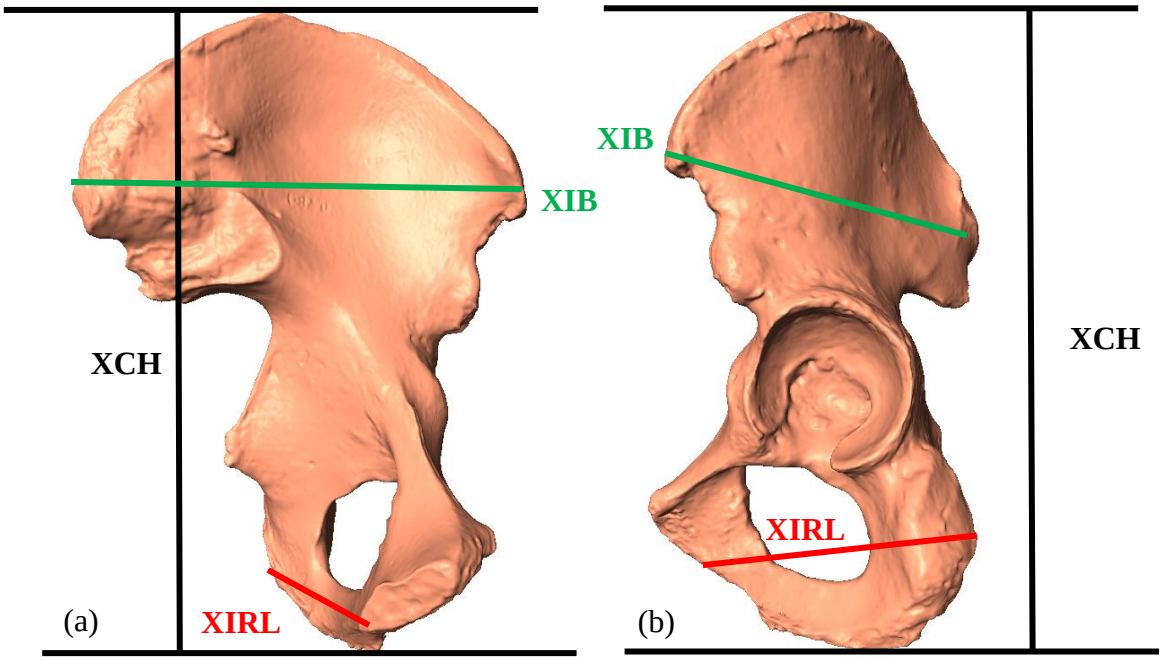


Figure 4.3 Location of linear measurements for Function 4; (a) anterior view, (b) lateral view

4.4.7 Function 5

Measurements included in Function 5 (Figure 4.4) combines XCH and XIB with minimum iliac breadth (WIB) measured between the most inward curvature of the greater sciatic notch to just below the anterior inferior iliac spine (Langley et al., 2016) and acetabular diameter (AcetD) measured between landmarks 16 and 17 in the 17-landmark configuration of Chapter 2 (Robertson et al., 2019). The additional measurements improved sex prediction in the study population by 3.3% compared to Function 4. Function 5 correctly predicted sex in 93.2% of Females (178/191) and 92.1% of Males (187/203). Using the same sectioning criterion as Function 4, 1 SD away from the centroid, as applied to the test population values above 0.282 are Male and values below -0.53 are Female. One White Female (age = 65) was misclassified as Male, and 16 individuals (Females = 3, Males = 13) fell within the area of undetermined sex. A less conservative sectioning point of Female < 0 > Male, correctly classified 98.3% of Females (59/60) and 95% of Males (57/60). Three Males (Black n = 1, American n = 2) and one White Female were misclassified using the less conservative sectioning point. This function could also be used on coxal bones between 50-25% complete.

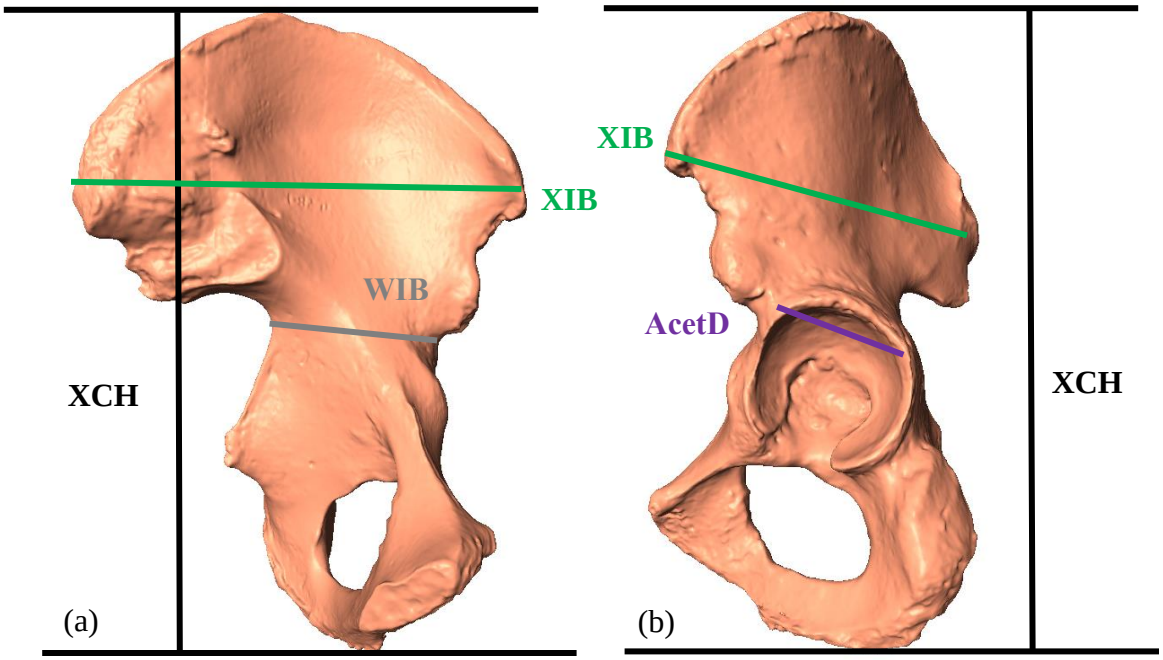


Figure 4.4 Location of linear measurements for Function 5; (a) anterior view, (b) lateral view

4.4.8 Function 6

To improve the sex prediction of function 5, on coxal bones between 50-25% compete, AcetD was removed as it does not measure maximum acetabular diameter and therefore is the least consistent measurement with previous studies (Albanese, 2003; Patriquin et al., 2005; Schulter-Ellis et al., 1983, 1985). Function 6 includes XCH, XIB, and WIB (Figure 4.5). Omitting AcetD from function 5 resulted in a less accurate function, contrary to the predicted effect. Sex prediction declined in the sample population to 90.9% (Females 172/191 = 90.1% accurate; Males 186/203 = 91.6% accurate) from 92.6% in function 5. However, in the test population, Function 6 performed slightly better at accurately predicting sex (86.7%) than Function 5 (85.8%). The ratio of number of Females to Males is also different in Function 6. Just

as many test specimens fell in the area of undetermined sex as in Function 5 (n = 104/120), however, more Females (54/60), and fewer Males (50/60) made up the undetermined individuals in Function 6. DF values above 0.288 are considered Male and values below -0.5 are Female, as in Function 5. A less conservative sectioning point applied to this equation would generate the same results as in Function 5 (93.3%, Males misclassified = 3, Females misclassified = 1).

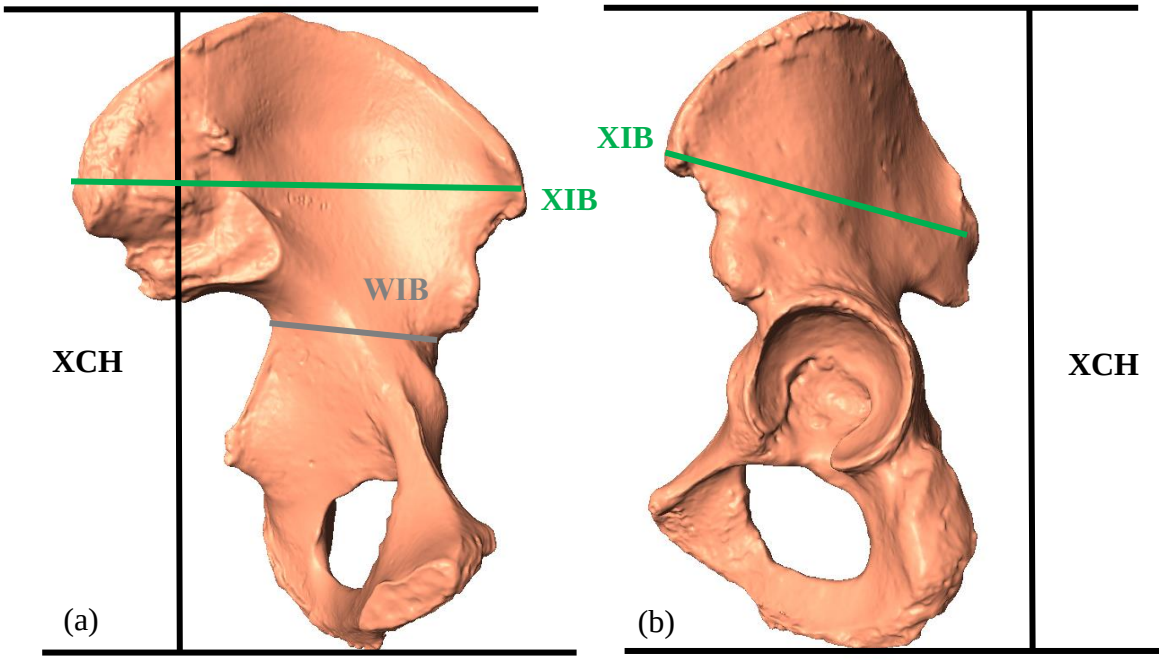


Figure 4.5 Location of linear measurements for Function 6; (a) anterior view, (b) lateral view

4.4.9 Function 7

Function 7 (Figure 4.6) included XPL, ISL, and two measurements from Langley et al., (2016): minimum pubic length (WPL), the distance between the most anterior inferior point on the lunate surface and the most superior point on the symphyseal face (landmarks 16 and 18

respectively in the 32-landmark configuration of Chapter 2) and the maximum ischiopubic ramus length (XIRL), the distance between the most inferior point on the symphyseal face and the most posterior point on the ischial tuberosity (landmarks 19 and 30 respectively in the 32-landmark configuration of Chapter 2). DFA correctly predicted sex with 98.7% accuracy (Females 189/191, Males 200/203) in the study population. In the test population using a sectioning point of 2 SD from the mean, values above -0.064 are Male and values below -0.278 are Female, one White Male (age = 71) fell between the sectioning points and three Males (White = 2, Black = 1) were misclassified as Female. All Females were correctly classified. This formula correctly predicted sex in 96.7% of the test population (Females 60/60, Males 56/60). This function could also be used on coxal bones between 50-25% complete.

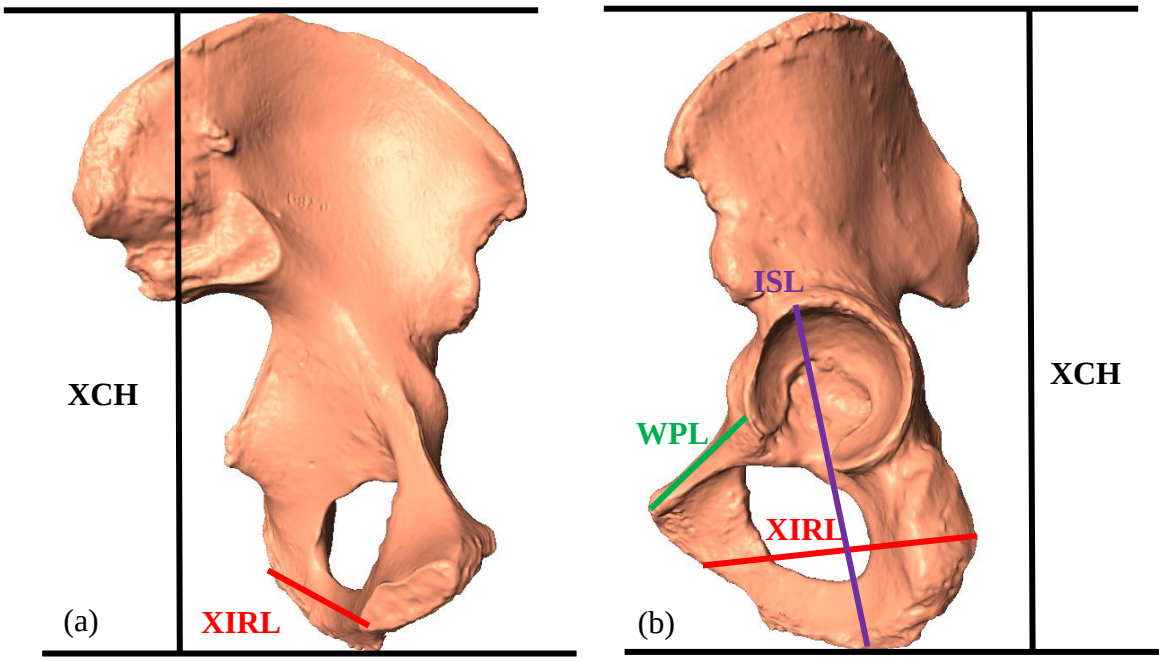


Figure 4.6 Location of linear measurements for Function 7; (a) anterior view, (b) lateral view

4.4.10 Function 8

Function 8 (Figure 4.7)included XCH, ISL, WAS (minimum distance from the most superior point on the symphyseal face to the anterior apex of the auricular surface) and ASISS (distance from the apex of the anterior superior iliac spine to the most superior point on the symphyseal face) from Chapter 3. This function can also be applied to coxal bones between 50-25% complete. DFA correctly predicted sex in the study population with 98.7% accuracy (Females 188/191, Males 201/203). Using two standard deviation from the mean as the sectioning criterion in the test population, values above -0.19 are Male and values below -0.196 are Female. The range of intermediate sex between sectioning points is by far the narrowest among all equations tested. Two Males were misclassified as Female, one White Male (age = 63) and one American Male (age = 27). If a less conservative sectioning criterion was used that did not allow for a classification of intermediate (Female < 0 > Male). the accuracy in the test population would not change, but it would fail the criterion set out at the outset of this study that a range of intermediate sex should be available.

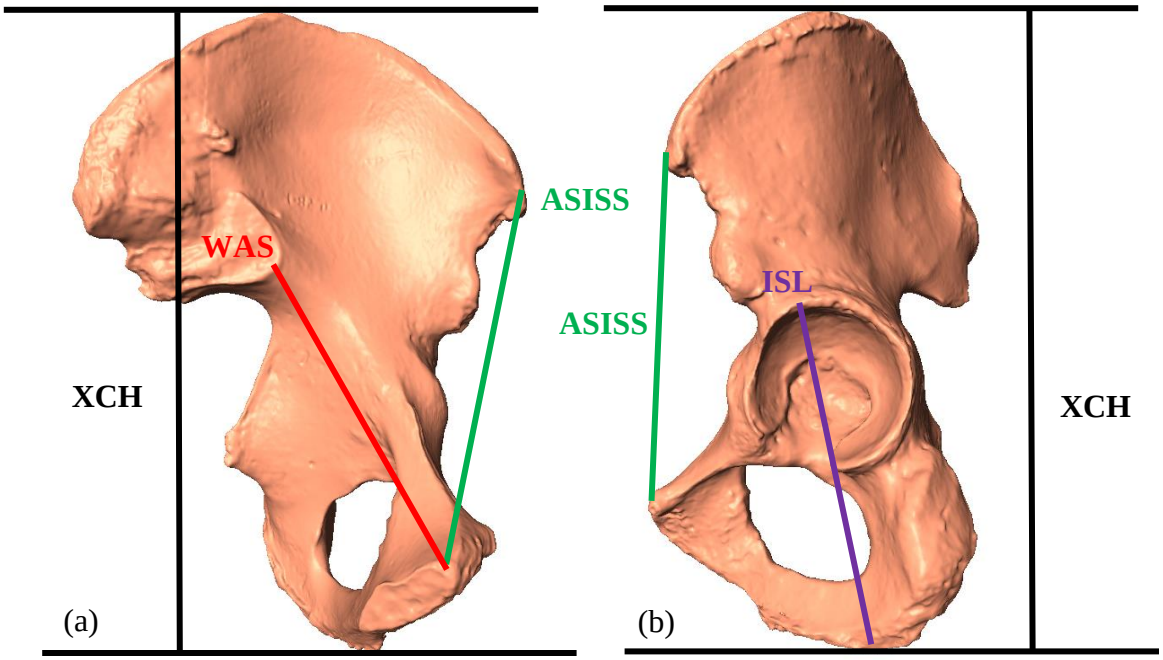


Figure 4.7 Location of linear measurements for Function 8; (a) anterior view, (b) lateral view

Function	DFA (n = 394)	Females (n = 191)	Males (n = 203)	Test Population (n = 120)	Females (n = 60)	Males (n = 60)
1	99.7%	100% (n = 0)	99.5% (n = 1)	99.2%	98.3% (n = 1)	100% (n = 0)
2	99.2%	98.4% (n = 3)	100% (n = 0)	99.2%	100.0% (n = 0)	98.3% (n = 1)
3	99.0%	99.0% (n = 2)	99.0% (n = 2)	100.0%	100.0% (n = 0)	100.0% (n = 0)
4	92.9%	93.7% (n = 12)	92.1 % (n = 16)	90.8%$^{+}$	90.0%$^{+}$ (n = 6)	91.7%$^{+}$ (n = 5)
5	92.6%	93.2% (n = 13)	92.1% (n = 16)	85.8%$^{+}$	93.3%$^{+}$ (n = 4)	78.3%$^{+}$ (n = 13)
6	90.9%	90.1% (n = 19)	91.6% (n = 17)	86.7%$^{+}$	90.0%$^{+}$ (n = 6)	83.3%$^{+}$ (n = 10)

7	98.7%	99.0% (n = 2)	98.5% (n = 3)	96.7%	100.0% (n = 0)	93.3% (n = 4)
8	98.7%	98.4% (n = 3)	99.0% (n = 2)	98.3%	100% (n = 0)	96.7% (n = 2)

+ 1 SD away from the centroid

Table 4.6 Summary of correct sex classification using discriminant functions on study and test populations at conservative sectioning points. Correct classifications represented as a percentage, number of misclassifications in parentheses.

Because Chapter 3 identified WAS as having a significant pattern of difference between right and left coxal bones within the study population, all equations representing fragmented remains were also subjected to a right and left coxal bone comparison (see Appendix D). Function 1 performed the best out of all functions for the left side (98.3%, 118/120), while Function 8 performed second best at 97.5% accuracy (117/120). Function 3 and Function 8 both performed the best on the right side (99.2%, 119/120). See Appendix D for complete results by side.

4.4.11 Logistic Regression Modeling

The first three LRM repeat the measurements from the first three DF (Figure 4.8). The first DF (Table 4.7) correctly classified all individuals in the study and misclassified one Black Female (LR = 39.22) in the test population (Table 4.8). The second DF equation repeated as a LR model, model 2, correctly identified specimens with the same level of accuracy, but when applied to the test population the LR model performed worse (96.7%) than the DF equation (99.2%). One White Female and three Males were misidentified in the test population using LRM2. DF3 and LRM3 performed comparably to each other.

Some measurements performed better in DFA than in LR, and some measurements performed better in LR than in DFA. For this reason, measurements used in Functions 4-8 were not duplicated in Models 4-7. Models 4-7 (Figures 4.9 – 4.12) represent the highest accuracies for LR obtained using coxal bone measurements outlined in Chapter 3 and eliminated measurements in Chapter 2. Model 4 improved the accuracy of Function 4 by including XIRL.

Model	XCH	XIB	XPL	ISL	WAS	XIRL	Constant
1	11.925		-39.871	50.359	-11.589		-1825.677
2			-1.264	2.512	-0.481		-65.977
3			-1.038	1.619			-52.942
4	0.413	-0.148				-0.407	-24.384
5		-0.037	-0.927	1.603		-0.127	-46.465
6	0.555				-0.475	-0.386	-22.925
7			-0.965	1.592		-0.124	-46.830

Table 4.7 Logistic regression of partial landmark configurations. Female values < 0.5, Male values > 0.5.

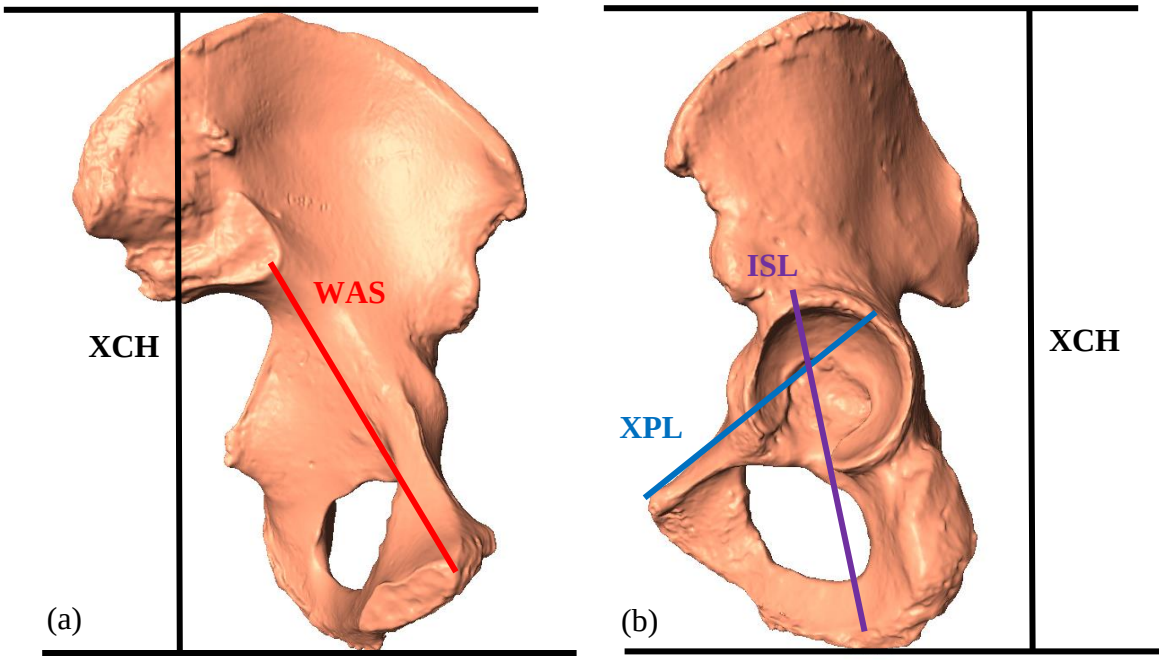


Figure 4.8 Location of linear measurements for Model 1; (a) anterior view, b) lateral view

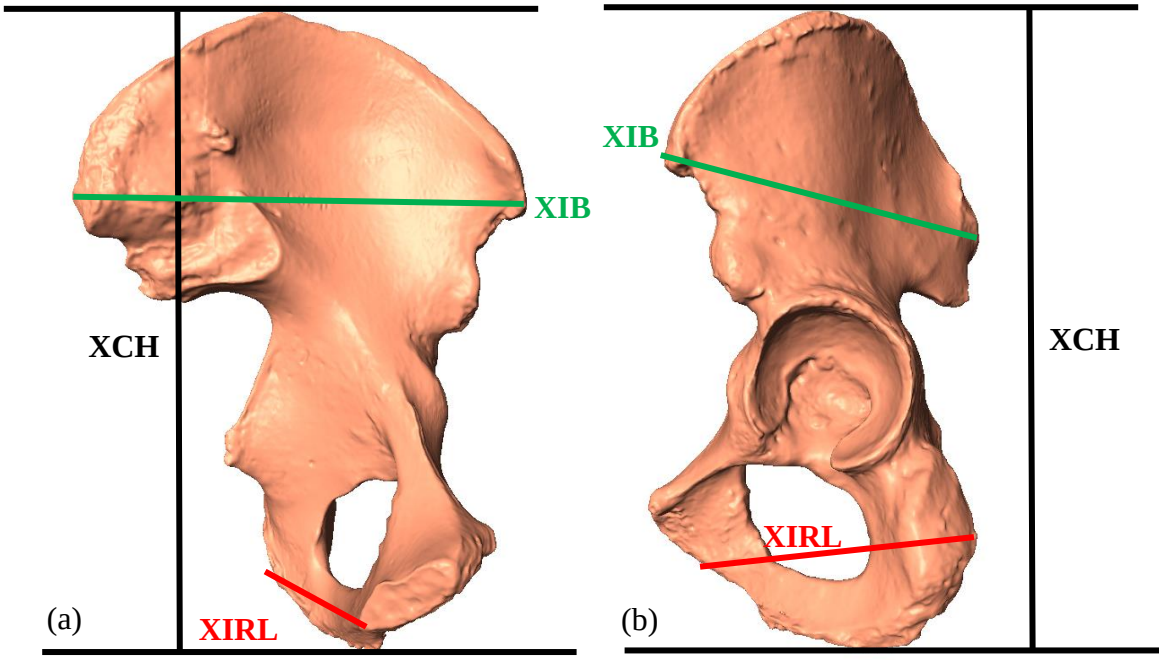


Figure 4.9 Location of linear measurements for Model 4; (a) anterior view, b) lateral view

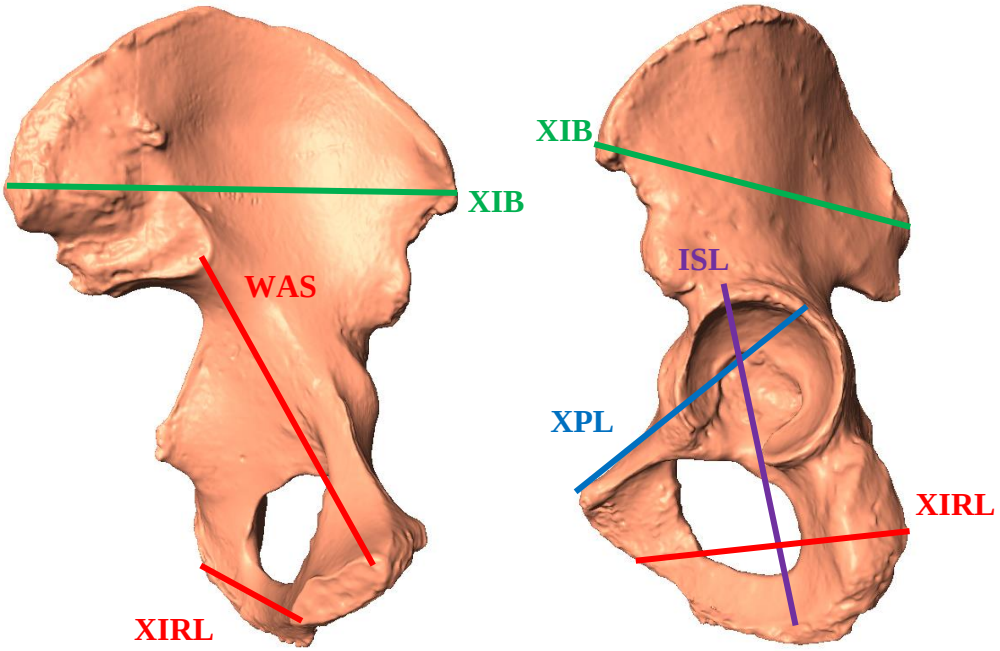


Figure 4.10 Location of linear measurements for Model 5; (a) anterior view, b) lateral view

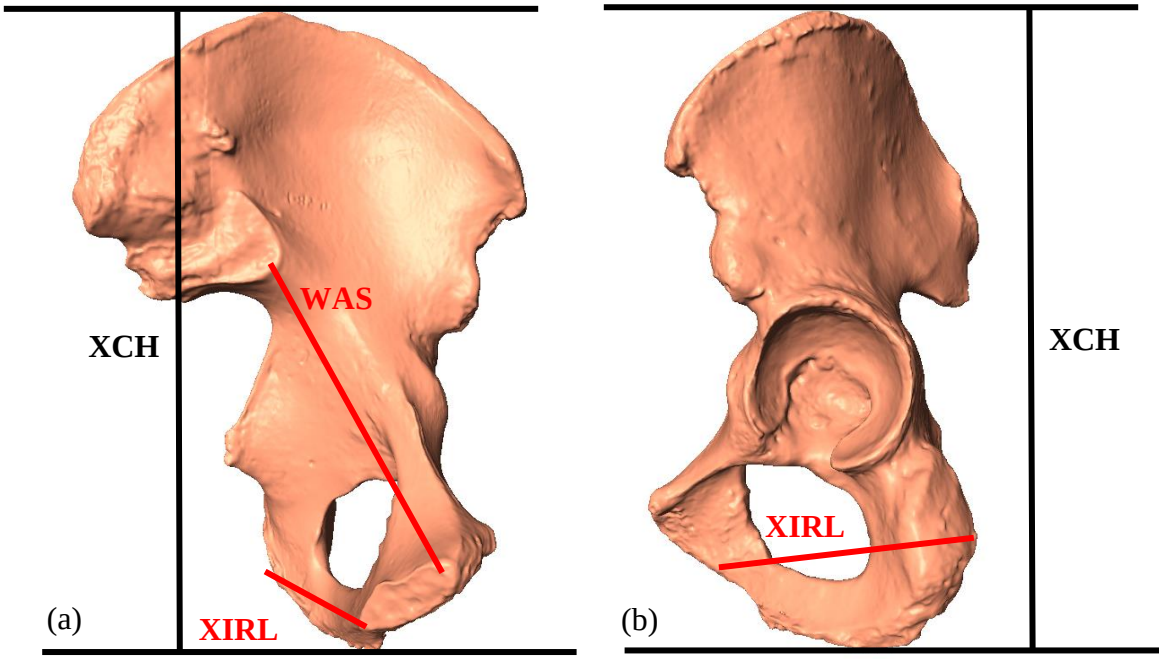


Figure 4.11 Location of linear measurements for Model 6; (a) anterior view, b) lateral view

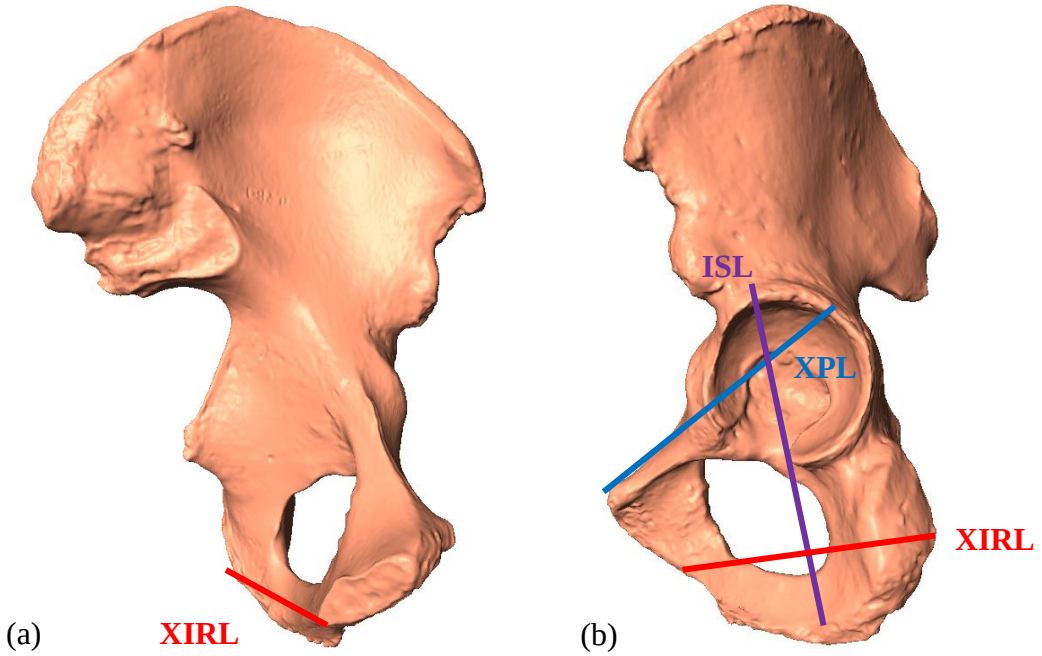


Figure 4.12 Location of linear measurements for Model 7; (a) anterior view, (b) lateral view

Model	LR (n = 394)	Females (n = 191)	Males (n = 203)	Test Population (n = 120)	Females (n = 60)	Males (n = 60)
1	100%	100% (n = 0)	100% (n = 0)	99.2%	98.3% (n = 1)	100% (n = 0)
2	99.2%	99.5% (n = 1)	99.2% (n = 2)	96.7%	98.3% (n = 1)	95.0% (n = 3)
3	99.2%	99.0% (n = 2)	99.5% (n = 1)	100%	100% (n = 0)	100% (n = 0)
4	92.8%	93.2% (n = 13)	92.6% (n = 15)	97.5%	98.5% (n = 1)	96.7% (n = 2)
5	99.0%	99.0% (n = 2)	99.0% (n = 2)	100%	100% (n = 0)	100% (n = 0)
6	98.0%	97.4% (n = 5)	98.5% (n = 3)	98.3%	100% (n = 0)	96.7% (n = 2)
7	99.0%	99.0% (n = 2)	99.0% (n = 2)	100%	100% (n = 0)	100% (n = 0)

Table 4.8 Summary of correct sex classification using logistic regression models on study and test populations. Correct classifications represented as a percentage, number of misclassifications in parentheses.

4.5 **Discussion**

In the best-case scenario LRM1, or the model from Chapter 3, is one of the best predictor of sex correctly predicting all members of the sample population and 99.2% of the test population (Females 59/60, Males 60/60). Equally useful is DF 1 or the DF model from Chapter 3; all but one Male individual was misclassified in the study population (99.7% accuracy) and one Female was misclassified in the test population (99.2% accuracy). The third-best option is LRM 3 where only the maximum pubis length and maximum ischial length are used. Model 3 correctly identified all but 2 Females in the study sample and all but 1 Male (99.2%), all individuals from the test population were correctly identified. The fourth best option is DF 3; it used the same measurements as LRM 3 but performed slightly lower in comparison. Two Females and two males were misidentified in the study population using DF 3 (99.0% accuracy), however all the individuals in the test population were correctly identified. Logistic regression models that performed equally well as DF 3 include LRM 5 and LRM 7. Both models used measurements involving the pubis and ischium (XPL, ISL, XIRL) and accurately predicted sex in 99.0% of individuals in the study population and 100% of individuals in the test population. The inclusion of maximum iliac breadth (XIB) as a measurement in LRM 5 did not impact the predictive power of the model. A similar effect could be said to occur among the discriminant functions. Functions 2 and 3 perform just as well without XIB as function 1 with XIB.

Logistic regression is often preferred over discriminant function analysis for group prediction because LR separates bivariate data into discrete groups without much room for “intermediate” or “unknown” sex classifications (Albanese et al., 2008). However, there are advantages to allowing for “unknown sex” classifications, as provided by DF sectioning points,

the most particular advantage is avoiding incorrect sex classification. For instance, DF 2 and LRM 2 use the same measurements (ISL, XPL and WAS) generate the same sex predictions in the study population (99.2% accurate), but when applied to the test population. The sectioning point for DF 3 was set 2 standard deviations from the group centroid. Females values were below -0.39 and Male values were above -0.08. The misclassified individuals in the study population using DF 2 were all Female (n = 3). When this same sectioning point was applied to the test population only one individual, a White Male, was misclassified. Logistic regression in SPSS has a fixed sectioning point of Females < 0.5 > Males. The LRM 2 sectioning point is different from the DF 2 sectioning point resulting in different individuals being misclassified by the two equations. Using LRM 2 one Female and two Males were misclassified. Although the accuracies between each equation remained the same (99.2%) the individuals being misclassified differed. Applying LRM 2 to the test population failed to correctly identify the sex of a White Female individual and two additional Males compared to DF 2, resulting in an overall accuracy of 96.7%. If the LR sectioning point is adjusted to Female < 0.0 > Male one fewer Male would be misclassified in the test population, however, the result in the test population for LRM 2 would remain less accurate than DF 2. This is the only instance in this study where LRM sectioning point was the least effective.

The results of this study demonstrate that DF and LRM can generate highly accurate sex estimation equations using the same or similar measurements. Discriminant function equations have an advantage over logistic regression modeling, however, in its ability to reduce the potential number of incorrect sex classifications by generating an “indeterminant sex” or “unknown sex” range between sectioning points. DF provides the option of having a category of indeterminate sex in sex estimation where LRM does not. Being able to choose between the

option to include or exclude an intermediate sex category provides flexibility to suite an array of biological archaeological needs. Forensic anthropological studies may wish to include a buffer against incorrect sex classification, in which case the DF equations presented in this study would be extremely useful. Outright misclassification can occur when landmarks are misidentified and measurements are inaccurate, or when pathological bone obscures the surface of the acetabulum and consequently obscures measurements for maximum ischium and pubis lengths. Evaluating sex estimation methods based on the distance between DF cut off points could indicate how likely or unlikely measurement error could lead to sex misclassification. Sectioning points that are closer together might encourage sex misclassification whereas sectioning points that are more widely spaced might pick up measurement error as an intermediate sex classification.

The landmark variation identified in Chapter 3 was greater for the ISL measurement (see Appendix D). Averaged measurements could lessen the effects of over or under measurements by observers. The discriminant fuctions presented in this study are applicable to both the right and left coxal bones, and it is recommended that many coxal bone measurements should be obtained, using protocols outlined by Langley et al. (2016) and applied to as many appropriate DF equations to narrow down potential error. The benefit of averaging right and left coxal bones measurements prior to using the discriminant function equations, as performed in this study, is to account for individual variation in the form of fluctuating asymmetry, account for sex-based variation, and buffer against possible observer error. Function 1 performed the best on the left side while Function 3 performed the best on the right side.

Differences in accuracies between the study population, test population, and inexperienced observers have as much to do with variations in sample size as they do with the ability of each formula to discriminate sex-based form. Regardless of sample size, the

measurements associated with Chapter 2's final landmark configuration perform better at correctly categorizing sex than osteometrics involving landmarks that have demonstrated redundancy and inefficiency at discriminating sex (Robertson et al., 2019). Function 1, from Chapter 3, is a more reliable function than the others in this study, yielding the highest sex prediction in DFA, the greatest number of correctly identified specimens in the test population when applied the left side, and the greatest number of correctly identified specimens among inexperienced observers. The superior performance of this function is likely because it uses all five measurements, which accounts for the complexity of variation as well as skeletal sex in the coxal bone, noted in Chapter 3.

The presence of the WAS measurement accounts for variations in sex, population, and fluctuating asymmetry, which serves to improve sex estimation when applied in conjunction with multiple sex-based measurements. Function 2, which also included WAS in the DF equation, was reliable and repeatable in the absence of coxal bone height (XCH) and iliac breadth (XIB). However, once WAS was removed from the DF equations in Functions 4-7, accuracies dropped by 9.9-8.3%. Once WAS was reintroduced into Function 8, accuracies elevated slightly in the test population, but not to the same level as Functions 1 and 2. The logistic regression models faired better with the inclusion of XIRL and the exclusion of XIB. Models 5 and 7 performed better than Model 2 in the test population without XIB. Function 3 and LRM 3 were the most successful at predicting sex using the maximum pubic length and ischial length and outperformed previous studies using these measurements or a transformation of these measurements into an index. To account for overall population variation in the coxal bone, several authors have taken to developing ancestry specific equations as an attempt to improve the accuracy of osteometric sex estimation methods (Dixit et al., 2007; Mahakkanukrauh et al.,

2017; Mahato, 2010; Vacca & Di Vella, 2012). Osteological theories of racialized differences in the skeleton have led to independent sex estimation formulas for individuals classified as White, Black or other. Not surprisingly, separated populations with varying maximum or minimum thresholds in body size yielded different discriminant equations, thus completing the tautological argument of racialized osteological variation, despite the evidence that the results of many racialized discriminant equations overlap. In Washburn's (1948) work on the ischiopubic index, Black Males were the benchmark for shortest pubic bone and White Females as the longest with White Males and Black Females filling in that range whereas Black Females had the shortest ischium length and White Males the longest, with the remaining two groups filling in the range. Despite the sexes in both groups clearly demonstrating overlapping ischiopubic index results, Washburn recommends treating each population separately in order to improve sex prediction rather than improving the index.

A common limitation found in sex estimation is retaining high levels of accuracy when estimating sex in fragmented coxal bones (Arun et al., 2012; Brůžek, 2002; Buikstra and Ubelaker, 1994; Rogers & Saunders, 1994). For example, Albanese (2003) applied a revised method of measuring the superior pubis ramus length, however, one of Albanese's measurement landmarks for pubis length and ischium length is located at the acetabulum on the infroanterior terminus of the lunate surface. This same landmark was tested in Chapter 3 representing minimum pubic length and was found not to contribute as much sex-based differential as the maximum pubic length, although Albanese's method improved intraobserver measurement error. Schulter-Ellis et al. (1983, 1985) improved the predictive power of Washburn's (1948) index by using measurement points for the anterior aspect of the ischium to a point on the superior rim of the acetabulum opposite the ischium. Additionally, individuals falling within overlapping Male

and Female ranges were re-estimated using femoral head diameter. This second step further improved accuracy of sex estimation to 97% in a population classified as Black and 98% in a population classified as White (Schulter-Ellis et al., 1983, 1985). Discriminant function equations for these two populations tested separately were, according to the authors, "essentially equal" (Schulter-Ellis et al., 1985:184). Schulter-Ellis et al.'s (1983, 1985) discriminant function equations were derived independently by testing both populations separately, however, their sample size was half that of Washburn's population. The results from both Schulter-Ellis et al.'s populations could be a result of their small sample size. An alternative hypothesis is that separate equations between individuals classified as Black and White might be redundant.

Schulter-Ellis et al. (1983, 1985) did not publish raw data for their acetabular and pubic measurements, but rather presented an index between these measurements in combination with ischial length in their DFA. For this reason, the sectioning points between Schulter-Ellis et al.'s studies and Function 3 from the current study were compared. Schulter-Ellis et al.'s 1983 study measured 100 individuals classified as Black and Schulter-Ellis et al.'s 1985 study measured 100 individuals classified as White, both from the Terry Collection at the Smithsonian. The DF equation generated by Schulter-Ellis et al. 1983 and 1984 that is most comparable to Function 3 of the current study is derived from the index between the diameter of the acetabulum and the distance between the pubic symphysis to the anteroinferior acetabular rim, and the maximum ischial length between the anterior of ischial tuberosity to the furthest extent of the acetabular rim. Using Schulter-Ellis's sectioning points from both Black and White DF equations, the difference between Male scores (> 0.3) and Female scores (< -0.1) is 0.4. Function 3 of the current study generated a similar range between Male and Female scores (m = > -0.138, f = < -0.554), with a difference of 0.42. The segregated studies of Schulter-Ellis et al. (1983, 1985) did

not improve cut off points between Male and Female categories and generated a lower prediction value than DF3 of the current study with a similar intermediate range and predictive value 2% among individuals classified as Black and 1% among individuals classified as White. LRM3 had a binary cut off point, Females < 0.5 > Males, and performed as well as DF3. Although it is acknowledged that the number of individuals classified as Black is lacking considerably in the current study compared to Schulter-Ellis et al. (1983), the functions and models presented here should call into question the scientific reasoning behind segregated sex estimation studies. Schulter-Ellis et al. claim that "[g]iven the range of human variation, it is unrealistic to expect ever to achieve 100% predictatility" (1983:169). This might be true, however, with meaningful landmarks and repeatable metrics, one can get pretty close using DFA and reach 100% predictability using LR.

Vacca and Di Vella (2012) did not use ischium length in their study and the reduced predictability of their method reaffirmed the idea that sex equations must differentiate between ancestry. However, upon closer examination of Vacca and Di Vella's work, if work was conducted on their model to determine where acceptable levels of human variation lie among their study sample, it would have alleviated the need to discriminate sex estimation formulas based on ancestry. When interpreting human variation, there is no consensus on what is considered acceptable levels of variation. Variation exists within the body as fluctuating asymmetry, as well as between people as either sex, or population variation. This study used fluctuating asymmetry as a benchmark for acceptable levels of intra- and inter-observer variation as reported in Chapter 3. The presence of variation among populations has not been demonstrated as statistically significant to warrant the need for separate sex estimation equations. Reflections on human adaptation, evolution and correlations with body form have highlighted a

more likely scenario that human variation is better understood as body size variation rather than ancestry (Arun et al., 2012; Betti et al., 2013; Kurki, 2011; Ridgeway et al., 2011).

Tague (1994) identified age-related contributors to variations in coxal bone form. Females younger than 24 years tend to have shorter *linea terminalii* and individuals older than 18 years accumulate osteophytes on the ischiopubic ramus, gradually narrowing the subpubic angle in both Males and Females as aging occurs. Unfortunatley, there were no Females younger than 30 in this study due to the age restrictions of the Bass Collection's body donation program, so differences in *linea terminalis* lengths were not a testable variable. Osteophytosis of the ischiopubic ramus could be a factor in this study, however, other authors have found no correlation between age-related changes in coxal bone form and sex misclassification (Mallard et al., 2017). There does not appear to be a correlation between age at death among Males or Females and sex misclassification in this study.

The specimens used in this study also account for a wide variation in body size, which is a better proxy for human variation than ancestry (Betti, 2017; Kurki, 2011; Steyn & Patriquin, 2009). The measurements associated with Function 1 are also found in studies measuring obstetric dimensions in the Female pelvis (Auerbach et al., 2018; Kurki, 2011, 2013). Auerbach et al.'s study suggests the measurement from the anterior apex of the auricular surface to the superiomedial surface of the pubic symphysis (WAS in this study, APL in Auerbach et al.) is not only useful for discriminating between sex, but also for discriminating between populations of high and mid latitudes. Because the discriminant function is set to discriminate against sex and not latitude, Function 1 can be said to be population inclusive. In the previous chapter, WAS also varied signficantly with fluctuating asymmetry.

4.6 **Conclusion**

This study demonstrates that coxal bone landmarks that best predict sex can be useful for improving sex estimation in fragmented coxal bones. High accuracy levels are maintained when only two measurements are used: maximum ischial length (ISL) and maximum pubic length (XPL) (DF 3: study population = 99.0%, test population 100%; LRM 3: study population = 99.2%, test population 100%), when ISL and XPL are accompanied by maximum ischiopubic ramus length (XIRL) (LRM 7: study population = 99.0%, test population = 100%) and maximum iliac breadth (XIB) (LRM 5: study population = 99.0%, test population = 100%), and ISL and XPL are accompanied by minimum apical boarder to the symphysion (WAS) (DF 2: study population = 99.2%, test population = 99.2%). The best measurements for discriminating sex are XPL and ISL using averaged coxal bone values, however ISL is at risk for inter-observer error. Special attention should be paid to taking this measurement. Testing these equations for repeatability on an independent test population validated their applicability for future studies.

Chapter 5: **Conclusion**

5.1 **Summary**

The os coxa is the most sexually dimorphic bone in the human skeleton (Meindl et al., 1985; Phenice, 1969; Singh and Potturi, 1978). The shape of the Female coxal bone reflects the morphological requirements for childbirth not seen in Males (Correia et al., 2005). Universal standards for estimating sex in human skeletons exist, but they are often criticized for being subjective and difficult to duplicate due to differences in population affinity (Armelagos, 2003; Brůžek, 2002; Luo, 1995). Population differences derived from climate adaptation inspired researchers to develop population-specific sex estimation techniques (MacLaughlin & Bruce, 1986; Listi, 2010; Rosenberg, 2002; Steyn & Patriquin, 2009). However, in order to apply these techniques, the population affiliation of skeletonized individuals must first be known or ascertained prior to conducting osteometric sex estimation methods, which is problematic when working with human skeletons from archaeological, fossil, or forensic contexts. Individual skeletal variation resulting from developmental, nutritional, or hormonal influences have also been known to contribute to size and shape differences in bone that have the potential to confound sex identification if not incorporated in sex estimations methodologies (Albert & Greene, 1999; Bateson & Gluckman, 2001; Larsen, 2005; Purifoy, 1981; Sitek et al., 2012; West-Eberhard, 2003).

This study included human skeletal variation in the form of allometric and isometric size, and fluctuating asymmetry to help isolate landmarks that contribute to greater sex estimation

accuracy (98.5%), and accurate (98.7-100%) and repeatable (98.3-100%) discriminant function and logistic regression equations to estimate sex in whole and fragmented coxal bones.

5.1.1 Landmarks

This dissertation outlines how to evaluate landmark points that address specific research questions about bone shape. The landmarks that best address the research question can be applied as linear measurements to evaluate and improve discriminant function equations for estimating sex. The 17-landmark configuration presented in Chapter 2 (Robertson et al., 2019) represented sex-based coxal shape for sex estimation better than previous studies (Bilfeld et al., 2012 - 15 landmarks, 87.62% accuracy; Bytheway & Ross, 2010 - 36 landmarks, 98% accuracy; González et al., 2017 – 28 landmarks, 80% accuracy). The landmarks that bests represent coxal bone sex-based shape are the four iliac spines, ischial spine, apex of the auricular surface, ischial tuberosity superior and inferior, distal ischial tuberosity, symphyseal face height, arcuate eminence, ischiopubic ramus, iliac crest, and acetabular rim posterior superior and anterior superior.

5.1.2 Discriminant Function in Complete Coxal Bones

This dissertation introduces a new method of estimating sex in the os coxa that is accurate to 99.7% in complete coxal bones and repeatable in 99.2% of cases.

$$DF = XCH\ (0.082) – XIB\ (0.030) – XPL\ (0.167) + ISL\ (0.233) – WAS\ (0.073) – 9.464$$

Values above 0.106 are Male and values below -0.554 are Female. This discriminant function can be applied to complete left and right coxal bones and accounts for variation in isometry, allometry, body size (a proxy for population variation), and fluctuating asymmetry (a proxy for individual variation). This discriminant function is also useful for coxal bones that are no less than 75% complete.

5.1.3 Discriminant Functions Using Fewer Measurements

This dissertation also introduces two discriminant functions that can estimate sex in fragmented coxal bones with an option to categorize intermediate sex. Function 2 is useful among coxal bones that have three of the five measurements available.

$$\text{Function 2} = \text{XPL } (-0.164) + \text{ISL } (0.328) - \text{WAS } (0.051) - 10.102$$

Function 2 is reliable and repeatable 99.2% of the time and can be used in sex estimation of fragmented coxal bones where the ischium, pubis, superior rim of the acetabulum, and auricular surface are intact. Values above -0.08 are Male, values below -0.39 are Female.

Function 3 can be used when only the ischial and maximum pubis length measurements are available. Function 3 is reliable 99% of the time and repeatable among all test specimens (100%). Values above -0.138 are Male and values below -0.288 are Female.

$$\text{Function 3} = \text{XPL } (-0.216) + \text{ISL } (0.339) - 11.402$$

5.1.4 Logistic Regression in Complete Coxal Bones

If categorization of intermediate sex is not desirable, one can use the following LRM that correctly identified all individuals in both the study and test populations. Values less than 0.5 are Female and greater than 0.5 are Male

LRM = XCH (11.925) – WAS (11.589) – XPL (39.871) + ISL (50.359) – 1825.677

5.1.5 Logistic Regression Using Fewer Measurements

When the coxal bone height measurement is not available LRM 3 can be used as an alternate measure with comparable results to the initial logistical regression model.

Model 3 = XPL (-1.038) + ISL (1.619) – 52.942

If coxal bone completeness requires the use of alternate measures, LRM 5 and LRM 7 both include maximum ischiopubic ramus length in the model to maintain high levels of accuracy.

Model 5 = XIB (-0.037) – XPL (0.927) + ISL (1.603) – XIRL (-0.127) – 46.465

Model 7 = XPL (-0.965) + ISL (1.592) – XIRL (0.124) – 46.830

All functions and models are variation-inclusive and could be applied to an array of skeletal populations. These methods demonstrate high predictive values and are considered repeatable and reliable for forensic and biological anthropological studies. These methods may

also be applicable among temporally distinct groups; however, more work needs to be done to confirm this hypothesis. Universal methods of sex estimation can be developed by taking individual and human shape variation into account.

5.2 Significance and Contributions

This dissertation outlines a method to evaluate coxal bone landmarks for sex-based shape analysis so only the most reliable landmarks can be applied to a highly accurate methods of sex estimation. These methods are most accurate (99.7 – 100%) when all equation measurements are included and ranged between 96-100% accuracy when osteometric measurements varied. The methods outlined in this dissertation can be applied to both biological anthropological and forensic anthropological estimation of sex. The osteometric methods were generated from datasets that include variables of individual and human variation, such as fluctuating asymmetry, stature and weight variation, and wide age at death ranges that should allow the models and functions to be used on a wide variety of forensic and biological anthropological skeletal populations. Future research could explore this outcome in greater detail by applying these methods to skeletal populations from around the world and from different time periods. This research provides great value to forensic anthropology by having a sex estimation method that can be applied to any individual. Currently, forensic anthropology methods are largely routed in population-specific methods of identifying sex and stature (İşcan, 2005), more accurate sex estimation methods that are also variation-inclusive is by far the best options in fields that rely on osteometrics to identify sex in partial and complete skeletons.

This research highlighted the limitations of conducting a geometric morphometric analysis of sex-based shape in the coxal bone without also incorporating fluctuating asymmetry and body size variation into the equation. Previous sex-based studies failed to control for or remove confounding factors of coxal bone shape in their sex estimation, which limited their reliability and repeatability. In order to improve other osteometric methods, they must be free from old biological anthropological assumptions that skeletal variation must be excluded. Human skeletal variation must be viewed from the perspective of all other forms of human skeletal variation to evaluate how those forms are being expressed, and ultimately interpreted, in the skeleton. New methods, like the ones presented in this manuscript, must account for variations in skeletal sex such as individual variation and population variation. This research addresses how methodological assumptions of population-specific sex estimation equations are necessary.

5.3 **Limitations**

5.3.1 Method

The specimens used in this research are contemporary and do not specifically reflect the changes in coxal bone shape related to nutritional deficiencies through time, although some individuals are suspected of having experienced nutritional deficiencies at some point during their youth, adolescence, or young adulthood having been born in the late 19th and early the 20th Centuries. The Bass Collection that made up both the study and test populations in this research is also primarily (98%) made up of individuals classified as White. Although this skeletal population contained the body size (stature and weight) and asymmetrical variation needed to contribute the human variation needed to satisfy the research goals of this study, the lack of

ancestral and temporal variation is insufficient to answer specific questions about the applicability of these methods outside of a contemporary White skeletal population. It is strongly encouraged that these functions and models be further tested on non-contemporary non-white skeletal material to validate their wider applicability. Their use in forensic analysis, however, is strongly encouraged.

Another limitation of this research is a lack of data from individuals with diverse diets. The prevalence of soy in some traditional diets can obscure the effect of testosterone on the developing Male coxal bone and render a more Female shape (Rosenberg, 2002). By testing these discriminant equations on skeletal individuals with a specific diet along with temporal variation will contribute greatly to validating these methods.

5.4 **Future Directions**

Biological anthropology, influenced by feminist theory (third wave feminism and queer theory specifically), can lead the way to changing the discipline and demystify sex. Using queer theory, Geller (2008) identifies the heteronormativity and inflexibility of sex. Biological anthropology takes its assumptions about sex from the biomedical matrix, which Butler (1993) and Fausto-Sterling (2000) have demonstrated, recreates heteronormativity. This dissertation does not destabilize ideas of two sexes in biological anthropology but supports the use of "skeletal sex" rather than "biological sex" to describe the skeletal manifestation of biological variants, supporting the theoretical approaches espoused by gender theorists: that biological sex is multifaceted and non-binary (Fausto-Sterling, 2000; Shapiro, 2005).

The landmark evaluation method outlined in Chapter 2 will be of interest to paleoanthropologists interpreting coxal bone changes independent from sex among fossil hominins, be they differences between temporally contemporary species or the morphological trajectory of non-habitual and habitual bipeds. This method could also inspire new techniques for evaluating biomechanical evolution of the femur and foot. Sex estimation from other bones, such as the femur, humerus, or cranium could also be initiated by the success of evaluating landmarks prior to discriminant function analysis, which would be useful in forensic anthropology. Evaluation of craniometric landmarks could lead to a revaluation of human variation. The coxal bone is unique in the skeleton in that it not only encompasses the biomechanical needs of parturition (childbirth) but also exhibits variations related to nutritional deficiencies and body size, that are imposed (external factors) or inherited (internal factors) from society, therefore this method could also be used to explore nutritional and temporal variation in coxal bone shape among human populations.

The research presented here contributes a new method to the scholarship of biological anthropology and new methodological approaches to evaluating landmarks for shape analysis. Because geometric morphometric analyses of shape are becoming a useful tool in biological anthropology, it is imperative that methodological practices become standardized in this field in order move the scholarship forward to applications in archaeological and forensic contexts (Zelditch et al., 2012). The more studies utilizing and validating the landmarks tested in this research, the more future studies will be comparable to one another, and more information can be gleaned from shape differences within other collections. Standardized landmark evaluation, and eventually the landmarks themselves, will unify methods using shape analysis for data comparisons.

Bibliography

Agarwal, S. C. (2012). The past of sex, gender, and health: bioarchaeology of the aging skeleton. *American Anthropologist, 114*(2), 322-335.

Albanese, J. (2003). A metric method for sex determination using the hipbone and the femur. *Journal of Forensic Science, 48*(2), 1-11.

Albanese, J., Eklics, G., & Tuck, A. (2008). A metric method for sex determination using the proximal femur and fragmentary hipbone. *Journal of Forensic Sciences, 53*(6), 1283-1288.

Anastasiou, E., & Chamberlain, A. T. (2013). The sexual dimorphism of the sacro-iliac joint: an investigation using geometric morphometric techniques. *Journal of Forensic Sciences, 58*(S1), S126-S134.

Arcini, C., Ahlström, T., & Tagesson, G. (2014). Variations in diet and stature: are they linked? Bioarchaeology and paleodietary Baysian mixing models from Linköping, Sweden. *International Journal of Osteoarchaeology, 24*, 543-556. doi:10.1002/oa.2247

Armelagos, G. J. (2003). Bioarchaeology as anthropology. *Archaeological Papers of the American Anthropological Association, 13*(1), 27-40.

Arnqvist, G., & Martensson, T. (1998). Measurement error in geometric morphometrics: empirical strategies to assess and reduce its impact on measures of shape. *Acta Zoologica Academiae Scientiarum Hungaricae, 44*(1-2), 73-96.

Arsuaga, J. L., & Carretero, J. M. (1994). Multivariate analysis of the sexual dimorphism of the hip bone in a modern human population and in hominids. *American Journal of Physical Anthropology, 93*, 241-257.

Arun, M., Nagesh, K. R., & Kumar, G. P. (2012). Estimation of sex from fragments of os coxa by metric analysis. *Australian Journal of Forensic Sciences, 44*(2), 145-153.

Auerbach, B. M., King, K. A., Campbell, R. M., Campbell, M. L., & Sylvester, A. D. (2018). Variation in obstetric dimensions of the human bony pelvis in relation to age-at-death and latitutde. *Journal of Physical Anthropology, 167*, 628-643.

Baab, K. L., Mcnulty, K. P., & Rohlf, J. (2012). The shape of human evolution: a geometric morpohometrics perspective. *Evolutionary Anthropology, 21*, 151-165.

Baumgarten, S. E., & Ousley, S. D. (2015, March 25-28). Estimating sex from the innominate utilizing new measurements. *Poster presented at the 84th Annual Meeting of the American Association of Physical Anthropologists.* St. Louis, Missouri.

Betti, L. (2014). Sexual dimorphism in the size and shape of the os coxae and the effects of microevolutionary processes. *American Journal of Physical Anthropology, 153*, 167-177.

Betti, L. (2017). Human variation in pelvic shape and the effects of climate and past population history. *The Anatomical Record, 300*, 687-697. doi:10.1002/ar.23542

Betti, L., von Cramon-Taubdel, N., Manica, A., & Lycett, S. J. (2013). Global geometric morphometrics analyses of the human pelvis reveal substantial neutral population history effects, even across sexes. *PLoS ONE, 8*(2), e55909. doi:10.137/journal.pone.0055909

Bierry, G., Le Minor, J.-M., & Schmittbuhl, M. (2010). Oval in males and triangular in females? A quantitative evaluation of sexual diimorphism in the human obturator foramen. *American Journal of Physical Anthropology, 141*, 626-631.

Bigoni, L., Velemínská, J., & Brůžek, J. (2010). Three-dimensional geometric morphometric analysis of cranio-facialsexual dimorphism in a Central European sample of known sex. *Homo-Journal of Comparative Human Biology, 61*, 16-32.

Bilfeld, M. F., Dedouit, F., Rousseau, H., Sans, N., Braga, J., Rougé, D., & Telmon, N. (2012). Human coxal bone sexual dimorphism and multislice computed tomography: geometric morphometric analysis of 65 adults. *Journal of Forensic Science, 57*(3), 578-588.

Biwasaka, H., Yasuhiro, A., Sato, K., Tanijiri, T., Fujita, S., Dewa, K., . . . Tomabechi, M. (2012). Analyses of sexual dimorphism of reconstructed pelvic computed tomography images of contemporary Japanese using curvature of the greater sciatic notch, pubic arch and greater pelvis. *Forensic Science International, 219*, 288.e1-288.e8.

Bookstein, F. L. (1978). *The measurement of biological shape and shape change.* New York: Springer-Verlag.

Bookstein, F. L. (1991). *Morphometric Tools for Landmark Data: Geometry and Biology.* Cambridge: Cambridge University Press.

Brown, K. M. (2015). Selective pressures in the human bony pelvis: decoupling sexual dimorphism in the anterior and posterior spaces. *American Journal of Physical Anthropology,157(3)*, 428-440.

Bruner, E. (2004). Geometric morphometrics and paleoneurology: brain shape evolution in the genus *Homo*. *Journal of Human Evolution, 47*, 279-303.

Brůžek, J. (1996). Degree of pelvic sexual dimorphism in human populations. A greene t-test application. *Human Evolution, 11*(2), 183-189.

Brůžek, J. (2002). A method for visual determination of sex, using the human hip bone. *American Journal of Physical Anthropology, 117*, 157-168.

Brůžek, J., & Murail, P. (2006). Methodology and reliability of sex determination from the skeleton. In A. Schmitt, E. Cunha, & J. Pinheiro, *Forensic Anthropology and Medicine: Complementary Sciences from Recovery to Cause of Death* (pp. 225-242). New Jersy: Humana Press.

Budinoff, L. C., & Tague, R. G. (1990). Anatomical and developmental bases for the ventral arc of the human pubis. *American Journal of Physical Anthropology, 82*, 73-79.

Buikstra, J. E., & Ubelaker, D. H. (1994). *Standards for data collection from human skeletal remains.* Fayetteville, Arkansas: Arkansas Archaeological Survey Research Series.

Butler, J. (1990). *Gender Trouble*. New York: Routledge.

Butler, J. (1993). *Bodies that Matter: On the Discursive Limits of "Sex"*. New York: Routledge.

Bytheway, J. A., & Ross, A. H. (2010). A geometric morphometric approach to sex determination of the human adult os coxa. *Journal of Forensic Sciences, 55*(4), 859-862.

Cabo, L. L., Brewster, C. P., Ousley, S. D., & Dirkmaat, D. C. (2012). Sexual dimorphism: Interpreting sex markers. In D. C. Dirkmaat, *A Compaion to Forensic Anthropology* (pp. 248-286). Malden: Wiley-Blackwell.

Cameron, N. (1991). Human growth, nutrition, and health status in Sub-Saharan Africa. *Yearbook of Physical Anthropology, 34*, 211-250.

Clark, A.-L. (2014). Health and sexual dimorphism in Bon Non Wat: The effects of the intensification of agriculture in prehistoric Southeast Asia. *BMSAP, 26*, 196-204.

Coleman, W. H. (1969). Sex differences in the growth of the human bony pelvis. *American Journal of Physical Anthropology, 31*, 125-152.

Corner, B. D., Lele, S., & Richtsmeier, J. T. (1992). Measuring precision of three-dimentional landmark data. *Journal of Quantitative Anthropology, 3*, 347-359.

Dixit, S. G., Kakar, D., Agarwal, S. C., & Choudhry, R. (2007). Sexing of human hip bones of Indian origin by discriminant function analysis. *Journal of Forensic and Legal Medicine, 14*, 429-435.

Djorojevic, M., Roldán, C., García-Parra, P., Alemán, I., & Botella, M. (2014). Morphometric sex estimation from 3D computed tomography os coxae model and its validation in skeletal remains. *International Journal of Legal Medicine, 128*, 879-888.

Doyle, J. A., & Paludi, M. A. (1991). *Sex and Gender: The Human Experience.* Dubuque: WMC Brown Publishers.

Đurić, M., Rakočević, Z., & Đonić, D. (2005). The reliability of sex determination of skeletons from forensic context in the Balkans. *Forensic Science International, 147*, 159-164. doi:10.1016/j.forsciint.2004.09.111

Fausto-Sterling, A. (2000). *Sexing the Body: Gender Politics and the Construction of Sexuality.* New York: Basic Books.

Fischer, B., & Mitteroecker, P. (2015). Allometry and secual dimorphism in the human pelvis. *The Anatomical Record, 300*, 698-705.

Franklin, D., Cardini, A., & Oxnard, C. E. (2010). A geometric morphometric approach to the quantification of population variation in Sub-Saharan African crania. *American Journal of Human Biology, 22*, 23-35.

Fuller, K. (1998). Adult females and pubic bone growth.

Geller, P. L. (2005). Skeletal analysis and theoretical complications. *World Archaeology, 37*(4), 597-609.

Geller, P. L. (2008). Conceiving sex: fomenting a feminist bioarchaeology. *Journal of Social Archaeology, 8*(113), 113-138.

Gilchrist, R. (1999). *Gender and Archaeology: Contesting the Past.* New York: Routledge.

Gómez-Valdés, J. A., Quinto-Sánchez, M., Menéndez Garmendia, A., Velemínská, J., Sánchez-Mejorada, G., & Brůžek, J. (2012). Comparison of methods to determine sex by evaluating the greater sciatic notch: visual, angular and geometric morphometrics. *Forensic Science International, 221*, 156.e1-156.e7.

Gómez-Valdés, J. A., Torres Ramírez, G., Báez Molgado, S., Herrera Sain-Leu, P., Castrejón Caballero, J. L., & Sánchez-Mejorada, G. (2011). Discriminant function analysis for sex assessment in pelvic girdle bones: sample from the contemporary Mexican population. *Journal of Forensic Sciences, 56(2),* 297-301.

González, N. C., Rascón Pérez, J., Chamero, B., Camvra-Moo, O., & González Martín, A. (2017). Geometric morphometrics reveals restrictions on the shape of the female os coxae. *Journal of Anatomy, 230*, 66-74.

González, P. N., Bernal, V., & Perez, S. I. (2009). Geometric morphometric approach to sex estimation of human pelvis. *Forensic Science International, 189,* 68-74.

González, P. N., Bernal, V., & Perez, S. I. (2011). Analysis of sexual dimorphism of craniofacial traits using geometric morphometric techniques. *International Journal of Osteoarchaeology, 21,* 82-91. doi:10.1002/oa.1109

Goodall, C. (1991). Procrustes methods in the statistical analysis of shape. *Journal of the Royal Statistical Society. Series B (Methodological), 53*(2), 285-339.

Guégan, J.-F., Teriokhin, A. T., & Thomas, F. (2000). Human fertility variation, size-related obstetrical performance and the evolution of sexual stature dimorphism. *Proceedings of the Royal Society of Biological Sciences, 267*, 2529-2535.

Hallgrimsson, B., Zelditch, M. L., Parsons, T. E., Kristensen, E., Young, N. M., & Boyd, S. K. (2008). Morphometrics and biological anthropology in the postgenomic age. In M. A. Katzenberg, & S. R. Saunders, *Biological Anthropology of the Human Skeleton* (pp. 207-236). Hoboken: Wiley-Liss.

Hiernaux, J. (1985). A comparison of the shoulder-hip-width sexual dimorphism in Sub-Saharan African and Europe. In J. Ghesquiere, R. D. Martin, & F. Newcombe, *Human Sexual Dimorphism: Symposia for the Society for the Study of Human Biology* (pp. 191-206). Philadelphia: Taylor & Frances.

Hollimon, S. E. (2011). Sex and gender in bioarchaeological research: theory, method, and interpretation. In S. C. Agarwal, & B. A. Glencross, *Social Bioarchaeology* (pp. 149-182). Malden: Wiley-Blackwell.

Humphries, J. M., Bookstein, F. L., Chernoff, B., Smith, G. R., Elder, R. L., & Poss, S. G. (1981). Multivariate discrimination by shape in relation to size. *Systematic Zoology, 30*(3), 291-308.

İşcan, M. Y. (2005). Forensic anthropology of sex and body size. *Forensic Science International, 147*, 107-112.

İşcan, M. Y. (1983). Assessment of race from the pelvis. *American Journal of Physical Anthropology, 62*(2), 205-208.

Jackson, D. A. (1993). Stopping rules in principal components analysis: a comparison of heuristical and statistical approaches. *Ecology, 8*, 2204-2214.

Jolliffe, I. T. (2002). *Principal Component Analysis* (2 ed.). New York: Springer.

Joyce, R. A. (2005). Archaeology of the body. *Annual Review of Anthropology, 34*, 139-158.

Klales, A. R., Ousley, S. D., & Vollner, J. M. (2012). A revised method of sexing the human innominate using Phenice's nonmetric traits and statistical methods. *American Journal of Physical Anthropology, 149*(1), 104-114.

Klingenberg, C. P. (2016). Size, shape, and form: concepts of allometry in geometric morphometrics. *Development Genes and Evolution, 226*, 113-137.

Klingenberg, C. P. (2011). MorphoJ: an integrated software package for geometric morphometrics. *Molecular Ecology Resources, 11*, 353-357.

Knudson, K. J., & Stojanowski, C. M. (2008). New directions in bioarchaeology: recent contributions to the study of human social identities. *Journal of Archaeological Resources, 16*, 397-432.

Kurki, H. K. (2017). Bilateral asymmetry in the human pelvis. *The Anatomical Record, 300*, 653-665.

Kurki, H. K. (2013). Bony pelvic canal size and shape in relation to body proportionality in humans. *American Journal of Physical Anthropology, 151*, 88-101.

Kurki, H. K. (2011). Pelvic dimorphism in relation to body size and body size dimorphism in humans. *Journal of Human Evolution, 61*, 631-643.

Kurki, H. K. (2007). Protection of obstetric dimensions in a small-bodied human sample. *American Journal of Physical Anthropology, 133*, 1152-1165.

Lance, R. F., Kennedy, M. L., & Leberg, P. L. (2000). Classification bias in discriminant function analysis used to evaluate putatively different taxa. *Journal of Mammalogy, 81*(1): 245-249.

Langley, N. R., Meadows Janz, L., Ousley, S. D., Jantz, R. L., & Milner, G. (2016). *Data Collection Procedures for Forensic Skeletal Material 2.0.* Knoxville, TN: Forensic Anthropology Centre, Department of Anthropology, The University of Tennessee.

Listi, G. A. (2010). The impact of racial metric variation in thr os coxa on the morphological assessment of sex. *Journal of Forensic Science, 55*(5), 1157-1161.

Listi, G. A., & Bassett, H. E. (2006). Test of an alternative method for determining sex from the os coxae: applications for modern Americans. *Journal of Forensic Sciences, 51*(2), 248-252.

Macaluso, P. J. (2010). Sex determination from the acetabulum: test of a possible non-population-specific discriminant function equation. *Journal of Forensic and Legal Medicine, 17*, 348-351.

MacLaughlin, S. M., & Bruce, M. F. (1986). Population variation in sexual dimorphism in the human innominate. *Human Evolution, 1*(3), 221-231.

Mahakkanukrauh, P., Ruengdit, S., Tun, S. M., & Case, T. D. (2017). Osteometric sex estimation from the ox coxa in a Thai population. *Forensic Science International, 271*, 127.e1-127.e7.

Mahato, M. (2010). Assessment of pelvic dimensions and evaluation of new morphometric indices for determination of sex in human hip bones. *Australian Journal of Forensic Sciences, 42*(2), 123-135. doi:10.1080/00450610903258094

Mallard, A. M., Savell, K. R., & Auerbach, B. M. (2017). Morphological integration of the human pelvis with respect to age and sex. *The Anatomical Record, 300*, 666-674.

Manzi, G., Saracino, B., Bruner, E., & Sassarello, P. (2000). Geometric morphometric analysis of mid-sagittal cranial profiles in Neandertals, modern humans, and their ancestors. *Rivista di Antropologia, 78*, 193-204.

Marcus, L. F. (1993). Some aspects of multivariate statistics for morphometrics. In L. F. Marcus, E. Bello, & A. García-Valdecasas, *Contributions to Morphometrics* (pp. 95-130). Madrid, Spain: Monographias del Museo Nacional de Ciencias Naturales.

Meskell, L. M. (2000). Writing the body in archaeology. In A. E. Rautman, *Reading the Body* (pp. 13-21). Philadelphia: University of Pensivlvania Press.

Mitteroecker, P., & Gunz, P. (2009). Advances in geometric morphometrics. *Evolutionary Biology, 36*, 235-247.

Nordbladh, J., & Yates, T. (1990). This perfect body, this vergin text: between sex and gender in archaeology. In I. Bapty, & Yates, *Archaeology After Structuralism* (pp. 222-237). New York: Routledge.

Papaloucas, C., Fiska, A., & Demetriou, T. (2008). Sexual dimorphism of the hip joint in Greeks. *Forensic Science International, 179*, 83.e1-83.e3.

Patriquin, M. L., Loth, S. R., & Steyn, M. (2003). Sexual dimorphic pelvic morphology in South African whites and blacks. *Homo, 53*(3), 255-262.

Patriquin, M. L., Steyn, M., & Loth, S. R. (2005). Metric analysis of sex differences in South African black and white pelves. *Forensic Science International, 147*, 119-127.

Pfeiffer, S. (2012). Conditions for evolution of small adult body size in southern Africa. *Current Anthropology, 53*(S6), S383-S394.

Phenice, T. (1969). A newly developed visual method of sexing in the os pubis. *American Journal of Physical Anthropology, 30*, 297-301.

Pinhasi, R., Teschler-Nicola, M., Knaus, A., & Shaw, P. (2005). Cross-population analysis of the growth of long bones and the os coxae of three early medieval Austrian populations. *American Journal of Human Biology, 17*, 470-488.

Plavcan, J. M. (2011). Understanding dimorphism as a function of changes in male and female traits. *Evolutionary Anthropology, 20*, 143-155.

Pretorius, E., Steyn, M., & Scholtz, Y. (2006). Investigation into the usability of geometric morphometric analysis in assessment of sexual dimorphism. *American Journal of Physical Anthropology, 129*, 64-70.

Quartrehomme, G., Ranmona, I., Mogueira, L., du Jardin, P., & Alunni, V. (2017). Sex determination using the DSP (probabilistic sex diagnosis) method on the coxal bone: Efficiency of method according to number of available variables. *Forensic Science International, 272*, 190-193.

Ridgeway, B. M., Arias, B. E., & Barber, M. D. (2008). Variation of the obturtor foramen and pubic arch of the female bony pelvis. *American Journal of Obstetric Gynecology*, 546.e1-546.e4.

Ridgeway, B. M., Arias, B. E., & Barber, M. D. (2011). The relationship between anthropometric measurements and the body pelvis in African American and European American women. *International Urogynecological Association, 22*, 1019-1024. doi:10.1007/s00192-011-1416-1

Robertson, H. I. (2013). *A Geometric Morphometric Study of Sexual Dimorphism in the Human Hip Bone.* Master's thesis, University of Victoria, Anthropology. Retrieved from http://dspace.library.uvic.ca/handle/1828/5080

Robertson, H. I., Pokotylo, D. L., & Weston, D. A. (2019). Testing landmark redundancy for sex-based shape analysis of the adult human os coxa. *American Journal of Physical Anthropology, 169*, 689-703.

Robinson, M. S., & Bidmos, M. A. (2009). The skull and humerus in the determination of sex: discriminant function equations. *Forensic Science International, 186*, 86.e1-86.e5.

Robinson, M. S., & Bidmos, M. A. (2011). An assessment of the accuracy of discriminant function equations for sex determination of the femur and tibia from a South African population. *Forensic Science International, 206*(1-3), 212.e1-212.5.

Rogers, T., & Saunders, S. (1994). Accuracy of sex determination using morphological traits of the human pelvis. *Journal of Forensic Sciences, 39*(4), 1047-1056.

Rosas, A., & Bastir, M. (2002). Thin-plate spline analysis of allometry and sexual dimorphism in the human craniofacal complex. *American Journal of Physical Anthropology, 117*, 236-245.

Rosenberg, K. (2002). A late Pleistocene human skeleton from Liujiang, China suggests regional population variation in sexual dimorphism in the human pelvis. *Variability and Evolution, 10*, 5-17.

Roth, V. L. (1993). On three dimentional morphometrics. In L. F. Marcus, E. Bello, & Garcia-Valdecasas, *Contributions to Morphometrics* (pp. 41-64). Madrid, Spain: Monographias del Museo Nacional de Ciencias Naturales.

San-Millán, M., Rissech, C., & Turbón, D. (2017). Shape variablilty of the adult human acetabulum and acetabular fossa related to sex and age by geometric morphometrics. Implications for adult age estimation. *Forensic Science International, 272*, 50-63.

Schulter-Ellis, F. P., Hayek, L. C., & Schmidt, D. J. (1985). Determination of sex with a discrimnant analysis of new pelvic bone measurements: Part II. *Journal of Forensic Sciences, 30*(1), 178-185.

Schulter-Ellis, F. P., Schmidt, D. J., Hayek, L.-A., & Craig, J. (1983). Determination of sex with discriminant analysis of new pelvic bone measurements: part I. *Journal of Forensic Sciences, 28*(1), 169-180.

Shapiro, J. (2005). Transsexualism: Reflections on the persistence of gender and the mutibility of sex. In J. Robertson, *Same-Sex Cultures and Sexualities: An Anththropological Reader* (pp. 138-161). Malden: Blackwell.

Sholts, S. B., Flores, L., Walker, P. L., & Wärmländer, S. K. (2011). Comparison of coordinate measurements precision of different landmark types on human crania using a 3D laser scanner and a 3D digitiser: implications for applications of digital morphometrics. *International Journal of Osteoarchaeology, 21*, 535-543. doi:10.1002/oa.1156

Singh, S., & Potturi, B. R. (1978). Greater sciatic notch in sex determination. *Journal of Anatomy, 125*(3), 619-624.

Sitek, A., Fijałkowska, M., Żądzińska, E., & Antoszewski, B. (2012). Biometric characteristics of the pelvis in female-to-male transsexuals. *Archives of Sexual Behaviour, 41*, 1303-1313.

Small, C., Bits, D. M., & Hemingway, J. (2012). Quantification of the subpubic angle in South Africans. *Forensic Science International, 222*, 395e1-395e5.

Sofaer, J. R. (2013). Bioarchaeological approaches to the gendered body. In D. BoLRer, *A Companion to Gender Prehistory* (pp. 224-244). Malden: Wiley-Blackwell.

Steyn, M., & Patriquin, M. L. (2009). Osteometric sex determination from the pelvis - does population specificity matter? *Forensic Science International, 191*, 113.e.1-113.e.5. doi:10.1016/j.forsciint.2009.07.009

Steyn, M., Pretorius, E., & Hutten, L. (2004). Geometric morphometric analysis of the greater sciatic notch in South Africans. *Homo, 54*(3), 197-206.

Stone, P., & Walrath, D. (2006). The gendered skeleton: anthropological interpretations of the bony pelvis. In R. Growland, & C. Knusel, *The Social Archaeology of Funerary Remains* (pp. 168-178). Oxford: Oxbow Books.

Sutherland, L. D., & Myers Suchey, J. (1991). Use of the ventral arc in pubic sex determination. *Journal of Forensic Sciences, 36*(2), 501-511.

Tague, R. G. (2000). Do big females have big pelves? *American Journal of Physical Anthropology, 112*, 377-393.

Tague, R. G. (1994). Maternal mortality or prolonged growth: age at death and pelvic size in three prehistoric Amerindian populations. *American Journal of Physical Anthropology, 95*, 27-40.

Vacca, E., & Di Vella, G. (2012). Metric characterization of the human coxal bone on a recent Italian sample and multivariate discriminant analysis to determine sex. *Forensic Science International, 222*, 401e1-401.e9.

Velemínská, J., Krajíček, V., Ján, D., Goméz-Valdés, J. A., Velemínský, P., Šefčáková, A., . . . Jaroslav, B. (2013). Technical note: geometric morphometrics and sexual dimorphism of the greater sciatic notch in adults from two skeletal collections: the accuracy and reliability of sex classification. *American Journal of Physical Anthropology, 152*, 558-565.

Vercellotti, G., Stout, S. D., Boano, R., & Sciulli, P. W. (2011). Intrapopulation variation in stature and body proportions: social status and sex differences in an Italian medieval population (Trino Vercellese, VC). *American Journal of Physical Anthropology, 145*, 203-214.

Viðarsdóttir, U. S., O'Higgins, P., & Stringer, C. (2002). A geometric morphometric study of regional differences in the ontogeny of the modern human facial skeleton. *Journal of Anatomy, 201*, 211-229.

von Cramon-Taubadel, N., Frazier, B. C., & Mirazón Lahr, M. (2007). The problem of assessing landmark error in geometric morphometrics: theory, methods, and modifications. *American Journal of Physical Anthropology, 134*, 24-35.

Walker, P. L. (2005). Greater sciatic notch morphology: sex, age, and population differences. *American Journal of Physical Anthropology, 127*, 385-391.

Walker, P. L. (1995). Problems of preservation and sexism in sexing: some lessons from historical xollections for palaeodemographers. (pp 31-48) in *Grave Reflections: Portraying the Past Through Cemetery Studies*. S. R. Saunders & A. Herring eds. Canadian Scholars' Press: Toronto.

Washburn, S. L. (1948). Sex differences in the pubic bone. *American Journal of Physical Anthropology, 6*(2), 199-208.

Webster, M., & Sheets, H. D. (2010). A practical introduction to landmark-based geometric morphometrics. *The Paleontological Society Papers, 16*, 163-188. doi:10.1017/S1089332600001868

Weiss, K. M. (1971). On the systematic bias in skeletal sexing. *American Journal of Physical Anthropology, 37*, 239-250.

Wescott, D. J. (2015). Sexual dimorphism in auricular surface projection and postauricular sulcus morphology. *Journal of Forensic Science, 60*(3), 679-685.

Wilson, L. A., Ives, R., Cardoso, H. F., & Humphrey, L. T. (2015). Shape, size, and maturity trajectories of the human ilium. *American Journal of Physical Anthropology, 156*, 19-34.

Zakrzewski, S. R. (2003). Variation in ancient Egyptian stature and body proportions. *American Journal of Physical Anthropology, 121*, 219-229.

Zakrzewski, S. R. (2011). Population migration, variation, and identity: an Islamic population in Iberia. In S. C. Agarwal, & B. A. Glencross, *Social Bioarchaeology* (pp. 183-211). Malden: Wiley-Blackwell.

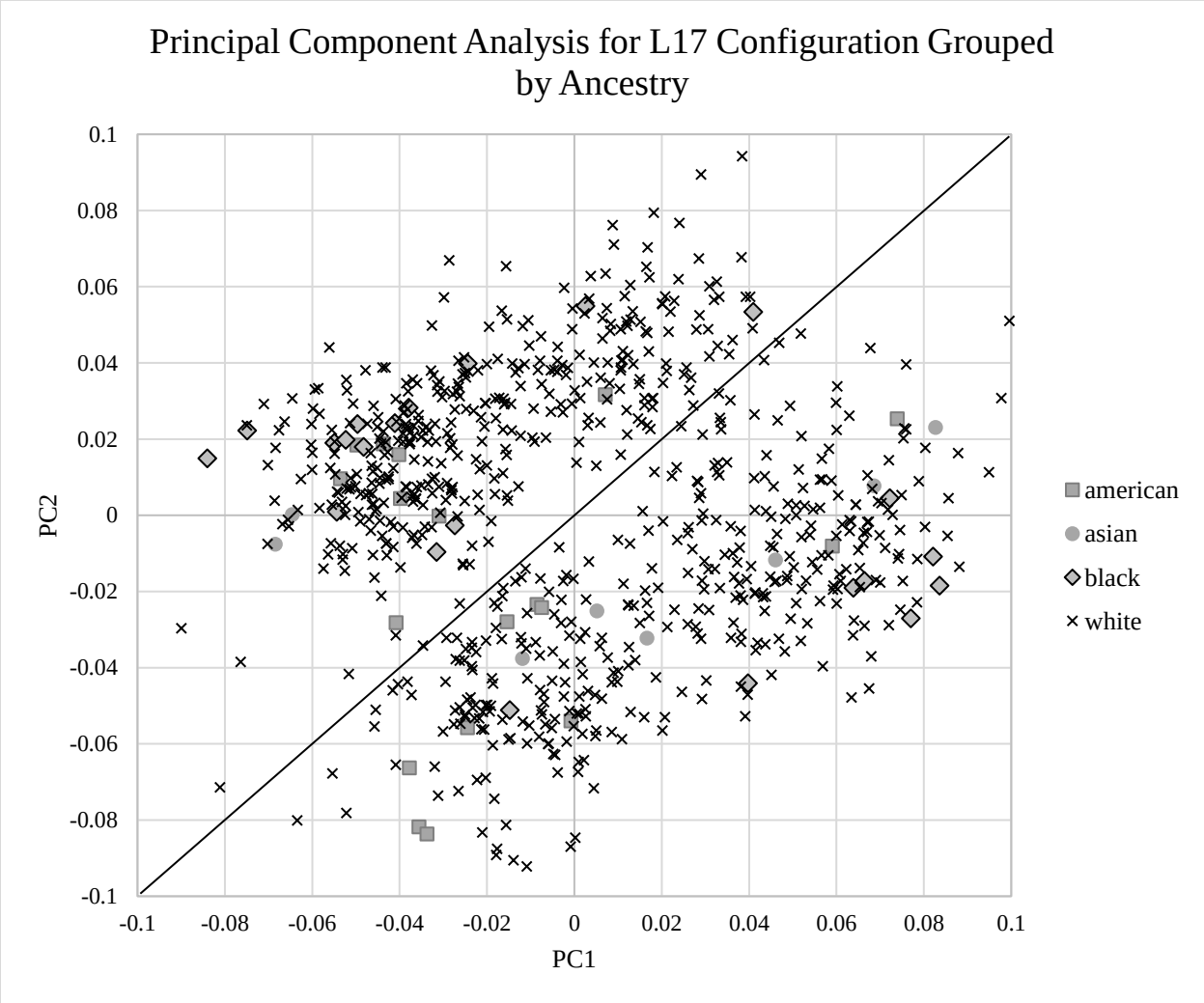
Zelditch, M. L., Swiderski, D. L., & Sheets, H. D. (2012). *Geometric morphometrics for biologists: a primer.* San Diego, California, USA: Academic Press.

Zeng, Y., Wang, Y., Zhu, Z., Tang, T., Dai, K., & Qiu, S. (2012). Differences in acetabular morphology related to side and sex in a Chinese population. *Journal of Anatomy, 220*, 256-262. doi:10.1111/j.1469-7580.2011.01471.x

Appendices

Appendix A

This scatter plot illustrates the division of individuals ladled according ancestry on PC1 and PC2 of the 17-landmark configuration. Linear equation $y = 1x$, Females are in the right half of the bisected square and Males are in the left half of the bisected square.



Appendix B

The following table outlines the results of the inter-observer repeatablility among coxal bone measurements. Coeficient of Variation greater than 2.5% was considered in this study to be too great a difference between observers, following the standards outlined in Corner et al. (1992), and are indicated in bold font. All measurements are in mm.

Specimen 1							
Left/Sex	**Male**	**Male**	**Male**	**Male**			
Measurement	**Obs[+] 1**	**Obs 2**	**Obs 3**	**Obs 4**	**Average**	**SD***	**CV[§]**
XCH	238.78	239.30	238.69	238.18	238.74	0.459	0.192
XIB	164.67	166.12	161.01	165.50	164.32	2.288	1.392
XPL	126.50	127.40	125.10	127.15	126.54	1.031	0.815
ISL	118.53	122.41	118.48	120.95	120.09	1.928	1.604
WAS	121.40	122.05	121.09	122.04	121.64	0.479	0.394
Right/Sex	**Male**	**Male**	**Male**	**Male**			
Measurement	**Obs 1**	**Obs 2**	**Obs 3**	**Obs 4**	**Average**	**SD**	**CV**
XCH	234.88	231.63	234.82	233.62	233.74	1.520	0.650
XIB	168.54	165.60	165.87	165.78	166.45	1.399	0.840
XPL	122.31	126.97	126.03	126.02	125.33	2.064	1.647
ISL	121.20	121.60	120.02	123.88	121.68	1.616	1.328
WAS	125.71	127.68	124.72	124.69	125.70	1.402	1.115

Specimen 2							
Left/Sex	**Female**	**Female**	**Male**	**Int**			
Measurement	**Obs 1**	**Obs 2**	**Obs 3**	**Obs 4**	**Average**	**SD**	**CV**
XCH	207.43	211.38	211.79	210.94	210.38	2.000	0.951
XIB	155.66	156.65	158.07	159.15	157.38	1.538	0.977
XPL	116.62	119.32	119.52	120.68	119.03	1.718	1.442
ISL	103.7	105.92	108.76	109.77	107.04	**2.758**	**2.577**
WAS	119.5	117.75	118.72	119.80	118.94	0.916	0.770

Right/Sex	Female	Male	Male	Male			
Measurement	**Obs 1**	**Obs 2**	**Obs 3**	**Obs 4**	**Average**	**SD**	**CV**
XCH	208.90	211.09	210.14	207.61	209.43	1.511	0.721
XIB	157.30	156.99	153.42	157.17	156.22	1.871	1.198
XPL	119.10	120.38	118.97	118.74	119.30	0.737	0.618
ISL	104.99	111.88	111.10	111.30	109.82	**3.235**	**2.946**
WAS	118.76	122.40	119.86	119.71	120.18	1.556	1.295

+ Observer; *Standard Deviation (mm); §Coefficient of Variation (%)

Appendix C

This table presents the results of an intra-observer test between two landmark placement events. Procrustes coordinates are grouped by sex and averaged. Significant differences between landmark placement events are indicated by bold font.

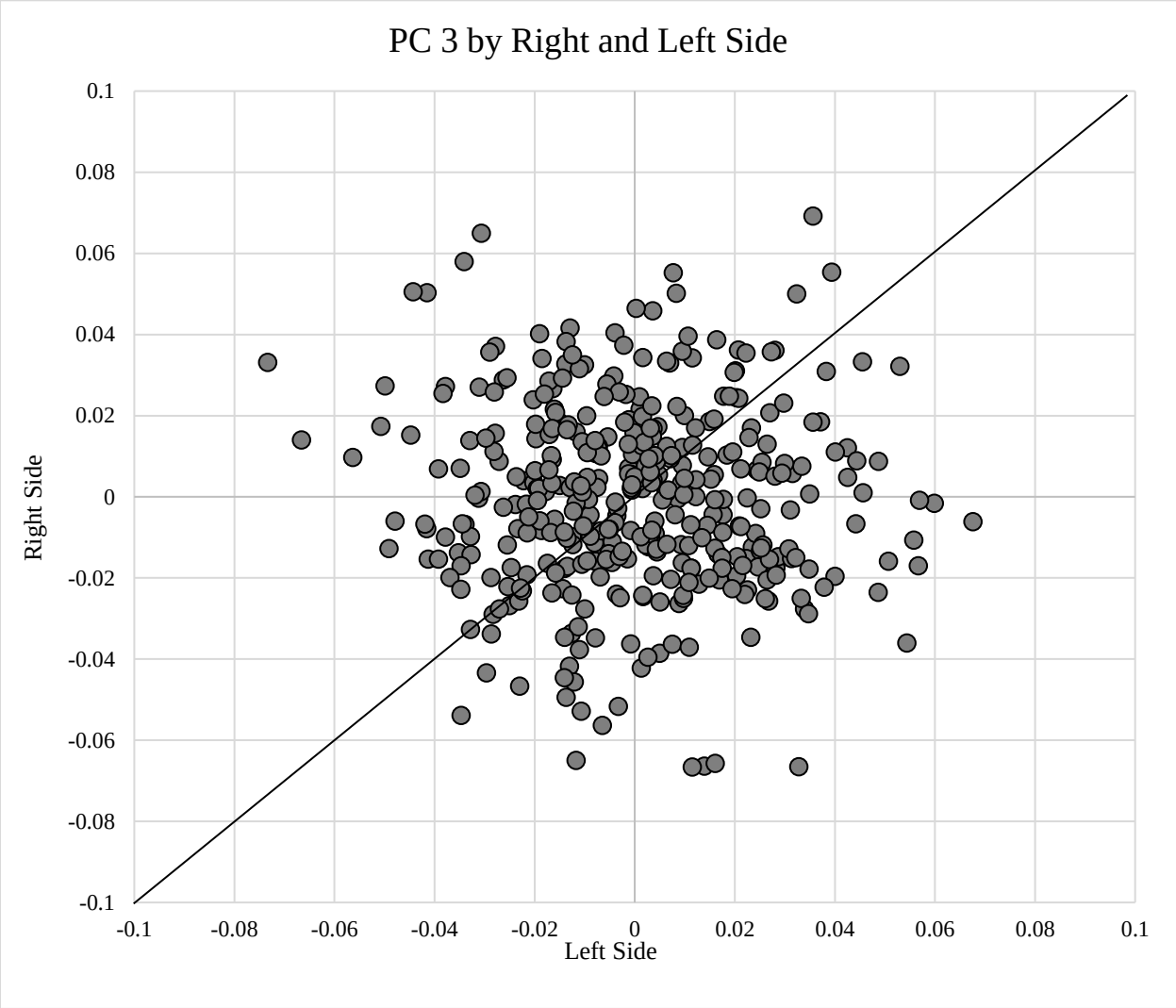
	Male_L		**Female_L**		**Male_R**		**Female_R**	
Landmark	**Mean (SD)1**	**Mean (SD)2**	**Mean (SD)1**	**Mean (SD)2**	**Mean (SD)1**	**Mean (SD)2**	**Mean (SD)1**	**Mean (SD)2**
1x	0.113 (0.009)	0.113 (0.008)	0.122 (0.018)	0.120 (0.016)	0.115 (0.012)	0.117 (0.010)	0.125 (0.016)	0.125 (0.016)
1y	0.256 (0.007)	0.256 (0.008)	0.248 (0.010)	0.247 (0.007)	0.253 (0.005)	0.255 (0.008)	0.248 (0.007)	0.249 (0.009)
1z	0.076 (0.023)	0.074 (0.024)	0.072 (0.026)	0.070 (0.027)	0.078 (0.022)	0.079 (0.024)	0.076 (0.030)	0.075 (0.029)
2x	0.201 (0.019)	0.202 (0.020)	0.207 (0.019)	0.213 (0.023)	0.203 (0.022)	0.205 (0.024)	0.222 (0.017)	0.222 (0.014)
2y	-0.140 (0.006)	-0.140 (0.008)	-0.144 (0.010)	-0.142 (0.005)	-0.138 (0.007)	-0.136 (0.010)	-0.143 (0.008)	-0.143 (0.008)
2z	-0.090 (0.014)	-0.092 (0.013)	-0.097 (0.004)	-0.098 (0.004)	-0.095 (0.013)	-0.094 (0.014)	-0.108 (0.013)	-0.107 (0.010)
5x	0.120 (0.009)	0.120 (0.009)	0.121 (0.009)	0.121 (0.011)	0.124 (0.007)	0.126 (0.008)	0.124 (0.011)	0.125 (0.013)
5y	0.019 (0.007)	0.019 (0.008)	0.006 (0.015)	0.005 (0.009)	0.016 (0.003)	0.017 (0.003)	0.010 (0.004)	0.008 (0.006)
5z	-0.045 (0.004)	-0.045 (0.004)	-0.056 (0.010)	-0.056 (0.006)	-0.049 (0.006)	-0.047 (0.006)	-0.054 (0.006)	-0.054 (0.006)
6x	-0.239 (0.010)	-0.237 (0.010)	-0.216 (0.013)	-0.219 (0.007)	-0.235 (0.006)	-0.235 (0.005)	-0.219 (0.008)	-0.219 (0.009)
6y	-0.076 (0.010)	-0.078 (0.013)	-0.084 (0.010)	-0.077 (0.008)	-0.078 (0.007)	-0.079 (0.013)	-0.078 (0.011)	-0.077 (0.012)
6z	0.030 (0.009)	0.033 (0.004)	0.045 (0.010)	0.038 (0.006)	0.036 (0.011)	0.034 (0.007)	0.040 (0.009)	0.039 (0.013)

8x	-0.188	-0.188	-0.198	-0.199	-0.192	-0.190	-0.197	-0.198
	(0.005)	(0.007)	(0.010)	(0.011)	(0.009)	(0.007)	(0.013)	(0.012)
8y	0.130	0.130	0.136	0.134	0.126	0.126	0.130	0.129
	(0.008)	(0.008)	(0.015)	(0.014)	(0.008)	(0.008)	(0.011)	(0.011)
8z	-0.121	-0.122	-0.127	-0.126	-0.121	-0.122	-0.131	-0.129
	(0.006)	(0.007)	(0.010)	(0.011)	(0.007)	(0.007)	(0.011)	(0.009)
14x	0.357	0.357	0.339	0.341	0.352	0.354	0.339	0.338
	(0.005)	(0.005)	(0.010)	(0.007)	(0.009)	(0.009)	(0.003)	(0.005)
14y	0.090	0.091	0.082	0.081	0.101	0.090	0.080	0.080
	(0.014)	(0.015)	(0.019)	(0.015)	(0.017)	(0.018)	(0.021)	(0.023)
14z	-0.0001	0.0005	0.017	0.017	0.012	0.004	0.021	0.015
	(0.010)	(0.013)	(0.030)	(0.023)	(0.018)	(0.007)	(0.024)	(0.018)
16x	0.023	0.023	0.012	0.009	0.028	0.024	0.010	0.009
	(0.006)	(0.007)	(0.008)	(0.007)	(0.011)	(0.006)	(0.010)	(0.006)
16y	0.017	0.013	0.024	0.019	0.017	0.013	0.023	0.020
	(0.007)	(0.010)	(0.014)	(0.013)	(0.004)	(0.008)	(0.008)	(0.008)
16z	0.109	0.109	0.120	0.119	0.106	0.106	0.120	0.122
	(0.008)	(0.008)	(0.013)	(0.011)	(0.005)	(0.005)	(0.012)	(0.012)
17x	0.016	0.018	0.006	0.007	**0.024**	**0.021**	0.012	0.013
	(0.007)	(0.007)	(0.007)	(0.009)	**(0.008)**	**(0.005)**	(0.011)	(0.010)
17y	0.107	0.104	0.104	0.101	0.103	0.105	0.101	0.099
	(0.007)	(0.008)	(0.005)	(0.005)	(0.006)	(0.006)	(0.008)	(0.009)
17z	0.103	0.104	0.106	0.107	0.099	0.100	0.104	0.106
	(0.007)	(0.007)	(0.010)	(0.009)	(0.004)	(0.005)	(0.010)	(0.011)

Appendix D

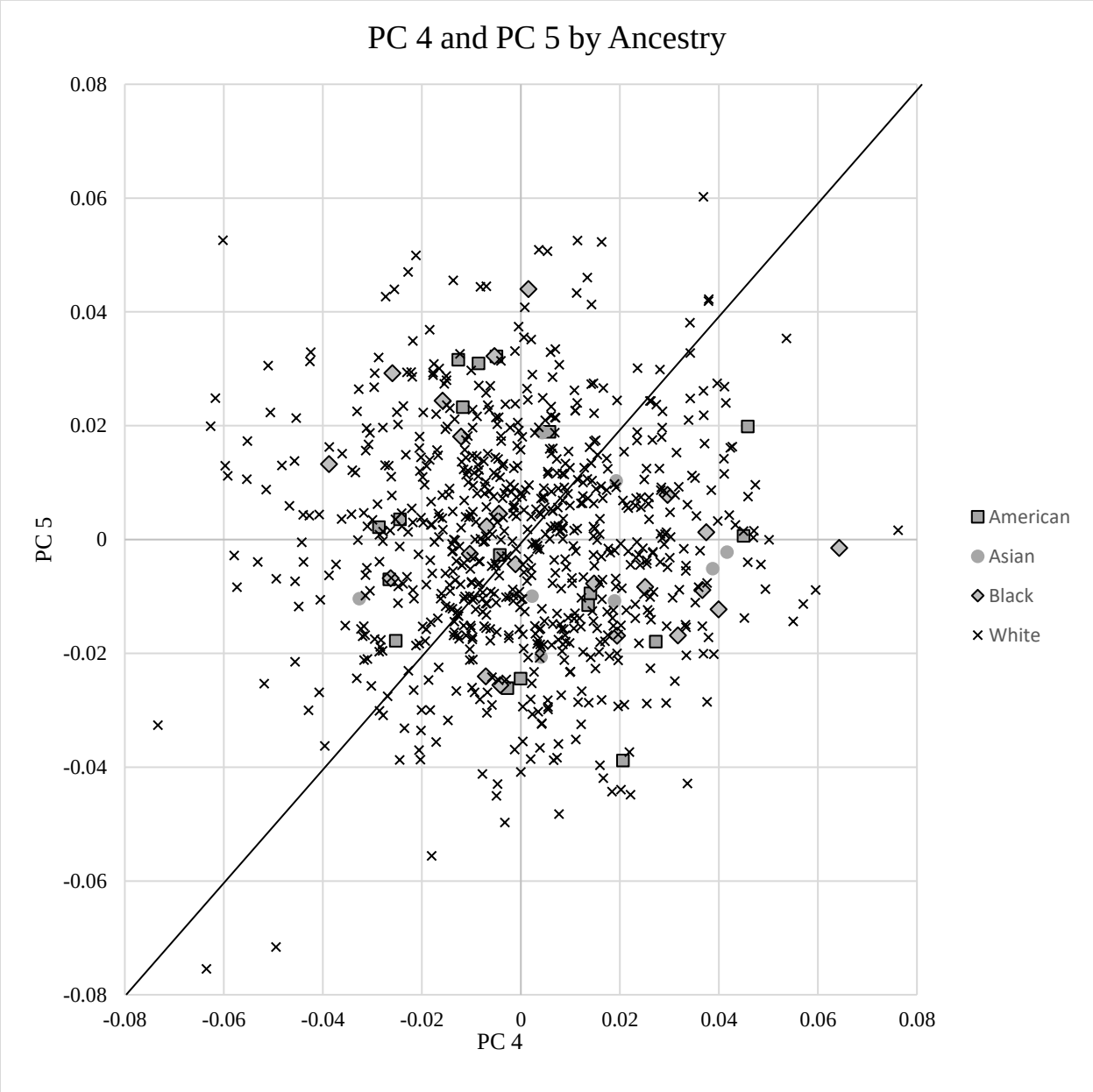
Scatter plot of principal component 3 values compared between right and left coxal bones.

Linear equation $y = 1x$. Independent t-test results $t = 2.341$, df = 786, $p = 0.019$.



Appendix E

Scatter plot of principal components 4 and 5 that correlate with ancestry. Linear equation $y = 1x$. Independent t-test results for PC4 ($t = 3.470$, df = 392, $p = 0.001$) and PC5 ($t = -3.647$, df = 392, $p = <0.0001$).



Appendix F

This table presents a Pearson Correlation of linear coxal bone measurements, Centroid Size, Geometric Mean, Stature, Weight, and anthroposcopic traits against the first five principal components containing meaningful sex-based shape data. Measurements were averaged between right and left coxal bones.

		PC1	**PC2**	**PC3**	**PC4**	**PC5**
XCH	Pearson Correlation	**-0.471**	**0.568**	**0.207**	-0.024	-0.067
	Sig. (2-tailed)	**<0.001**	**<0.001**	**<0.001**	0.639	0.187
	Covariance	-0.240	0.272	0.066	-0.007	-0.017
XIB	Pearson Correlation	**-0.178**	**0.209**	**-0.108**	-0.007	0.052
	Sig. (2-tailed)	**<0.001**	**<0.001**	**0.032**	0.885	0.299
	Covariance	-0.056	0.061	-0.021	-0.001	0.008
XPL	Pearson Correlation	-0.062	**0.135**	0.011	-0.008	**-0.160**
	Sig. (2-tailed)	0.220	**0.007**	0.834	0.870	**0.001**
	Covariance	-0.013	0.027	0.001	-0.001	-0.017
ISL	Pearson Correlation	**-0.487**	**0.680**	**0.097**	-0.035	**-0.080**
	Sig. (2-tailed)	**<0.001**	**<0.001**	**0.055**	0.487	**0.112**
	Covariance	-0.142	0.185	0.018	-0.006	-0.012
XIRL	Pearson Correlation	**0.259**	**-0.224**	**-0.144**	-0.054	-0.066
	Sig. (2-tailed)	**<0.001**	**<0.001**	**0.004**	0.288	0.190
	Covariance	0.049	-0.040	-0.017	-0.006	-0.006
ASISS	Pearson Correlation	0.028	0.033	**0.421**	**0.507**	**-0.273**
	Sig. (2-tailed)	0.583	0.513	**<0.001**	**<0.001**	**<0.001**
	Covariance	0.010	0.011	0.095	0.112	-0.043
PSISS	Pearson Correlation	-0.051	**-0.110**	**0.335**	**-0.273**	**-0.178**
	Sig. (2-tailed)	0.309	**0.029**	**<0.001**	**<0.001**	**<0.001**
	Covariance	-0.019	-0.039	0.078	-0.062	-0.034

WAS	Pearson Correlation	**0.194**	**-0.370**	**0.333**	**-0.167**	**-0.380**
	Sig. (2-tailed)	**<0.001**	**<0.001**	**<0.001**	**0.001**	**<0.001**
	Covariance	0.054	-0.102	0.058	-0.028	-0.054
FHD	Pearson Correlation	**-0.193**	**0.655**	0.022	-0.062	-0.084
	Sig. (2-tailed)	**<0.001**	**<0.001**	0.660	0.217	0.097
	Covariance	**-0.066**	**0.082**	0.002	-0.005	-0.006
FML	Pearson Correlation	**-0.412**	**0.447**	**0.178**	-0.079	-0.067
	Sig. (2-tailed)	**<0.001**	**<0.001**	**<0.001**	0.121	0.187
	Covariance	-0.423	0.431	0.114	-0.049	-0.035
Centroid Size	Pearson Correlation	**-0.254**	**0.286**	**0.278**	-0.048	**-0.10**
	Sig. (2-tailed)	**<0.001**	**<0.001**	**<0.001**	0.343	**0.030**
	Covariance	-0.143	0.152	0.098	-0.017	-0.031
Geometric Mean	Pearson Correlation	-0.058	0.086	**0.160**	-0.059	**-0.149**
	Sig. (2-tailed)	0.253	0.090	**0.001**	0.242	**0.003**
	Covariance	-0.014	0.020	0.025	-0.009	-0.019
Stature	Pearson Correlation	**-0.406**	**0.520**	**0.187**	-0.075	-0.057
	Sig. (2-tailed)	**<0.001**	**<0.001**	**<0.001**	0.139	0.266
	Covariance	-0.147	0.179	0.042	-0.017	-0.010
Weight	Pearson Correlation	0.089	**0.167**	-0.002	0.031	0.032
	Sig. (2-tailed)	0.080	**0.001**	0.964	0.545	0.526
	Covariance	0.209	0.368	-0.003	0.044	0.038
Age at Death	Pearson Correlation	0.016	-0.009	**-0.137**	0.079	-0.019
	Sig. (2-tailed)	0.751	0.861	**0.006**	0.118	0.710
	Covariance	0.008	-0.004	-0.043	0.024	-0.005