Summary Sex: A Multivariate Approach to Sex Estimation from the Human Pelvis

Samuel Richard Rennie

A thesis submitted in partial fulfilment of the requirements of Liverpool John Moores University for the degree of Doctor of Philosophy

August 2018

Contents

Acknowledgements

I would like to begin with thanking my Director of Studies Prof. Silvia Gonzalez. She has guided and supervised throughout the research and has been a calming force for when panic had ensued. Her support and encouragement during many challenges has helped me understand that setbacks are a normal part of this career path, but the key is to find an alternative route and to continue moving forward.

To my co-supervisor Dr Margaret Clegg who has played an integral role in my postgraduate studies. Alongside her supervision, she has encouraged me to attend and present at conferences at an international level for which I am extremely grateful for. She has also given boosts of morale and constructive criticism when needed.

I would also like to thank my friends/colleagues who I got to work beside over the past few years at LJMU. I think of us more as a family rather than an assortment of people who like to study dead people. You have all played many roles spanning from an intricate support system to setting up the “#postgradsontour” hashtag for whenever we go to conferences en masse. Without the likes of Ele Dove, Carla Burrel, Cal Davenport, Sarah Canty, and countless more, I doubt that my experience during the PhD would have been as eventful as it has been.

A big thankyou to my second academic family at the University of Pretoria, South Africa. My six months would have been completely different without your generosity and friendship. I would like to personally thank Charlotte Theye, François Marin, and Clément Zanolli for adopting me into their French family whilst living in Pretoria. I would also like to thank Gabi Krüger and Clarisa Sutherland for letting me invade the osteology lab and distract them with pointless conversations. Thanks to Rachel Holgate who helped keep the homesickness at bay by being the other “Brit” in the lab and for playing ‘Infinite Monkey Cage’ whilst we were all working in the lab. And finally, thanks to Franci Dorfling and Kate Leigh Strachan for showing me their home country of South Africa.

My research in South Africa would not have been possible without the funding from the Erasmus Mundus programme: AESOP. During my degree, I was also fortunate enough to attend and present at conferences both nationally and internationally, which would not have been possible without funding from LJMU. To help with research costs and tuition fees, I thank the Queen Elizabeth Foundation, a local funding body for which I would not have been to continue in further education without their assistance.

I express extreme gratitude towards my family and friends (outside of academia). My parents, Carole and Geoffrey, and my sister, Alexandra, who have had to endure the lows of being associated with a doctoral student and helped celebrate the successes. They have persevered alongside the research and had to listen to my non-stop talk about my research. This thanks also extends to my TNF family, who again have had to tolerate my ramblings during the entire time I worked in retail.

A PhD has always been described as being an emotional rollercoaster and true to its description it has been. As I leave behind my identity of being student, I endeavour to create a successful career in science. I hope that as I make this transition that all those that I have acknowledged (named or not) recognise that they have all made a lasting impression to my life.

List of Tables

List of Figures

Abstract

With the progression of multivariate statistics, the creation of population specific equations is on the rise. Multivariate analysis generally revolves around metric methods or geometric morphometrics, not on morphoscopic features.

A total of eight samples were analysed spanning from prehistoric American to modern day South African and ranged between pygmy populations from the Andaman Islands to medieval British populations. With a sample size of more than 1100 individuals, each *os coxa* was scored using eight morphoscopic features most commonly used by physical anthropologists and osteoarchaeologists.

Trait frequencies were compiled and compared between each of the eight samples. Then, the samples were placed into two groups: a known age and sex group (Christ Church Spitalfields, South African White, South African Black, and South African Coloured), and an unknown archaeological group (Poulton, St. Owens, Chumash, and Andaman). When comparing trait frequencies, slight differences between the samples could be seen.

Ordinal Logistic Regressions (OLR) were applied onto each of the four samples from the known age and sex group to create population specific sexing equations (cross-validated). Results from these four equations ranged from 90.24% (South African Black population specific equation) to 96.38% (Christ Church, Spitalfields population specific equation). Population specifity was tested by applying all of the equations onto each sample in this group. In an attempt to reduce this, two new equations were created by combining samples together resulting in a South African specific equation (92.54% accuracy) and a "Summary Sex" equation (92.98% accuracy). After applying each of the six new OLR equations onto the four archaeological samples, high percentage accuracies (ranging from 92.59% to 100.00%) were found when comparing them to the previous records. The only sample that did not produce as high of an accuracy was the Chumash sample with 82.35%.

In the attempt to analyse fragmented remains, three avenues were taken. Firstly, all missing values were replaced by the median score. Secondly, the original six OLR equations were 'sectioned' to make three smaller sets of equations. Lastly, to mirror the sectioned equations, three new sets of OLR equations were generated.

This study shows that when using morphoscopic traits for sex estimation, applying multivariate techniques can be used to obtain a high accuracy even when dealing with fragmented samples.

1. Introduction

1.1 Rationale

Sex estimation, be it for forensic or bioarchaeological purposes, is one of the most important aspects of creating a biological profile (Kjellström 2004, Murail *et al.* 2005). Correct classification can lead to either a positive identification of an individual or a more accurate portrayal of a past population. Many studies have been performed on the human skeleton, however the *os coxae* is considered the most sexually dimorphic bony element (Tague 1992, Kurki 2007). This difference is said to be due to locomotion and parturition (Abitbol 1987). Because of these known differences, the *os coxae* has been studied extensively both quantitatively (Steyn & Patriquin 2009, Betti 2014) and qualitatively (Phenice 1969, Bruzek 2002).

When performing sex estimations from human skeletal material, visual assessment of the cranium and *os coxae* are the most often used when available. When assessing the skull, Walker's (2008) descriptions and scoring techniques are normally used. The cranial traits that Walker described were based on the traits found in Buikstra & Ubelaker (1994). Walker (2008) found that the morphology of the glabellar and mastoid process was the most effective when used as a single trait to determine sex (greater than 70% accuracy). However, when combining the five traits, his accuracy increased to 89%. From there, Krüger *et al.* (2015) re-evaluated the traits in relation to South African remains and found that when applying the scores to a logistic regression, accuracies as high as 93% could be achieved.

Holobinko (2012) stated that cranial traits are secondary if the pelvis is present. This can be explained by the pelvis as having a much greater amount of sexual dimorphism than other skeletal features (Tague 1995, 2005). With accuracies as high as 96% being achieved from three traits from the *os pubis* (Phenice 1969), it is understandable why it is preferred over the cranium (Spradley & Jantz 2012). Bruzek (2002) proposed a new method of sex estimation using visual traits by analysing two skeletal collections (with known sex): one from France (University of Paris) and the other from Portugal (Coimbra Collection). Bruzek (2002) examined five characters: the preauricular surface, the greater sciatic notch, form of the composite arch, morphology of the inferior pelvis and the ischiopubic proportion. Using three conditions (0 = indeterminate, M= male, F= female) for each trait, the accuracy rate was close to 98% and only 3% were classified into the indeterminate category. Listi & Bassett (2006) applied the Bruzek (2002) methodology on a sample of American Whites and Blacks, which resulted in an accuracy of 90-92%. Meindl *et al*. (1985) found that when sexing with both the *os coxae* and cranium, an accuracy of 98% is achievable.

With the progression of multivariate statistics, the creation of population specific equations is on the rise. Multivariate analyses generally revolve around metric methods or geometric morphometrics, not on morphoscopic features. This study will be the first extensive look at applying several morphoscopic sexing traits from the human pelvis with multivariate statistics. Alongside the creation of population specific equations, an attempt at a universal sexing equation will be made. Fragmented remains and missing data are one of the shortfalls to these multivariate equations, therefore an investigation into the handling of these equations for fragmented skeletal remains will be undertaken. Eight morphoscopic traits from the pelvis were chosen specifically for this study: ventral arc, subpubic concavity, medial aspect of the ischiopubic ramus, shape of the pubic body, subpubic angle, obturator foramen, greater sciatic notch, and the preauricular sulcus. These eight were chosen because of their popularity among practitioners and researchers for estimating sex for both archaeological and forensic remains (The Workshop for European Anthropologists 1980, Rogers & Saunders 1994, White & Folkens 2000, Durić *et al*. 2005, Rösing *et al*. 2007, Sidler *et al*. 2007). They were also chosen because of their location; not one specific area of the pelvis was taken into account, the whole pelvis was taken into consideration for sex estimation.

1.2 Aims & Objectives

The aim of this PhD research is to determine if it is possible to create a universal equation for sex estimation using morphoscopic traits of the pelvis. This will be achieved by observing eight morphoscopic features on the human pelvis and analysing them using Ordinal Logistic Regression. Alongside this, attempts will be made to investigate how to analyse remains that are fragmented using the median as a missing data imputation technique. The main objectives are to:

- Create population specific multivariate equations for sex estimation for four known age and sex samples (Christ Church, Spitalfields, South African White, South African Black, and South African Coloured) and test the idea of population specificity.
- Apply each multivariate equation to four archaeological samples (British Medieval Poulton, British Medieval St, Owens, Chumash, and Andaman) to see if accuracies differ from previous sex estimation results.
- Investigate the use of using the median score to replace missing data and to explore the use of smaller equations to help assess fragmented skeletal remains.

This study hopes to provide a plausible universal equation to sex estimation in humans and also, to create population specific equations using the human samples studied.

1.3 Thesis Structure

This thesis has a total of nine chapters excluding the appendices and bibliography. In Chapter 2, the relevant literature will be discussed regarding the anatomy of the human pelvis, biological profiling techniques and the statistical techniques currently used to determine sex. Chapter 3 will provide information on the samples used for this research, including history, and why these particular samples have been chosen for analysis. Chapter 4 will describe the methodology and framework for the statistical analysis to be undertaken. Chapters 5 – 7 present the results from the statistical analysis and if they reject or fail to reject the null hypotheses. Chapter 8 formulates the discussion of the results and how they fit in with the current discussions surrounding sex estimation and missing data imputation. The final chapter (Chapter 9) states the final conclusions of the study and what further work can be done.

2. Literature Review

2.1 Sex vs Gender

The terms "sex" and "gender" have been used interchangeably, however within biological anthropology and other sciences they have different meanings. Sex, or biological sex, is a term that refers to the binary genetic make-up of an individual; Male (XY) or Female (XX). It includes the genetic and anatomical properties such as the sex chromosomes, hormones and genitalia. With this taken into account, all size and shape differences found between the sexes are put under the term of sexual dimorphism (Vanputte *et al.* 2010). Differences between males and females are most apparent in soft tissue, yet it is also evident in the human skeleton because of the production of sex hormones during puberty (Vanputte *et al.* 2010). These differences allow biological anthropologists/archaeologists/osteologists to estimate biological sex from skeletal material due to the morphology of skeletal elements. Because these estimations/assessments are based on biological differences, practitioners use the term biological sex, or sex, in the creation of a biological profile.

Gender, on the other hand, is a socially constructed term that has several different types of classifications and is not based on the individual's biological sex (explained briefly above). An individual's gender depends on their culture and how they classify themselves within that culture's definition of masculinity and femininity (Delphy 1993). In the UK, there are currently 14 different genders that an individual can identify as which will become part of the next national survey (Kelly 2016). The topic of gender vs sex is outside the realm of this research and therefore will not be further discussed. Because gender is not based on biological attributes it will not be referred to in this thesis; only biological sex will be discussed.

2.2 Anatomy of the bony pelvis

The human pelvis is constituted by three main elements, these being the two *os coxa* and the sacrum. The *os coxa* has three different regions being the ilia (superior); ischium (posterior inferior) and pubis (anterior). These three regions fuse at the triradiate cartilage (forming the acetabulum) during puberty (Scheuer & Black 2000) to form one bone.

These three bones have their own centre of ossification during their development. The ilium is first to form with its centre of ossification located within the vicinity of the greater sciatic notch. Ossification continues in a 'fan-like' manner. The second is the ischium with its centre of ossification located inferior to the acetabulum. Finally, the pubis starts to develop with its primary site of ossification found within the superior pubic ramus. Between the ages of 5-8 years, the fusion of the ischiopubic ramus begins, and by 11 years the posterior acetabular epiphysis starts to ossify. Fusion of the acetabulum (tri-radiate) is complete between the ages of 12-14 years for females and 14-17 years for males. This fusion finally unites all three bones into one bone. Following this, the ischial epiphysis completes its fusion between the ages of 16-18 years for both males and females. The last epiphysis to fuse is that of the iliac crest which commences ossification between the ages of 17-20 years and finishes around the age of 23 years.

The two *os coxa* articulate at the pubic symphysis anteriorly and covered by an intrapubic fibro-cartilaginous disk between layers of hyaline cartilage. The superior and inferior pubic ligaments help support this articulation (Figure 1). This articulation is a nonsynovial amphiarthrodial joint, which means that the joint has minimal movement. The auricular surface of the ilium articulates to the corresponding auricular surface of the sacrum. These sacroiliac joints are classed as amphiarthroses meaning that they are virtually fixed in place. This joint is supported by the ventral, interosseous and dorsal sacroiliac ligaments.

With these three bones articulated, it forms the pelvic cavity which primarily holds the individual's reproductive organs and the rectum. With this, the bony articulated pelvis is separated into two sections: the true (lesser) pelvis and the false (greater) pelvis. The true pelvis includes the ischium and pubis primarily. The false pelvis is superior and includes the iliac blades, iliac fossa and both the superior posterior and inferior posterior iliac spines. The boundary between the true and false pelvis is the arcuate line which runs anteroinferiorly from the apex of the auricular surface on the ilium along the iliopubic (superior pubic) ramus and finishes on the superior portion of the pubic symphysis.

The pelvic girdle articulates with the rest of appendicular skeleton via the twelfth thoracic vertebra and follows the same anatomy as the spine being; an intervertebral disc, anterior and posterior ligaments and the supra/interspinous ligaments. However, to help reinforce this joint, the iliolumbar ligament connects the transverse process of the fifth thoracic vertebra to the iliac crest, and the lateral lumbosacral ligament connects the transverse process of the fifth thoracic vertebra to the ala of the sacrum.

The 'hip' joint is the articulation between acetabulum of the pelvis and the femoral head of the proximal femur. Three main ligaments help ensure that they remain articulated being the ilio/ischio/pubofemoral ligaments. The iliofemoral ligament transcends from the anterior inferior iliac spine and the acetabulum towards the intertrochanteric line on the anterior portion of the femoral head. The ischiofemoral ligament starts on the inferoposterior portion of the acetabulum and helps form the joint capsule which attaches to the intertrochanteric line of the femur. Lastly, the pubofemoral ligament attaches the obturator crest and superior pubic ramus to the intertrochanteric line of the femur.

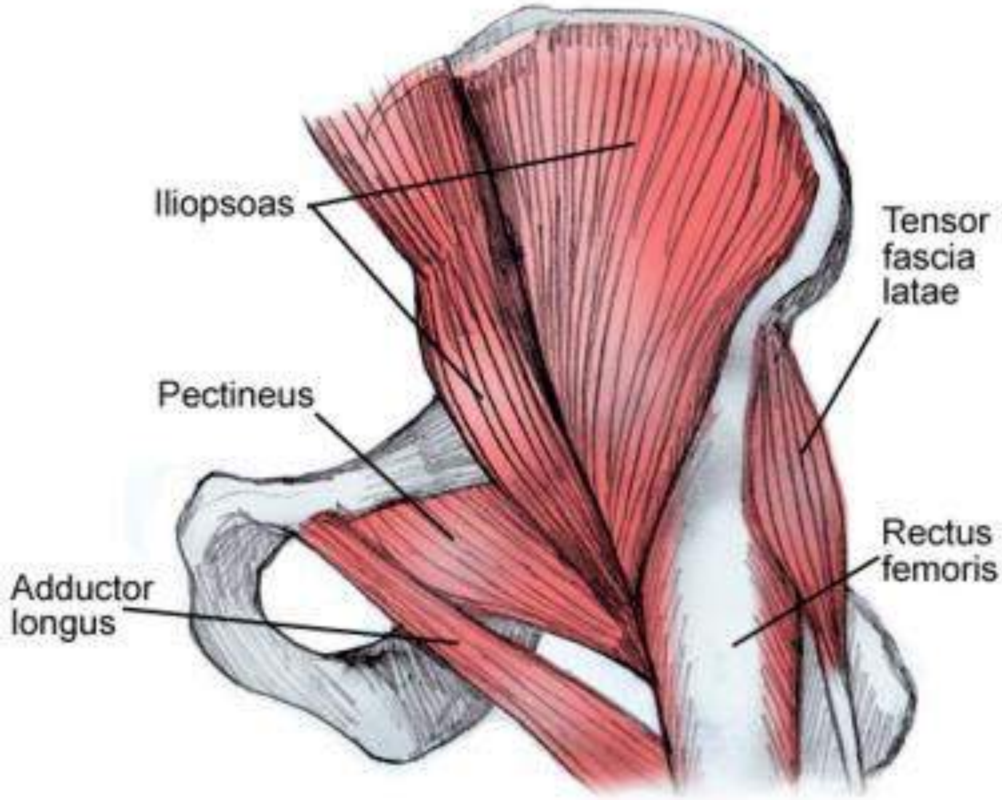


Figure 1: Lateral view of the bones and muscles for the human pelvis and proximal femur. (Taken from http://emedicine.medscape.com/article/1898964-overview#a2).

2.3 Biological Profile

When it comes to creating a biological profile for unknown skeletal remains, one of the main questions asked is 'What is the sex of this individual?', alongside matters relating to age-at-death, ancestry, and stature. Sex estimation techniques are normally only applied to adult remains, with juvenile sexing still holding a large amount of uncertainty. Adult sexing is split into two categories, metric measurements of osseous material, or a visual assessment scored along an ordinal scale. Both of these methods come with their own strengths and weaknesses, but as long as consideration and caution are taken into account, then an accurate estimation can be made. The majority of the human skeleton has been examined for sexual dimorphic traits (both metric and non-metric/morphoscopic) with the majority of the research on the skull (cranium and mandible) (Walker 2008; Krüger *et al*. 2015), pelvis (*os coxae*) (Phenice 1969; Steyn & Patriquin 2009; Klales et. al 2012), and long bones (Albanese 2013).

2.3.1 Juvenile Sex and Age Estimation

Even though sexing is not normally performed on juvenile skeletal remains, attempts have been made with varying success. Schutkowski (1993) observed and described the morphology of the mandible and ilium of the juvenile skeleton using the known juveniles from the coffin plate sample from Christ Church, Spitalfields. He noted that there was a difference in the angle of the greater sciatic notch between males and females alongside how the mental eminence in males was “flatter” than females which tended to protrude. He states that mandibular morphology, as a whole, is a better indicator for juvenile sex determination than pelvic morphology (Figure 2). One of the suggestions as to why there are signs of sexual dimorphism in juvenile skeletal material is due to the higher level of androgen testosterone hormones present in foetal males than that of their female counterparts (Grumbach & Conte 1992; Saunders 1992). With his descriptions, males were sexed correctly at a much higher rate than females for both mandible and pelvic traits, with the mandible being a better sex discriminator for males (age-at-death pooled for ages 0-5 years). With the popularity of Schutkowski’s (1993) method, Cardoso & Saunders (2008) applied the morphoscopic descriptions on a known age and sex juvenile sample from Portugal (Bocage Museum/National Museum of Natural History, Lisbon, Portugal) and tested its observer error and accuracy rates. They found that there were major difficulties in not only applying the traits but also how reliable the technique is.

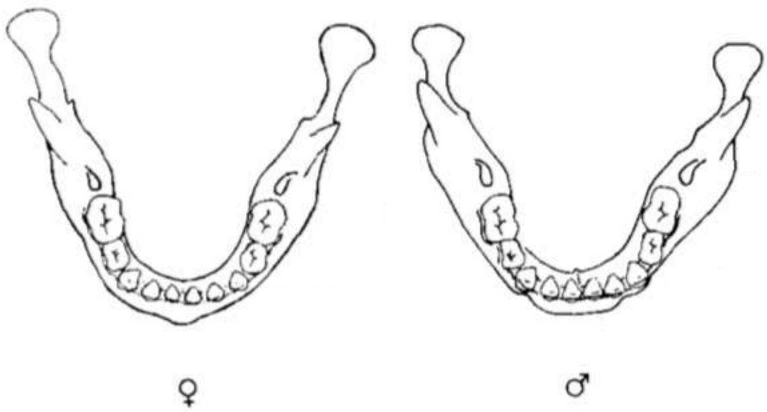


Figure 2: Sex differences between juvenile female (left) and male (right) mandibles. (Image adapted from Schutkowski 1993 pg, 200).

Wilson *et al*. (2008) re-evaluated the juvenile’s ilia morphology and analysed the original descriptions by Schutkowski (1993) using a geometric morphometric approach. After analysing the same Spitalfields sample, they found that the greater sciatic notch shape criterion can be used for sexing human juvenile remains; however, they state that caution must be taken if applied onto other samples/populations. They had three observers with a range of expertise spanning from “inexperienced” to a “senior osteologist” to test Schutkowski’s (1993) method. Stull (2013) aimed at developing advanced statistical modelling in ageing and sexing juvenile remains (between neonate and 12 years) in South Africa using metric measurements of long bones. Long bones were chosen because sex determination for juveniles originally focussed on what is sexually dimorphic for adults (like the pelvis and skull). Flexible Discriminant Function Analysis (MDA) was used for sex estimation with the option of using a bootstrap to help with the classification accuracy.

Multiple Adaptive Regression Splines (MARS) were utilised for estimating age-at-death for non-adults (Stull *et al*. 2014). With these methods, estimations could be undertaken even if not all six long bones were present. From this, they developed a statistical package within "R" called "KidStats" which allowed quick entry of measurements and instant results of the sub-adult's estimated age and sex (with associated probabilities). Her dataset was based off roentgenograms taken from hospitals from across South Africa, so age and sex were known (Stull 2013). Stull found that when it came to sex estimation, the width of the distal femur and posterior tibia were the most effective (Pers. Comms. Kyra Stull 2015). With this in mind, caution must be taken, and the probabilities must be properly understood before stating the sex of a sub-adult individual. Stull *et al*. (2017) looked at the how classification rate changed at the 95% confidence value when applying different statistical models depending on the number of variables used. Overall, single bone models achieved low classification rates (which is to be expected) while multivariable models resulted in a much higher classification rate. They found that Linear Discriminant Analysis (LDA) averaged around the 75% classification mark when age was excluded from the analysis, whereas Logistic Regression (LR) and Flexible Discriminant Analysis (FDA) achieved accuracies in the 80% margin. Because of the issues arising with the reproducibility, reliability, and the need for population specific standards, sexing sub-adult skeletal remains is still a controversial topic especially when DNA analysis is not available.

Because sexing juvenile remains is not entirely accepted by the anthropological/archaeological community, an age-at-death estimation must be completed to warrant a reliable sex assessment. Normally, sex estimations will be conducted on individuals that have been aged as 18 years or older. To determine whether an individual is 18 years or older, several observations and measurements can be made. Dental development is an excellent assessor of sub-adult remains and can determine the age of a juvenile within a six-month error margin (especially for neonates) (Alqhatani 2008; Alqhatani *et al*. 2010). For the older ages of sub-adults, the eruption of the third molar is normally an indicator that the individual was at least 18 years of age; however, a large amount of variation in the timing of eruption is present (Turner *et al*. 2013). Alongside dental development, long bone lengths and epiphyseal fusion of long bones are good indicators of sub-adult age-at-death estimations. The former is primarily used for younger individuals from foetal to twelve years while the latter for the adolescent years. Fazekas & Kosá (1978) published their work on skeletal dimensions of foetal remains which are still used as a standard for today. Schaefer *et al*. (2009) later reviewed Fazekas & Kosá's measurements with other studies to give a comprehensive view of the variation seen across populations/samples giving practitioners an idea of what their error of margin could be for an unknown juvenile individual. For adolescent ageing, epiphyseal fusion is normally observed. The three bones of the pelvis (ilium, ischium, and pubis) and distal radius fuse between the ages of 11- 16 for females and 14-16 for males and the fusion of the proximal ulna between the ages of 12-15 for females and 14-18 for males (Schaefer *et al*. 2009). However, when trying to ensure that the individual is 18 years or older, the coracoid process should be fully fused (Coqueugniot & Weaver 2007, Schaefer 2008), full fusion of distal femur (Cardoso 2008) including the femoral head (Jit & Singh 1971), fusion of the iliac crest (Web & Suchey 1985, Schaefer 2008), anterior inferior iliac spine (Coqueugniot & Weaver 2007), and fusion of the anterior portion of the ischial tuberosity (Jit & Singh 1971, Cardoso 2008) should all be noted. If remains were to have the complete fusion of the medial portion of the clavicle (Jit & Kulkarni 1976, Webb & Suchey 1985, Schultz *et al*. 2005), the spheno-occipital synchondrosis (Coqueugniot & Weaver 2007), and/or the complete fusion of the S1-S2 and S2-S3 sacral segments (Coqueugniot & Weaver 2007) then the individual will be classed as being over the 18 years old threshold.

2.3.2 Adult Age-at-Death Estimation

For ageing adult remains, metric measurements are very few and far between as the focus shifts onto the degenerative changes the skeleton undertakes. Normally the cranium and the pelvis are taken into consideration when performing an age-at-death estimation; however, research has been undertaken into using the mandible, femur, and dentition.

One of the major age assessors is the degenerative change seen on the auricular surface of the ilium. One of the first thoughts that the degeneration of the auricular surface was due to age related changes were proposed by Sashin in 1930. However, Schunke (1938) opposed the idea and stated that it was in fact due to the movement of the joint. Described by Kobayashi in 1967, he noted that there seemed to be differences present on the auricular surface between individuals of different ages. Kobayashi (1967) examined 97 male and 45 female human skeletons ranging in age of 10-81 years and described the skeletal changes that were appearing. After observing the auricular surface, he noted the degenerative changes and placed them into five age groups (18-21, 22-30, 30-40, 40-50, and 50+ years) and produced images to help with the identification of these new age ranges. One thing to note from this would be its small sample size and many age groups. This small sample size would mean that each group would have less than 20 males and ten females present which does not create a robust method statistically. He later used this new ageing system to help with the palaeodemographic analysis he was conducting on several time periods in Japan (Kobayashi 1967). In 1985, Lovejoy and colleagues (Lovejoy *et al*. 1985b) reviewed the descriptions created by Kobayashi (1967) and developed a more detailed view of the metamorphic changes from reviewing material from the archaeological Libben sample and Todd collection. Instead of the five groups proposed by Kobayashi (1967), Lovejoy *et al*. (1985b) expanded it to have eight categories that pushed the older age category from “50+” to “60+”. With this further age category, it was possible to get a better idea of what senescent life would have been like for the older generations within a population (for an archaeological viewpoint) and would further help forensic practitioners to prove a more refined age-at-death estimation for the individual. Buckberry & Chamberlain (2002) re-assessed Lovejoy *et al*. (1985b) descriptions of the auricular surface as issues with their methodology had arisen, especially with the confusion of some age stages. They stated that Lovejoy *et al*. (1985b) had oversimplified the age groups which caused difficulty in underpinning a single age group when features overlap. After they developed a new scoring system, they tested the technique on the Spitalfields sample held at the Natural History Museum, London and found that it had a better accuracy for sex and a higher correlation to known age than previously outlined by Lovejoy *et al*. (1985b). Mulhern & Jones (2005) tested the revised method by Buckberry & Chamberlain (2002) on the Terry and Huntington Collections held at the Smithsonian Institute (USA) and tested if there were observable differences in not only males and females but also between white and black Americans. They found that, similar to Buckberry & Chamberlain (2002), that no differences were found between males and females and also that the revised technique worked well for both white and black Americans. However, they found that the original method proposed by Lovejoy *et al*. (1985b) performed better for young and middle adult stages while the revised method yielded more accurate results for the older individuals in their sample. A re-examination of the auricular surface morphology was undertaken by Falys *et al*. (2006) using the coffin plate sample of St. Brides (held at St. Brides Church, UK). They agreed with the revised method proposed by Buckberry & Chamberlain (2002), stating that it is more practical to use than the Lovejoy *et al*. (1985b) method, however, they put forward recommendations for the revision of some age categories to help refine the large amounts of variation found. Moraitis *et al*. (2014) had the same view point of Falys *et al*. (2006) concerning the easier application of the revised method (Buckberry & Chamberlain 2002) and added that because of this, it might be preferable to use for inexperienced practitioners. They also agree with the idea raised by Hens & Belcastro (2012) stating that the revised method is more in line with general stages in an individual's life rather than a specific age category.

The morphology of the pubic symphysis has been highly regarded and used proficiently as an indicator of an individual's age-at-death. The metamorphic changes were first developed into an ageing method by Todd in 1920 and subsequently further described by him in 1921 (Todd 1920, 1921a, 1921b, 1921c). From these, he created a ten-phase system for age estimation which was capped to a maximum age group of "50+" which included both black and white males and females. Brooks (1955) undertook a test of reliability over the Todd method to which he found that his method overaged individuals normally within their 30's and 40's. Mckern & Stewarts (1957) approach to the pubic symphysis was highly criticised when they created a three-component system from analysing individuals from North Korea. The major criticisms put forward was that their method was only created and tested on the male population and only had a small age range. After these were made apparent, Gilbert & Mckern (1973) increased the original sample of 349 North Korean males to include 103 females from the same collection. Even with the authors addressing the initial concerns, Suchey (1979) tested their new method on a different sample with little accuracy. This little accuracy can be explained by only testing the method on 11 known age-at-death pubes. Meindl *et al*. (1985a) reviewed and revised the methods that were currently using the pubic symphysis for age-at-death estimations. They used the skeletal sample from Hamann-Todd (held at the Cleveland Museum of Natural History, OH, USA) to test the original Todd (1921) method, the revised McKern & Stewart (1957)/Gilbert & McKern (1973) method, and the regression model that Hanihara & Suzuki (1978) created. They found that they were in agreement with a statement that McKern & Stewart (1957) made about the Todd method, which is similar to what Hens & Belcastro (2012) mentioned about the auricular surface; that the metamorphosis seen are more related to general life stages than actual chronological age. However, with this they revised Todd's ten stage method with a more in-depth description of a better application for both the forensic and archaeological scenarios. Brooks & Suchey (1990) evaluated and modified the morphology of the pubic symphysis into six main stages that were sex specific ranging to the upper limit of "60+" years (Figure 3). There have been more modern attempts to create new standards for using the pubic symphysis, for instance, Dudzik & Langley (2015), who created a component based system that not only included the morphology of the pubic symphyseal face (i.e. billowing) but also features surrounding it like the appearance of the pubic tubercle. Even though their method achieved a high percentage accuracy (cross validated), they only had three age categories with the last category having the age range of 33-40 years. With this limited age range, it is understandable why methods such as the Brooks & Suchey (1990) and Todd (1921) have remained popular with practitioners.

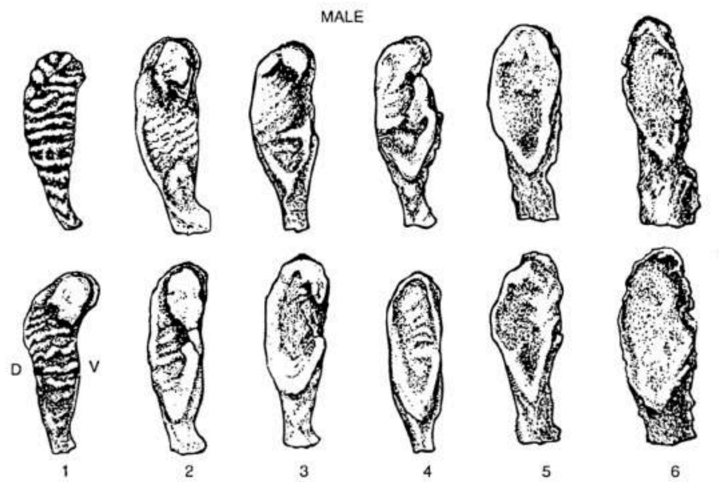


Figure 3: Male pubic symphysis showing the changes according to the Suchey-Brooks (1990) method for estimating age-at-death. (Image adapted from Buikstra & Ubelaker 1994 pg. 24).

Other age-at-death techniques that do not involve pelvis have also been developed, one being ectocranial suture closure. With similarities to the pubic symphysis ageing method, Todd & Lyon (1924, 1925a, 1925b, 1925c) conducted observations on both ectocranial and endocranial suture closure to aid in the age-at-death estimations using human crania, and to see if a person's sex or ancestry caused any variation. In their first article (Todd & Lyon 1924), they stated that a trend did exist in the progress of suture closure relating to the individual's age even though they rejected a large number of crania from the analysis. In Todd & Lyon's later papers (1925a, 1925b, 1925c) they planned to compare endocranial sutures to their ectocranial counterparts and found that the endocranial were less erratic and that the timing of the closures was more predictable than ectocranial sutures. Considering the great popularity of this technique, it was not until 1953 when Singer reported on how unreliable the reports produced by Todd & Lyon were (Singer 1953). This related to the odd selection criteria that Todd and Lyon had created to where they excluded crania that had an irregular suture closure and that females were also excluded from their original sample. Further to this, Brooks (1955) produced an article using the methods created by Todd & Lyon (1924, 1925a, 1925b, 1925c) on native American remains from California. He found that closure of the sutures for females had a discrepancy of up to 25 years and for males up to 8 years. One major disadvantage with Brooks' 1955 study was that he was testing cranial suture closure on an archaeological sample where known age was not available. In 1985, Meindl & Lovejoy (1985a) published a new method of scoring the ectocranial suture closure based on crania from the Hamann-Todd skeletal collection (Figure 4). They chose only to include ectocranial sutures instead of the endocranial (which was preferred by Todd & Lyon) because "…ectocranial activity is far more closely associated with extreme age (for which new forensic standards are most needed)…" (Meindl & Lovejoy 1985a pp58). From analysing 17 sites from 236 crania, they found that the sutures located along the lateral-anterior portion of the crania were the best for estimating age-at-death and state that no sex or ancestry bias was observed. With the improvement of the suture closure method, it is still a method that is considered to be highly variable and inconsistent with its results. Therefore, should only be utilised with other age estimation techniques (Singer 1953, Brooks 1955, Meindl & Lovejoy 1985a). Key *et al*. (1994) reviewed three cranial suture techniques by applying them to the known sex and age-at-death Spitalfields collection held at the Natural History Museum. They first applied the Acsádi & Nemeskéri (1970) method of endocranial suture closure and found that it was a viable method for age estimation, however only for determining if an individual is "young", "middle-aged", or "old"; not for it to be used for specific age categories. From this, Key *et al*. (1994) applied the Meindl & Lovejoy (1985) method and found a large difference between the scores of males and females from the Spitalfields sample. Alongside the significant difference between the sexes, bilateral asymmetry and a large error rate were noted when comparing the results to the known age of the individuals. Because of these findings, Key and colleagues (1994) highlight that Meindl & Lovejoy's (1985) method using ectocranial sutures was not appropriate for use on this specific sample. The third method was created by Perizonius (1984), where by two different scoring techniques were formulated; one for individuals younger than 50 years, and the other for individuals older than 50 years. They found that this technique was insufficient in its application when concerning the Spitalfields sample and found that the technique for 50+ years showed no relationship with age for the older individuals. Only 47.5% of individuals younger than 50 years were recognised as being younger than 50 by the system designed for the younger age group. However, of the 47.5% that was recognised, a strong relationship between the Perizonius inferred age and actual age-at-death was noted. This correspondence was mainly due to the assessment of the endocranial sutures rather than the ectocranial sutures. After reviewing the three methods, they state that cranial suture closure is a viable method for age-at-death estimation, however it is best to only use them for broad age categories, i.e. young, middle-aged and old individuals.

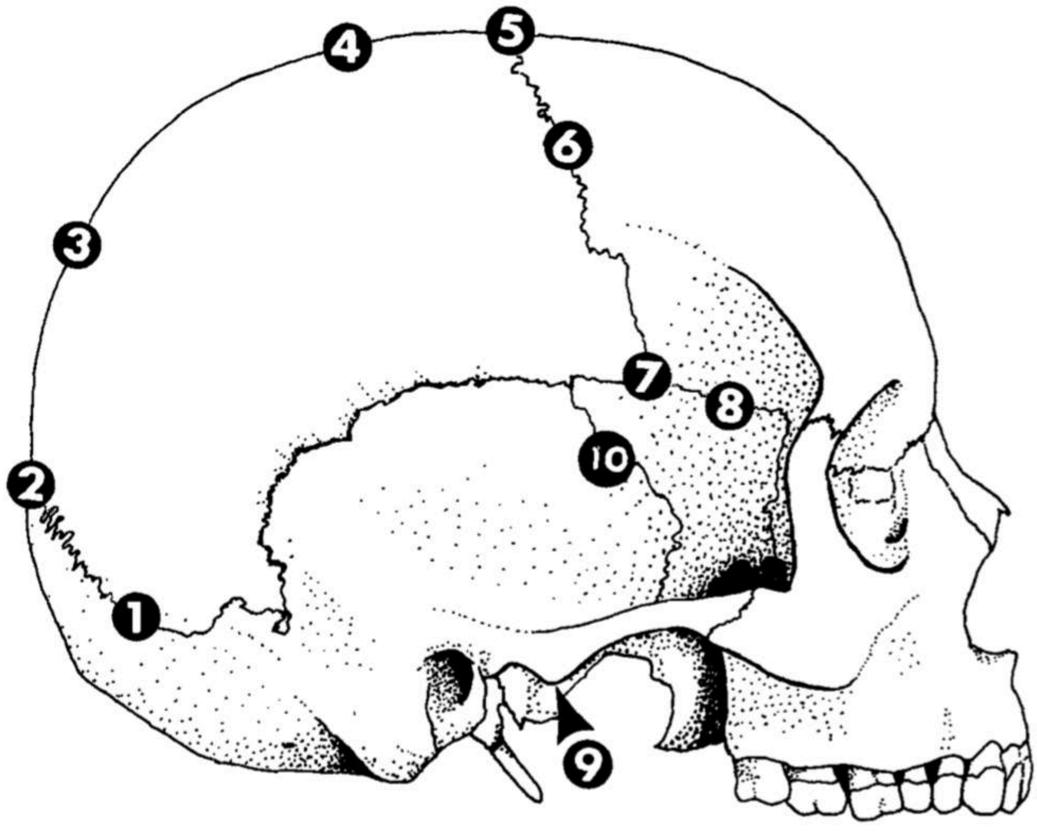


Figure 4: Cranium in lateral view showing ten of the 17 ectocranial suture sites. (Image adapted from Meindl & Lovejpy 1985a pg. 60).

With the overall view of using as many indicators as possible for age-at-death estimations, Lovejoy *et al*. (1985a) proposed a multivariate statistical approach that combined several methods to estimate a skeleton of unknown age. Their technique utilised the auricular surface (Lovejoy *et al*. 1985b), pubic symphysis (Meindl et al. 1985a), trabecular involution of the proximal femur (Walker & Lovejoy 1985), cranial suture closure (Meindl & Lovejoy 1985), and dental wear (Lovejoy 1985a). They compared the multivariate model alongside seriation results. They found that after the analysis of the Hamann-Todd collection, Summary Age (multifactorial age estimation technique) had a high correlation with known age-at-death and only presented a small amount of bias. They hypothesised that the bias was caused by the different sample sizes between males and females which resulted in different mean age-at-death for both groups skewing the data set (Lovejoy *et al*. 1985a). After the results of this test, Lovejoy and colleagues (1985a) conclude that their multifactorial ageing technique "Summary Age" outperforms any of the age's based on a single indicator. In 1993, Bedford and colleagues applied the Summary Age method onto 55 known age and sex skeletons from the Grant Collection (Toronto CA). However, this article was followed by a string of comments regarding their claim for being able to estimate age that was "statistically indistinguishable from those of real age" (Bedford *et al*. 1993 pp297) and raised the issue with their small sample size and statistical reasoning (Fairgrieve & Oost 1995). Boldsen *et al*. (2002) created a new approach to using multiple indicators for age-at-death estimation and dubbed it "Transition Analysis". They focused on three main and well-used areas being: the auricular surface, pubic symphysis and cranial sutures (reviewed above). The two pelvic indicators were broken down into several components which follow a new scoring system while scoring the ectocranial sutures followed a similar standard that has been performed for the past 50 years. A total of five cranial sutures, five components on the pubic symphysis, and nine components of the auricular surface are to be scored for each skeleton for analysis. However, normally when missing variables are present, multivariate analyses become difficult to use as not a full data set is present. In the case of Transition Analysis and its basis on Bayesian statistics and the ADBOU computer programme, analysis can still be conducted if variables are missing. Milner & Boldsen (2012) produced a validation report using the Bass Donated Collection and individuals from the Mercyhurst University. They found, like with Summary Age (Lovejoy *et al*. 1985a), using single indicators can be detrimental to the estimation of age, especially when using cranial sutures. Milner and Boldsen also state that their method does not work as well as they would originally hope, but with further work (i.e. refining component descriptions for scoring) and analysis of new skeletal samples, improvements are more than likely.

2.3.3 Ancestry Estimation

The concept of race and its validity to define human populations has been debated and scrutinised for over a century (Ousley *et al*. 2009). In the past, the classification of race used arbitrary traits such as skin colour, skeletal morphology, facial characteristics and sometimes even behavioural traits to divide humans into three distinct biological groups: "Caucasoid", "Negroid", and "Mongoloid" (Stein & Rowe 1989). Because of these distinctions of racial groups, it inspired the idea that there existed an ideal set of characteristics which inevitably lead to discrimination, segregation, and racist thinking between groups (Brace 1995). Human populations are more diverse than just the biological traits they have; they are also influenced by culture, religion and language including their geography, making race, not just a biological concept, but also a social concept as well (Edgar & Hunley 2009; L'abbe *et al*. 2013a).

Biological anthropologists are frequently asked to estimate the ancestry of skeletal elements by those in law enforcement to aid in the identification of the unknown individual. They are also asked to help identify ancestry in cases of repatriation of skeletal remains within both forensic and archaeological practices. Typically, the cranium will be used to make this estimation using metric and morphoscopic analyses and has been stated as being the best element to use for this task (Isçan & Steyn 1999; Hefner 2009; L'Abbé *et al*. 2011). Giles and Elliot (1962) applied multivariate analyses – DFA – to classify different populational groups using craniometrics to which then lead to Howells (1973) creating a worldwide databank of cranial measurements to aid in population variation. The most current system for ancestry estimation from cranial measurements is the FORDISC 3.0 program (Jantz & Ousley 2005). This program takes measurements outlined by Buikstra and Ubelaker (1994) and attempts to place the skull in one of the many population groups the Forensic Databank has. The FORDISC 3.0 program also has access to the Howells databank to look at a much wider geographic distribution. Even though it has the huge popularity for ancestry estimation, it does not come without its faults. The faults can range from an incorrect interpretation of the results or that the individual's population is not part of the databank (Ousley, 2015, Pers. Comms.). These faults came to attention in 2016 when South Western Hispanics were being classified as Asian, or more accurately as Japanese (Dudzik & Jantz 2016). Alongside the metric analysis of the skull, morphoscopic observations have also been under scrutiny and rigorous testing for their applicability to ancestry estimation. The majority of cranial traits that are taught and used in investigations will be those outlined by Rhine (1990). However, when observing the data, traits were not necessarily attributed to a particular population. Hefner (2009) analysed 11 traits and created new, more refined, descriptions to help categorise each trait and found that once applying a good statistical framework, the use of morphoscopic cranial traits became more promising for ancestry estimation (Figure 5). Following Hefner (2009), L'Abbé (2011) investigated the use of morphoscopic cranial traits for South African populations. They found that using only a visual and experienced based assessment the results were not reliable, only when under the scrutiny of a statistical framework can they be used for estimating ancestry.

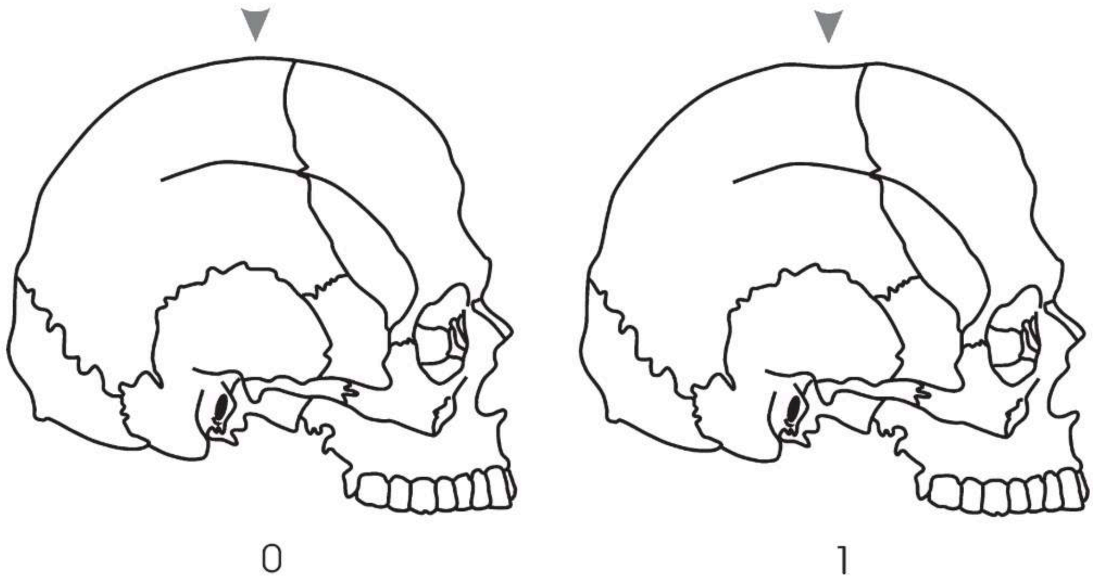


Figure 5: Lateral view of crania showing the absence (left) and presence of a bregmatic depression (right). (Image adapted from Hefner 2009 pg.990).

After the cranium, the mandible has been used for ancestry estimation using both metric and non-metric analyses. Kile (1983) (found in Berg 2008) applied DFA onto 25 metric variables from the mandible using the Terry collection to help differentiate between American Whites and Blacks. He found that accuracies ranged between 38.5% and 76.9% depending on variables used and state that the mandibular corpus and ramus had the best discriminatory power. Rhine (1990) also stated mandibular traits as well as cranial traits but with the same issues as before; the conclusions made were based on small samples sizes which created sweeping generalisations about population affinity. Berg (2001) examined seven morphoscopic mandibular traits from three populations but only assessed males. From this analysis, it was found that it was incorrect to substitute Native Americans with Asians as their morphology differed greatly. This was followed several years later when Berg (2006) who applied these ordinal scores of non-metric traits to linear discriminant functions to determine ancestry with a leave-one-out cross-validation method. Berg found that a two-group comparison had the best result (as expected) but when a new comparative group was added, accuracy declined. An average of 74% accuracy was obtained when observing two closely related groups for determining ancestry. Buck & Vidarsdottir (2004) attempted to estimate ancestry for subadult mandibles using three-dimensional geometric morphometrics. After obtaining 17 landmarks from five different sample populations, the cross-validated linear discriminant functions produced ~70% accuracy for the five-way comparison. To mimic fragmentation, Buck & Vidarsdottir (2004) split their data into two sets: the corpus and the ramus, and found that the ramus was a better predictor of ancestry than the corpus (73% and 67% respectively).

The assessment of postcranial elements has not been in favour for ancestry estimation as researcher's state that they provide inconsistent and unreliable results (Stewart 1979; Albanese & Saunders 2006). Because of this, limited work has been conducted on how well they can perform. Morphological characteristics of the femur have been used (Stewart 1979) to how many talar articular facets are present on the superior portion of the calcaneus (Bidmos 2006). One of the most well used postcranial metric analyses for ancestry estimation is the creation of the platymeric index of the femur (Brown 2006). The platymeric index is created by taking the anteroposterior sub-trochanteric diameter and dividing it by the transverse subtrochanteric diameter and multiplying it by 100. The index creates "shape" information where individuals are said to be more rounded (eurymeric) or flat (platymeric) (Brown 2006). Multivariate analysis of postcranial elements has been undertaken; however, they are limited. Patriquin *et al*. (2002) and Bidmos (2006) applied multivariate analysis on metric measurements from the pelvis and the femur which resulted in accuracies averaging around 80%. Even though accuracies this high have been generated, these studies are limited by their sample size and unclear definitions of measurements taken.

2.3.4 Stature Estimation

Anthropologists and archaeologists estimate living stature of individuals using skeletal measurements and applying them to regression formulae. When a complete skeleton is present, the most accurate method for stature estimation is the revised Fully method (Raxter *et al*. 2006). This revised method followed the same principle as the original created in 1956 (Fully 1956) however it included some correction factors (e.g. the interaction of soft tissue) to the linear equation (Raxter *et al*. 2006). Raxter *et al*. (2006) found that when applying the revised method, estimates were within 4.5cm of documented stature for 95% of individuals in their sample. However, complete skeletal remains are rare, so when incomplete remains are all that is to offer for analysis, long bones such as the femur, tibia, and fibula can be used to provide estimates. The long bones of the lower limb have a strong correlation between overall living stature and limb bone length (Trotter 1970, Jantz 1992). When bone fragments are only available, researchers have investigated the use of metacarpals (Meadows & Jantz 1992) and metatarsals (Byers *et al*. 1989); however, their accuracy is much lower than those of the lower limb.

When applying most stature estimation formulae, they come with a caveat stating they are population specific, sex specific, or both, and when neither ancestry or sex can be determined then a 'less powerful' formula must be used (Feldesman & Fountain 1996). Albanese *et al*. (2011) tested this presumption by applying stature estimation equations within the FORDISC 3.1 computer program on American Whites and Blacks. They found that the sex specific and population specific stature equations were comparable to those that were not population or sex specific. To continue testing the caution of equation specificity, Albanese *et al*. (2016) developed new regression formulae based on American White and Black populations from the Terry collection and applied them to individuals within the Forensic Data Bank and the Lisbon Collection. They found that regardless of knowing the individual's geographic origin or sex, that living stature could be estimated for over 95% of their study sample (Albanese *et al*. 2016). With this being said, caution must be taken, until further research on this has been undertaken using different population samples.

2.3.5 Adult Sex Estimation

Conventional sexing of adult skeletal remains can be separated into two main categories being metric or morphoscopic analysis. Both categories have the same sub-categories also, where they split into either cranial or postcranial assessment. When a complete skeleton is present, a multifactorial approach is best (Meindl *et al*. 1985b) by taking into account both metric and morphoscopic/ cranial and postcranial features from the skeleton which can have accuracies as high 98% (Meindl *et al*. 1985b).

When observing the skull, the Workshop for European Anthropologists published a report named "Recommendations for Age and Sex Diagnoses of Skeletons" (1980). They table ten cranial and four mandibular morphoscopic characteristics for sex estimation with (very) brief descriptions of how to score each trait. From this article, the presence and expression of the glabella, mastoid process, nuchal crest, and zygomatic process are all given equal and strong weightings for sex estimation. As useful as having weightings attached to the morphoscopic traits, having one-word descriptions for scoring (e.g. smooth or marked) opens up a minefield of interpretation errors and no standard towards scoring. Another issue with this article is the lack of visual descriptions. The Workshop for European Anthropologists (1980) are "describing" morphoscopic traits yet only included two sketches, therefore, the remaining eight traits are left for personal interpretation, and all four of the mandibular characteristics are also left with the same issue. In 1994, Buikstra & Ubelaker produced a document as a set of standards for recording human skeletal material ranging from creating a biological profile to an introduction to skeletal pathologies (Buikstra & Ubelaker 1994). In their section for cranial morphology for sex estimation (Buikstra & Ubelaker 1994 pp 19), they focus on only five traits with descriptions and images for how to score each, unlike the recommendations made 14 years prior. These five traits (nuchal crest, mastoid process, supra-orbital margin, supra-orbital ridge/glabella, and the mental eminence) each were given a five-scale scoring system where a score of "1" would be defined as definite female, "3" as being ambiguous, and "5" as definite male (Figure 6). In 2008, Walker (2008) took these descriptions and illustrations and applied them to multivariate analysis to see how well these traits were at discriminating sex. Walker (2008) collected data from 304 individuals from the Hamann-Todd collection (American White and Black), Terry Collection (American White and Black), and Saint Brides Church (UK). He found that applying the descriptions from Buikstra & Ubelaker (1994) to logistic regressions; accuracies ranged between 84%-88% for correct classification depending on the number of variables used. With the need to create population specific equations for South African individuals, Krüger *et al*. (2015) applied the descriptions from Buikstra and Ubelaker (1994) to a sample of South African White and Black skulls (N=245). After the data had been collected, they were applied to Walkers (2008) logistic regression equations. Using the equations from Walker (2008), an accuracy of 68%-84% was achieved, however, when both sex and ancestry was taken into account, females were misclassified more often than males. With this issue, Krüger *et al*. (2015) created new population specific equations using ordinal logistic regressions where accuracies were 80% or higher depending on the variables used. Alongside the ordinal logistic regressions to estimate sex, Krüger *et al*. (2015) collected 19 craniometric variables and applied a polyserial correlation to assess the relationship between the craniometric variables and the five morphoscopic variables. They found that scores for the glabella were highly correlated (strong and positive) with the craniometric variable BNL and the mastoid process was highly correlated (strong and positive) with the measurement MDH.

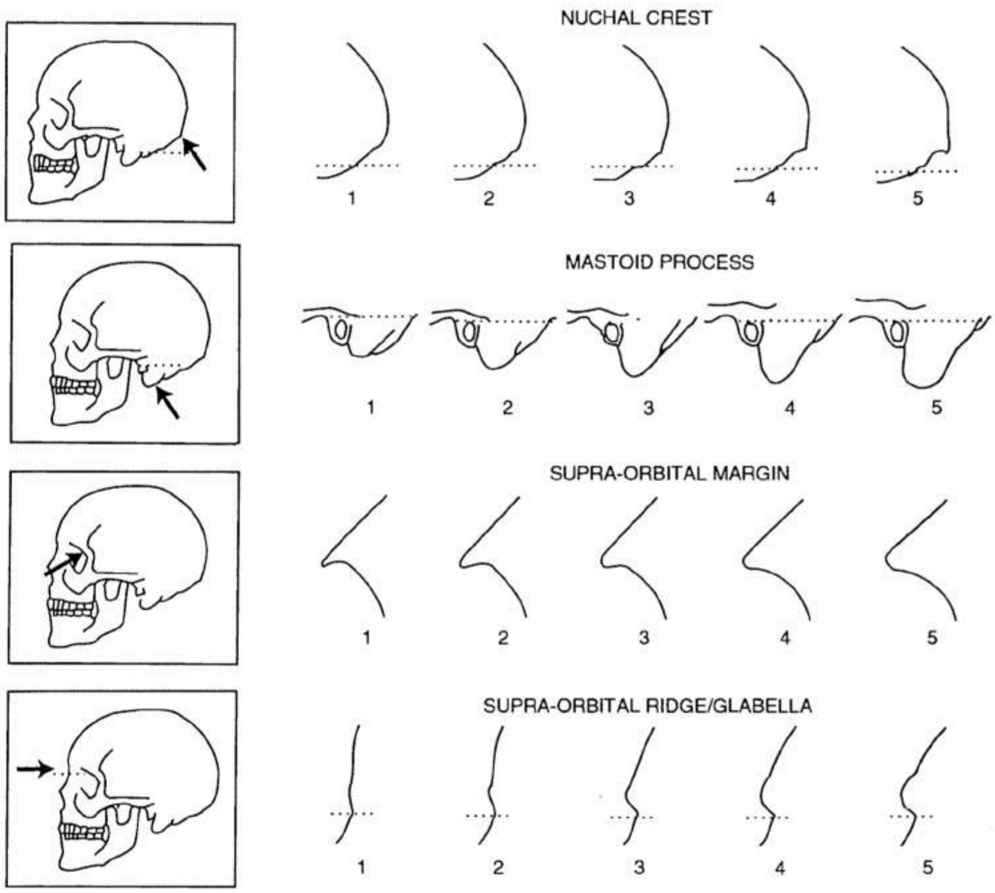


Figure 6: Morphoscopic scoring of the cranium. (Image adapted from Buikstra & Ubelaker pg. 20).

When only using metric analysis of the skull, the Workshop for European Anthropologists (1980) reported two equations that were derived from DFA. They provide the cranial DFA equation created by Giles & Elliot (1963) and Giles (1966, 1968) which used eight variables and was created from White American individuals from the Terry Collection (USA). The report states that the misclassification percentage as 13.4% and also provides the sectioning points for males and females. The Workshop for European Anthropologists (1980) provided one metric illustration but included the definitions for each measurement that needed to be taken. Even though they only provided one equation which can only be used for one population, they do provide information on who had created multivariate equations for other populations. Alongside the cranial equation, the Workshop for European Anthropologists (1980) highlight an equation that can be used for mandibular measurements that were formulated by Giles (1964). This equation was also created using White American individuals from the Terry Collection (USA) and includes three variables that need to be collected. Spradley & Jantz (2011) questioned the idea that the skull is the second best indicator for sex estimation, with the pelvis coming in first place. They evaluated over 500 individuals from the Forensic Databank and created a different multivariate equation for several individual skeletal elements. They found that when using the cranium, eight variables were chosen from the stepwise-discriminant analysis for American Whites and 11 variables were chosen for the American Black equation. For the White American sample, an accuracy of 90.64% was found and 90.01% for Black individuals. For Black Americans, the cranium was fifth best out of 14 equations, where the humerus resulted in an accuracy of 93.84%. The cranium was eighth best for American Whites, and the best skeletal element was the radius for sex estimation with an accuracy of 94.34%.

Postcranial sex estimation normally involves metric analysis of long bones (e.g. humerus and femur). The most highly used metric measurements taken for sex estimation using the femur are on the femoral head (Asala 2001). One of the first examples of how the femur could be used as a sex indicator in Britain was done by Parsons (1914). Parsons (1914) used a British-Medieval sample (n=300) and took multiple measurements including maximum femoral length, the diameter of the femoral head and maximum antero-posterior diameter of the femoral shaft. Parsons (1914) found that the maximum diameter of the femoral head was the most useful due to its availability in archaeological remains. Parsons (1915) refined his previous method using a modern population. He found that 65% of the femora could be correctly sexed with just the vertical measurement of the femoral head; the rest would need further femoral measurements taken to aid in the estimations. He concluded that this updated method should give an accuracy of 91% - 92%. Steel (1964) took Parsons (1915) research and attempted to create new formulae into sexing using long bones. He concluded that even though Parsons states that the femoral head is the best indicator of sex, it is, in fact, the bicondylar width of the femur that is more sexually dimorphic. Steyn & Iscan (1997) thought it was best to use not only the femur but also the tibia to estimate sex using osteometric analysis. The analysis that included measurements based on "width" and "circumference" dimensions yielded better sex classification than measurements that only incorporated "length" dimensions (Steyn & Iscan 1997). Circumstances in South Africa resulted in the need to identify individuals using only fragmentary remains. Asala (2001) found that the femoral head had symmetrical benefits and did not show any bilateral asymmetry. Asala (2002) looked at the demarking points for males and females on the femoral head to see if a higher accuracy could be achieved than individuals from south eastern Nigeria with original accuracies ranging between 29.6 to 55.1% (Asala 2002). Sadly, Asala (2002) reported accuracies of 32% was found for both white and black South Africans. Purkait (2003) suspended the femur in "life position" and collected four measurements from the proximal portion. Purkait (2003) managed to achieve an accuracy as high as 92.1% for identifying sex by using the vertical and horizontal diameter of the femoral head. The femur can be used not only in modern forensic investigations but for a prehistoric context also (Murphy 2005). Murphy's (2005) article showed that accuracies ranging from 80.9 – 82.4% can still be achieved with the use of discriminant function analyses; however, this is based off other assessments rather than known sex. The use of radiographs has now been incorporated into sex identification (Harma & Karakas 2007; Kranioti *et al*. 2011). Harma & Karakas (2007) used radiographs and managed to correctly sex 77% of the sample just by observing the vertical diameter of the femoral head, with the use of the student t test.

The human pelvis has been highlighted as being the best skeletal element for sex estimation. Metric analysis of the pelvis normally relies on the principals that females have bigger pelves than males due to the additional functional pressure's females have (Tague 2000). Patriquin *et al*. (2005) studied 400 specimens from the Raymond A. Dart and Pretoria Skeletal collection to understand the sex differences between South African black and whites pelves. They took nine measurements across the os coxa and were chosen for their: expected sexual dimorphism, landmarks that can be easily identified for measuring, and for their likeliness to be present when fragmentation has occurred. From the nine measurements, they applied stepwise and direct DFA and created six functions for South African Whites and South African Blacks. When all measurements were used, accuracy ranged from 90% -98% of the two samples once the cross-validated procedure had been implemented. Accuracies did drop quite dramatically when the other five functions were used. These five functions were created by grouping measurements that were close to each other, i.e. Function 2: Sciatic Notch – Sciatic Notch Width and Sciatic Notch Depth. This second function had a much lower accuracy of ~75% for both sample populations (Patriquin *et al*. 2005). Even though Patriquin and colleagues created these equations for sex estimation, they did not investigate how the equations would perform if ancestry were unknown.

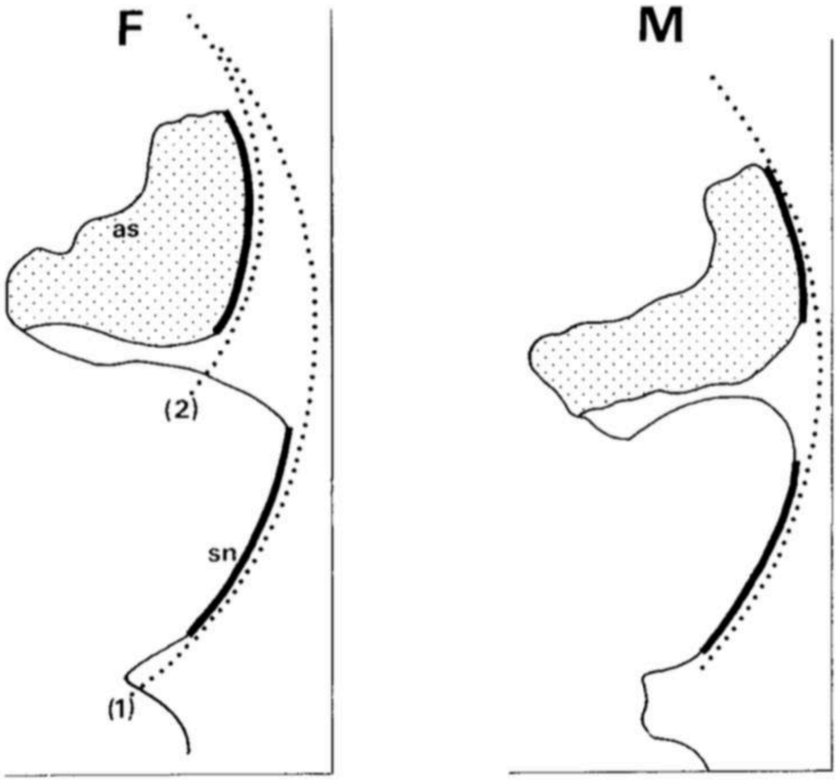
Steyn & Iscan (2008) measured 192 os coxa from Heraklion (Crete) to try and establish a specific metric sexing equation for a contemporary Greek population. They originally had 15 variables that were taken from the os coxa and sacrum and applied a step-wise DFA to establish the measurements that were the 'most sexually dimorphic'. The remaining six functions were created by focussing on certain areas of the pelvis, e.g. the Sciatic Notch by applying only certain variables to a "direct" DFA analysis. The best function for sex estimation was from the stepwise DFA where six variables were chosen (Acetabular Diameter, Sciatic Notch Breadth, Total Height, Pubic Length, Ischial Length, and Pubic Tubercle-Acetabulum Length) and resulted in a cross-validated accuracy of 93.5%. From the remaining six functions, "Function 2" had the greatest accuracy of 89% (variables from the pubis and ischium) and "Function 5" had the lowest accuracy of 59.4% (variables taken from the sacrum). To test the idea of population specificity, Steyn & Patriquin (2009) assessed seven pelvic measurements from two modern day South African samples and one contemporary Greek sample from Crete. Steyn & Patriquin (2009) created five discriminant function equations for each of the three samples and then combined all three samples together and repeated the analysis to compare the results. From the first round of analysis (when the samples were evaluated separately) they achieved average accuracies of 93.5% - 94.1% when all measurements were included. This result falls within previous accuracies found by Patriquin *et al*. (2005). As fewer variables were included, accuracies did decrease. These results again mirror that of Patriquin *et al*. (2005) where accuracies were between 73% - 79%. When all three samples were pooled together, accuracies did not vary greatly between the individual analyses. From these articles, when all measurements were not present, using variables from the pubis and ischium were better at discriminating sex than those located near the acetabulum and greater sciatic notch (Patriquin *et al*. 2005; Steyn & Patriquin 2009). Gómez-Valdés *et al*. (2011) created several discriminant function equations for sex estimation based on a contemporary Mexican sample from the National Autonomous University of Mexico. A total of 24 metric measurements were taken from both coxal bones and the sacrum resulting in three groupings of equations (left os coxa, right os coxa, and sacrum). From the step-wise DFA analysis, four equations were created for each os coxa where if all (four) variables were used then accuracy could range between 97.9% (left os coxa) and 99.1% (right os coxa). These equations included variables from across the os coxa such as total pelvic height, transverse acetabular diameter, pubic length, and iliac width. When observing the sacrum, three variables were chosen: anterior-posterior diameter of the base, transverse diameter of the base and anterior superior breadth of the sacral body, resulting in an accuracy of 86.8%. Accuracy dropped by 11% when the anterior-posterior diameter of the base was not used alongside the other two variables. Moving away from dry remains, Decker *et al*. (2011) evaluated the use of collecting measurements from 3D CT pelvic models from living patients. They originally took 12 distance measurements and two angles from the reconstructed pelves and ran a Pearsons Correlation Coefficient to determine which variables would be the best at predicting sex. From this analysis, four variables were chosen: innominate height, greater sciatic notch angle, subpubic angle, and transverse pelvic outlet. From these four variables, a binary logistic regression was used to predict sex which resulted in a 100% accuracy for both males and females. Instead of ending the research at that point, Decker *et al*. applied their sample (60 males and 40 females from patients at the University of South Florida College of Medicine) to the inbuilt postcranial sex estimation package in Fordisc 3.0. For this to run, seven variables were chosen and an accuracy of 86% was achieved where females were classified more successfully than males (98.3% and 67.5% respectively). Considering that Decker and colleagues achieved a perfect result, it can be said that they 'cherry picked' the best variables for the job – it would be interesting to see how well the other variables would estimate sex. Another issue would be the comparison between their binary regression and the results from the Fordisc 3.0 output. Only one variable remained the same between both analyses (the innominate height), which begs the question "Can you compare these results?"

Nagesh *et al*. (2007) looked at investigating an index that took into consideration the pubis (seen previously to be a good sex assessor) and the acetabulum for sex estimation in a South-Indian sample. They took the pubis length and acetabulum diameter and created an index value. Once the index was calculated, an identification point and "demarking point" was established. When a sectioning point was used to estimate sex, an accuracy of 81% for females and 83% males were found. Accuracy did decrease when using the identification point where the majority of accuracies were found between 53% for males and 48% for females. Accuracy took an even more dramatic decline when the demarking point was used where only 17% of males and 26% of females were correctly identified. There are some issues with this article being the terminology used (gender and sex are used interchangeably) and also descriptions/illustrations created for data collection (Page 306 Figure 1. Photograph of a penned outline of an os coxa on creased paper). Nagesh *et al*. (2007) also did not provide an inter-observer error analysis.

Chapman *et al*. (2014) applied the Diagnose Sexuelle Pobabiliste (DSP) tool, created by Murail *et al*. (2005), onto 49 dry pelves from the Body Donation Programme of the Université de Bruxelles however sex was not known. They compared visual assessments to the measurements defined by DSP to assess how well the DSP tool estimated sex. When comparing the visual assessment to the manual measurement taking for DSP variables, there was a 94% agreement between the two. From here, they compared the manual taking of DSP measurements to virtually measuring CT scans of the same individuals and found a 100% agreement in sex assignment. One of the benefits of the DSP tool is that its database contains 2040 individuals from 12 different reference populations (Murail *et al*. 2005).

Benazzi *et al*. (2008) used a novel approach of using the profile of the acetabulum to estimate sex for forensic cases due to its robusticity. They used AutoCAD to draw the profile of the acetabular rim and created four variables: area of the acetabular profile's internal surface, acetabular perimeter, the difference in max and min values along the y-axis, and the difference in max and min values along the x-axis. After applying step-wise discriminant function analysis and using all four variables, a 94.9% accuracy for males and a 97.7% accuracy for females was achieved. Even though such a high accuracy was achieved for both males and females, the amount of time it would take to extract the relevant information would be large. Macaluso Jr. (2011) used a very similar method for assessing pelvic remains from the George Olivier Collection housed at the Musée de l'Homme (France). The acetabular rim was recorded using ImageJ, and three variables were collected: Acetabular diameter, area, and perimeter. From here, both DFA and Logistic Regressions were created to assess how well these three variables were for estimating sex. Macaluso Jr. (2011) only used univariate models for both the statistical tests as accuracies did not increase when the other variables were used. With this is mind accuracies did not exceed the 90% margin like the Benazzi *et al*. (2008) study. From applying photogrammetry to collect metric variables, Gonzalez *et al*. (2009) used 2D images of dry pelvic bones and applied Geometric Morphometric analysis to estimate sex. They used a sample of 121 left os coxa from the Coimbra Collection (Portugal) and focussed their analysis on two areas: the greater sciatic notch, and the ischiopubic complex. For the greater sciatic notch, two landmarks were placed (the base of ischial spine and the piriform tubercle) to which 14 semi landmarks were placed. The same procedure was used for the ischiopubic complex where two landmarks were used to place another 25 semi landmarks. Using this data, a k means cluster analysis was used where accuracies ranged between 90% to 95% depending on if the two areas were analysed separately or combined. When shape and size were both taken into account, a much greater range in accuracy was found, ranging from 87.6% to 95.86%. Overall accuracy increased for the "shape" and "shape and size" variables when both areas were applied to a DFA (94.2% and 90.09% correct classification, respectively).

When estimating sex, visual observations are widely used in archaeological and forensic cases and are often termed as morphoscopic, morphological, or non-metric analysis. This type of analysis has been argued to be very subjective and needs a high amount of user experience to be properly used. Even with these negatives, there has been a drive in trying to make them less subjective with clear definitions of how each trait will look along the scale of male to female (Klales *et al*. 2012). The most famous and well used non-metric technique for sex estimation is the Phenice technique (1969). Phenice (1969) utilised three traits, the subpubic concavity, medial aspect of the ischiopubic ramus, and the ventral arc) From its simplicity, other techniques did become apparent like the Bruzek method (Bruzek 2002). Bruzek developed a three-score procedure for five areas of the pelvis using two European collections (France and Portugal). The five areas focussed on the morphology of the true pelvis, to which three assessments (bar one) were to be made (Figure 7). Bruzek achieved an overall accuracy of 95% with 3% of the sample being classed as unknown, and the remaining 2% being incorrectly sex. This technique states that it uses a three-point scale system rather than a five-point ordinal scale because it is less subjective and shows the “inherent variation seen in pelvic morphology” (Bruzek 2002: pg. 167). Sadly, no observer error was reported. Listi (2010) observed 21 morphoscopic sexing variables using three osteological collections in the USA (William M. Bass Donated Collection, Robert J. Terry Anatomical Skeletal Collection, and the Donated Collection at Louisiana State University) in the hope to assess if there were any racial differences in how the traits presented themselves and if that would affect sex estimation. The 21 morphoscopic traits included variables outlined by Phenice (1969) and Bruzek (2002). From using two observers, they found that there was a slight difference in scoring between males for two of the traits (shape of the pubic bone and external eversion) and four traits for female specimens (ventral arc, negative relief of the preauricular sulcus, presence/absence of the preauricular, and the existence of grooves/pits in the



preauricular area). For their main hypothesis, “Does ancestry affect sex assessment?”, they found that these traits showed very little difference between American whites and blacks.

Figure 7: Scoring of the composite as described by Bruzek (2002). Left image shows female morphology and the right shows the male morphology. (Image adpated from Bruzek 2002 pg. 161).

Novak *et al*. (2012) used the William M. Bass Donated Skeletal Collection and the Terry Collection to analyse individuals from European and African ancestry to assess how well the posterior ilium is for sex estimation. They used the elevation of the auricular surface, the presence of the preauricular sulcus, and the width of the greater sciatic notch. Sadly, no description (except a brief definition of how to score) for the elevation of the auricular surface was provided therefore repeating the study would be difficult. They compared accuracy results to a logistic regression formula to see if combining the traits would increase the accuracy. They found that using the logistic regression did increase accuracy compared to the univariate analysis by 6%, resulting in an overall 94.9% accuracy. The elevation of the auricular surface had the third highest accuracy (79.3%), yet it was nearly 10% lower than the sciatic notch (88.4%). Wescott (2015) defined the elevation of the auricular surface and was scored as either being completely elevated, partially elevated or nonelevated; complete elevation was a female expression, and nonelevation was a male expression. When only this trait was used, an accuracy of 83.1% was achieved however a large bias between the sexes was apparent. Male accuracy was near perfect with only one male being incorrectly sexed while females were only sexed ~67% of the time.

The following section describes the morphoscopic traits of the pelvis that are most well used by researchers and practitioners alike.

2.4 Morphoscopic Sexing Traits of the Pelvis

2.4.1 Greater Sciatic Notch

The greater sciatic notch starts at the sacroiliac joint and finishes distally on the ischium forming the ischial spine. This area is home to several ligament attachments such as the anterior sacroiliac ligament and the sacrospinous ligament. Throughout the literature, it has been described as showing a high degree of sexual dimorphism and is often used by anatomists, anthropologists and archaeologists alike to aid them in the diagnosis of sex.

The sexual dimorphism for the notch is described to be narrow in males and wide in females (Lazorthes & Lhez 1940, Jovanovich & Zivanivich 1965, Glanville 1967, Singh & Potturi 1978) (Figure 3). From the general description of the notch's angle, Acsádi & Nemeskéri (1970) created a 5-point grading system to assess the sexual dimorphism between males and females and integrated it into their 'complex method' for sex estimation. From there, the European meeting of Anthropologists (1980) and Buikstra & Ubelaker (1994) adapted Acsádi & Nemeskéri's (1970) descriptions in an attempt to standardise osteological techniques to sex estimation. Walker (2005) published his results on the descriptions from Buikstra & Ubelaker (1994) and found that 88% of females exhibited the extreme morphology of "Grade 1", whilst males showed more variation with 91% ranging from the ambiguous score of "3" to the hyper masculine score of "5". Gómez-Valdéz *et al*. (2012) applied the Walker (2005) descriptions on a modern Mexican population to find similar results. The majority of females were deemed to be in the hyper feminine score of 1 while males showed more variation and spanned the ordinal scale.

From applying visual methods, researchers have attempted to 'metricise' the greater sciatic notch by calculating the angle and linear measurements (MacLaughlin & Bruce 1986, Singh & Potturi, 1978). Also, researchers have investigated its morphology using geometric morphometric methods (Gómez-Valdés *et al*. 2012, Velemínská *et al*. 2013). When applying a geometric morphometric approach, Gómez-Valdéz *et al*. (2012) found an accuracy of more than 80%. However, further work needs to be done for this method to become normal practice. One of the issues with this approach is the length of time it takes to analyse each specimen; hence the 'classic' visual technique of scoring is preferred in both archaeological and forensic cases.

The morphology of the greater sciatic notch is also one of the preferred sexing techniques because of it robusticity and likelihood of survival. Waldron (1987) and Stojanowski *et al*. (2002) observed the rate of preservation of skeletal elements from different archaeological assemblages and found that when focussing on the pelvic bones, the sacroiliac joint was preserved at a much higher frequency than the os pubis.

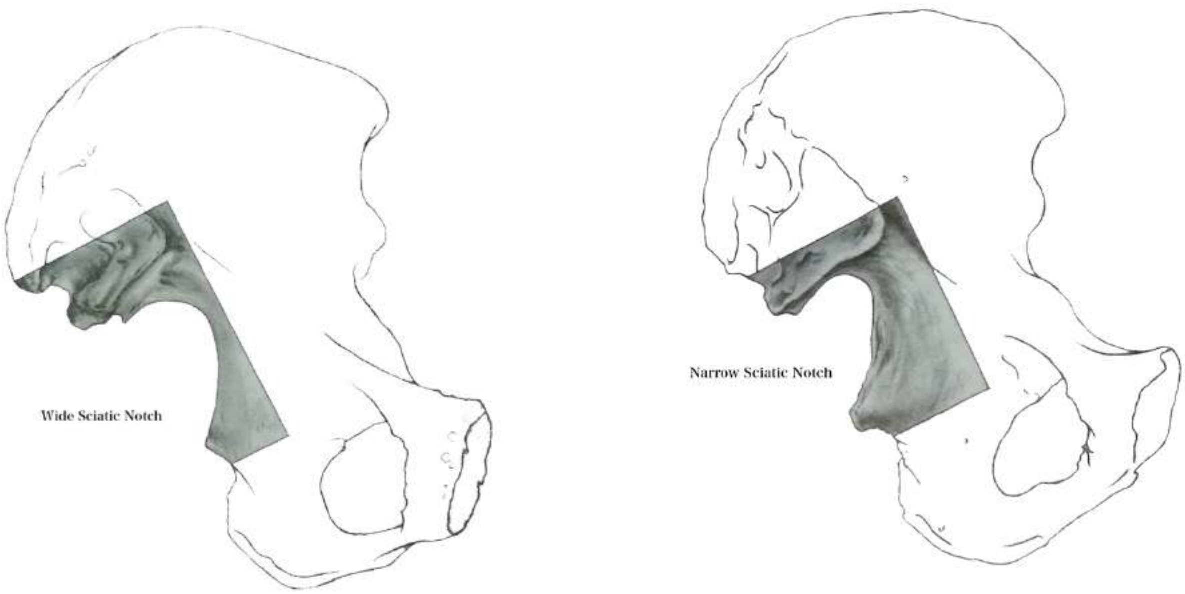


Figure 8: Left: Os coxa of female showing a wide greater sciatic notch. Right: Os coxa of male showing a narrower greater sciatic notch. (Images adapted from Forensic Anthropology Training Manual 2nd ed. 2007. Page 122).

2.4.2 Preauricular Sulcus

The preauricular sulcus, or paraglenoid sulcus, is located on the inferior portion of the sacroiliac joint of the pelvis. The main ligament attachment is the anterior sacroiliac ligament. The sulcus is an important landmark for sex assessment due to its preservation potential in both archaeological and forensic material. Because of its good preservation, it is commonly used in palaeodemographic studies to help in sex estimations of skeletons (Ubelaker & De La Paz 2012). Even though the sulcus has been greatly researched and scrutinised (Zaaijer 1866, Derry 1911, Hoshi 1961, Angel 1969, Acsádi & Nemeskéri 1970, Houghton 1974, Houghton 1975, Ubelaker & De La Paz 2012) its aetiology is still unknown. Also, researchers are still unsure as to why there is a difference in morphology between males and females and if it is a sign of parity in females.

It was first described in 1866 by Zaaijer (1866) when analysing and describing known skeletal material from Java. Zaaijer (1866) found that the sulcus preauricularis was present in 23 of the 26 females he analysed and hypothesised the cause of the sulcus was due to the anterior sacroiliac ligaments. Derry (1911) stated that the sulcus preauricularis was found more commonly in females than in males.

The preauricular sulcus has primarily been studied to see if it is a marker for parity/childbirth or not. Angel (1969) suggested that preauricular sulcus can not only give information on if the female had given birth but also allude to how many children the woman had bore. With this publication, its popularity as a parity indicator grew and Acsádi & Nemeskéri (1970) created an ordinal scale to assess the severity of the sulcus. The more the severe the sulcus, the more likely it would be female. With this, Houghton (1974, 1975) continued research into the preauricular sulcus and described that there were two types of sulci. The groove of parturition (GP) which were only found in parous females; and the groove of ligamentous attachment (GL) which were found in both nulli/parous females and males. Individuals that did not have a GP or GL groove were more likely to be male. With the research into the question of whether the sulcus was a parity indicator, Tague (1988) investigated the hormones and enzymes that cause bone resorption and pitting in both human and non-human mammals. He compared the resorption between three Amerindian populations and samples of *Gorilla gorilla* and *Pan troglodytes* and concluded that it is highly unlikely to be a parity indicator. Spring *et al*. (1989) looked at abdominal radiographs from living individuals to not only observe if the sulcus can be seen radiographically but also its possible relationship to parity. They found that only 15% of females showed signs of having a groove and for those that had a groove, a deep sulcus was present in both parous and nulliparous females.

When looking at the preauricular sulcus for sex diagnosis, Hoshi (1961) concluded from his sample of modern Japanese specimens that the mere presence of a sulcus is not indicative of sex. However, when observing the morphology of the grooves, he found two distinct types of groove.

The 'Cavity Type' being more frequent in females and the 'Furrow Type' found more frequently in males. These grooves were later described by Houghton (1974) as GP and a GL grooves. When observing the morphology of the sulcus, Bruzek (2002) observed three different characteristics of the preauricular area. Even though the Bruzek (2002) approach is considered more complex, it yields a better classification of the sexes and doesn't fall into the controversial subject of parity. Figure 4 illustrates the scoring method set out by Milner (1992).

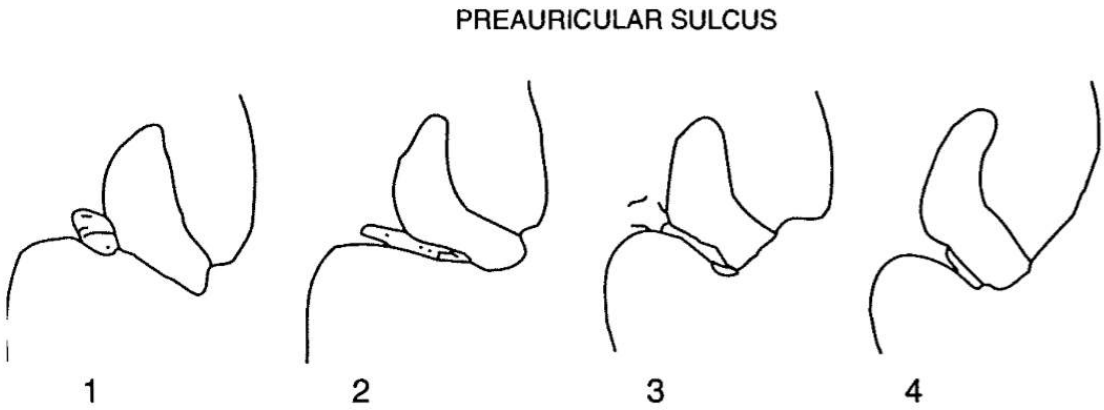


Figure 9: Illustrations of the different types of preauricular sulcus. (Image adapted from Standards for data collection from human skeletal remains: proceedings of a seminar at the Field Museum of Natural History. 1994. Page 19).

2.4.3 Subpubic Angle

The subpubic angle is created when both pubic bones are placed together in correct anatomical position. While in the living person they are separated by hyaline cartilage, in dry skeletal analysis both sides touch. The angle is then formed by the inferior pubic rami.

Normally, the subpubic angle is used as a metric assessment and defined as being greater than 90° in females and less than 90° in males (Rösing *et al*. 2007). However, when assessing modern day Egyptians, Abd-El-Hameed *et al*. (2012) found that the angles were much larger. If an individual had an angle less than 111.64°, it was a male (with a 74% accuracy), and those greater than 127.31° were female (with an 86.7% accuracy). This is a much higher accuracy than those seen in the study conducted by Igbigbi & Igbigbi (2003). They calculated the demarcation point for the angles at <80.53° for males and >136.10° for females. Given this wide range left in the indeterminate range, accuracies were as low as 31.82% for males and 10.53% for females, which can be described as being worse than 'flipping a coin'. Igbibgi & Igbigbi (2003) explain that this is due to the high amount of variation found within the Ugandan sample, another factor could be the use of a demarking point method. Msamati *et al*. (2005) study resulted in a much higher accuracy which averaged at ~65%. This vast improvement over Igbibgi & Igbigbi (2003) is most likely due to a better-defined sectioning point for males and females alongside a much narrower indeterminate range.

Despite the work done on the metric analysis of the angle, anthropologists still visually score the angle. In the publication from the Workshop of European Anthropologists (1980), it is described as being the third best morphoscopic feature to use for sex diagnosis (Figure 5). They recommend scoring the angle using a similar method devised by Acsádi & Nemeskéri (1970) on a five-point ordinal scale. Rogers & Saunders (1990) applied a simplified version of the grading system and established that if the angle were more of a 'U' shape, the individual would be female, and if it was more 'V' shaped, then it was male. After analysing 49 os coxae from a 19th-century cemetery, they found that the subpubic angle had an overall accuracy of 83.8% (N=37) when they used the simplified version. The 'simplified' criteria for the subpubic angle has also been used in forensic cases and has been recommended as one of the best traits to use for sex identification (Durić *et al*. 2005, Sidler *et al*. 2007).

The vague descriptions for the subpubic angle mean that its analysis is open to interpretation; and yet it is still one of the most applied morphoscopic techniques to estimate sex. Researchers have concluded that the main difference between male and female morphology is due to females having the ability to give birth. Obstetrical studies have shown that a wider angle is 'better' adapted to childbirth (Tague 1992, Kurki 2007, 2011, 2013a, 2013b). From this, researchers have used the subpubic angle for sex estimation of fossil hominid remains, such as AL-228-1 (*Australopithecus afarensis*) (Tague & Lovejoy 1986) and BSN49/P27 (*Homo erectus*) (Simpson *et al*. 2008) and LB1 (Homo floresiensis) (Jungers *et al*. 2009). The five-point scoring method was applied to fossil, and modern anatomical human remains for sex estimation to aid in a multivariate analysis for sex estimation (Rennie *et al*. 2015). Rennie *et al*. 2015 found that the subpubic angle was one of the most influential sex indicators from the other seven morphoscopic traits studied.

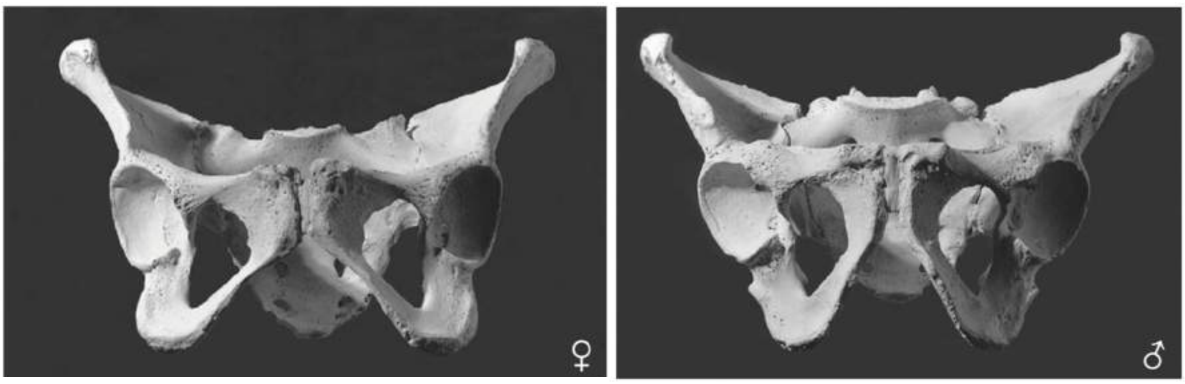


Figure 10: Left: Female os coxae showing a wide "U" shaped subpubic angle. Right: Male os coxae showing a narrow "V" shaped subpubic angle. (Image adapted from The Human Bone Manual. 2005. Pages 415-416).

2.4.4 Obturator Foramen

The obturator foramen is formed by the interior borders of the pubis and ischium. The obturator foramen houses several soft tissue structures such as the obturator membrane which allows the obturator artery, vein and nerve to pass through the canal. Superiorly to the membrane lies the obturator groove to which the vessels pass. The obturator membrane attaches to two tubercles on the foramen, the posterior and anterior obturator tubercle. The posterior tubercle is located along the medial border of the ischium, while the anterior tubercle lies in the superior ramus of the pubis.

Sexual differences have been described by many authors (i.e. Schaeffer 1953, White & Folkens 2000, Standring 2008). The differences include the foramen tending to be large and oval in males and relatively smaller and more triangular for females (Figure 6). These descriptions have not changed in nearly 200 hundred years of medical/anatomical research since it was first described in the late 18th Century by Ackermann (1788).

Considering this long history, very little work has been undertaken into why the obturator foramen is sexually dimorphic, and there have been little to no updates on the descriptions of the oval and triangular classification.

The Workshop for European Anthropologists (1980) gave the recommendation of a five-point grading system that could be used to score the sexual variation a given skeleton has along with other morphoscopic features. Rösing *et al*. (2007) also recommended the visual assessment of the obturator foramen for forensic identification as well as an intermediate weight as to how well it performs as a sex indicator. Rogers & Saunders (1994) used a multitude of morphoscopic sexing traits to assess which combinations diagnose sex with the highest accuracy. They found that combining the obturator foramen with the ventral arc or the true pelvis shape the accuracy was as high as 98%. This ranking of the obturator foramen within the top five traits used for sex determination contradicts results by St. Hoyme (1984) who stated that the foramen had very little value when considering it a sex indicator.

Metric analysis has also been conducted to help create a more 'standard' approach to assessing the shape of the foramen. Quantitative analysis was first attempted by Martin (1914) who used four reference points to create an index for sexual differences. Martin (1914) found accuracies of ~65%.

With the advancement of statistical tests, Bierry *et al*. (2013) applied a Fourier analysis to assess if there was any truth behind the obturator foramen shape differences between males and females. They used Computerised Tomography (CT) images from the virtual collection at the University Hospital in Strasbourg (France). When applying this multivariate approach with the addition of a Discriminant Function Analysis, an overall accuracy of 84.6% was achieved.

For this technique to be used further, better images of the shape change with detailed descriptions are needed so that it can be used more effectively by anthropologists, archaeologists, and anatomists alike.

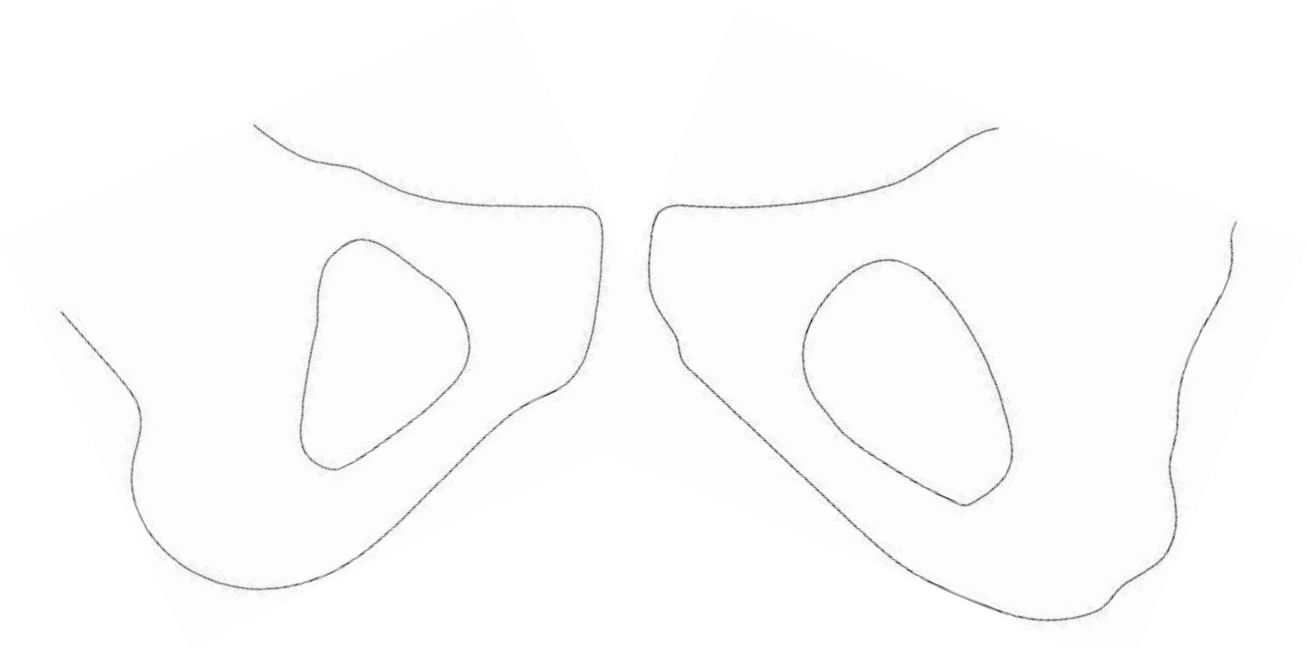


Figure 11: Left: Female os coxa showing a triangular obturator foramen; Right: Male os coxa showing a rounder obturator foramen. (Illustrations created by author).

2.4.5 Shape of the Pubic Body

The pubic body (Corpus ossis pubis) is the attachment site for several ligament attachments along with the union of the two pubes with the pubic symphysis. This feature has been described as being high in sexual dimorphism between males and females. It is described as being wide and rectangular in females and smaller with more of a triangular/wedge shape in males (Figure 7).

The shape of the pubic body was investigated by Rogers & Saunders (1994) who followed the descriptions set by St. Hoyme (1984) on the corpus morphology. Rogers & Saunders (1994) found that as a stand-alone sex indicator, it could correctly identify sex with an 86.2% accuracy (N=36). This trait was ranked sixth out of the 17 that were tested by Rogers & Saunders (1994). Listi (2010) extracted the information from the St Hoyme (1984) and the Rogers & Saunders (1994) publications and applied the descriptions to American White and Black skeletal populations from the William M. Bass Donated Collection (University of Tennessee, USA), Robert J. Terry Anatomical Skeletal Collection (National Museum of Natural History, USA) and the Donated Forensic Collection (Louisiana State University, USA). Listi aimed to assess differences between American White and Black individuals and found that the morphology of the pubic body was stable between the two populational groups.

Differences between males and females have been noted in other species as well, and not just in primates. Zammit *et al*. (2013) described the sex difference in pelvic morphology within two otariid species (Sea Lions). They discovered that the reverse morphology was evident when focusing on the pubic body. Males presented a rather large square body while its female counterpart had a much smaller body. However, they never concluded why a difference was present in the species.

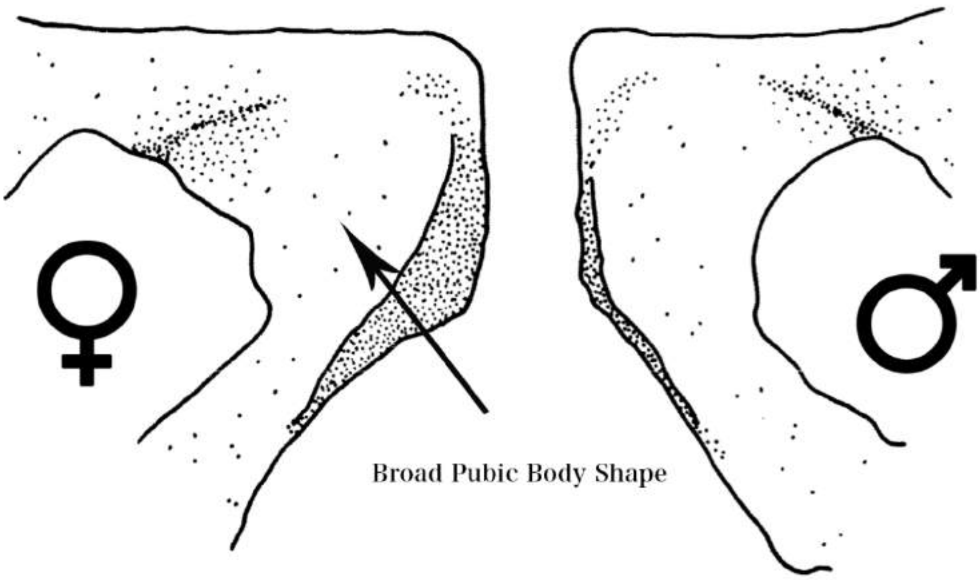


Figure 12: Left: Female os pubis with a broad corpus; Right: Male os pubis with a smaller, wedge shaped corpus. (Image adapted from Anderson 1990, page 450).

2.4.6 Ventral Arc

The ventral arc has been used for the basis of sex determination in human skeletal remains for over 50 years. Located on the ventral surface of the Corpus pubis it can be seen medially, close to the pubic symphysis. Considering its common use, the anatomical description of the ventral arc and its attachments were not fully outlined until 1990 (Anderson 1990, Budinoff & Tague 1990). Before then, Todd (1920) stated that the ventral arc was the line of attachment for the gracilis muscle, followed by Buikstra & Mielke (1985) who said that the ventral arc was related to the crus penis and crus clitoris. Bass (1987) described it as the attachment of the arcuate ligament. From their dissections, Budinoff & Tague (1990) concluded that the tendons from the gracilis and adductor brevis attach to the ventral arc. From there, Naňko *et al*. (2007) and Šedý *et al*. (2008) concluded that the reason for different yet similar morphology is due to how the nervus dorsalis penis/clitoridis are situated along the os pubis.

Phenice (1969) used the ventral arc to create his technique for sex estimation from the os pubis (with two other anatomical features). He described the female expression of this feature as a "slightly elevated ridge of bone which extends from the pubic crest and arcs inferiorly across the ventral surface to the lateral most extension of the subpubic concavity where it blends with the medial border of the ischiopubic ramus." (Phenice 1969: 298) (Figure 8). The male expression was the absence of this ridge or presence of something similar. The latter was described as the ridge of bone extending from the pubic crest inferiorly and run parallel to the medial border of the symphysis. Rogers & Saunders (1994) evaluated the trait, and pre-emptively placed it as one of the most reliable traits to be used for sexing. They found that when it was used as a single trait, it yielded an accuracy of 86.9%. When combined with the obturator foramen it could estimate sex at a 98% accuracy. Klales *et al*. (2012) redefined the descriptions originally set out by Phenice (1969) and also created a new scoring system for the ventral arc morphology. Instead of it being male, female, or ambiguous; it was scaled to a 1 to 5 system that described the morphoscopic changes from "typical male" to "typical female". With the new descriptions, an accuracy of 88.5% was achieved when used on its own. However, when combined with either of the other two Phenice traits, accuracy increased to greater than 90%. Because of the new description, the confusion between the 'ridge of bone' that can be present in both males and females has been addressed and incorporated into the Klales *et al*. (2012) ventral arc description.

An investigation into if the trait could be observed on digitised remains and still yield a high accuracy like on real remains was conducted by Gamble *et al*. (2011). Gamble *et al*. (2011) used CT scans, surface laser scans and real specimens to assess morphoscopic sexing techniques. They found that the intra and inter observer accuracies did not change drastically between the digital and real specimens.

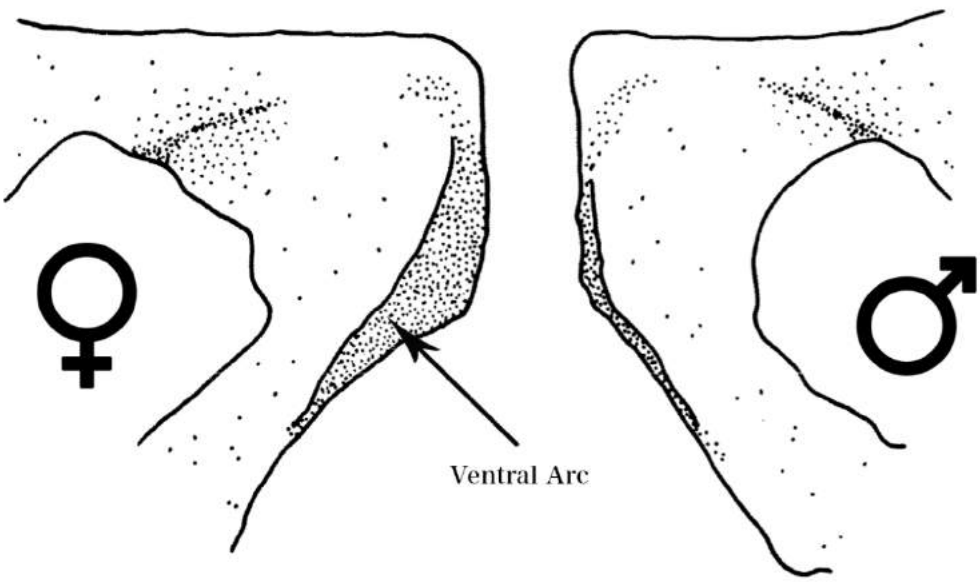


Figure 13: Left: Female os pubis showing the presence of a ventral arc; Right: Male os pubis showing the absence of a ventral arc. (Image adapted from Anderson 1990, page 450).

2.4.7 Subpubic Concavity

The morphology of the subpubic concavity was first described as being sexually dimorphic by Phenice (1969) when he combined this trait with the ventral arc and the medial aspect of the ischiopubic ramus. The subpubic concavity refers to the area lateral to the pubic symphysis on the ischiopubic ramus (Buikstra & Ubelaker 1994). Phenice (1969) described the female morphology as having a lateral recurve below the pubic symphysis when viewing the pubis and ischiopubic ramus dorsally. Males tend not to have this "recurve" present but, like the ventral arc, males can present a small amount of "recurve" however this morphology is uncommon (Figure 9).

When looking how the subpubic concavity should be weighed in accordance to other traits, Rogers & Saunders (1994) placed it as being one of the most effective based on the results from Phenice (1969). When analysed, it achieved an accuracy of 83.8% (N=37) and was ranked ninth best trait from an original 17 traits analysed. When paired with sacrum shape, accuracy dramatically increased to 95%. Similar to Rogers & Saunders (1994) weighting of the trait, Kjellström (2004) noted it as one of the strongest, alongside the greater sciatic notch and ventral arc.

Arsuaga & Carretero (1994) analysed ten non-metric indicators of sex from 418 pelves from the known age and sex Coimbra Collection (Portugal) and found that when using Phenice's (1969) description, the overall correct classification was 88.6%. The trait performed much better for males (95.8%) than it did for females (77.6%).

Klales *et al*. (2012), updated the original description made by Phenice (1969), referring to it as the subpubic contour and created a scoring system from 1 to 5 to help explain the "recurve" of the concavity as it goes from most male to most female. When applying the new scoring system to a logistic regression analysis, an overall accuracy of 86.6% was found. Contrary to Arsuaga & Carretero (1994), they found that the subpubic contour was better at correctly classifying females (90.4%) than males (82.8%). This flip in which sex performed better between Klales *et al*. (2012) and Arsuaga & Carretero (1994) could be explained by the new grading system that Klales *et al*. (2012) implemented and/or the different sample used (one American, the other European). Gamble *et al*. (2011) assessed the observer error between scoring the trait on real specimens and its virtual counterpart. They found that the subpubic concavity performed the best with an almost perfect repeatability score.

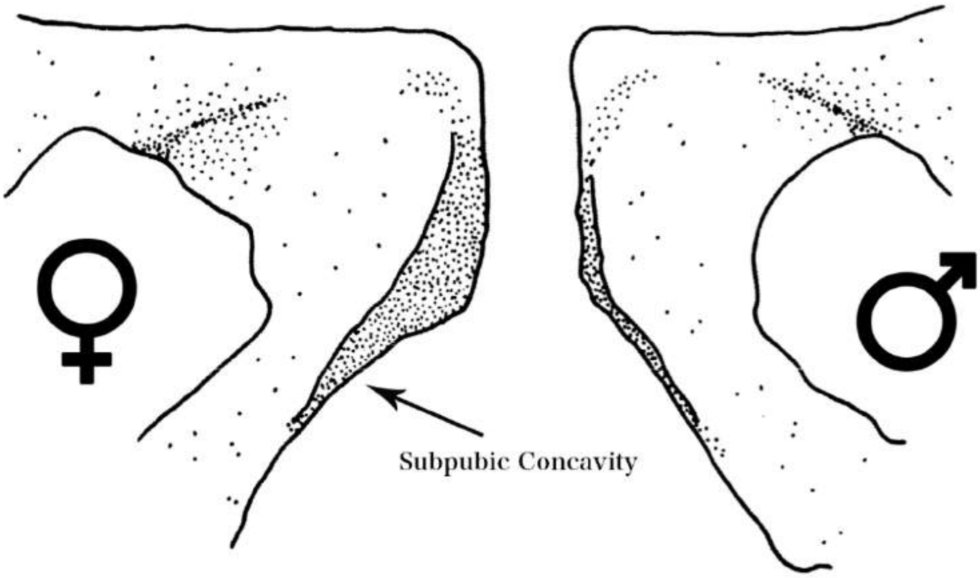


Figure 14: Left: Female os pubis presenting a concave Ischiopubic ramus; Right: Male os pubis with no subpubic concavity. (Image adapted from Anderson 1990, page 450).

2.4.8 Medial aspect of the Ischiopubic Ramus

The ischiopubic ramus is located on the dorsal surface of the pelvis. Phenice (1969) described the male form to have a broad surface on the ramus directly beneath the symphyseal surface of the pubis. In contrast, females were described to have a ridge along the ramus that presents more of a 'pinched' look (Figure 10). It is important to note that Phenice (1969) does state that there is a slight overlap between the male and female forms.

Using the descriptions set by Phenice (1969), Arsuaga & Carretero (1994) gained an overall accuracy of 76.7% when they applied them to the Coimbra Collection. They found that the trait works equally well for both males and females; 77.4% and 76% respectively. Rogers & Saunders (1994) assessed the ischiopubic ramus and classed it as a trait that should only be used in conjunction with others based on the results by Coleman (1969). Once analysed, it achieved an overall accuracy of 80% and was ranked 11th best out of the 17 traits that they analysed. When tested for intra-observer error, Rogers & Saunders (1994) found that it had the highest percentage error (11.3%) from the 17 traits.

Issues were raised by several validation studies (Lovell 1989; Maclaughlin & Bruce 1990; McBride *et al*. 2001) regarding the reported accuracies from the Phenice (1969) study. Klales *et al*. (2012) tackled this issue by redefining the descriptions and images and fully testing both inter and intra observer errors. From the new definition and applying the 1-5 scoring method to logistic regression analysis an overall accuracy of 75.8% was found. The trait performed better for males (79.3%) than females (72.3%).

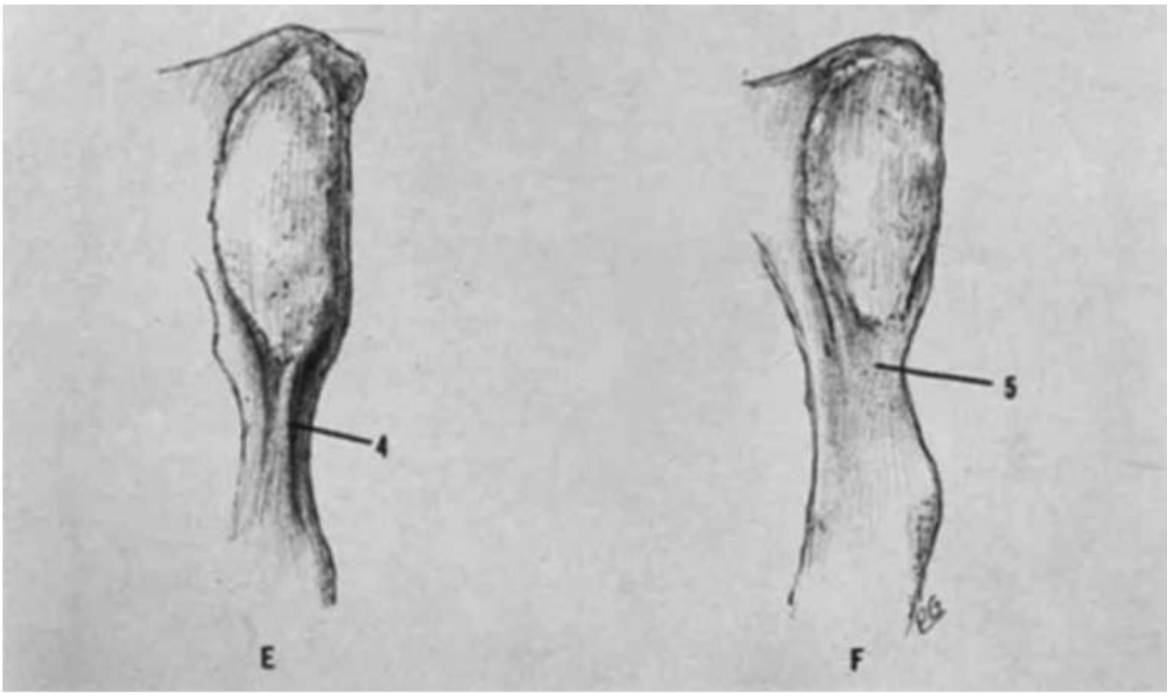


Figure 15: Left: Female os pubis with a pinched ischiopubic ramus (indicated by no. 4); Right: Male os pubis with a broad ischiopubic ramus (indicated by no. 5) (Image taken from Phenice 1969, page 299).

2.5 Statistical Approaches in Biological Anthropology

The use of multivariate statistics within biological anthropology has grown exponentially over the last couple of decades. Normally the focus for biological anthropology has been to try and develop new ways to classify individuals, for both extant and extinct humans (Berge 1984, Steudel 1978, Arsuaga & Carretero 1994).

The main approach for classification is the Discriminant Function Analysis (DFA). Morrison (1969) described the approach as a classification tool using a set of independent variables to classify objects into x amount of categories. With this as its main function, it is ideal for creating population specific equations for age, sex and ancestry determination. Calcagno (1981) used 29 measurements from the mandible and mandibular dentition to assess how well a DFA would be for determining sex from three separate sample groups in the USA. Calcagno (1981) obtained an accuracy ranging between 91.7% to 98.6%. Luo (1995) used 122 individuals to create a Discriminant Function equation from four measurements on the os pubis. When applied to specimens from the Human Identification Laboratory, Department of Anthropology, University of Arizona (USA), an accuracy of 100% was achieved. DFA has been used on nearly every skeletal element within the human body from the femur (Asala 2004), skull (Johnson *et al*. 1989, L'Abbé *et al*. 2013, Krüger *et al*. 2015) and the pelvis (Dixit *et al*. 2007, Gómez-Valdés *et al*. 2011, 2012). Due to the nature of the statistical analysis, a priori knowledge is needed to create the linear equations (Mitteroecker & Bookstein 2011). Alongside the need for the prior knowledge of groups for individuals to be classified into, another issue that has to be taken into account are the differences found between data points. This exaggeration of data can be seen as a good thing in the case of classifying however because of this, information relating to the samples actual variation can be seen as being skewed (Mitteroecker & Bookstein 2011). Another issue of DFA would be that the analysis only allows individuals to be classified into x amount of groups that have been. If the individual does not have a represented group to be classified into, it will be forced into the closest group. An example of this would be when an unknown cranium is analysed by the program FORDISC 3.0. The unknown crania can be “forced” into a group that is not related to therefore rendering the analysis useless. Because this can happen frequently, it is important to take a closer inspection of the descriptive statistics associated with the analysis.

A less extreme version of DFA is Principal Components Analysis (PCA). Prior knowledge of the groups is not needed because the analysis searches for variability within the sample rather than specific groups or clusters (Zar 2010, Sokal & Rohlf 2012). This approach has primarily been used to look at variation differences between and within groups. PCA has been used in geometric morphometrics, linear measurements and molecular data (e.g. DNA Haplotype) (Key & Jantz 1981, Kennedy *et al*. 1985, Lockwood *et al*. 2004). It also it has been used to explain shape changes in the human skull associated with geographic affinities (Powell & Neves 1999), shape changes in the humerus concerning human evolution (Bacon 2000) and odontometric differences between human populations (Harris & Bailit 1988).

Very little work has been produced using non-metric variables when using multivariate statistics. This does not mean PCA has not been attempted. Lovejoy *et al*. (1985) used PCA to create a multifactorial ageing technique using degenerative age changes observed on the skull and pelvis and radiographic data from the femoral neck. PCA has also been used to analyse ordinal/categorical data from Likert scales. Likert scales are mainly used in behavioural studies or market research. Because of the amount of information that is collected in this format, researchers choose to apply PCA as a data reduction technique and as a first step for analysis (Zani & Berziari 2008). Stull *et al*. (2013) used a wide array of multivariate statistics on non-metric sex indicators from the skull and pelvis such as Linear Discriminant Analysis, Flexible Discriminant Analysis and Logistic Regressions. They found that Logistic Regressions outperformed the other four multivariate statistical analyses.

When analysing non-metric/ordinal data, the statistical approach would be to conduct either an Ordinal Logistic Regression or Correspondence Analysis. However, researchers have attempted at using other methods, as mentioned previously, like Stull *et al*. (2013) who used DFA on non-metric data. Another example of this would be Berg (2008). In his doctoral dissertation, Berg (2008) applied linear discriminant functions to estimate ancestry and sex from morphoscopic traits of the mandible. To help increase accuracy, he then went on to combine metric and morphoscopic variables using the same statistical approach to produce a “morphometroscopic” analysis. To combine ordinal and continuous variables is said to violate the assumptions that a DFA adheres to; being that the variables must be continuous. Berg (2008) states that even though the analysis violates this assumption, the results are still meaningful and the method can be applied to forensic work. This is because researchers have argued that ordinal scoring is an extension of binary data (see Berg 2008 pg. 46 Ousley & Hefner 2005). Hefner & Ousley (2014) applied ordinal data for ancestry estimation to 10 different classification methods including Quadratic Discriminant Function Analysis (QDFA), Linear Discriminant Function Analysis (LDFA), and Logistic Regression. They state that even though they were "…bending or even breaking the rules… The classification performance can be very good" (Hefner & Ousley 2014 pg. 885). With this, they found that LDFA outperformed the Logistic Regressions because LDFA has stricter assumptions (considering they were broken) than the regression analysis.

Correspondence Analysis (CA) is used as an alternative to PCA as it is a data reduction technique that helps understand the variation within the sample(s). The main differences between the two are that CA uses categorical/ordinal data while PCA assumes the data are continuous and that instead of the inter-correlation matrix the PCA produces, CA uses a chi-square distance matrix. Similar to PCA, each axis that the CA calculates relates to the amount of variation present, with the x-axis explaining the largest amount of variation within the sample (Irish 2005). With this analysis, it is possible to observe which of the variables are the most influential in causing the variation. Irish (2005) applied CA to 36 dental traits from 12 dental samples (eight samples from Lower Nubia and four samples for Upper Nubia). He found that the Jebel Sahaba sample (Lower Nubia) differed from the other samples indicating possible genetic discontinuity by population replacement. Wirth *et al*. (2004) applied CA onto genetic data to try and distinguish human populations via Helicobacter pylori bacteria. Wirth *et al*. (2004) found that after analysing microsatellite genotypes with CA, two main clusters could be found that separated Buddhists from Muslims from the Ladakhis sample.

Ordinal Logistic Regression (OLR) has several assumptions that must be met; being that the dependent variable must be dichotomous, the data does not need to be normally distributed, the independent variables are measured without error, and that the independent variables are not linear combinations of each other (Press & Wilson 1978). To gain accurate results, it is said that a minimum of at least 50 cases is needed per predictor variable (personal communication Kenyhercz 2015). OLR is based on odds ratio which can be defined as the ratio of the probabilities of success and the probabilities of failure. In the context of this thesis, it is the probability of being correctly classified as male or female (success) and the probability of being misclassified (failure). Klales *et al*. (2012) applied OLR to a newly revised method of the Phenice technique (Phenice 1969), extending the "M, ?, F" classification to a five point ordinal scale ranging from most feminine to most masculine. The success of this approach has led to recalibrating the original equation Klales *et al*. (2012) created for other sample populations, e.g. Mexico (Gómez-Valdes *et* *al*. 2017). Gómez-Valdes et al. (2017) found that even though the original equation worked well, the recalibrated equation was said to not only improve classification accuracy but also helped eliminate population bias. The technique has also been adapted and applied to juvenile sex estimation using OLR to calculate the probability of correct sex (Klales & Burns 2016). Klales & Burns (2016) found that sex could be accurately predicted in the oldest cohort at 97.2% accuracy with accuracy decreasing the younger the individual. Krüger *et al.* (2015) applied OLR to non-metric sexing traits from the skull in an attempt to create population specific equations for South African Whites and Blacks. Not only are logistic regressions used for sex estimation techniques, but they have also been used for age and ancestry estimation. Hefner & Ousley (2014) and Klales & Kenyhercz (2015) applied several cranial morphoscopic traits to determine ancestry to OLR. Hefner & Ousley (2014) found that out of ten classification methods used, that OLR was the least effective at classifying all three sample populations (American White, American Black, and Hispanic) with Hispanics having a misclassification rate as low as 58.2%. Klales & Kenyhercz (2015) applied a variation of cranial morphoscopic traits that differed from Hefner & Ousley (2014) study, which resulted in a higher classification accuracy for both American Black and White samples. Also, OLR was the highest performing statistical approach compared to the other four analyses that Klales & Kenyhercz (2015) tested (kNN, LDFA, Random Forest Models, and Naïve Bayesian Statistics). Dudzik & Langley (2015) applied logistic regression on age-at-death scores to create a new system for age-at-death estimation. They took scoring systems from previous authors and created a three phase scoring system relating to the age ranges. Dudzik & Langley (2015) found that with the regression formulae, the highest accuracies were for the first two age categories at 91% for females and 88% for males. However, after analysis, they state that there may be a need for an overlapping phase for categories 2 and three as that is where the majority of the misclassifications were.

One of the largest drawbacks to using these multivariate statistics is that all variables must be present/observable. When dealing with human remains, complete skeletons are a rarity for both forensic and archaeological cases. Therefore, as powerful as these techniques can be, they are rendered useless if one variable is not present. Little research has been conducted in how to overcome fragmentation that doesn't involve “just guessing”. Kjellström (2004) investigated the use of calculating weighted means to determine sex from a skeletal collection in Sweden. Kjellström followed previous research by Acsádi & Nemeskéri (1970) and calculated the mean, rather than the medium, for ordinal data to assess the amount of sexualisation present in each skeletal individual. For missing data, a neutral score was used in its place, therefore reducing the amount of “sexualisation” that the weighted means approach is trying to portray. With this weighted means approach, Kjellström found that individuals were likely to be identified as feminine if more than five traits were originally missing. Even though using a neutral score for missing data allows for analysis, it does dampen the amount of sexual dimorphism present in the skeletal remains and does not portray the amount of, as she terms, “sexualisation” correctly. Kjellström (2004) did highlight however that in biological terms a skeleton is either male or female, but the range of variation between the two is vast and requires analysis that is not strictly binary (in a sense), but data should be collected that helps account for the variation between the sexes. Kenyhercz *et al*. (2016) presented several approaches towards missing data imputation for morphoscopic features. Kenyhercz and colleagues used six cranial morphoscopic traits associated with ancestry estimation and randomly deleted scores from their sample of 688 individuals (Black = 292, White = 210, Hispanic = 186). From there, they applied four different methods for missing data: Hot Desk, Iterative Robust Model-based, K Nearest Neighbour (kNN), and Trait Median Replacement. Hot Desk replaces the missing variables by observing a similar individual whose known variables match that of the individuals in question. This individual is known as a donor, and each individual that has missing variables have their own donor. Iterative Robust Model-Based imputations (IRMI) is a multivariate model that uses the missing data as a response variable, and with that, known variables (regressors) are used to input the missing data. Because this method assumes a normal distribution, the predicted values can be ridiculously large or small which sometimes are not biologically meaningful for analysis. K Nearest Neighbour imputations use the kNN algorithm to observe individuals that are most similar to the individual with missing data and calculates what that score could be based in its most similar “neighbour(s)”. For this to be a viable method, there must be a good representation of the missing data variable within the rest of the sample. These three descriptions were taken from Kenyhercz & Passalacqua (2016). Kenyhercz *et al*. (2016) found that the median achieved the highest classification when 50% of the data had to calculated (59.2% classification when compared to 65.5% from the complete dataset) and that Hot Desk had the lowest classification. With this being said, Hot Desk achieved the highest accuracy when only 25% of data was missing. Kenyhercz *et al*. (2016) state that the method for missing data imputation depends on how much data are originally missing from the dataset and to continue from there. This thought process of missing data imputation was also applied to metric variables from the skull by Kenyhercz & Passalacqua (2016). They found that when dealing with metric variables that kNN imputation had the smallest difference in accuracy when compared to the original dataset where 25% of data were missing. When presented with 50% of the data needing to be calculated, Hot Desk was found to be the best method when compared to the original dataset, followed by using the variable mean (Kenyhercz & Passalacqua 2016).

3. Materials

Throughout the literature, there is very little diversity/or the same type of diversity when evaluating sex differences and creating multivariate equations; such as American research only focussing on American Black and White. Only recently is there an integration of Hispanic data. The same occurs in South Africa, where the majority of osteological research is conducted on South African Whites and Blacks with very little produced for South African Coloureds. However, these are normally very localised pieces of research. With this in mind, the samples chosen for this project ranged from prehistoric American to modern day South African giving the project a large span both in terms of geography and in time. The geographic spread of the samples studied here has only been seen a couple of times (see Berg 2008, Klales *et al.* 2016, and Betti *et al.* 2017), but sadly only have a breadth in geography, not in time. In this chapter, we review the histories of each of the eight human samples chosen for the PhD study.

a) The Christ Church, Spitalfields sample. This is a rare example of a British historic collection of known age and sex individuals and allows the methods to be checked for accuracy against the known sex of the individuals.

b) The two Medieval British samples (Poulton and St. Owens). These were chosen for two reasons. Firstly, neither of the Poulton or St. Owens samples have been studied in great detail before, and this is one of the first glimpses at what these collections have to offer (other current PhD studies at Liverpool John Moores University include palaeodemographic analysis (Davenport 2017), assessing non-metric traits (Burrell 2017), evaluation of stress markers (Dove 2017), and cranial variation (Valoriani 2018). Secondly, because of the ease of access being held at Liverpool John Moores University.

c) The prehistoric Chumash and historic Andaman samples. They were chosen for similar reasons to the medieval British samples. Barely any anthropological research has been undertaken in these collections held at the Natural History Museum, London. So not only would choosing to study these two additional samples highlight some of the lesser known collections at the institution, but they also added a fantastic temporal and geographical depth to the project.

d) The three South African samples (South African White, Black, and Coloured). They were selected after successfully acquiring funding to study in South Africa (Erasmus Mundus: AESOP Scholarship). In the past many studies in South Africa had a tendency to focus on white and black South African groups. Only recently are the coloured South African group being studied in light of biological/forensic anthropology; with this project being one of a handful to incorporate the population in its analysis (others include biodistance analysis using cranial discrete traits (Sutherland 2015), and postcranial metric analysis (Liebenberg 2014)).

3.1 Christ Church Spitalfields, U.K.

Christ Church, Spitalfields is located east of Bishopsgate within the city of London (U.K.) (Figure 11). Created from the Manor of Stepney, the hamlet gained its parish as construction of Christ Church began in 1714 and was finally completed 15 years later. The church was later modified to 'fit in' with the ecclesiastical ideologies that were present in mid-1850. The area in which the parish was built derived its name 'Spitalfields' from the meaning 'belonging to a hospital' because the site of St. Mary's hospital also was located within the area. After its initial construction, the church was consecrated on the 5th June 1729 and their first burial was three days later. It is estimated that 68,000 burials followed, with the last internment being the 23rd February 1859. Christ Church also had many vault burials, some of which were family burials and other members of the parochial church. These vaults were sealed in 1867. During the sealing process, each coffin was covered with ash and sand, which acted as a sanitizing layer between the coffins stacked on top of each other (Cox 1989, Reeves & Adam 1993, Molleson & Cox 1993).

The occupants of Spitalfields were mostly of French origin, whose ancestors may have settled there after the St. Bartholomew massacre in France and through subsequent migrations from 1572. As time progressed the area was dubbed 'Petty France' due to the high percentage of French Huguenots. With the craft skills that the Huguenots brought to the area, Spitalfields was thriving and prosperous during the early-mid 1800's.

Plans to restore the church were approved in 1965. During the initial inspection of the church grounds and vaults, the church council agreed that the restoration was needed. Excavations of the crypts began 19 years later in 1984 under the direction of Jez Reeve and were completed in 1986. It was the first scientific excavation of a church vault within the U.K. As well as providing osteological material, the excavation also retrieved material for those who specialise in funerary/burial practices. The human remains were removed in varying degrees of decomposition. This is largely due to the differential preservation caused by the individuals being interred in lead lined coffins or not. Nine hundred and sixty-eight individuals were exhumed from the crypts and out of these, 387 individuals are included in the named coffin-plate collection. This collection was selected by strict criteria; each skeleton had to be retrieved in a secure association with the legible coffin plate. Each of the skeletons kept the original identification numbers assigned by the archaeologists and the abbreviation 'CAS' was given, for example CAS 2956 (Cox 1989, Reeves & Adam 1993, Molleson & Cox 1993). The collection is currently housed at The Natural History Museum, London (UK).

After the excavation of the crypt, analysis of the remains began which identified 296 adults (18+ years old) and 91 non-adults (<18 years old). By using archival evidence, it was possible to associate place of birth to 206 of the individuals. They found that 7 were born in France and, as this was a Huguenot community, there was a high chance that they came from Normandy, Picardy or Poitou (Cox 1989, Molleson & Cox 1993).

Figure 16: Ordnance Survey Map of Christ Church, Spitalfield, First Series, Sheet 1. 19th century. (Taken from http://www.visionofbritain.org.uk/maps/). Christ Church, Spitalfield is marked in red.

3.2 South Africa

Three skeletal collections were used to create samples that best describe the diversity of the country's demography. The three samples were modern day white South African, black South African, and coloured South African. The locations of each of the human skeletal collections are shown in Figure 12. The different population groups within South Africa all come from different origins and with past segregations and social barriers (e.g. Apartheid regime), this resulted in limiting the gene flow between the groups. The Union of South Africa was established in 1910 and is labelled as one of the precursors of the Apartheid regime. This period in time is also referred to as the segregation era where the start of large racial divides began, which lay in favour of white South Africans. This era forced all individuals who resided in the country to be classified by race, and then depending on their race, be designated where in the country they shall live, and by what standard. Alongside this, it was made illegal for 'inter-group' marriages further restricting the gene flow between groups (Ross 1999, Morris 2012, Thompson 2013). The Apartheid regime, itself, was also another detriment to gene-flow. The regime was enforced in 1948 and was later abolished in 1991. The apartheid regime created even tighter laws regarding the classification of its citizens (as either black, white, coloured, Asian/Indian, or 'other'), where they could live and who they could marry. This resulted in a large amount of variation not just between groups, but within them. Below is a brief history of three modern South African population groups that have been chosen for study.

3.2.1 Black South Africans

The ancestral route of the black South African people has been said to be a result of a southward migration of Bantu speakers from western-central Africa (Ribot 2004, Franklin *et al* 2010). The term 'Bantu' refers to the near 500 languages which are spoken in the western and sub-Saharan African countries (Ribot 2004). According to the Bantu-expansion theory, they originated in modern day Nigeria and Cameroon with the first migration wave starting ~2000 BC, with branches heading both east and west (de Filippo *et al.* 2012). The separate groups arrived in modern day South Africa around 300 AD. The migration of the Bantu speakers led to further divisions/groups/tribes which all have different cultural systems (Hall & Morris 1983, Herbert 1990). There are currently nine Bantu subgroups/tribes residing in South Africa which include Nguni (Ndebele, Swazi, Xhosa, and Zulu), Sotho (Tswana, Pedi, Southern Sotho, and Tswana), Shangaan-Tsonga and Venda (Franklin *et al*. 2007). Because of the southern migration and separation of the Bantu speaking tribes, researchers have attempted test if there are differences, on a tribal level, to try and further aid in skeletal identification (Franklin *et al*. 2007, Siddiqi 2013). The phenotypic differences between the tribes are thought to be due to a restricted gene flow resulting from segregational regimes, however some mixing with other indigenous groups such as the Khoesan is apparent (Franklin *et al*. 2007, Franklin *et al*. 2010).

3.2.2 White South Africans

There were two periods of colonisation in South Africa during the 17th Century. The Dutch East Indian Company (also known as the Veerenigde Oost-Indische Compagnie or VOC) established the Cape colony which saw the migration of firstly Dutch immigrants and later the British, German and French immigrants (Greef 2007, Thompson 2013). With South Africa's history regarding race and the laws surrounding it, led to a uniform white populace which can be distinguished quite easily from the other South African groups. Considering their European descent, South African White's differ skeletally from there other European and North American white individuals. This can be explained by the founder's effect and the adaptation to the new environmental conditions (Boas 1931, Steyn & Iscan 1997, Steyn & Iscan 1999, Steyn & Smith 2007).

3.2.3 Coloured South African

The term “coloured”, in South Africa, refers to an individual from a highly diverse social group which has a large amount of phenotypic variability (Adhikari 2005). The group consists of people of mixed ancestry, from descendants of the Khoesan subgroup and slaves later brought to South Africa from Malaysia and Indonesia by the VOC (Adhikara 2005). The Khoesan is a collective term which consists of two separate groups originally being the Khoe (pastoralists) and the San (hunter-gatherers) (also referred to as “Hottentots” or “Bushmen” by European colonists) (Barnard 1992). The Khoesan originally occupied southern Africa around 2000 years ago (Sampson 2010) before the southward migration of the Bantu tribes. The San tribes, in comparison the Khoe, were smaller in stature and more petite due to their hunter-gatherer lifestyle (Thompson 2013). Even though they are from different ways of life, they are recognised as different components of the same single genetic population (Stynder 2009). During the colonial period, the Khoesan had their land repossessed and were forced into slavery by the European settlers. Present day Khoesan are mainly living in the Kalahari regions of Botswana and Namibia (Peterson *et al*. 2013). A large contribution to the modern day coloured South Africans are from various other groups that were brought to the Cape colony from the European settlers. Some of the contributing populations include Indonesia, Malaysia and India (Patterson *et al*. 2010, de Wit *et al*. 2010).

The Khoesan are related however to some of the Bantu speaking tribes mentioned previously. A study using geometric morphometrics links the Khoesan to the Xhosa, Venda, Zulu and Southern Sotho tribes (Franklin *et al*. 2007) and their relationship can also be seen both genetically and linguistically (Wood *et al*. 2005).

Figure 17: Map of South Africa with its neighbouring countries. (Taken from www.africaguide.com) Location of the Pretoria Bone and Raymond A. Dart Collections are highlighted in red and the Kirsten Collection is highlighted in purple.

3.2.4 Pretoria Bone Collection, RSA

The Pretoria Bone Collection (PBC) is housed within the Department of Health Sciences at the University of Pretoria, located within the northern portion of the Gauteng Province (South Africa). The PBC started in the 1942 when the University created its medical school within the Department of Anatomy. All of the material housed was originally used to teach students of medicine, nursing, dentistry and associated healthcare fields. The collection is comprised of donated cadavers following the regulations of the South African Human Tissue Act 2003. It states that if a body is unclaimed within a 24-hour period then it can be donated. For an individual to claim a body (even after donation) they must show proof of relationship or a court order. If the body has been claimed, then all expenses lie with the claimant. Approximately 50-100 bodies are reclaimed *per annum* (L'Abbé *et al.* 2005)

The procedure of an individual being accessioned into the PBC once it has complied with Act 61 of the South African Human Tissue Act 2003, is that they must be 'known'. 'Known' in this case means the individuals sex, age at death, geographic origin and population affinity, height, weight, and cause of death was noted during autopsy. Once the individual had been transferred from the hospital to the University, an accession number is generated and then embalmed. After a one to two year period, the individual will then be used as a teaching aid for medical and clinical anatomy students for dissection. Once finished with dissection, they are macerated and accessioned into the PBC. The same accession number is kept from arrival to the University to placement within the skeletal collection. This lowers the amount of transcription errors. After maceration, if an individual is not complete and/or certain information is not known (i.e. sex or age at death) they are accessioned into the Student-Teaching Collection. Each skeleton is stored in an acid free box.

The PBC is separated into five research areas being Complete Skeleton, Complete Skull, Incomplete Skull, Complete Postcranial and Incomplete Postcranial. For these five research areas, demographics can be provided. Not only can geographic origin be given but also an individual's population affinity. There are several main population groups within this collection; white South African, black South African, and coloured South African. Coloured South African refers to individuals with mixed ancestry (white, black, and/or Asian). There are noticeable differences in the composition of the collection regarding the population groups. These variances can be explained by the differences in religion, socioeconomic circumstances and the regulations within the South African Human Tissue Act 2003.

3.2.5 Raymond A. Dart Collection of Human Skeletons, RSA

The Raymond A. Dart Collection (RDC) is housed within the School of Anatomical Sciences at the University of the Witwatersrand, Johannesburg, Gauteng Province (South Africa). The collection was established in the early 1920's by Dr Raymond A. Dart who was a lecturer in anatomy. The collection is comprised of donated cadavers, which are firstly used as teaching aids for dissection for the medical and clinical anatomy students. After this, the individual is macerated and accessioned within the collection. The RDC currently holds over 2,600 human skeletal remains that are used solely for research purposes (Dayal *et al.* 2009).

The collection is one of the most well-known and established collections not only in South Africa but also in the world, being the second largest modern skeletal collection in the world. Because of this, it is used as both a comparative collection and it also helps to create population standards within biological/forensic anthropology. It is primarily used as a research collection within the fields of forensic anthropology, population biology, dentistry/dental anthropology and clinical anatomy.

Sadly, in 1959, the basement where the collection was kept was flooded which ended up with a large amount of the known skeletal material being mixed. This occurred before the skeletal material had been fully labelled with the individuals generated accession number. This mixing of the remains was sadly never fully rectified and researchers have noted problems with the specimen's accessioned pre 1959.

The collection contains three main groups being white South African, black South African and coloured South African. As previously mentioned in the PBC history, “coloured” refers individuals of mixed ancestry. Not only does the collection state what main ancestral group they belong too, but in the case of black South Africans information of the individuals “tribe” is known. South Africa is rich in diversity having 11 official languages to which many are attached to specific indigenous South African tribes such as Zulu, Khoisan or Soto. Similar to the PBC, there are differences in the number of males and females, and age groups within the three main ancestral groups. Again, this is due to differences in religious practices etc.

3.2.6 Kirsten Skeletal Collection, RSA

The Kirsten Skeletal Collection (KSC) is housed within the Division of Anatomy and Histology at Stellenbosch University, Western Cape Province (South Africa). The collection consists of 892 cadaver derived skeletal remains, which are representative of the population located within the northern suburbs of Cape Town and its surrounding areas in the Western Cape.

Each skeleton goes through a similar accession process seen by the collection housed at the University of Pretoria and University of the Witwatersrand. Information on each cadaver includes their name, last residence, date of birth, age at death, sex, ancestry and cause of death. Some records are incomplete due to a lack of information for unclaimed bodies. The medical histories of all individuals are unknown. Males represent a larger proportion of the skeletons (54%) than females. Age at death ranges from 15 to 100 years, with the majority (87%) of individuals dying between the ages of 40 and 69. The collection mainly represents three population groups, black (14.3%), white (13.9%) and coloured (60.9%) South Africans. The remaining individuals in the KSC are from archaeological excavations with the majority having their ancestry estimated to being of the historic Khoe and San tribes (pers. coms. Mandi Alblas 2015).

As a result of damage and loss due to student use, the collection currently has 454 complete individuals (54 of which are articulated) and 363 incomplete individuals; 91 individuals have mainly post cranial skeletal elements and 272 skulls only. The rest of the individuals are on loan to other institutions.

3.3 Medieval Britain, U.K.

Two human skeletal collections were chosen for analysis which have been dated to be from the British Medieval period. There were two events in British history that helped define the Medieval period, the Norman conquest in 1066 AD and the battle of Bosworth in 1485 AD (Whittock 2009). This period saw a transition in political mind set and had several large scale events such as the development of the church, evolution of the language, industry, and the change in social hierarchy (Whittock 2009). Even though there was a division between men and women in this period, the treatment of the dead was very similar (Whittock 2009). The vast majority of individuals were buried in a Christian style which was on an east-west alignment (Daniell 1998) although some deviations can be found. A change from the norm can be dictated by a person 'honour' or/and wealth, warranting them a burial closer or even within the chancel or church. This meant that the churchyard was left for the 'general person' or layman.

Two samples were chosen for this project being the Medieval Poulton Chapel, close to Chester, and St. Owens Church in Gloucester. The two sites are located along the English-Welsh border with both following a Christian-like burial practice of east-west alignment. The main differences between the two are that they are from different ways of living, individuals from the Poulton Chapel having a rural lifestyle and the individuals from the St. Owens having an urban lifestyle.

3.3.1 Poulton Skeletal Collection

The Poulton Skeletal Collection originates 3 miles south of Chester along the border of Wales (Figure 13). This is a multi-period site and it can be dated between the Mesolithic period 4,300 BC to World War II (1939 AD-1945 AD). The sites primary focus is on locating the 'lost' Cistercian abbey (built 1153 AD) with the chapel itself 'vanishing' in the 17^{th} Century (Emery 1996, 2000). At present, the full outline of the chapel has been excavated along with the tower, nave and chancel. Approximately 770 burials have been identified with roughly 550 housed at Liverpool John Moores University (LJMU). Approximately 30 individuals were repatriated at Mt. St. Bernard and the remaining being housed onsite or at the University of Liverpool (UoL). From the skeletons excavated so far, it is thought that these 770 individuals represent only a small proportion of the burials that are within the chapel's limit. More burials are expected once the chapel has been fully excavated (pers. coms. Mike Emery 2013). The excavation is still ongoing but with a shift in focus from the medieval cemetery to the Roman ring ditches.

The majority of individuals were buried in the typical east-west Christian burial format (Daniell 1998), however a few have been noted to be buried in a west-east orientation. This can be due to several reasons such as limited space within the churchyard - causing head to toe packing of bodies, carelessness, or deliberate ill treatment of the dead (Burrell & Carpenter 2013). There is currently no evidence of any coffin burials at this site, however large amounts of shroud pins have been identified (Burrell & Carpenter 2013). Apart from the shroud pins, there has been a small number of associated finds for the burials at Poulton. The only other notable finds that have been recorded were the metal arrowheads found within the thorax of two skeletons (Canavan 2014), a 'belt buckle', and a small knife blade (Burrell & Carpenter 2013). Because of the geology of the site, finding a complete set of preserved human remains is rare with a large portion of skeletons excavated suffering from a high degree of taphonomic damage leading to fragmentation (Burrell 2017, Davenport 2017). Several individuals from the collection were sent for AMS Radiocarbon

Dating. Two specimens (SK53 and SK535) were analysed in 2012 which provided dates ranging from 1280-1320 AD and 1560-1630 AD. These dates are in agreement with additional analysis of SK535 which was associated with a M7 Bodkin arrowhead found within the ribcage of the individual (Canavan 2014). Two additional AMS Radiocarbon dates were obtained after analysis of two individuals with Paget's disease of bone (Burrell *et al.* 2016). Specimen SK463 gave two possible dates being 1275-1310 AD and 1360-1385 AD and specimen SK750 being dated as 1285-1330 AD and 1340-1395 AD. These dates and the material culture found on the archaeological site suggest that Poulton Chapel was in use between, and at the latest, 1285 AD and 1630 AD.

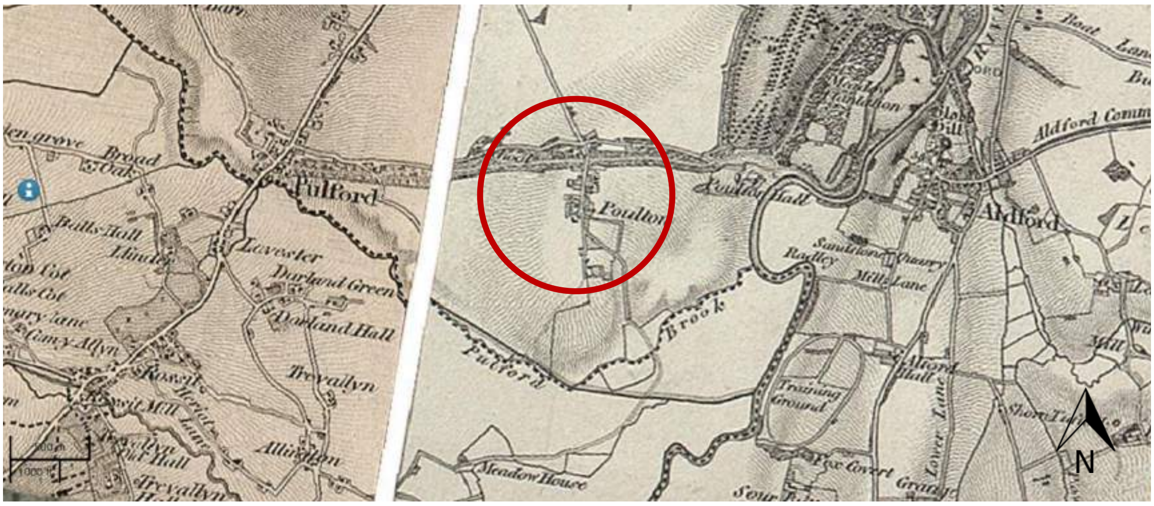


Figure 18: Ordnance Survey Map of Poulton, First Series, Sheet 80. 19th century. (Taken from http://www.visionofbritain.org.uk/maps/). The Poulton site is marked in red.

3.3.2 St. Owens Skeletal Collection

The St. Owens skeletal collection is also housed at Liverpool John Moores University. This collection derives from excavations at Southgate Street in Gloucester (U.K.) with the site running close to the Welsh border along the River Severn. This sample is from a multi-period archaeological site (Figure 14), where there were two stages of excavation (in association with Bristol and Gloucester Archaeology), the first excavation in 1983 now referred to as 13/83, and the second in 1989 referred to as 3/89. Rawe & Hodsdon (1990) suggested that the site was originally a Roman suburb that was later acquired by the Llanthony Priory in 1137. Many other excavations in this area have been undertaken which reveal information regarding the earliest phases of the Roman Conquest. As well as the archaeological evidence, the earliest archival record of Gloucester/Gloucestershire can be found in the Doomsday Book in 1086 (Morgan 1978). After the flourishing Roman settlement between the 2nd and 5th Centuries, the area was attacked by Anglo-Saxons leading to the site being abandoned for a small amount of time until the 10th and 11th Century (Atkin & Garrod 1990). Because of this new Anglo-conquest, the foundations of St. Peters Abbey and Llanthony Priory were built south of the now fortified Roman city wall. The St. Owens Church was founded in 1100 and became possession of the Llanthany Priory in 1137

(Fullbrook-Leggatt 1945). The church was demolished in 1847 for the extension of dockland (Atkin & Garrod 1990).

Figure 19: Ordnance Survey Map of Gloucester, First Series, Sheet 43. 19th century. (Taken from http://www.visionofbritain.org.uk/maps/). The Llanthony Priory site is marked in red.

From these two excavations, ~300 burials were located within the priory's surrounding churchyard. It is believed that they are of Christian origin due to burial orientation (Atkin & Garrod 1990) with only one individual being buried in the reverse west-east orientation. Research on this collection has been limited especially when concerning the study of the skeletal material. This will be one of the first pieces of research undertaken on the collection. The initial excavation in 1983 resulted in 71 individuals being exhumed and the 1989 excavation in exhuming 225 individuals. From archival records, it seems that a large proportion of the burials were shroud burials with suggestions that some individuals were coffin burials. Further archival research is needed to validate this idea.

3.4 Chumash, U.S.A.

The Native American tribe known as the Chumash originates from the Pacific coast of Northern America, in the state of California (Figure 15). Before the European contact, there were more than 60 spoken languages with complex socioeconomic systems (described as chiefdoms) in this area (Arnold 2001). Even though it is unknown when or where the first people entered, what is now California, it is known that it was inhabited as early as 13,000 years ago. This date was suggested by AMS radiocarbon dates from a femur found in the Arlington Canyon on the Island of Santa Rosa (Orr 1962, Johnson *et al*. 1999, Reeder *et al*. 2008). Even though direct dating was used on the femur, very little information could be further gleaned due to the amount of erosion. Walker

(2006) stated that it is possible to divide the region into five periods of occupation being Paleo-Indian (pre-9,000 BC), Early (9,000 – 1,000 BC), Middle (1,000 BC – 1,100 AD), Late (1,100 – 1769 AD), and Colonial (1,100 AD – present). Using Walker's (2006) periods of occupation, the femur mentioned previously sits easily within the Paleo-Indian period for California. With the early Chumash (9,000 BC – 1,000 BC), there was evidence that the those inhabiting the Channel Islands predominantly relied on a high carbohydrate diet with a higher prevalence of dental caries found in females than males (Walker & Erlandson 1986). Even though they were relying on a high carbohydrate diet, isotopic analysis (of Nitrogen and Carbon ratios) show that the Islanders consumed a high proportion of marine life in contrast to their mainland counterparts (Walker & DeNiro 1986). There is also isotopic evidence as well to suggest that as time progressed and entered the Middle phase of occupation, the southern islands on the Californian Coast traded foods with the mainland (Goldberg 1993).

The Californian Channel Islands were excavated under the direction of Roy Lee Moodie in 1932 as part of a Wellcome Institute grant to study health and disease in past populations. The remains were then transported to the U.K. to be studied, sadly, Roy Moodie died in 1934 never completing his work on the Chumash skeletal material. The skeletal collection remained at the Wellcome Institute until 1981/1982 where they were sent to the British Museum (Natural History) for curation (pers. coms. Margaret Clegg 2015). The collection holds individuals from the Santa Cruz, San Miguel, Santa Rosa, Santa Barbara, San Nicolas, Santa Catalina and San Clemente islands. The collection, now housed at the Natural History Museum, London (U.K.) formerly part of the British Museum, contains ~500 human skeletons varying from complete specimens to only a metacarpal being present. There are approximately 350 adult and 150 sub-adult human skeletons present. However, caution must be taken when only using the collections catalogue as some errors are apparent, i.e. sub-adults being misrepresented as adults and vice versa. Even though little research has been done on this collection, preliminary studies have been undertaken into how old the skeletons are. From unpublished AMS radio carbon dates, the collection ranges from 4,490±40BP to 1,030±40BP (per. Coms. Silvia Gonzalez, 2014), placing it within the Early and Middle periods of Californian prehistory (Walker 2006).

Figure 20: Map showing the location of the Channel Islands in relation to the United States of America (Taken from http://www.montroserestoration.noaa.gov/case-document/ssettlement/ map-sc-bight-and-chan-is-incl-white-pt-bmp/)

3.5 Andaman Islands, India

The Andaman and Nicobar Islands are located in southern Asia and they are declared as one of the seven union territories of India. The islands fall between the Bay of Bengal and the Andaman Sea (Table 16). The Andamanese are part of the Negrito groups and they are classed as one of few pygmy populations in the world. Originally there were five major tribes living on the islands; Great Andamanese, Jarawa, Jangli, Sentinelese, and Onge. However, the population declined drastically when the British started to colonise South East Asia. European colonisation began in the 16th Century however, the British did not colonise the Bay of Bengal until 1783. With this invasion, introduction of European mainland diseases nearly wiped the entire population out. Transportation of the human skeletal remains from the Andaman Islands started in the 18th Century and continued throughout the 19th and early 20th Centuries.

The Andaman collection is currently held at the Natural History Museum, London and contains approximately 100 specimens. Some of the specimens in the collection hold information regarding their age at death, sex and which city they lived in on the Islands, with the large majority from Port Blair. Due to the current closure of the collections at the Natural History Museum, information of when the collection was curated and also the demographics of the collections are unavailable. It is expected that they will be available in early 2017.

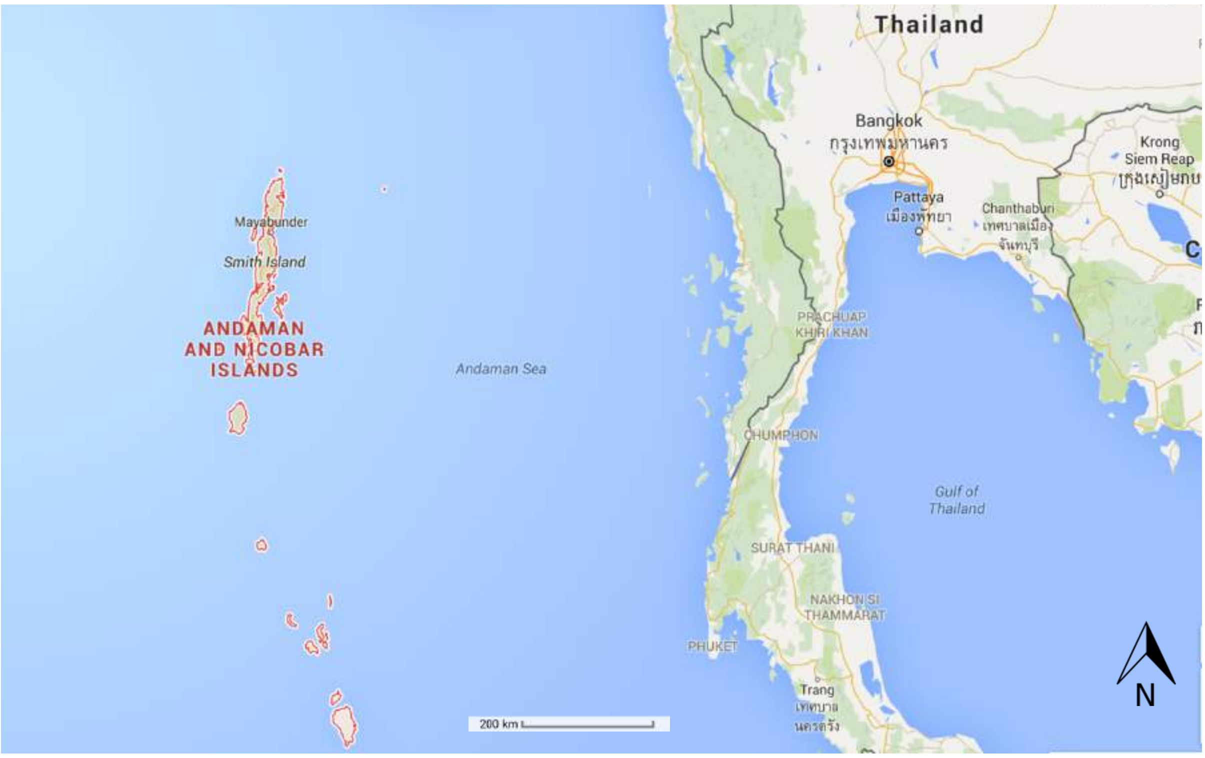


Figure 21: Map showing the location of the Andaman and Nicobar Islands in relation to Thailand. (Taken from Google Maps, 20/12/2015).

3.6 Selection Criteria and Sampling

For individuals to be selected for analysis, only those with a partial or complete pelvis who were aged 18 years and above were chosen. The 18 years old 'cut of point' was chosen as this is generally seen as the point of adulthood. For the four known age and sex samples (Christ Church, Spitalfields, South African White, South African Black, and South African Coloured), age was taken from collection/cadaver records. The Spitalfields age and sex is from the associated coffin plate they were found with. The three south African samples are from cadaver based human skeletal collections, thus sex and age is determined from physical examination before dissection and medical records. The two British Medieval samples (Poulton and St. Owens) underwent individual skeletal aging and sex was estimated using a technique known as seriation. These estimation techniques for both age and sex are stated in the Chapter 4. The Chumash and Andaman samples age and sex were based from museum records and previous research. Both samples would have undergone a similar method seen to the medieval samples but due to time and lack of physical space at the Natural History Museum, their records had to be used. Because some issues had arisen when collecting data from the Native American sample, age and sex estimations had to be undertaken to make amendments to the museum records but only in extreme circumstances (i.e. individual referred to as sub-adult aged 9-10 years when in fact it was a fully matured adult skeleton). Age and sex for these individuals were estimated using the same methods for the sexed based seriation and age estimations for the Poulton and St. Owens samples.

Below are the sample sizes created using the selection criteria described above (Table 2). An overall sample of 1,128 individuals were analysed which included 27 Andaman, 38 Chumash, 205 black South Africans, 205 coloured South Africans, 193 white South Africans, 132 from Medieval Poulton, 100 from Medieval St. Owens, and 228 from Christ Church, Spitalfields. Figure 17 shows the geographic distribution of the samples with Table 1 highlighting their place in time.

Table 1: Breakdown of the eight human skeletal samples analysed.

Population	Location	Sample Size (N)	Time Period
Andamanese	Andaman Islands	27	$18^{th}/19^{th}$ Century
Chumash	North America	38	4,490 BP-1,090 BP
RSA Black	South Africa	205	$20^{th}/21^{st}$ Century
RSA Coloured	South Africa	205	$20^{th}/21^{st}$ Century
RSA White	South Africa	193	$20^{th}/21^{st}$ Century
Poulton	U.K.	132	13^{th} Century
St. Owens	U.K.	100	15^{th} Century
Spitalfields	U.K.	228	$18^{th}/19^{th}$ Century
Total		1128	

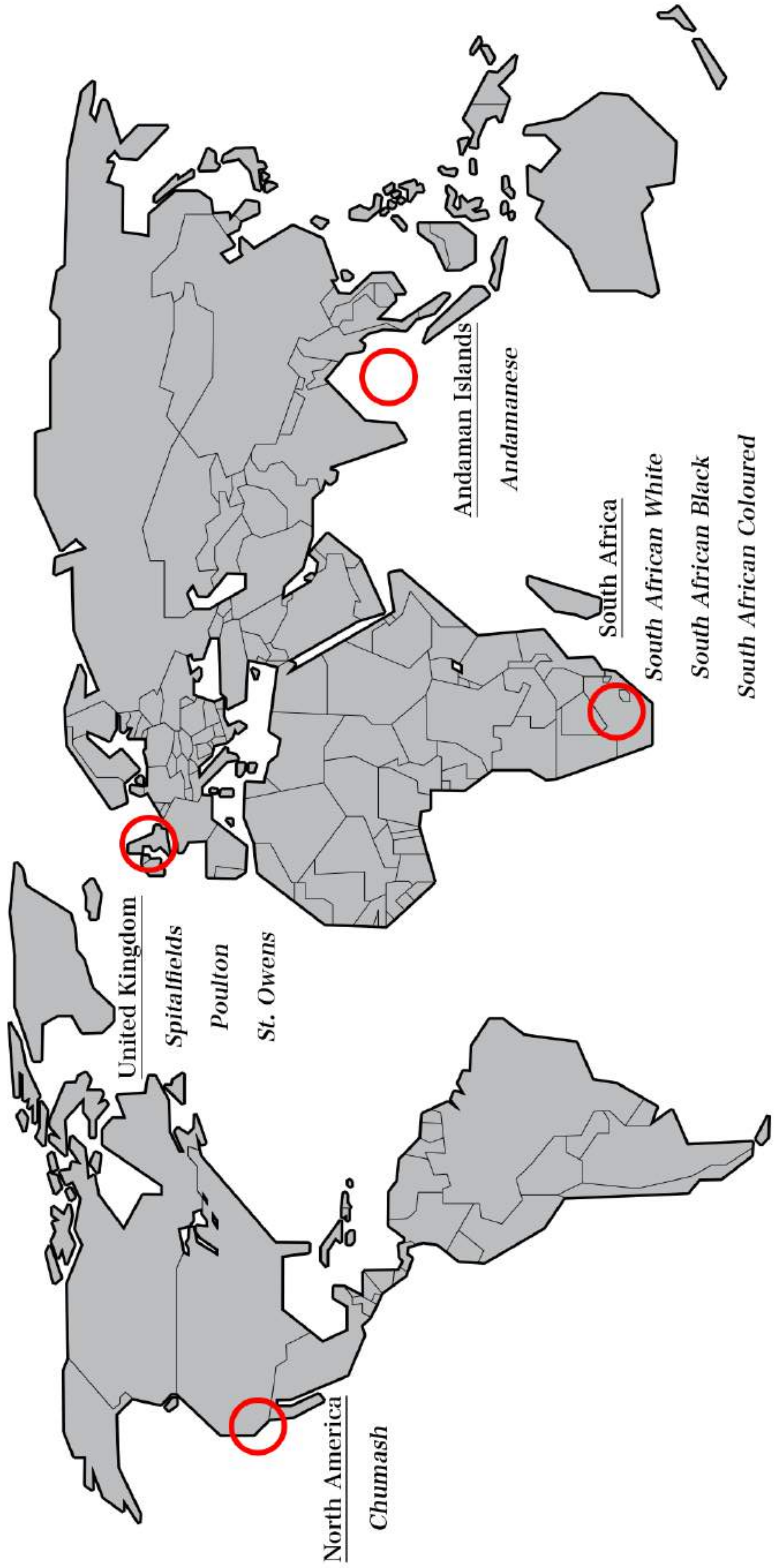


Figure 22: World map showing the location of the eight populations analysed. (Image adapted from https://upload.wikimedia.org/wikipedia/commons/e/ec/World_map_blank_without_borders.svg)

4. Methodology

This chapter will be explaining the methods used for data collection and data analysis. To help move the analysis forward, the overall sample of 1,128 individuals was split into two groups. Firstly, the known age and sex samples (Spitalfields and the three South African samples) and secondly the four archaeological samples (Chumash, Andaman, Poulton, and St. Owens). The primary analysis will be performed onto the first group and then later applied onto the second group.

4.1 Scoring System & Data Collection

Eight observations were made on each individual in the sample. Eight morphoscopic traits from the pelvis were chosen specifically for this study: ventral arc, subpubic concavity, medial aspect of the ischiopubic ramus, shape of the pubic body, subpubic angle, obturator foramen, greater sciatic notch, and the preauricular sulcus. These eight were chosen because of their popularity among practitioners and researchers for estimating sex for both archaeological and forensic remains (The Workshop for European Anthropologists 1980, Rogers & Saunders 1994, White & Folkens 2000, Durić *et al*. 2005, Rösing *et al*. 2007, Sidler *et al*. 2007). They were also chosen because of their location; not one specific area of the pelvis was taken into account, the whole pelvis was taken into consideration for sex estimation. Table 2 summarises the eight morphoscopic traits with their associated accuracies.

Each trait on the pelvis was scored along a five-point ordinal scale ranging between “-2” (hyper feminine) to “+2” (hyper masculine), with “0” being classed as ambiguous. For the Greater Sciatic Notch, Walker’s (2005) descriptions were used (Figure 18) and the descriptions and illustrations from Klales *et al*. (2012) were used for scoring the Phenice triad (1969) (Figure 19). The Preauricular Sulcus (Figure 20) was scored using the descriptions made by Milner (1992) (taken from Buikstra & Ubelaker 1994). The scoring of the obturator foramen and subpubic angle were based from the Workshop of European Anthropologists (1980). Basic descriptions for the shape of the pubic body were found in Roger & Saunders (1994).

Because the obturator foramen, the shape of the pubic body, and subpubic angle do not have any illustrations and/or descriptions for a five-scale ordinal scale, new illustrations and expanded descriptions were created for this study. Figures 21 - 23 show the new scoring method for the obturator foramen (Figure 21), shape of the pubic body (Figure 22), and the subpubic angle (Figure 23). These definitions have been previously tested on individuals with varying experience of morphoscopic analysis and have been modified to best suit all levels of expertise. Results from the inter/intra observer can be found in Chapter 5.

The scoring system described for the Phenice triad (Klales *et al*. 2012), Greater Sciatic Notch (Walker 2005), and Preauriucular Sulcus (Milner 1992 found in Buikstra & Ubelaker 1994) was adapted to fit a scale described by Acsadi & Nemskeri (1979). So, instead of a "1" to "5" scale, scoring ranged from "-2" to "+2", where "-2" was female and "+2" was male. The remaining three traits (described below) follow this new scoring scale.

Table 2: Morphoscopic traits used in this study and associated accuracies.

Trait	Source	Accuracy
Ventral Arc	Rogers & Saunders 1994; Klales *et al*. 2012	86-95%
Subpubic Concavity	Rogers & Saunders 1994; Klales *et al*. 2012	80-95%
Ischiopubic Ramus	Rogers & Saunders 1994, Klales *et al*. 2012	75-95%
Obturator Foramen	Workshop for European Anthropologists, 1980; Rogers & Saunders 1994	93.80%
Shape of Pubic Body	Rogers & Saunders 1994	86.20%
Preauricular Sulcus	Workshop for European Anthropologists 1980; Houghton 1974, Buikstra & Ubelaker 1994	76%
Subpubic Angle	Workshop for European Anthropologists 1980; Rogers & Saunders 1994	74%
Greater Sciatic Notch	Walker 2005; Gómez-Valdes *et al*. 2012	65-70%

The descriptions for the scoring of the greater sciatic notch described by Walker in Buikstra & Ubelaker (1994) are as follows:

"The notch should be held six inches above the diagram [Figure 18] so that the greater sciatic notch has the same orientation as the outlines, aligning the straight anterior portion of the notch that terminates at the ischial spine with the right side of the diagram. While holding the bone in this manner, move it to determine the closest match. Ignore any exostoses that may be present near the preauricular sulcus and the inferior posterior iliac spine. Configurations more extreme than "1" and "5" should be scored as "1" and "5" respectively. [Figure 18], illustration "1" present a typical female morphology, while the higher numbers show masculine conformations." (Buikstra & Ubelaker 1994 page 18). For this study, a score of "1" has been changed to a score of "-2", and a score of "5" is now a score of "+2".

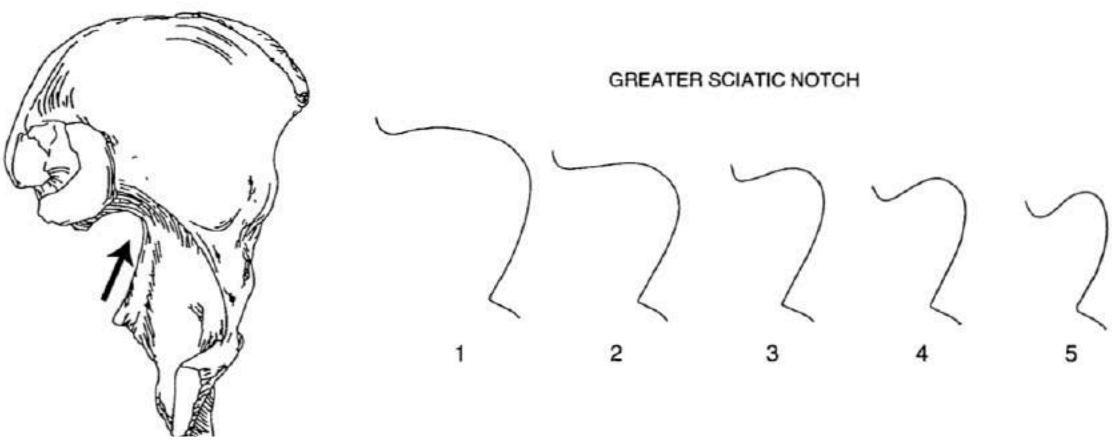


Figure 23: Morphoscopic scoring for the Greater Sciatic Notch ranging from most female (1) to most male (5). Image taken from Buikstra & Ubelaker page 18

The descriptions for the Phenice triad described by Klales *et al.* (2012) are as follows:

- Ventral Arc

    1. Arc present at approximately or at least a 408 angle in relation to the symphyseal face with a large trian- gular portion of bone inferiorly placed to arc. Scored as -2.

    2. 2. Arc present at approximately a 25–408 angle in rela- tion to the symphyseal face with a small triangular portion of bone inferiorly placed to arc. Scored as -1.

    3. 3. Arc present at a slight angle (less than 258) to the symphyseal face with a slight, nontriangular portion of bone inferiorly placed to arc. Scored as 0.

    4. 4. Arc present approximately parallel to the symphyseal face with hardly any additional bone present inferior to arc. Scored as +1.

    5. 5. No arc present (therefore, no additional bone present inferior to the arc). Scored as +2.

- Subpubic Concavity (described as Subpubic Contour in Klales *et al.* 2012)

    1. Well-developed concavity present inferior to symphy- seal face and along length of inferior ramus. Scored as -2.

    2. 2. Slight concavity present inferior to face extended par- tially down inferior ramus. Scored as -1.

    3. 3. No concavity present, bone is nearly straight (may be a very slight indentation just below the symphyseal face). Scored as 0.

    4. 4. Small convexity, especially pronounced along inferior pubic ramus. Scored as +1.

    5. 5. Large convexity, especially pronounced along inferior pubic ramus. Scored as +2.

- Medial Aspect of the Ischiopubic Ramus

    1. Ascending ramus is narrow dorso-ventrally with a sharp ridge of bone present below the symphyseal face. Scored as -2.
    2. 2. Ascending ramus is narrow dorso-ventrally with a plateau/rounded ridge of bone present below the sym- physeal face. Scored as -1.
    3. 3. Ascending ramus is narrow dorso-ventrally with no ridge present. Scored as 0.
    4. 4. Ascending ramus is medium width dorso-ventrally with no ridge present. Scored as +1.
    5. 5. Ascending ramus is very broad dorso-ventrally with no ridge present. Scored as +2

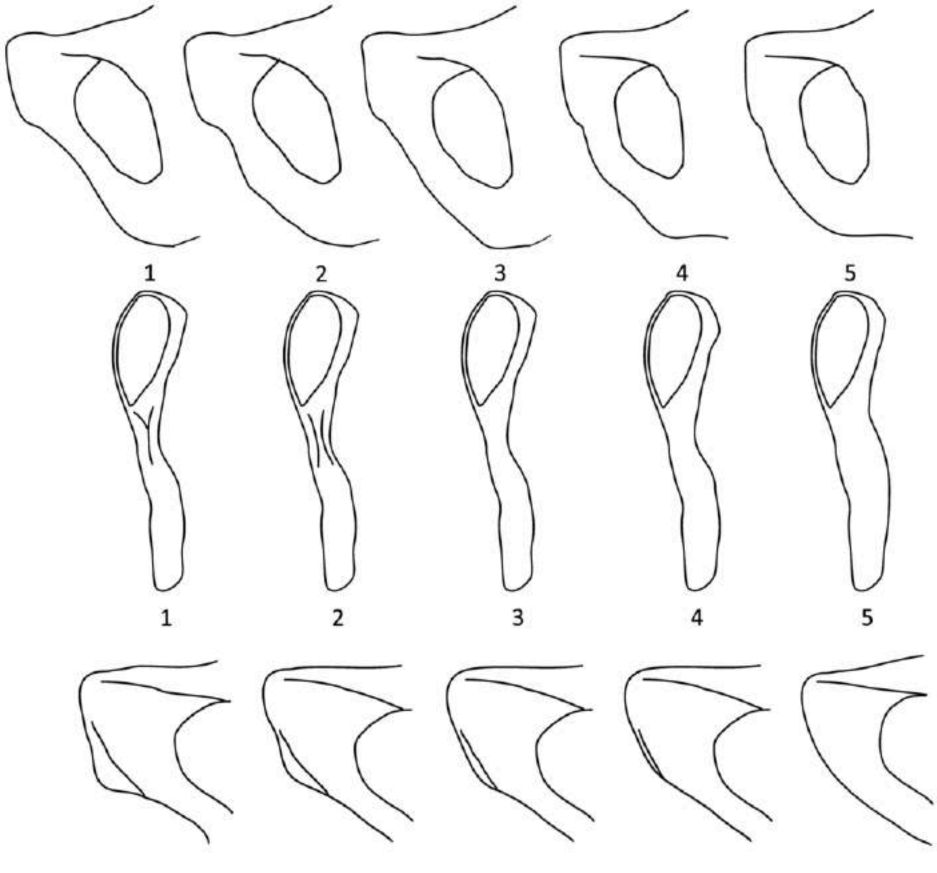


Figure 24: Phenice triad taken from Klales *et al*. 2012. Scoring for the Subpubic Concavity, medial aspect of the Ischio-pubic Ramus, and Ventral Arc. Female (1), ambiguous (3), and male (5).

The descriptions for the preauricular sulcus described in Buikstra & Ubelaker (1994) are as follows:

1. The preauricular sulcus is wide, typically exceeding 0.5cm, and deep. Scored as -2.
2. The preauricular sulcus is wide (usually greater than 0.5cm) but shallow. Scored as -1.
3. The preauricular sulcus is well defined but narrow, less than 0.5cm deep. Scored as 0.
4. The preauricular sulcus is a narrow (less than 0.5cm), shallow, and smooth walled depression. Scored as +1.
5. Absence of preauricular sulcus (not illustrated). Scored as +2.

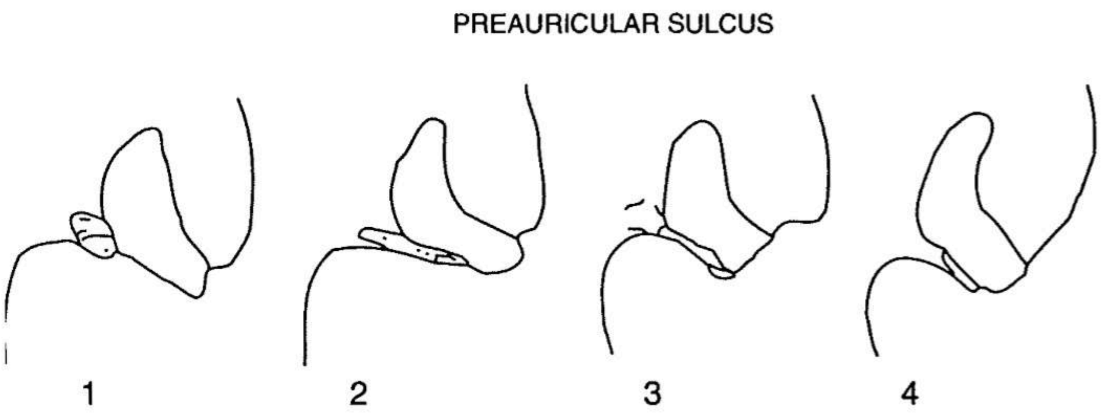


Figure 25: Morphoscopic scoring of the Preauricular Sulcus ranging from female (1) to male? (4). Absence of sulcus not illustrated. Image taken from Buikstra and Ubelaker 1994 page 19.

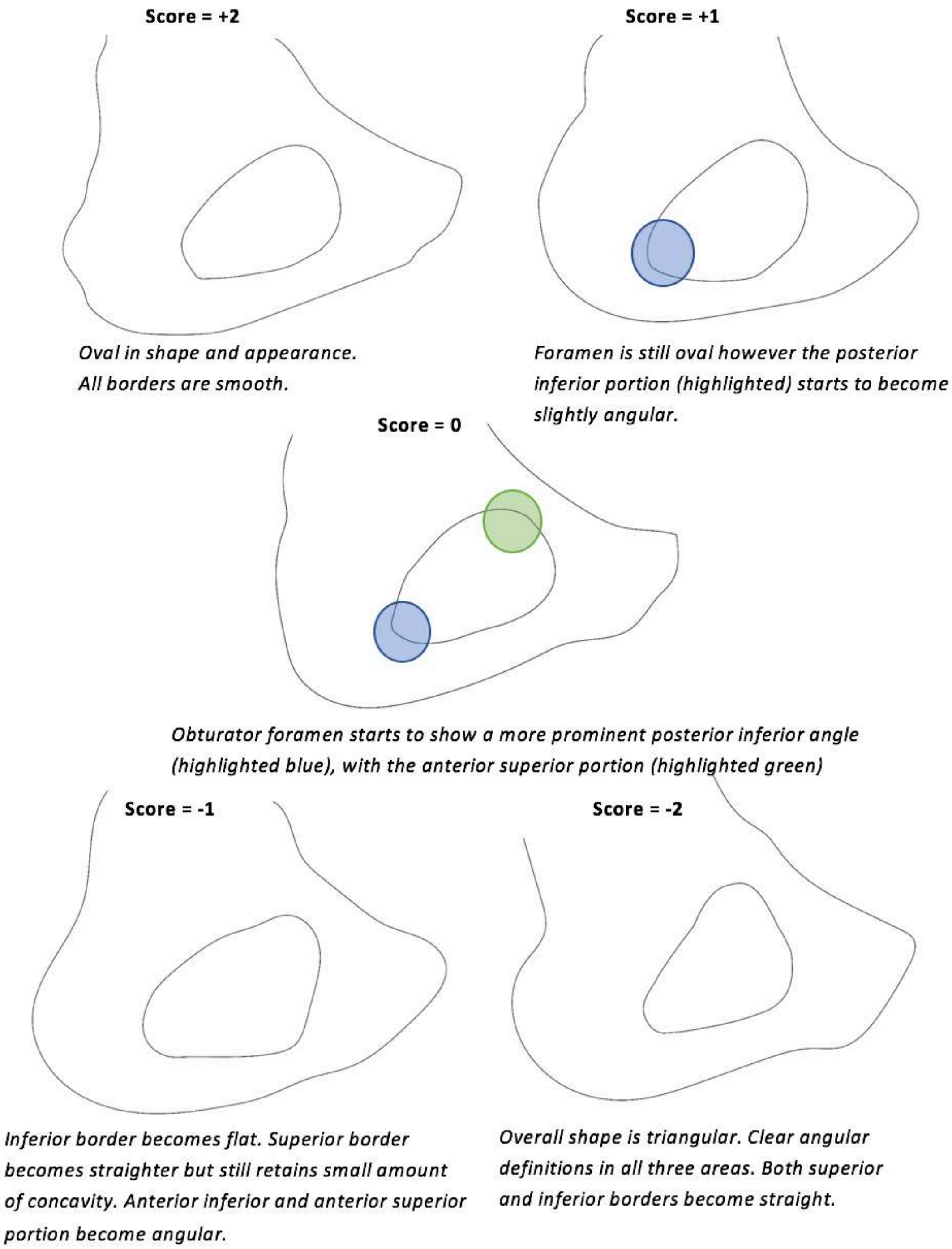


Figure 26: Morphoscopic scoring of the Obturator Foramen (dorsal view) ranging from male (+2) to female (-2). Drawn by author.

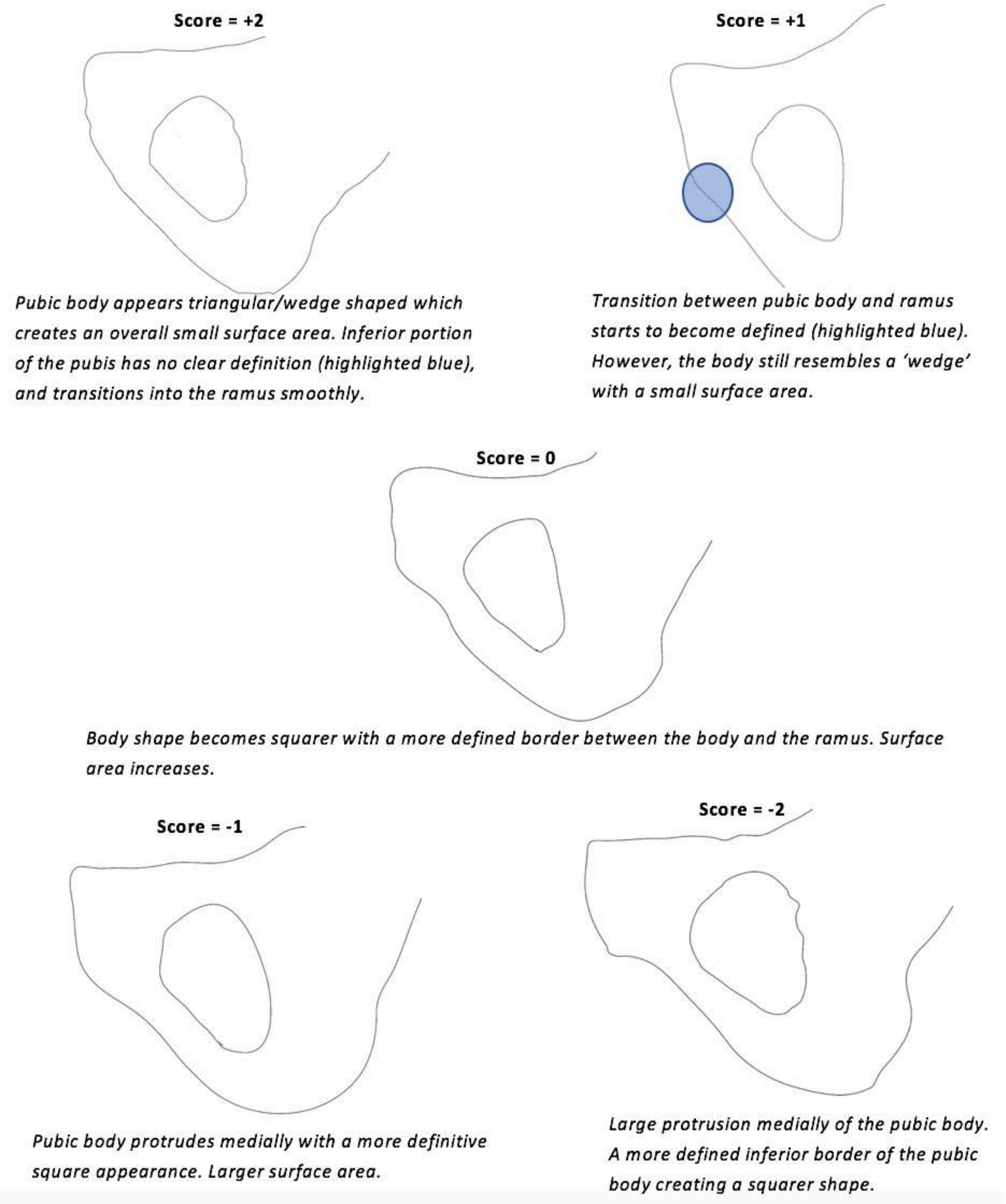
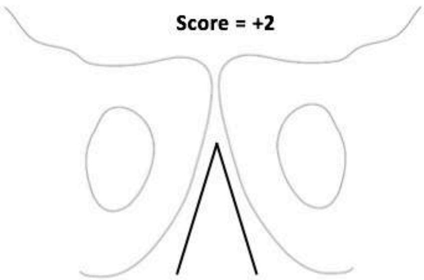
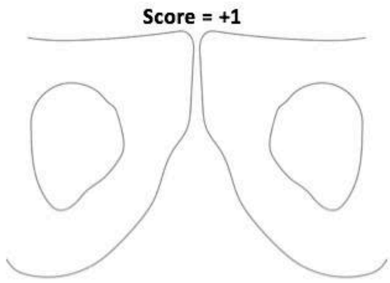


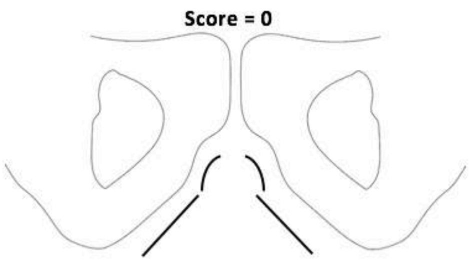
Figure 27: Morphoscopic scoring for the Shape of the Pubic Body (anterior view/ ventral surface of the pubic body) ranging from male (+2) to female (-2). Drawn by author.



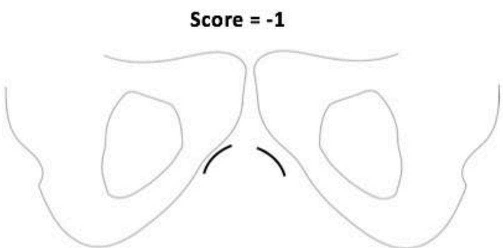
Angle between the os pubis is sharp and narrow. Forms an inverted 'V' shape. No concavity on the ischio-pubis.



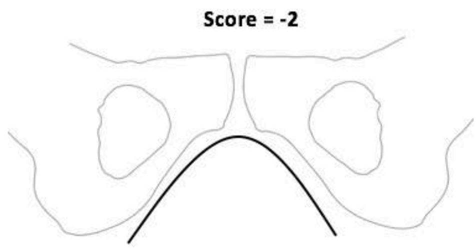
Angle is still narrow however a slight concavity can be seen along the ischio-pubic ramus. Still forms an overall 'V' shape but not as exaggerated as previous.



Subpubic concavity is present creating more of a 'U' shape in the superior portion of the ischio-pubic ramus but straightens inferiorly.



Angle is considered to be wide with an overriding 'U' shape. Softer transition between the curvature seen superiorly to the straightening inferiorly.



Angle is wide which creates a wide, soft 'U' shape. The curvature extends inferiorly towards the ischium.

Figure 28: Morphoscopic scores for the Subpubic Angle (dorsal view) ranging from male (+2) to female (-2). Drawn by author.

4.2 Statistical Analysis

Each of the four known age and sex samples were analysed separately. Firstly, trait frequencies were calculated to look at the spread of the five scores for each sex within a sample. From there, accuracies were calculated to see how well each morphoscopic trait was for estimating sex.

Inter and intra observer analyses will be conducted on the ordinal data using Kappa Cohens test of repeatability following Landis & Koch's (1977) definitions (Figure 24).

Kappa Statistic	*Strength of Agreement*
<0.00	Poor
0.00–0.20	Slight
0.21–0.40	Fair
0.41–0.60	Moderate
0.61–0.80	Substantial
0.81–1.00	Almost Perfect

Figure 29: Statistics and definitions for the Kappa Cohens test of repeatability. Taken from Landis & Koch (1977).

From here, Ordinal Logistic Regression analysis will be performed on each of the four populations using R-Studio (R Core Team 2015, RStudio Team 2015). The *glm()* function using the binomial distribution will be applied to create the OLR using all eight morphoscopic traits. A Wald Test from the package {aod}, function wald.test(), will be conducted on the OLR which will state if four models (each sample) were significant, meaning that the variables used all contribute to the model. Following this, *cv.glm()* function, from the package {boot}, will be performed to cross validate the regression analysis. Therefore, all results reported will be once they have undergone the cross-validation procedure. ROC curves will be generated for each of the OLR analyses. These curves will produce a visual representation of 'how well the model fits for its intended task'.

Once these results have been generated, each equation will be applied onto the other three to assess how accuracies differ when population specific equations are used on differing samples. To attempt a 'universal' equation, two cluster equations will be created. This will be achieved by combining the three South African samples to generate a collective South African OLR equation, and a Summary Sex equation will be created using all four samples. These will be created using the same functions seen above.

4.3 Archaeological Sex and Age-at-Death Estimation

Once the accuracy rates for each of the, now six, OLR equations have been tested, they will each be applied onto the four archaeological samples. Because these samples are not of known age or sex (except seven individuals from the Andaman samples), the results from applying the OLR equation will be compared to the either the museum records (for the Chumash and Andaman samples) or the results from the sexed based seriation (Poulton and St. Owens samples).

4.3.1 Age-at-death estimation

One of the selection criteria for analysis was that each individual had to be considered to be of 18 years or older, termed 'adult'. For the known age and sex samples, this was taken from the medical/museum records, however, for the medieval samples age-at-death had to be estimated. To make sure each individual was at least 18 years old at death, fusion of the ischial tuberosity and iliac crest was examined. If both epiphyses were at a stage of "later partial fusion" to "complete fusion" with the fusion line obliterated, then the individual was considered for analysis. However, if the iliac crest was only partially fused onto the iliac blade with the partial fusion of the ischial tuberosity then they were not considered. Descriptions from Webb & Suchey (1985) and Cardoso (2008) were used for epiphyseal fusion alongside Schaefer *et al*. (2009). After the initial observation, age-at-death estimations were undertaken using morphology of the auricular surface (Lovejoy *et al*. 1985a, Buckberry & Chamberlain 2002) and the pubic symphysis (Meindl *et al.* 1985a, Suchey & Brooks 1990), and cranial suture closure (Meindl & Lovejoy 1985) if a skull was present. Individuals were placed into three broad categories being young (18-35), middle (35 – 60), and old (60+) adult. These methods were also used to estimate age-at-death for several individuals from the Chumash collection which were originally aged younger than 18 years but were much older on inspection. Other than these individuals from the Chumash sample, the remaining Chumash and the Andaman sample, age-at-death was taken from the museum records.

4.3.2 Sex Estimation

Both the rural and urban medieval skeletal samples were seriated to estimate sex. To do this, each individual in the sample was ranked according to how 'male' or 'female' their morphology was. Several alternative sexing techniques from the cranium and mandible (if present), femur, and the pelvis were employed to aid in the seriation. For the cranium and mandible, Walker's (2008) descriptions of five morphoscopic traits were used (Mental Eminence, Orbital Margin. Nuchal Area, Mastoid Process, and the Glabellar area). These data were collected if they were present with no attempt to reconstruct missing data. Dimensions of the femoral head were taken for estimating sex using Stewart (1979) using a Mitsitoyo digital calliper. The overall robusticity of the pelvis was taken into account as well for determining where in the seriation an individual would be placed. The final set of variables that were taken were also from the pelvis; Bruzeks (2002) morphoscopic descriptions were applied to all individuals. When trying to rank which traits perform best, inspiration was taken from Kjellström (2004) article on sexing fragmented human skeletal remains. The dimensions from the femoral head held the lowest ranking, cranial traits were classed as 'mid-ground', and the pelvic traits were of the highest ranking. Only once each individual had been analysed and placed within the seriation was sex then assigned. This procedure follows what was originally done in 1985 by Lovejoy and colleagues (Lovejoy *et al.* 1985a) who seriated skeletons for age-at-death estimations using the Hamann-Todd collection at the Cleveland Museum of Natural History (USA). These methods were also used to estimate sex for several individuals of the Chumash collection but did not undergo seriation. Other than these individuals from the Chumash sample, the remaining Chumash and the Andaman sample, sex was taken from the museum records.

4.4 Applying the Known onto the Unknown

After collecting the data from the four archaeological samples, the six OLR equations will be applied onto them to assess how the assignment of sex differs from the museum records/seriation sex estimates. Only individuals with all eight traits present were chosen for analysis. Figure 25 illustrates how each archaeological sample will be analysed.

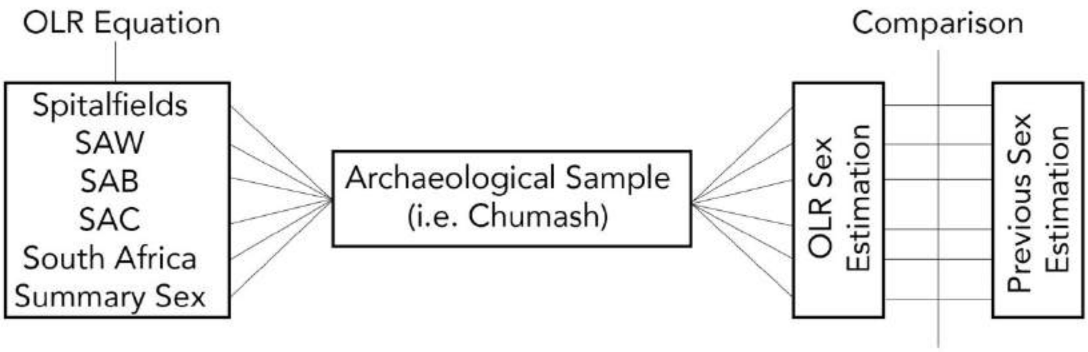


Figure 30: Framework for applying the OLR equations onto the archaeological samples. Drawn by author.

4.5 Fragmentation

As with many osteological collections, fragmented remains are common. In order to not disqualify specimens from further study, three possible routes were taken to increase sample sizes.

1. Use the individuals median value to replace all missing scores to complete the OLR equations for fragmented samples.
2. Apply only sections of the already generated OLR equations onto individuals who are fragmented to increase sample size:
    a. OLR Equations with only the ventral arc, subpubic concavity, and ischiopubic ramus variables ("Phenice OLR equation").
    b. OLR Equations with only the subpubic angle and shape of the pubic body ("Anterior Posterior OLR equation").
    c. OLR Equations with only the greater sciatic notch and preauricular sulcus variables ("Posterior Pelvis OLR equation").

3. Apply new and shorter OLR equations onto individuals who are fragmented:
    a. OLR Equations with only the ventral arc, subpubic concavity, and ischiopubic ramus variables ("Phenice OLR equation").
    b. OLR Equations with only the subpubic angle and shape of the pubic body ("Anterior Pelvis OLR equation").
    c. OLR Equations with only the greater sciatic notch and preauricular sulcus variables ("Posterior Pelvis OLR equation").

The first approach will be applied onto four samples that originally excluded individuals with missing data (Spitalfields, Poulton, St. Owens, and Chumash). To calculate the median value, imagine an individual that only had the variables greater sciatic notch and preauricular sulcus present and they were scored "-1" and "-2" respectively. Then the remaining six variables, that were originally missing, would have the value of "-1.5" in place of "N/A". Once an individual's median value had been calculated the new samples were ready for the six OLR equations to be tested.

The second and third routes would be an attempt to increase sample sizes without having to estimate missing values (unlike the first approach). The second approach is to create smaller OLR equations by taking sections of previously created OLR equations. Three sectioned equations will be created for each sample which best represent groups of variables. The first sectioned equation will be comprised of three traits originally used by Phenice (1969) and later adapted by Klales *et al*. (2012). The second set of sectioned equations, termed "Anterior Pelvis", will be created by pairing the subpubic angle and the shape of the pubic body. Lastly, the "Posterior Pelvis" equation will be made by lifting the greater sciatic notch and preauricular sulcus sections from the six previously created OLR equations.

To see if sectioning previously generated OLR equations will yield high accuracies and is a viable option for including individuals who are fragmented, a third approach will be created. This route will mirror the types of equations created for the second approach, but instead of sectioning OLR equations, they will be newly created OLR equations. This third route will hopefully show that sectioning equations have the same discriminatory power than new equations.

5. Results I

5.1 Introduction

In this chapter, we will be investigating the use of eight morphoscopic sexing traits of the human pelvis from four known age and sex samples using a variety of methods. Firstly, we observe the sexing accuracy for only one morphoscopic trait. Secondly, we apply all eight sexing traits to an Ordinal Logistic Regression (OLR) and compare the results to the single sexing trait. Thirdly, we attempt to create a universal OLR sexing equation.

This chapter tackles three hypotheses.

$1H_1$: The results from the OLR will provide a greater accuracy than that of a single morphoscopic trait for sex estimation.

$1H_0$: The results from the OLR will not differ from the sexing accuracy from a single morphoscopic trait for sex estimation.

After this, we assess the applicability of the using an OLR equation created from one sample and using it on the other three known age and sex samples. This will assess if the OLR equations are population specific.

$2H_1$: There is a difference in sexing accuracy when a different OLR equation is applied onto a sample.

$2H_0$: There is no difference in sexing accuracy when a different OLR equation is applied onto a sample.

If differences are found, and we reject the second null hypothesis, then an attempt into creating a universal equation that still yields a high sexing accuracy will be investigated. This will be achieved by combining samples together and observing if they reduce the inaccuracies seen in the previous analysis.

$3H_1$: The creation of a "universal" OLR sexing equation will reduce the misclassification rate seen when applying different OLR sexing equations on populations they were not created from.

$3H_0$: The creation of a "universal" OLR equation will not reduce the misclassification rate seen when applying different OLR sexing equations on populations they were not created from.

Once these three hypotheses have been tested, a summary of the results, considering each hypothesis, will be reviewed and a small discussion into what the results tell us will be stated.

5.2 Trait Frequency

To give an overview of trait frequencies for each population, all individuals (complete and incomplete pelves) were considered for analysis from each of the four known sex samples. Each sample was analysed separately, and percentage accuracies were calculated for each trait as a single sex assessor. If a minus score was given, the individual was deemed female and if a positive score it would be deemed male.

5.2.1 Christ Church, Spitalfields UK

The distribution of scores for males and females shows that, the majority of individuals have typical morphology with the scores of "-2" for females and "+2" for males or the Ventral Arc, Ischiopubic Ramus, and the Pubic Body Shape. There was very little variation in distribution with regards to the other scores given for both males and females for these traits. Males were shown to have a higher distribution of scores ranging from "+2" – "-1". Female individuals were shown to have a the more typical form of the subpubic angle whilst males had a much wider distribution ranging from the ambiguous "0" score to the typical male "V" shape angle with a score of "+2" score. The opposite can be seen for the preauricular sulcus; males were observed to have predominantly no sulcus present whilst females were found to have a range of all five scores. The obturator foramen and greater sciatic notch shown the largest amount of variation between the scores for males and females. This can be seen in the percent accuracy these traits have as single sex assessors (Table 3).

Table 3 shows the results from the macroscopic assessment of each trait. The shape of the pubic body performed the best as a single sex assessor with an accuracy of 90.42% and had the lowest rate of misclassification. This was followed by, the subpubic angle, and preauricular sulcus with accuracies of 83.95% and 79.91%, respectively. The remaining five traits all had accuracies ranging between 70 – 75%, with the subpubic concavity having the lowest accuracy of 70.59% and the highest inaccuracy of 18.82%.

5.2.2 South Africa: White

The distribution of scores for the ventral arc, subpubic concavity, and preauricular sulcus are very similar, in that the majority of males show the "+2" morphology whilst females show more of a spread of scores "-2" – "0". The shape of the pubic body and subpubic angle have near identical spread of scores between males and females. Predominantly females shown a "-2" score for both traits whilst males were split between the hyper-masculine and masculine scores of "+2" and "+1". The greater sciatic notch and ischiopubic ramus both show an equal spread of "masculine" and "feminine" scores for their associated sex, more so for the ischiopubic ramus than the greater sciatic notch. The obturator foramen shows a wide distribution and a large overlap between scores for both sexes.

For the White population of South Africa, the shape of the pubic body followed closely by the subpubic angle were the traits that had the highest correct sex accuracy (88.60% and 88.08% respectively). The three Phenice (1969) traits and greater sciatic notch had accuracies ranging between 79 - 87%. The remaining two traits had the lowest accuracies being 69.43% for the obturator foramen and 68.39% for the preauricular sulcus (Table 4).

Table 3: Accuracy rates for each morphoscopic trait for the Christ Church, Spitalfields sample.

Christ Church, Spitalfields (UK)			
	Correct	Unknown	Incorrect
VA	120/163 (73.62%)	8/163 (4.91%)	35/163 (21.47%)
SPC	120/170 (70.59%)	18/170 (10.00%)	32/170 (18.82%)
IPR	125/173 (72.25%)	28/173 (16.18%)	20/173 (11.56%)
PBS	151/167 (90.42%)	8/167 (4.79%)	8/167 (4.79%)
SPA	136/162 (83.95%)	17/167 (10.18%)	9/167 (5.39%)
OF	115/154 (74.68%)	14/154 (9.09%)	25/154 (16.23%)
GSN	159/220 (72.27%)	39/220 (17.73%)	22/220 (10.00%)
PAS	171/214 (79.91%)	8/214 (4.68%)	35/214 (16.36%)

Table 4: Accuracy rates for each morphoscopic trait for the South African White sample.

South Africa: White Sample			
	Correct	Unknown	Incorrect
VA	153/193 (79.27%)	18/193 (9.33%)	22/193 (11.40%)
SPC	162/193 (83.94%)	11/193 (5.70%)	20/193 (10.36%)
IPR	167/193 (86.53%)	11/193 (5.70%)	15/193 (7.77%)
PBS	171/193 (88.60%)	14/193 (7.25%)	8/193 (4.15%)
SPA	170/193 (88.08%)	11/193 (5.70%)	12/193 (2.12%)
OF	134/193 (69.43%)	24/193 (12.44%)	35/193 (18.13%)
GSN	159/193 (82.38%)	16/193 (8.29%)	18/193 (9.33%)
PAS	132/193 (68.39%)	32/193 (16.58%)	29/193 (15.03%)

5.2.3 South Africa: Black

The distribution of scores of males and females for South African Black individuals are very similar for the shape of the pubic body, subpubic concavity, subpubic angle, ischiopubic ramus, and greater sciatic notch. All five have clear distinctions between males and females. The preauricular sulcus and ventral arc have near identical distributions, with more males having more of the "typical male" morphology whilst females showing more of a spread of scores ranging mainly between "-2" and "0". The obturator foramen shows a large overlap between the sexes.

Table 5 shows the results for the South African Black population. The shape of the pubic body and subpubic angle have the highest accuracies being 86.83% and 85.85%. The ischiopubic ramus, subpubic concavity, greater sciatic notch, and preauricular sulcus range in accuracies between 71 – 80%. The last two traits, the ventral arc and obturator foramen, had the lowest success rate at 65.37% and 57.56% respectively.

5.2.4 South Africa: Coloured

For South African Coloured individuals, the subpubic concavity, ischiopubic ramus, pubic body, and the subpubic angle all have similar distribution of scores for males and females. Primarily, the sexes are observed to follow their associated morphology of "+2" and "+1" for males and "-2" and "-1" for females with a very small amount of overlap. The ventral arc has a similar pattern but the female scores are more shifted towards to the "male end of the spectrum" with the majority of scores being "-1". The greater sciatic notch shows the greatest amount of distinction between the sexes with the vast majority being scored on both extremes of the ordinal scale. The lack of any sulcus present is found in the vast majority of males however the presence in females is quite variables with scores ranging from "-2" to "+2". The obturator foramen has the largest overlap between the sexes present in all of the samples studied. A near equal distribution across the ordinal scale is found in males and with females showing morphology ranging from "-2" to "0".

The subpubic angle, pubic body shape, and subpubic concavity have the highest percentage accuracy of >80% with the greater sciatic notch and ischiopubic ramus having accuracies ranging between 76-79% (Table 6). Both the ventral arc and preauricular sulcus had correct classifications between 70-69%. The lowest correct classification was from the obturator foramen, for which had an accuracy <50% for diagnosing the correct sex.

Table 5: Accuracy rates for each morphoscopic trait for the South African Black sample.

South Africa: Black Sample			
	Correct	Unknown	Incorrect
VA	134/205 (65.37%)	33/205 (16.10%)	38/205 (18.54%)
SPC	159/205 (77.56%)	23/205 (11.22%)	23/205 (11.22%)
IPR	164/205 (80.00%)	14/205 (6.89%)	27/205 (13.17%)
PBS	178/205 (86.83%)	11/205 (5.37%)	16/205 (7.80%)
SPA	176/205 (85.85%)	13/205 (6.34%)	16/205 (7.80%)
OF	118/205 (57.56%)	44/205 (21.46%)	43/205 (20.98%)
GSN	155/205 (75.60%)	18/205 (8.78%)	32/205 (15.60%)
PAS	146/205 (71.22%)	31/205 (15.12%)	28/205 (13.66%)

Table 6: Accuracy rates for each morphoscopic trait for the South African Coloured sample.

South Africa: Coloured Sample			
	Correct	Unknown	Incorrect
VA	140/205 (68.29%)	27/205 (13.17%)	38/205 (18.54%)
SPC	166/205 (80.98%)	27/205 (13.17%)	12/205 (5.85%)
IPR	156/205 (76.10%	30/205 (14.63%)	19/205 (9.27%)
PBS	169/205 (82.44%)	23/205 (11.22%)	13/205 (6.34%)
SPA	170/205 (82.93%)	24/205 (11.71%)	11/205 (5.37%)
OF	95/205 (46.34%)	49/205 (23.90%)	61/205 (29.76%)
GSN	160/205 (78.05%)	27/205 (13.17%)	18/205 (8.78%)
PAS	133/205 (64.88%)	13/205 (6.34%)	59/205 (28.78%)

5.3 Intra/inter Observer Error

Intra/inter observer errors were analysed using the Kappa Cohen's test for repeatability following Landis & Koch's (1977) definitions. The samples held at Liverpool John Moores University (Poulton and St. Owens sample) and the University of Pretoria (South African White and Black sample) were used for the observer analysis. These samples were chosen because of the ease of getting individuals for the inter-observer error analysis. A random sample was generated using random sample generator built in SPSS (IBM Corp. 2012) to create a sub-sample for re-evaluation (N=113). The sample for re-evaluation included 20 specimens from the Poulton sample, 20 from the St. Owens sample and the remaining 73 were from the three South African Samples (30 South African White, and 43 South African Black).

The intra-observer analysis was conducted two weeks after the initial data were collected at both institutions.

All of the eight morphoscopic traits scored >0.65 in regards to the Kappa Cohen's value when testing for intra observer error (Table 7). Using Landis & Koch's (1977) definitions of the Kappa values, four traits (Ischiopubic Ramus, Pubic Body Shape, Obturator Foramen, and Greater Sciatic Notch) have substantial agreement between the intra observer scores. The remaining four traits all have values >0.8 meaning that there is almost perfect agreement between the intra observer scores.

Table 7: Intra observer results for the eight morphoscopic traits (N=113).

Intra Observer Error			
Trait	Kappa Value	p Value	Interpretation (Landis & Koch 1977)
Ventral Arc	0.823	<0.001	Almost Perfect Agreement
Subpubic Concavity	0.844	<0.001	Almost Perfect Agreement
Ischiopubic Ramus	0.724	<0.001	Substantial Agreement
Pubic Body Shape	0.741	<0.001	Substantial Agreement
Subpubic Angle	0.861	<0.001	Almost Perfect Agreement
Obturator Foramen	0.676	<0.001	Substantial Agreement
Greater Sciatic Notch	0.690	<0.001	Substantial Agreement
Preauricular Sulcus	0.883	<0.001	Almost Perfect Agreement

Two observers were chosen to conduct the inter-observer analysis. The first observer (Observer 1) holds a M.Sc. in Anatomy and has four years of experience assisting in forensic anthropology casework. Observer 1 has conducted similar research on morphoscopic sexing traits but from the human skull.

The second observer (Observer 2) holds a B.Sc. (Hons) in Forensic Anthropology and is currently a PhD Candidate in Biological Anthropology and has 4 years of experience of skeletal analysis on archaeological skeletal remains. The second observer has conducted research on cranial non-metric/discrete traits from the human skeleton.

Both conducted the inter-observer error with no assistance, only the descriptions of the scoring techniques described in Chapter 4. Observer 1 rescored 73 individuals from South Africa and Observer 2 rescored the 40 individuals from the Poulton and St. Owens samples.

When analysing the inter observer data, all traits showed a minimum of having a moderate agreement between the scoring for both observers (Table 8 and Table 9).

For Observer 1, three traits scored between 0.5 and 0.6 (Pubic Body Shape, Obturator Foramen, and Greater Sciatic Notch) showing that there was a moderate amount if agreement between the author and Observer 1. The Ventral Arc, Subpubic Concavity, Ischiopubic Concavity and Subpubic Angle all had a substantial agreement between scorers. The only trait to remain in almost perfect agreement was the Preauricular Sulcus with a Kappa value of 0.861 ($p<0.001$).

For the second observer, three of the eight traits resulted in a Kappa Value ranging between 0.5 – 0.6 (Obturator Foramen, Pubic Body Shape, and Ischiopubic Ramus) stating a moderate agreement between observers was apparent. The Ventral Arc, Subpubic Concavity, Greater Sciatic Notch and Subpubic Angle all had a substantial agreement between scorers. The Preauricular Sulcus resulted in a Kappa value of 0.848 meaning the trait achieved a better than accepted error rate between the author and Observer 2.

There is only a small difference between the two observer results being the scoring relating to the Greater Sciatic Notch and the Ischiopubic Ramus. The scores of Observer 2 were more in agreement in regards to the Greater Sciatic Notch whilst the Ischiopubic Ramus was less accurate, being interpreted as having only a “moderate agreement”. The opposite can be seen for Observer 1, where the Ischiopubic Ramus was scored more accurately than the Greater Sciatic Notch.

Table 8: Inter observer results for the eight morphoscopic traits (Observer 1) (N=73).

Inter Observer Error (Observer 1)			
Trait	Kappa Value	p Value	Interpretation (Landis & Koch 1977)
Ventral Arc	0.799	<0.001	Substantial Agreement
Subpubic Concavity	0.704	<0.001	Substantial Agreement
Ischiopubic Ramus	0.644	<0.001	Substantial Agreement
Pubic Body Shape	0.567	<0.001	Moderate Agreement
Subpubic Angle	0.620	<0.001	Substantial Agreement
Obturator Foramen	0.529	<0.001	Moderate Agreement
Greater Sciatic Notch	0.577	<0.001	Moderate Agreement
Preauricular Sulcus	0.861	<0.001	Almost Perfect Agreement

Table 9: Inter observer results for the eight morphoscopic traits (Observer 2) (N=40).

Inter Observer Error (Observer 2)			
Trait	Kappa Value	p Value	Interpretation (Landis & Koch 1977)
Ventral Arc	0.701	<0.001	Substantial Agreement
Subpubic Concavity	0.724	<0.001	Substantial Agreement
Ischiopubic Ramus	0.600	<0.001	Moderate Agreement
Pubic Body Shape	0.550	<0.001	Moderate Agreement
Subpubic Angle	0.705	<0.001	Substantial Agreement
Obturator Foramen	0.437	<0.001	Moderate Agreement
Greater Sciatic Notch	0.733	<0.001	Substantial Agreement
Preauricular Sulcus	0.848	<0.001	Almost Perfect Agreement

5.4 Sexing using Ordinal Logistic Regression

A more formal approach to analysing ordinal data for classifications would be to use an Ordinal Logistic Regression (OLR). For this round of analysis, only individuals with all eight morphoscopic traits present were considered for analysis. Table 10 highlights the demographic of the four samples used to create four population specific OLR equations.

Table 10: Breakdown of the four Known Sex samples (N=741).

Sample		N	Mean Age (Years)	Min-Max (Years)
Christ Church, Spitalfields				
	Male	69	57.29	22-91
	Female	69	58.43	23-89
	Total	*138*	*57.86*	*22-91*
South African White				
	Male	95	63.16	28-94
	Female	98	67.51	19-88
	Total	*193*	*65.37*	*19-94*
South African Black				
	Male	108	45.44	20-86
	Female	97	45.05	20-80
	Total	*205*	*45.26*	*20-86*
South African Coloured				
	Male	102	53.34	19-101
	Female	103	46.17	18-89
	Total	*205*	*49.74*	*18-101*

The *glm()* function using the binomial distribution was applied to create the ordinal logistic regression using all eight morphoscopic traits. A Wald Test from the package {aod}, function wald.test(), was conducted on the OLR which stated that all of the four models were significant, meaning that the variables used all contribute to the model (Table 11). Following this, *cv.glm()* function, from the package {boot}, was performed to cross validate the regression analysis. All results that are reported are after the cross-validation procedure.

Table 11: Results from the Wald Test for the four Known Sex samples.

Sample	x^2	df	p Value
Christ Church, Spitalfields	17.3	5	<0.01
South African White	16.6	5	<0.01
South African Black	21.5	5	<0.001
South African Coloured	24.6	5	<0.001

OLR Generated Equations

$$\begin{aligned} &Christ\ Church, Spitalfields\ \ Equation \\ &\quad = (0.962 * Ventral\ Arc\ Score) + (0.931 * Subpubic\ Concavity\ Score) \\ &\quad + (0.986 * Ischiopubic\ Ramus\ Score) + (3.255 * Pubic\ Body\ Shape\ Score) \\ &\quad + (2.464 * Subpubic\ Angle\ Score) + (1.185 * Obturator\ Foramen\ Score) \\ &\quad + (1.130 * Greater\ Sciatic\ Notch\ Score) \\ &\quad + (1.163 * Preauricular\ Sulcus\ Score) + \mathbf{0.000} \end{aligned}$$

$$\begin{aligned} &SA{:}\ White\ Equation \\ &\quad = (0.980 * Ventral\ Arc\ Score) + (1.147 * Subpubic\ Concavity\ Score) \\ &\quad + (1.449 * Ischiopubic\ Ramus\ Score) + (2.141 * Pubic\ Body\ Shape\ Score) \\ &\quad + (0.998 * Subpubic\ Angle\ Score) + (1.539 * Obturator\ Foramen\ Score) \\ &\quad + (1.547 * Greater\ Sciatic\ Notch\ Score) \\ &\quad + (1.819 * Preauricular\ Sulcus\ Score) + \mathbf{0.000} \end{aligned}$$

$$\begin{aligned} &SA{:}\ Black\ Equation \\ &\quad = (1.702 * Ventral\ Arc\ Score) + (0.750 * Subpubic\ Concavity\ Score) \\ &\quad + (1.378 * Ischiopubic\ Ramus\ Score) + (1.453 * Pubic\ Body\ Shape\ Score) \\ &\quad + (2.211 * Subpubic\ Angle\ Score) + (1.272 * Obturator\ Foramen\ Score) \\ &\quad + (0.993 * Greater\ Sciatic\ Notch\ Score) \\ &\quad + (3.014 * Preauricular\ Sulcus\ Score) + \mathbf{0.000} \end{aligned}$$

$$\begin{aligned} &SA{:}\ Coloured\ Equation \\ &\quad = (0.828 * Ventral\ Arc\ Score) + (1.468 * Subpubic\ Concavity\ Score) \\ &\quad + (1.016 * Ischiopubic\ Ramus\ Score) + (4.181 * Pubic\ Body\ Shape\ Score) \\ &\quad + (1.134 * Subpubic\ Angle\ Score) + (0.659 * Obturator\ Foramen\ Score) \\ &\quad + (1.713 * Greater\ Sciatic\ Notch\ Score) \\ &\quad + (1.482 * Preauricular\ Sulcus\ Score) + \mathbf{0.000} \end{aligned}$$

Table 12: Classification results for the four Known Sex samples with their specific OLR analysis.

Sample	N	Correct Female	Correct Male	Overall Accuracy
Christ Church, Spitalfields	138	68/69 (98.55%)	65/69 (94.20%)	96.38%
South African White	193	93/98 (94.90%)	91/95 (95.79%)	95.34%
South African Black	205	85/97 (87.63%)	100/108 (92.59%)	90.24%
South African Coloured	205	95/103 (92.23%)	98/102 (96.08%)	94.15%

From the OLR analysis, all samples achieved an overall accuracy that was greater than 90% (Table 12). The highest accuracy found was the British sample from Christ Church, Spitalfields with an accuracy of 96.38%. A total of five individuals were misclassified being four males and one female meaning that females were scored with a higher accuracy than males (98.55% and 94.20% respectively). South African White's gained an overall accuracy of 95.34% from their population specific OLR equation. Nine individuals were misclassified, five females incorrectly sexed as male and four males incorrectly sexed as females. Males had a slightly higher percentage accuracy of 95.79% than females (94.90%). The South African Black OLR analysis resulted in the lowest overall accuracy from the four samples with an accuracy of 90.24%. Females were misclassified at a higher rate with a total of 12 misclassifications, which resulted in an accuracy of 87.63%, whilst males had a higher success rate of 92.59% with only eight being misclassified as female. An overall accuracy of 94.15% was achieved when analysing the South African Coloured individuals. Eight females and four males were misclassified (92.23% and 96.08% respectively), totalling in 12 individuals overall.

After the analysis, the 46 individuals were cross compared to original records to see if any of the misclassifications were in fact false negatives. No clerical errors were made hence the 46 individuals remained as true misclassifications.

Figures 26 – 29 are ROC curves produced from each of the four individual OLR equations. As we can see, all four show that they have an extremely good discriminatory power. Slight variations can be seen between the four curves, the most extreme between the Christ Church, Spitalfields (Figure 26) and South African Coloured (Figure 29) samples. The Christ Church, Spitalfields OLR produced a near perfect model from the onset whilst the South African Coloured OLR produced more 'steps' earlier. With this being said, only a 2.23% difference in accuracies (cross validated) was seen (Table 12).

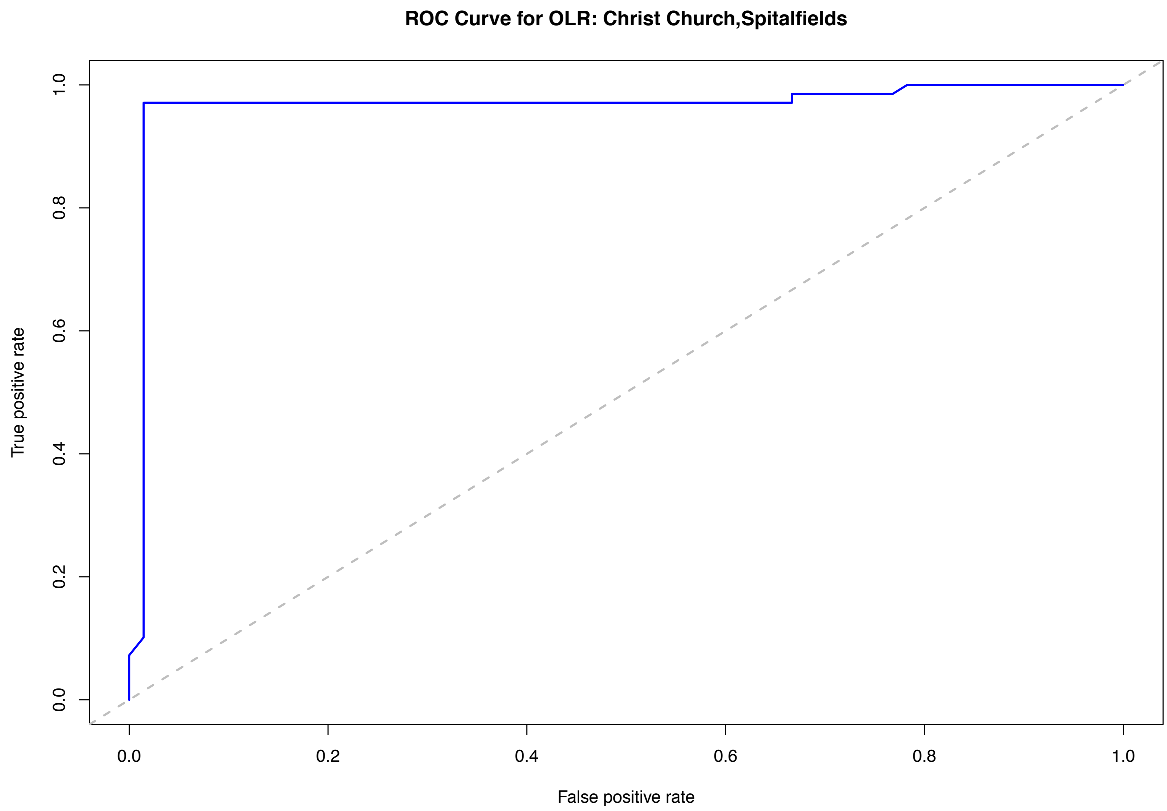


Figure 31: ROC curve for the Christ Church, Spitalfields OLR equation.

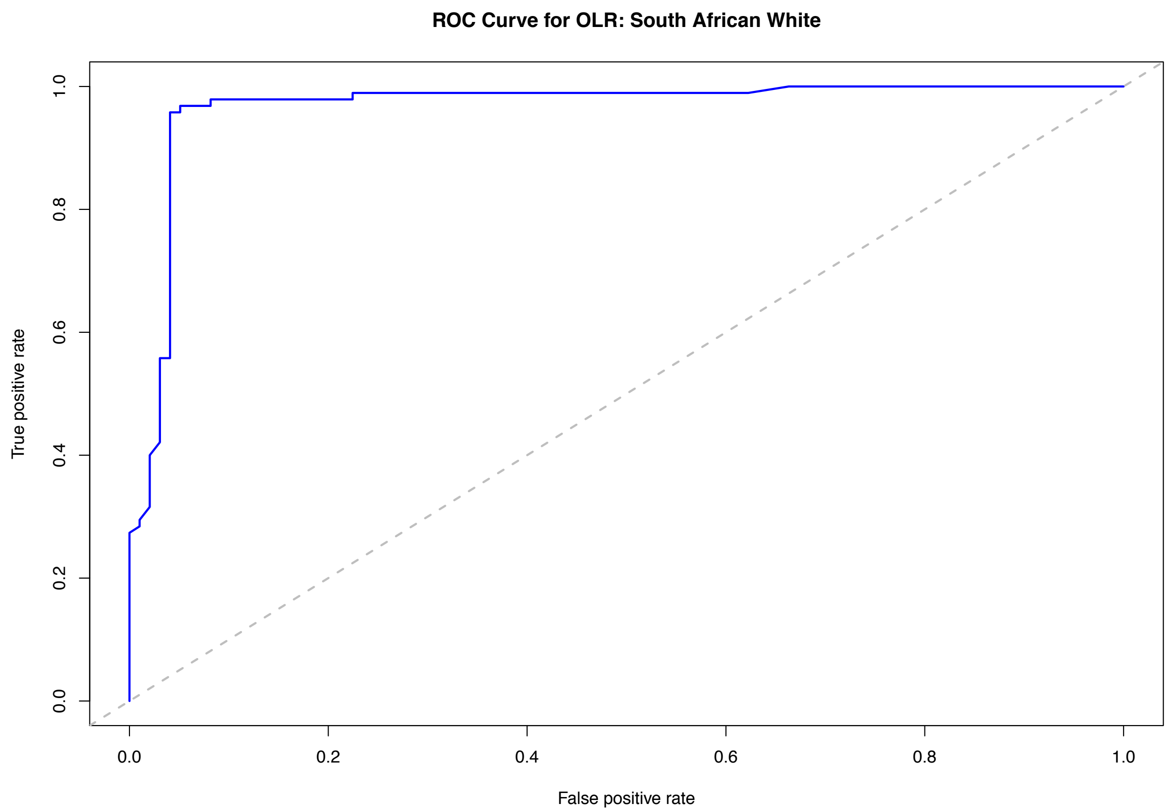


Figure 32: ROC curve for the South African White OLR equation.

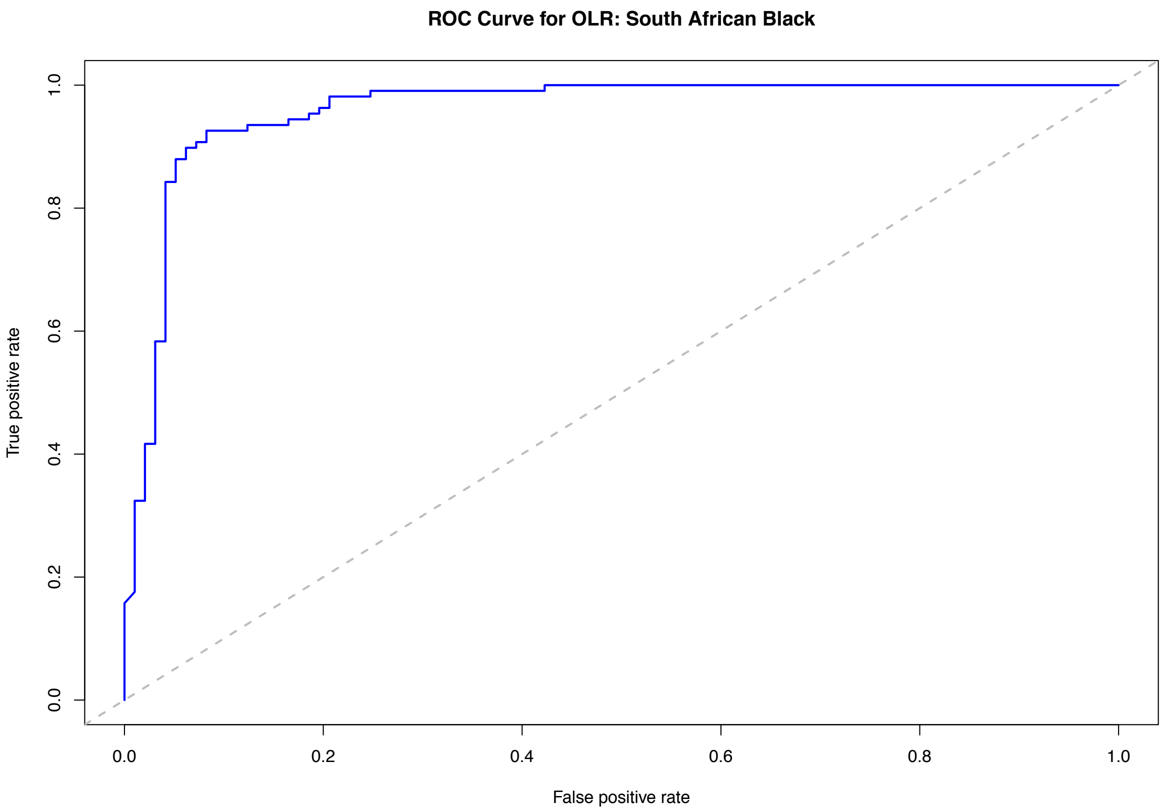


Figure 33: ROC curve for the South African Black OLR equation.

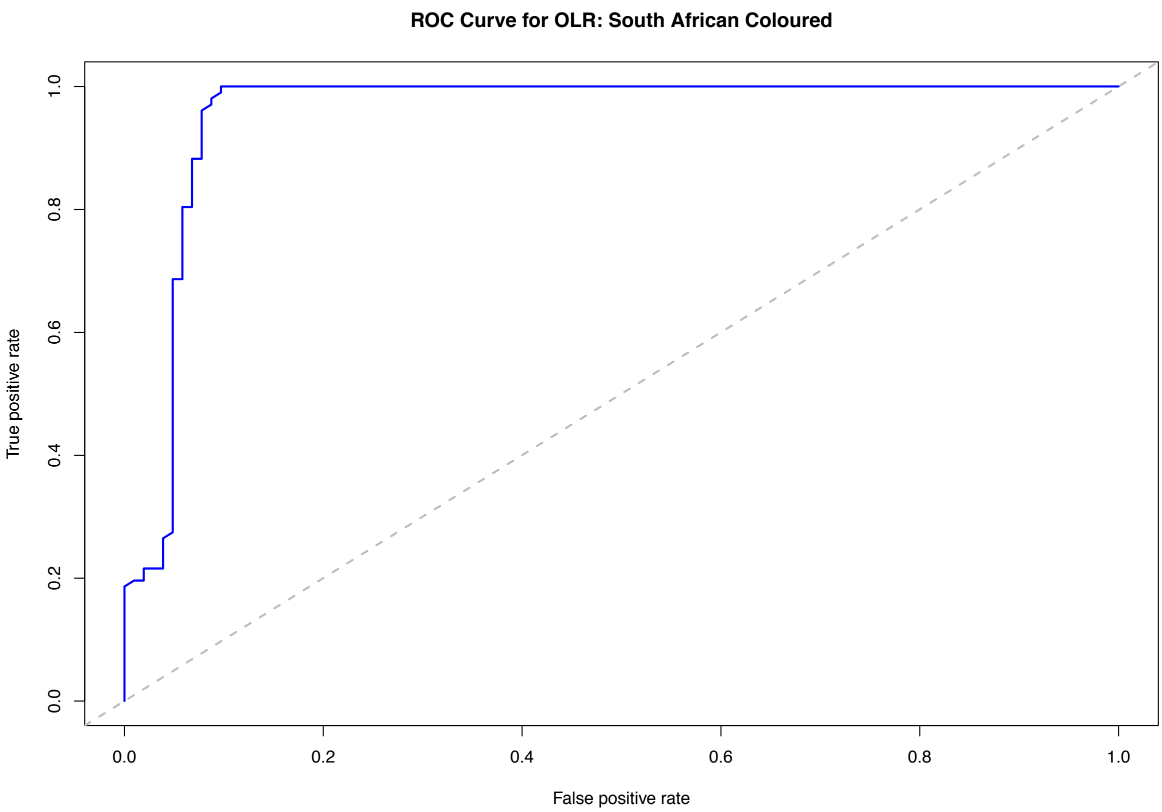


Figure 34: ROC curve for the South African Coloured OLR equation.

5.5 Cross comparing samples

The argument for population specific equations are given in several articles, however, many of these articles do not apply their population specific equation onto other populations such as Steyn & Patriquin (2009) (See Chapters 2 and 8 for further discussion).

The following section applies each of the four OLR equations to the other three samples. This will allow any differences in percentage accuracy to be noted between the population specific OLR equation and the three others for each sample. Tables 13 – 16 show the classification results for the cross comparisons.

When applying the three South African OLR equations onto the 138 individuals from Christ Church, Spitalfields sample, the same percentage accuracy was achieved from using the South African White and Coloured OLR equations (96.38%). However, when using the South African Coloured equation different individuals were misclassified. An additional female and one less male was misclassified meaning that females had a lower accuracy at 97.10% and males had a higher accuracy at 95.65%. The South African White equation resulted in the same individuals being misclassified that were found from the original Christ Church, Spitalfields OLR equation. A slight decrease in overall accuracy was noted when the South African Black OLR equation was applied to the British sample. Two females and four males were misclassified (98.55% and 94.20% accuracies respectively) which resulting in the overall accuracy of 95.65% (Table 13).

Table 13: Classification results for the Christ Church, Spitalfields sample using each of the four OLR equations (N=138).

OLR Equation	Correct Female	Correct Male	Overall Accuracy
Christ Church, Spitalfields	68/69 (98.55%)	65/69 (94.20%)	96.38%
South African White	68/69 (98.55%)	65/69 (94.20%)	96.38%
South African Black	67/69 (97.10%)	65/69 (94.20%)	95.65%
South African Coloured	67/69 (97.10%)	66/69 (95.65%)	96.38%

Using the South African White sample (N=193) as a base to apply the remaining two South African equations and the 18th/19th Century British equation, it was found that accuracy did not differ greatly overall. The same individuals were misclassified after applying the Christ Church, Spitalfields OLR equation resulting in the same percentage accuracy (95.34%). The South African Coloured equation resulted in a slightly lower percentage accuracy with one additional female being misclassified when compared to the results from the South African White and Christ Church, Spitalfields OLR equations. Females had a correct classification of 93.88% and males remained the same with a correct classification of 95.79% creating an overall accuracy of 94.82%. The South African Black equation resulted in an accuracy greater than the population specific OLR equation. One additional male was correctly classified which resulted in a 0.51% increase in the overall accuracy (Table 14).

Table 14: Classification results for the South African White sample using each of the four OLR equations (N=193).

OLR Equation	Correct Female	Correct Male	Overall Accuracy
Christ Church, Spitalfields	93/98 (94.90%)	91/95 (95.79%)	95.34%
South African White	93/98 (94.90%)	91/95 (95.79%)	95.34%
South African Black	93/98 (94.90%)	92/95 (96.84%)	95.85%
South African Coloured	92/98 (93.88%)	91/95 (95.79%)	94.82%

The South African Black OLR equation resulted in the lowest percentage accuracy when applied onto its own sample. When applying the other two South African equations, the same accuracy was not maintained. For both South African equations, two additional males were correctly identified increasing the male accuracy from 92.59% to 94.44%. However, five additional females were misclassified which lowered their accuracy from 87.63% down to 82.47%. This dropped the overall accuracy for the South African Black sample to 88.78%. The Christ Church, Spitalfields OLR equation proved to work better than the South African White and Coloured equations by correctly identifying one additional female. This increased the female accuracy to 83.51%, which in turn increased the overall accuracy to 89.27%. However, these three equations still did not achieve the same accuracy as the original set of results created by the South African Black OLR equation.

Table 15: Classification results for the South African Black sample using each of the four OLR equations (N=205).

OLR Equation	Correct Female	Correct Male	Overall Accuracy
Christ Church, Spitalfields	81/97 (83.51%)	102/108 (94.44%)	89.27%
South African White	80/97 (82.47%)	102/108 (94.44%)	88.78%
South African Black	85/97 (87.63%)	100/108 (92.59%)	90.24%
South African Coloured	80/97 (82.47%)	102/108 (94.44%)	88.78%

The last comparison was using the South African Coloured sample and applying the South African White and Black, and Christ Church, Spitalfields OLR equations to it. The original overall accuracy was 94.15% which sadly was not recreated by the other three equations. The South African White and Christ Church, Spitalfields equation resulted in the same overall accuracy of 93.17%, however a different amount of males and females were classified correctly. Three females were misclassified when using the Christ Church, Spitalfields equation lowering its accuracy to 89.32% when compared to the South African Coloured equation. From this, a further two females were incorrectly sexed when applying the South African White equation, lowering the female percentage accuracy even further to 87.37% (Table 16). On the other hand, male classifications increased when using the South African White and Christ Church, Spitalfields equations. One additional male was when using the Christ Church, Spitalfields equation (97.06% male accuracy) and three males were additionally correctly identified using the South African White equation (99.02% male accuracy). The South African Black equation achieved the lowest accuracy of all four equations that were applied onto this sample. It did classify the same amount of males as the South African White equation with an accuracy of 99.02%, however 19 females were misclassified resulting in the lowest female accuracy for this sample (81.55%). This lowered the overall accuracy to 90.24%.

Table 16: Classification results for the South African Coloured sample using each of the four OLR equations (N=205).

OLR Equation	Correct Female	Correct Male	Overall Accuracy
Christ Church, Spitalfields	92/103 (89.32%)	99/102 (97.06%)	93.17%
South African White	90/103 (87.37%)	101/102 (99.02%)	93.17%
South African Black	84/103 (81.55%)	101/102 (99.02%)	90.24%
South African Coloured	95/103 (92.23%)	98/102 (96.08%)	94.15%

5.6 Attempting to find a Universal Equation

Following the cross comparison of the population specific OLR equations onto the four known sex samples, some slight differences can be found in the percentage accuracies in sexing. Applying the correct equation to an individual or individuals would require knowledge of their ethnicity or geographic origin. This becomes an issue when this is unknown, and no estimation of ancestry can be made, especially if certain skeletal elements are missing (e.g. the skull). This section will attempt to create a universal OLR equation to try and tackle the issue of population specificity but first creating (and cross validating) a South African OLR equation and applying it onto the 18th/19th Century British population. This will hopefully create a standard equation for South Africa that can be used on three of the four main population groups for that country and also test its possible use on a geographically and temporally different sample. From there, an overall equation will be created (and cross validated) by combining all four known samples to create a potentially universal OLR equation which will be dubbed "Summary Sex".

The glm() function using the binomial distribution was applied to create the ordinal logistic regression using all eight morphoscopic traits. A Wald Test from the package {aod}, function wald.test(), was conducted on the OLR which stated that all of the four models were significant, meaning that the variables used all contribute to the model (Table 17). Following this, cv.glm() function, from the package {boot}, was performed to cross validate the regression analysis. All results that are reported are after the cross-validation procedure.

Table 17: Results from the Wald Test for the clustered samples.

Sample	x^2	df	p Value
South African	61.2	8	<0.001
Summary Sex	87.4	8	<0.001

OLR Generated Equations

$$\begin{aligned} \textit{South African Equation} &= (0.985 * \textit{Ventral Arc Score}) + (1.053 * \textit{Subpubic Concavity Score}) \\ &+ (1.171 * \textit{Ischiopubic Ramus Score}) + (1.867 * \textit{Pubic Body Shape Score}) \\ &+ (1.551 * \textit{Subpubic Angle Score}) + (1.075 * \textit{Obturator Foramen Score}) \\ &+ (1.142 * \textit{Greater Sciatic Notch Score}) \\ &+ (2.137 * \textit{Preauricular Sulcus Score}) + \mathbf{0.000} \end{aligned}$$

$$\begin{aligned} \textit{Summary Sex Equation} &= (0.942 * \textit{Ventral Arc Score}) + (1.018 * \textit{Subpubic Concavity Score}) \\ &+ (1.196 * \textit{Ischiopubic Ramus Score}) + (2.127 * \textit{Pubic Body Shape Score}) \\ &+ (1.521 * \textit{Subpubic Angle Score}) + (1.149 * \textit{Obturator Foramen Score}) \\ &+ (1.425 * \textit{Greater Sciatic Notch Score}) \\ &+ (1.755 * \textit{Preauricular Sulcus Score}) + \mathbf{0.000} \end{aligned}$$

Table 18: Classification results for the two clustered samples from their specific OLR equations.

Sample	N	Correct Female	Correct Male	Overall Accuracy
South African	603	269/298 (90.28%)	289/305 (94.75%)	92.54%
Summary Sex	741	334/367 (91.01%)	355/374 (94.99%)	92.98%

After combining the South African White, Black, and Coloured samples together to create a single South African sample and applying the new population specific OLR equation on the dataset, an overall accuracy of 92.54% was achieved. A total of 29 females were misclassified which resulted in a percentage accuracy of 90.28% and 16 males were incorrectly sexed creating a male accuracy of 94.75%. The same female individuals were misclassified in this analysis and in the original OLR results (Table 18 and Table 12) with an additional four females, these being three South African Black females and one South African Coloured female. For males, the same individuals were misclassified as using the South African OLR equation and from the three South African population specific OLR equations.

When applying the South African OLR equation to the Christ Church, Spitalfields sample a higher overall percentage accuracy (97.10%) was achieved when compared against the samples specific OLR equation. Similar to the original result, only one female was misclassified creating a female accuracy of 98.55% whilst only three males were misclassified unlike the four that were in from the original analysis. This resulted in the males having a 95.65% accuracy when applying the South African OLR equation.

A total of 52 individuals were misclassified when the Summary Sex OLR equation was applied to the four known sex samples (N=741). Females were correctly sexed with an accuracy of 91.01% meaning a total of 33 were misclassified (Table 18). The same females were misclassified that were found in the combined South African OLR analysis and the single female from the Christ Church, Spitalfields sample, with the addition of three more South African Black females. For males, a total of 19 were misclassified. These males were the same males misclassified from the combined South African OLR analysis where 16 males were misclassified from South Africa and three were misclassified from the Christ Church, Spitalfields sample.

Similarly, to the sample specific OLR analyses, ROC curves were created for the South African and Summary Sex samples. Both ROC curves are practically identical and indicate that they both have a great discriminatory power (Figure 30 and Figure 31).

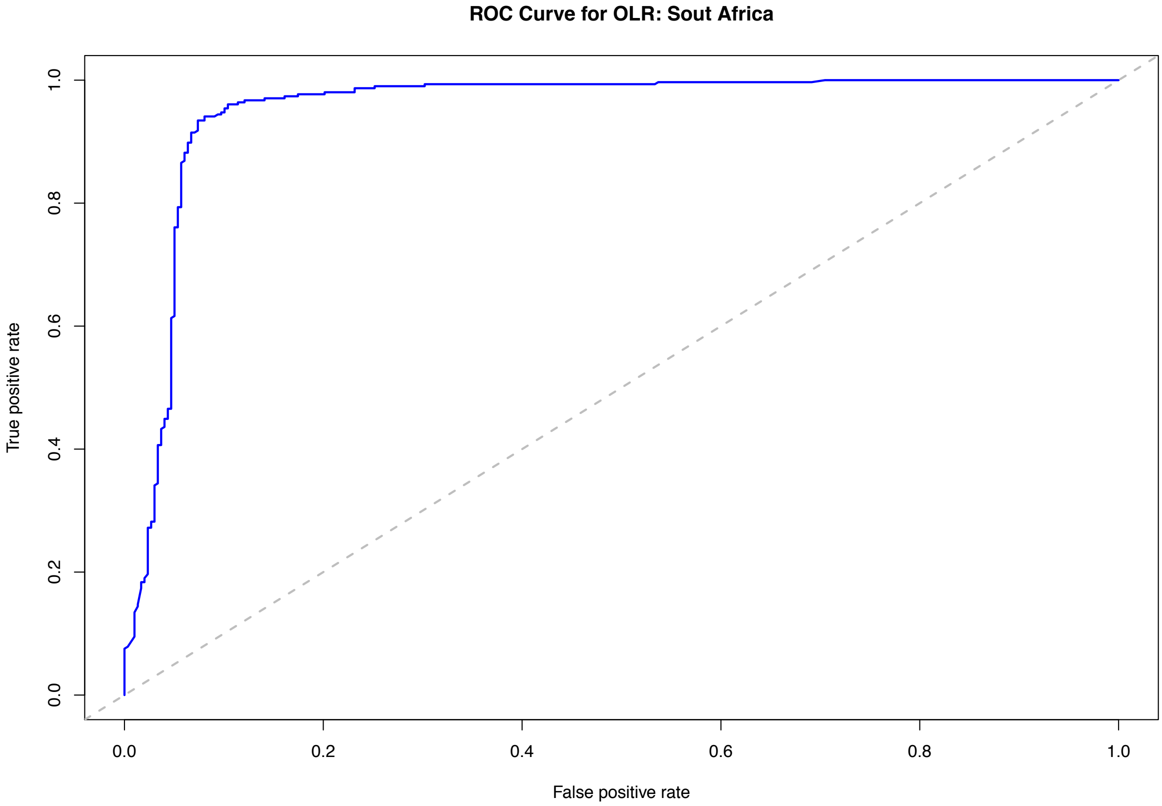


Figure 35: ROC curve for the South African OLR equation.

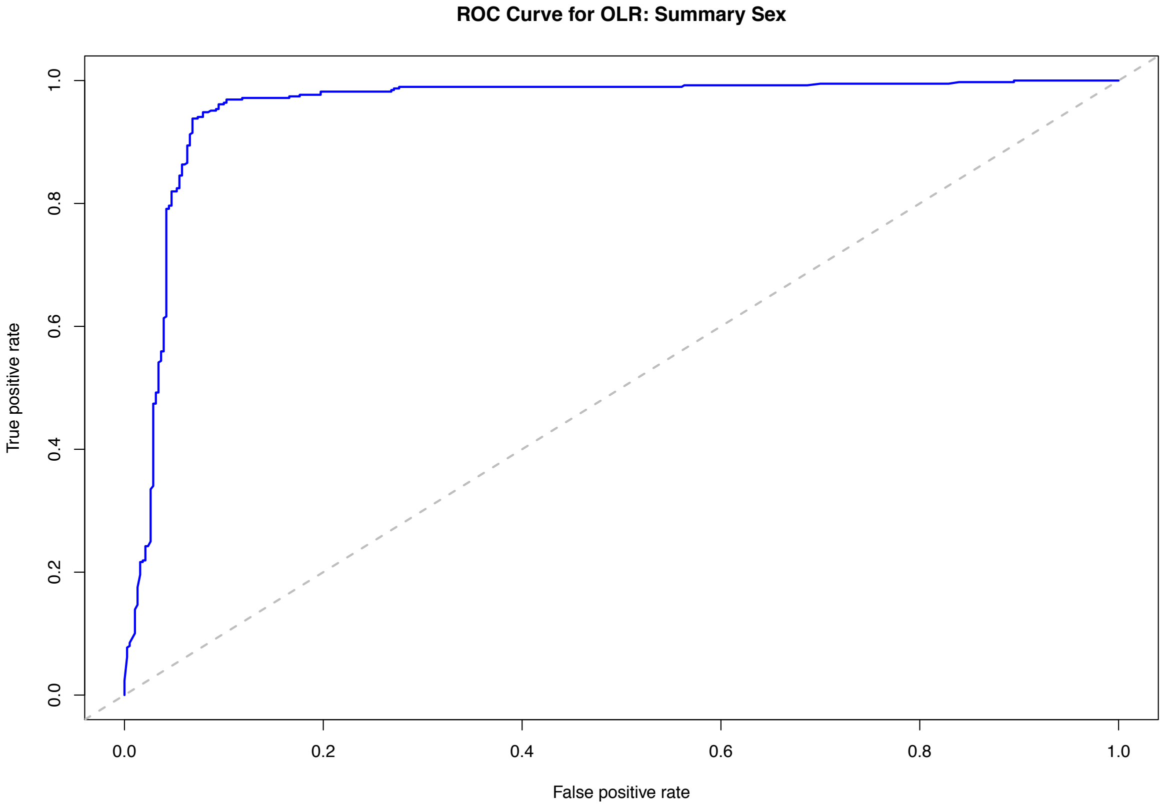


Figure 36: ROC curve for the Summary Sex OLR equation.

5.7 Summary

The results presented in this chapter show that multivariate analyses, that combine several sexing morphoscopic traits, can be used to determine sex at a much higher accuracy than only using a single morphoscopic trait. With this in mind, it is still good to note that the trait with the greatest accuracy was the shape of the pubic body (Christ Church, Spitalfields: 90.42%; South African White: 88.60%; South African Black: 86.83%) (Tables 4, 5, and 6) and the subpubic angle having the greatest accuracy for the South African Coloured sample (82.93%) (Table 7). Accuracies for the OLR equations ranged from 90.24% to 96.38%.

They later show that the four OLR equations created show signs of population specifity. This was determined by applying each equation onto the remaining three data sets and to compare if accuracies differ, which they inevitably did.

The final hypothesis was to create and assess the use of a universal equation to see if it would have solved the issue of population specifity seen with the four OLR equations. Two equations were created, the South African equation (which contained all three South African samples) and the Summary Sex equation (which contained all four samples). Accuracies were lower than the specific OLR results, but higher than when they were individually applied onto samples that were not their own.

6. Results II

6.1 Introduction

In this chapter, we investigate the use of the OLR equations, created and assessed in the previous chapter (Chapter 5) on archaeological samples and if they are applicable for archaeological skeletal collections. Because sex is not known in these samples, we are basing the results from these analyses from predetermined sex. For the Andaman and Chumash samples, sex was based from records held at the Natural History Museum, London (UK). For the Poulton and St. Owens samples, sex was based from the individuals being seriated for sex using cranial morphoscopic assessment (Walker 2008), femoral head diameter (Stewart 1979) and general size and robusticity of the pelvis (Acsádi & Nemeskéri 1970, European meeting of Anthropologists 1980)

Similarly, to the previous chapter, accuracy of each trait will be observed for each sample and from there each OLR equation will be applied to the samples. This chapter has one testable hypothesis:

H_1: The results from the OLR will provide a greater accuracy than that of a single morphoscopic trait for sex estimation.

H_0: The results from the OLR will not differ from the sexing accuracy from a single morphoscopic trait for sex estimation.

6.2 Trait Frequencies for Archaeological Samples

A seriation for sex was used for the Poulton and St. Owens samples. This meant assessing overall pelvic morphology and robustness following the descriptions from European meeting of Anthropologists (1980), skull morphology (Walker 2008), and the diameter of the femoral head (Stewart 1979). The latter was used primarily because these two samples shown signs of a high degree of fragmentation. This issue of fragmentation will be fully explored in Chapter 7. Individuals were not assigned sex until all individuals in the sample had been seriated and the spread of sexual variability was assessed.

For the Chumash sample, estimated sex was based in information made available from the museum's database and from work of previous researchers. There were some issues surrounding this, for example some individuals in the database were described to be 12 years therefore they were not assigned a sex even though the physical remains were that of a fully mature adult.

Walker's (2008) skull morphology and diameter of the femoral head (Stewart 1979) was used to establish sex.

Sex for the Andaman sample was based on information from previous researchers and the database records. Some of the individuals in the sample had death certificates and written notation of who the individual was, detailing their name and sex.

To give an overview of trait frequencies for each archaeological sample, all individuals (complete and incomplete pelves) were considered for analysis from each of the four archaeological samples. Each sample was analysed separately and percentage accuracies were calculated for each trait as a single sex assessor. If a minus score was given, the individual was deemed female and if a positive score was given then it would be male.

6.2.1 Poulton UK

The female individuals from the Poulton sample seem to show a small amount of spread throughout the five possible scores with the majority showing the hyper-feminine morphology "-2". Males on the other hand have a very different story. A much larger spread is seen with the majority of males either not showing any ventral arc (+2) or showing a similar morphology of an arc which is indicative of females with a score of "-1". The shape of the pubic body, subpubic concavity and the greater sciatic notch all have a similar spread of scores for males and females. The majority of males show the hyper-masculine form of "+2" with a lower proportion with a score of "+1" and an even lower amount with an ambiguous score "0". Females show that the majority have the hyper-feminine form with only a slightly lower proportion with a "-1" morphology. Both males and females were observed to show extreme morphologies for the ischiopubic ramus, more so for females than males. However, males were slightly split between both male scores of "+2" and "+1". There were slight differences evident between males and females when reviewing the obturator foramen. The morphologies for males was more prominent than for females. Females show a much more even distribution from the "-2" to the "0" many of the males stay within the masculine scores of "+2" and "+1. There is a clear distinction between males and females in regards to the preauricular sulcus with the vast majority of both sexes showing their expected extreme forms.

Table 19 shows the results for the medieval Poulton collection. The four traits that had the highest accuracy were the ischiopubic ramus, shape of the pubic body, the subpubic angle, and the preauricular sulcus. All of these four traits had accuracies greater than 80%. The subpubic concavity and obturator foramen achieved accuracies of 79.41% and 72.34% respectively. The remaining two traits, ventral arc and the greater sciatic notch showed the lowest accuracies ranging between 60-65%.

6.2.2 St. Owens, Gloucester UK

The subpubic concavity, ischiopubic ramus, and pubic body shape all show a similar spread of scores for males and females. Males tend to show a more extreme morphology with a "+2" whilst females show more of a slight spread between the female scores of "-2" and "-1". Nearly all males have no preauricular sulcus whilst females show a range of morphology. This is similar to the morphology of the ventral arc but to a lesser extent with males showing more variability in the presence of the arc. The obturator foramen and the greater sciatic notch show similar distribution of scores for males and females with males revealing a much larger amount of variability then their female counterparts. The subpubic angle shows a clear distinction between males and females and is the opposite of what was seen for the preauricular sulcus. Nearly all females were observed with the hyper-feminine score of "-2" and males were observed with a nearly equal distribution of scores ranging from "+2" to "0". Table 20 shows the results for the medieval St. Owens collection. The two traits with the highest accuracy rates were the shape of the pubic body and the subpubic angle (83.33% and 82.98%). Following these, the three Phenice traits achieved accuracies ~80% followed closely by the greater sciatic notch with an accuracy of 78.79%. The last two traits had the lowest overall accuracies; preauricular sulcus (74.74%) and obturator foramen (68.29%).

Table 19: Accuracy rates for each morphoscopic trait for the Poulton sample.

	Poulton, (UK)		
	Correct	Unknown	Incorrect
VA	51/80 (63.75%)	3/80 (3.75%)	26/80 (32.5%)
SPC	54/68 (79.41%)	10/68 (14.71%)	4/68 (5.88%)
IPR	60/74 (81.10%)	10/74 (13.51%)	4/74 (5.41%)
PBS	66/82 (80.49%)	9/82 (10.98%)	7/82 (8.54%)
SPA	64/76 (84.21%)	5/76 (6.58%)	7/76 (9.21%)
OF	34/47 (72.34%)	7/47 (14.89%)	5/47 (10.63%)
GSN	83/129 (64.34%)	31/129 (24.03%)	15/129 (11.63%)
PAS	107/126 (84.92%)	0/126 (0%)	19/126 (15.08%)

Table 20: Accuracy rates for each morphoscopic trait for the St. Owens sample.

	St. Owens, Gloucester (UK)		
	Correct	Unknown	Incorrect
VA	38/47 (80.85%)	4/47 (8.15%)	5/47 (10.64%)
SPC	38/48 (79.17%)	6/48 (12.50%)	4/48 (8.33%)
IPR	39/48 (81.25%)	6/48 (12.50%)	3/48 (6.25%)
PBS	40/48 (83.33%)	5/48 (10.42%)	3/48 (6.25%)
SPA	39/47 (82.98%)	6/47 (12.77%)	2/47 (4.26%)
OF	28/41 (68.29%)	9/41 (21.95%)	4/41 (9.76%)
GSN	78/99 (78.79%)	15/99 (15.12%)	6/99 (6.06%)
PAS	74/99 (74.74%)	12/99 (12.12%)	13/99 (13.13%)

6.2.3 Chumash USA

The three Phenice traits all had similar male and female score distributions with a large proportion of males scoring more the along the extreme morphology of "+2" than any other. Females, on the other hand, had a greater distribution throughout the scores from "-2" to "0". The obturator foramen and preauricular sulcus had the greatest overlap between males and females, having distribution scores from "+2" and "-2". This is the first sample where a "-2" morphology was observed on a male individual. The shape of the pubic body and the subpubic angle both showed that males and females were found to have more of their extreme morphologies present. The greater sciatic notch also presented a good distinction between the sexes, however males were found to have a greater spread from "+2" to "0" whilst females predominantly remained with the negative scores.

The results for the Chumash population can be found in Table 21. In general, the accuracies were much lower than those seen previously. The trait that correctly estimated sex was the subpubic angle (86.49%) followed closely by the ischiopubic angle (85.71%). The ventral arc, subpubic concavity and shape of the pubic body gained accuracies between 70-80%. From these, the greater sciatic notch and obturator foramen had the sixth and seventh highest percentage accuracies (68.42% and 62.16% respectively). The preauricular sulcus had the lowest accuracy, correctly classifying just over 50% of individuals in the population.

6.2.4 Andaman Islands

The results from the trait frequencies can be seen in Table 22. Only the shape of the pubic body achieved an accuracy of greater than 90% with only one unknown and incorrect classification. The subpubic angle and the subpubic concavity gained accuracies just lower than 90% (88.89% each) with the ventral arc getting the fourth highest accuracy of 81.48%. The remaining four traits had accuracies ranging between 70-80% apart from the preauricular sulcus who achieved a correct classification rate of 66.67%.

The ventral arc, pubic body shape, subpubic angle, and the preauricular sulcus all show the same distribution for males scores. The majority of all males having a “+2” morphology whilst females showed a greater distribution of scores which ranged predominantly between “-2” and “0”. For the subpubic concavity males and females were generally spread between their scores (positive for males and negative for females) with one male observed with a feminine morphology. Males were shown to have the widest variability in greater sciatic notch scores ranging between the “+2” to “-1” which slightly overlapped with the females which typically ranged in the negative scores. Females had a more equal distribution of negative scores for the obturator foramen whilst the majority of males were shown to have a “+1” morphology. The male ischiopubic ramus had a greater amount of variation than females scoring from (with the majority) hyper-masculine “+2” to ambiguous “0” whilst females were more equally distributed between “-1” and “-2”.

Table 21: Accuracy rates for each morphoscopic trait for the Chumash sample.

	Chumash (USA)		
	Correct	Unknown	Incorrect
VA	27/35 (77.14%)	2/35 (5.72%)	6/35 (17.14%)
SPC	28/37 (75.68%)	3/37 (8.11%)	6/37 (16.21%)
IPR	30/35 (85.71%)	1/35 (2.86%)	4/35 (11.43%)
PBS	29/37 (78.38%)	1/37 (2.70%)	7/37 (18.92%)
SPA	32/37 (86.49%)	0/37 (0.00%)	5/37 (13.51%)
OF	23/37 (62.16%)	7/37 (18.92%)	7/37 (18.92%)
GSN	26/38 (68.42%)	8/38 (21.05%)	4/38 (10.53%)
PAS	20/38 (52.63%)	7/38 (18.42%)	11/38 (28.95%)

.

Table 22: Accuracy rates for each morphoscopic trait for the Andamanese sample.

	Andaman (India)		
	Correct	Unknown	Incorrect
VA	22/27 (81.48%)	3/27 (11.11%)	2/27 (7.41%)
SPC	24/27 (88.89%)	2/27 (7.41%)	1/27 (3.70%)
IPR	21/27 (77.78%)	5/27 (18.52%)	1/27 (3.70%)
PBS	25/27 (92.60%)	1/27 (3.70%)	1/27 (3.70%)
SPA	24/27 (88.89%)	1/27 (3.70%)	2/27 (7.41%)
OF	21/27 (77.78%)	3/27 (11.11%)	3/27 (11.11%)
GSN	20/27 (74.07%)	3/27 (11.11%)	4/27 (14.82%)
PAS	18/27 (66.67%)	6/27 (22.22%)	3/27 (11.11%)

6.3 Applying Known onto the Unknown

The use of population specific equations is fundamental to forensic investigations however, knowing how applicable they are for archaeological material and/ or the unknown is still undetermined. The following chapter shows the results from applying the four population specific and two clustered population equations onto four unknown archaeological samples. Percentage accuracy will be changed to percentage agreement as sex for the archaeological samples have all been estimated or have circumstantial evidence for that sex. This chapter will only deal with individuals that had all eight morphoscopic traits present. Individuals that were fragmented and missing variables will be discussed and analysed in Chapter 7. Table 23 gives the breakdown of sexes for these four samples.

Table 23: Breakdown of complete male and female skeletons in the four archaeological samples.

Sample	N	Females	Males
Poulton	47	24	23
St. Owens	30	18	12
Chumash	34	16	18
Andaman	27	13	14

From the Poulton sample, 47 individuals had all eight morphoscopic sexing traits present. This breaks down as 24 adult females and 23 adult males (Table 23). These 47 individuals had each of the OLR equations from Chapter 5 applied to them and a percentage agreement was calculated. When applying the four OLR equations, the overall agreement to the sex based from the seriation was 93.62%. One female and two males were sexed differently (95.83% and 91.30%, respectively). All three misclassifications were the same individuals in all four analyses. The same percentage agreement was also achieved when applying the combined equation "Summary Sex" with the same individuals being sexed differently. One additional male was sexed correctly when using the South African equation increasing the overall percentage agreement to 95.74% and increasing the male accuracy to 95.65% (Table 24).

Table 24: Classification results for the Poulton sample using each of the six OLR equations (N=47).

OLR Equation	Correct Females	Correct Males	Overall Agreement (%)
Christ Church, Spitalfields	23/24 (95.83%)	21/23 (91.30%)	93.62%
South African White	23/24 (95.83%)	21/23 (91.30%)	93.62%
South African Black	23/24 (95.83%)	21/23 (91.30%)	93.62%
South African Coloured	23/24 (95.83%)	21/23 (91.30%)	93.62%
South African	23/24 (95.83%)	22/23 (95.65%)	95.74%
Summary Sex	23/24 (95.83%)	21/23 (91.30%)	93.62%

A total of 30 individuals had a complete set of eight morphoscopic traits from the St. Owens archaeological sample. From the seriation analysis, it was found that there were 18 females and 12 males in this sample (Table 23). After applying the four population specific OLR equations, all individuals were sexed the same as from the seriation analysis with the exception of the South African Coloured equation (Table 25). When the South African Coloured equation was applied, one male was miss-identified as female thus creating a male agreement accuracy of 91.67%. The overall agreement for this equation was 96.67% for the St. Owens sample. The two combined OLR equations, South African and Summary Sex, resulted in all individuals being sexed the same as the seriation analysis, just like the three population specific OLR equations.

Table 25: Classification results for the St. Owens sample using each of the six OLR equations (N=30).

OLR Equation	Correct Females	Correct Males	Overall Agreement (%)
Christ Church, Spitalfields	18/18 (100%)	12/12 (100%)	100.00%
South African White	18/18 (100%)	12/12 (100%)	100.00%
South African Black	18/18 (100%)	12/12 (100%)	100.00%
South African Coloured	18/18 (100%)	11/12 (91.67%)	96.67%
South African	18/18 (100%)	12/12 (100%)	100.00%
Summary Sex	18/18 (100%)	12/12 (100%)	100.00%

There were 16 females and 18 males that had all eight sexing traits preserved in the Chumash sample, with a complete sample size of 34 individuals (Table 23). After applying all six OLR equations, the same four females and two males were misclassified in all instances. This created an overall agreement accuracy of 82.35% with a 75% accuracy for females and an 88.89% accuracy for males (Table 26).

Table 26: Classification results for the Chumash sample using each of the six OLR equations (N=34).

OLR Equation	Correct Females	Correct Males	Overall Agreement (%)
Christ Church, Spitalfields	12/16 (75.00%)	16/18 (88.89%)	82.35%
South African White	12/16 (75.00%)	16/18 (88.89%)	82.35%
South African Black	12/16 (75.00%)	16/18 (88.89%)	82.35%
South African Coloured	12/16 (75.00%)	16/18 (88.89%)	82.35%
South African	12/16 (75.00%)	16/18 (88.89%)	82.35%
Summary Sex	12/16 (75.00%)	16/18 (88.89%)	82.35%

A total of 27 individuals created the Andaman sample with 13 females and 14 males (Table 23). All these 27 individuals had all eight sexing traits present for analysis. Only one female was misclassified when applying the OLR equations for Spitalfields, South African White, South African Coloured samples This was also the same for the two combined South African and Summary Sex equations, which created an overall accuracy of 96.30% (92.31% accuracy for females and 100% agreement for males). One additional female was misclassified when the South African Black equation was applied onto the Andaman sample. This lowered the female accuracy to 84.62% and the overall accuracy to 92.59% (Table 27).

Table 27: Classification results for the Andaman sample using each of the six OLR equations (N=27).

OLR Equation	Correct Females	Correct Males	Overall Agreement (%)
Christ Church, Spitalfields	12/13 (92.31%)	14/14 (100%)	96.30%
South African White	12/13 (92.31%)	14/14 (100%)	96.30%
South African Black	11/13 (84.62%)	14/14 (100%)	92.59%
South African Coloured	12/13 (92.31%)	14/14 (100%)	96.30%
South African	12/13 (92.31%)	14/14 (100%)	96.30%
Summary Sex	12/13 (92.31%)	14/14 (100%)	96.30%

When applying all six OLR equations onto the archaeological samples with very minor percentage differences between them, normally only one individual is further misclassified (Tables 24 – 27). The only OLR equation that came close to matching the results from the seriation analysis (Poulton and St. Owens samples) and the catalogue records (Chumash and Andaman samples) was the combined South African OLR equation. The results from the South African equation had the same percentage agreement for St. Owens, Chumash, and Andaman samples but increased accuracy for the Poulton sample by classifying one additional male (Table 24). This result is surprising because the South African individuals are completely different geographically and temporally. For these samples, the South African OLR equation attains the highest accuracy which reveals that it is the least biased both geographically and temporally.

6.4 Summary

Four archaeological skeletal samples were used to evaluate the possible use of pre-made OLR equations on known age and sex skeletal samples. First, the accuracies of each morphoscopic trait were explored, resulting in the subpubic angle having the greatest accuracy for the Poulton and Chumash samples (84.11% and 86.49%) (Tables 19 and 21) as well as the shape of the pubic body for the St. Owens and Andaman samples (83.33% and 92.60%) (Table 20 and 22). To reject the null hypothesis stating that the OLR equations will not yield greater accuracies compared to single morphoscopic traits, each OLR equation would have to result in an equal to, or lower accuracy to those listed above. For Poulton, St. Owens, and Andaman samples, the null hypothesis can be rejected with accuracies ranging from 92.59%-100.00% (Tables 24, 25, and 27).

H_1: The results from the OLR will provide a greater accuracy than that of a single morphoscopic trait for sex estimation.

The Chumash sample however achieved a lower accuracy with all the OLR equations which was 4.14% lower than using just subpubic angle (Table 26). Caution must be taken when viewing the results from the Chumash sample due to issues with the records being noted. This will be fully discussed in Chapter 8.

H_0: The results from the OLR will not differ from the sexing accuracy from a single morphoscopic trait for sex estimation.

7. Results III

7.1 Dealing with Fragmentation

In an ideal scenario, all skeletal elements would be preserved and undamaged from taphonomic processes such as bio erosion and/or fire/burning. This scenario is rare, and what is normally found in both archaeological and forensic cases are fragmented pieces of the human skeleton, if not fragmented, then damaged to an extent where analysis is not possible.

As we can see from the sample sizes for the archaeological sample (Chapters 3 and 6), very few preserved all eight morphoscopic traits on the pelvis. In this chapter, we explore three different ways in how to deal with fragmented remains and sex estimation:

1. Use the individual's median value to replace all missing scores to complete the OLR equations for fragmented samples.
2. Apply only sections of the OLR equations onto individuals who are fragmented to increase sample size:
    a. OLR Equations with only the ventral arc, subpubic concavity, and ischiopubic ramus variables ("Phenice OLR equation").
    b. OLR Equations with only the subpubic angle and shape of the pubic body ("Anterior Posterior OLR equation").
    c. OLR Equations with only the greater sciatic notch and preauricular sulcus variables ("Posterior Pelvis OLR equation").
3. Apply new and shorter OLR equations onto individuals who are fragmented:
    a. OLR Equations with only the ventral arc, subpubic concavity, and ischiopubic ramus variables ("Phenice OLR equation").
    b. OLR Equations with only the subpubic angle and shape of the pubic body ("Anterior Pelvis OLR equation").
    c. OLR Equations with only the greater sciatic notch and preauricular sulcus variables ("Posterior Pelvis OLR equation").

The second and third method of potentially dealing with fragmented remains will be attempted on known samples first, then applied onto the four archaeological samples. Applying sections firstly onto the knowns samples will provide known accuracies for when the sectioned equations will be used for the archaeological samples.

This chapter has several hypotheses that will be tested:

$1H_1$: A difference in accuracy will be found when using median values for missing variables for fragmented individuals and the original accuracies from the previous chapter.

$1H_0$: No difference in accuracy will be observed when using median values for missing variables for fragmented individuals and the original accuracies from the previous chapter.

$2H_1$: A difference in accuracy will be found when using sectioned OLR Equations to that of the original accuracies from the previous chapter.

$2H_0$: No difference in accuracy will be observed when using sectioned OLR Equations to that of the original accuracies from the previous chapter.

$3H_1$: A difference in accuracy will be found when using the new 'smaller' OLR Equations to that of the original accuracies from the previous chapter.

$3H_0$: No difference in accuracy will be observed when using the new 'smaller' OLR Equations to that of the original accuracies from the previous chapter.

7.2 Median value replacing missing scores

To include individuals that were fragmented, the median value for each specimen was calculated and replaced all missing values for that specimen. Four samples had individuals with a fragmented pelvis. From the Christ Church, Spitalfields population, 90 specimens did not preserve all eight morphoscopic traits. From the 90 individuals, there were 464 missing values with the majority of individuals only having two (N=42), five (N=9) or eight (N=15) morphoscopic traits present. From the Poulton collection, 85 individuals had at least one missing trait. A total of 207 missing values were replaced with individuals only having two (N=49) traits present. The St. Owens collection had 70 individuals that had a fragmented pelvis, to which most of the individuals only had two traits present (N=50) similar to the Poulton collection however, there were 238 missing values that were replaced overall. The last population that had fragmented remains was the Chumash population. Four individuals had a fragmented pelvis, with only 22 missing values overall. Two individuals had five traits missing, whilst the remaining individuals had four and seven traits replaced with their median score. An overview of the four samples can be seen in Tables 19 - 22.

Table 28: Breakdown of fragmented individuals

Sample		N	Mean Age (Years)	Min-Max (Years)
Christ Church, Spitalfields				
	Male	43	56.37	21-83
	Female	47	57.89	26-87
	Total	*90*	*57.17*	*21-87*
Poulton				
	Male	41	Adult	Adult
	Female	44	Adult	Adult
	Total	*85*	*Adult*	*Adult*
St. Owens				
	Male	44	Adult	Adult
	Female	26	Adult	Adult
	Total	*70*	*Adult*	*Adult*
Chumash				
	Male	1	Adult	Adult
	Female	3	Adult	Adult
	Total	*4*	*Adult*	*Adult*

Table 29 summarises the results for the 90 individuals from the Spitalfields sample that had their median value replace missing scores. After applying all six OLR equations (found in Results Chapter 1) the lowest overall accuracy was achieved by the South African Black and the South African equation at 90.00%. Both these equations achieved the same male-female classification with six female misclassifications and three male misclassifications. However, the three male misclassifications were calculated to be ‘unknown’ or truly ambiguous. The remaining four equations achieved an overall accuracy of 91.11% however, differences in the male-female classification were seen. The South African Coloured equation had a higher classification of females with 93.62% being correctly sexed whilst the other three equations (Spitalfields, South African White, and Summary Sex) achieved a lower accuracy of 89.36%. With this being said, these three equations had a higher accuracy for males being 93.02% where the three males were deemed as ambiguous rather than male or female. The South African Coloured equation misclassified an additional two males as females unlike previously.

After analysing the 85 individuals from the Poulton sample that had missing values replaced by their median score, the equation with the greatest percentage agreement was the South African Coloured equation. With an overall agreement of 88.24%, only five females and five males were sexed differently. The Spitalfields equation generated the lowest agreement by sexing five females and eight males differently (88.64% and 80.49%) resulting in an overall agreement of just 84.71% (Table 30). The remaining four equations yielded the same overall agreement (85.88%) and male-female classification (80.49% and 90.91% respectively). No individuals were calculated to be ambiguous.

Table 29: Classification results for the fragmented Christ Church, Spitalfields sample using each of the six OLR equations (N=90).

OLR Equation	Correct Females	Correct Males	Overall Agreement (%)
Christ Church, Spitalfields	42/47 (89.36%)	40/43 (93.02%)*	91.11%
South African White	42/47 (89.36%)	40/43 (93.02%)*	91.11%
South African Black	41/47 (87.23%)	40/43 (93.02%)*	90.00%
South African Coloured	44/47 (93.62%)	38/43 (88.37%)*	91.11%
South African	41/47 (87.23%)	40/43 (93.02%)*	90.00%
Summary Sex	42/47 (89.36%)	40/43 (93.02%)*	91.11%

* 3 male individuals were classified as unknown and deemed incorrect

Table 30: Classification results for the fragmented Poulton sample using each of the six OLR equations (N=85).

OLR Equation	Correct Females	Correct Males	Overall Agreement (%)
Christ Church, Spitalfields	39/44 (88.64%)	33/41 (80.49%)	84.71%
South African White	40/44 (90.91%)	33/41 (80.49%)	85.88%
South African Black	40/44 (90.91%)	33/41 (80.49%)	85.88%
South African Coloured	39/44 (88.64%)	36/41 (87.80%)	88.24%
South African	40/44 (90.91%)	33/41 (80.49%)	85.88%
Summary Sex	40/44 (90.91%)	33/41 (80.49%)	85.88%

Table 31 shows the results for the fragmented medieval St. Owens sample (N=70). The South African Coloured equation achieved the greatest agreement with 88.75% with only two females and six males being sexed differently. From the six males, two were deemed to be truly ambiguous whilst the remaining four were sexed as female. Four equations achieved the same overall agreement and male-female classification. The Spitalfields, South African White, the clustered South African, and Summary Sex equations achieved an 82.86% agreement with five females and seven males being sexed differently. The male misclassifications from these four equations include two males being deemed ambiguous. The South African Coloured equation achieved a slightly greater agreement than the previous four with one less male being sexed as female resulting in an overall agreement of 84.26%.

From the Chumash sample, only four individuals had missing values that were replaced by their median score. After applying the six OLR equations all three females and one male was sexed correctly (Table 32).

Table 31: Classification results for the fragmented St. Owens sample using each of the six OLR equations (N=70).

OLR Equation	Correct Females	Correct Males	Overall Agreement (%)
Christ Church, Spitalfields	21/26 (80.77%)	37/44 (84.09%)*	82.86%
South African White	21/26 (80.77%)	37/44 (84.09%)*	82.86%
South African Black	21/26 (80.77%)	38/44 (86.36%)*	84.26%
South African Coloured	24/26 (92.31%)	38/44 (86.36%)*	88.57%
South African	21/26 (80.77%)	37/44 (84.09%)*	82.86%
Summary Sex	21/26 (80.77%)	37/44 (84.09%)*	82.86%

* 2 male individuals were classified as unknown and deemed incorrect

Table 32: Classification results for the fragmented Chumash sample using each of the six OLR equations (N=4).

OLR Equation	Correct Females	Correct Males	Overall Agreement (%)
Christ Church, Spitalfields	3/3 (100.00%)	1/1 (100.00%)	100.00%
South African White	3/3 (100.00%)	1/1 (100.00%)	100.00%
South African Black	3/3 (100.00%)	1/1 (100.00%)	100.00%
South African Coloured	3/3 (100.00%)	1/1 (100.00%)	100.00%
South African	3/3 (100.00%)	1/1 (100.00%)	100.00%
Summary Sex	3/3 (100.00%)	1/1 (100.00%)	100.00%

7.3 Sectioned OLR Equations

Another route for dealing with fragmented remains is to use sections of an already created equation and considering what variables are actually present in the sample. To highlight the use if this method, each equation was sectioned three times which best shown the possible combinations an osteologist might encounter. The first combination is recreating the Phenice method (1969) and Klales *et al*. (2012) method which use three traits, Ventral Arc, Subpubic Concavity, and the medial aspect of the Ischiopubic Ramus. The second combination includes the two traits from the posterior pelvis, the Greater Sciatic Notch, and the Preauricular Sulcus. The final combination includes two traits from the anterior pelvis being the shape of the pubic body and the subpubic angle. Even though there are countless more combinations that could be created, these three are the most likely combinations to appear. The obturator foramen was not included in any of the equations because normally when fragmentation occurs, the break occurs along the inferior and superior ramus of the pubis. Each of the sectioned OLR equations can be found in Appendices A – C. Firstly, the equations will be tested on the four known age and sex samples then applied to the four archaeological samples (Tables 10, 23, and 28).

7.3.1 Sectioned Phenice OLR

After applying the sectioned Phenice OLR equations onto the Christ Church, Spitalfields sample, the sample size increased from 138 individuals (breakdown can be seen in Chapter 5, Table 10) to 158 individuals. This is an addition of 8 females and 18 males. As we can see in Figure 32 (Appendix A.2 Table 1) no equation achieved an accuracy greater than 90%. Three equations (South African White, clustered South African, and Summary Sex) nearly achieved an overall accuracy of 89.87%. Both of these equations achieved better accuracies for females (94.81%) than for males (85.19%) with 4 female and 18 male misclassifications. Both the South African Black and Coloured equations misclassified one further female which decreased the accuracy to 89.24%. Surprisingly, the Spitalfields sectioned Phenice OLR equation achieved the lowest accuracy of 88.61%. Two additional females were misclassified when compared against the three equations resulting in the greatest accuracy (Appendix A.2 Table 1).

No additional individuals were added to the South African White sample because all individuals had all traits present. Regardless of this, the sectioned OLR equations were applied onto this sample and the results are summarised in Figure 32 (Appendix A.2 Table 2). As expected the South African White sectioned equation performed the best with an overall accuracy of 94.30% alongside the two clustered equations (South African and Summary Sex). A total of eight females (91.84%) and three males (96.84%) were misclassified. Both the Spitalfields and South African Black equation resulted in a near similar accuracy however, one additional female was misclassified. This resulted in a lower accuracy of 93.78%. The lowest accuracy was found when applying the South African Coloured equation. This equation resulted in eight females (91.84%) and seven males (92.63%) being misclassified. No individuals were deemed as being ambiguous.

When applying the sectioned OLR equation to just include the Phenice traits onto the South African Black sample (N=205), no equation exceeded an accuracy of 85% (Appendix Table A.2 Table 3). Four of the six equations, Summary Sex, South African, South African Coloured, and South African White, achieved the greatest accuracy of 84.44% with nine males (91.67%) and 22 females (76.53%) being incorrectly sexed (one female was classed as ambiguous). The Spitalfields equation resulted in a near similar accuracy of 84.39% but had a completely different male-female misclassification accuracy. Twenty-four females (75.26%) and eight males (92.59%) were misclassified which includes one female being classed as ambiguous. Similar to the results seen in Figure 32, the South African Black equation resulted in the lowest accuracy on its own sample. Nine males were misclassified, similar to the four of the other equations, but a total of 27 females were incorrectly sexed. This created an overall accuracy of 82.44%.

The fourth clustered column in Figure 32 (Appendix A.2 Table 4) summarises the results from sectioning the OLR equations and applying them onto the South African Coloured sample (N=205). The South African Coloured equation resulted in the highest accuracy of 92.20% with only three males (97.06%) and 13 females (88.24%) being misclassified. The South African White, South African, and Summary Sex equations had a slightly lower accuracy (91.22%) than previous, with one additional female and one additional male being misclassified. The Spitalfields and South African Black equations resulted in the lowest accuracies with 88.29% and 87.80% respectively (Appendix A.2 Table 4).

After testing the Phenice OLR equation onto the known skeletal samples, Appendix Tables 5 – 8 summarise the results of applying them onto the four archaeological samples and can also be seen on the right hand side of Figure 32.

The Poulton sample has an increased sample of 67 individuals (previously 47). The South African Black equation resulted in the lowest agreement with 82.90% with more females being sexed correctly (91.18%) than males (70.58%). The Spitalfields Phenice OLR equation had a greater accuracy than the South African Black by nearly 7%. An overall agreement of 89.55% was calculated with only two females and five males being sexed differently. The remaining four equations (South African White, South African Coloured, South African, and Summary Sex) all ended with the same overall percentage agreement of 92.54% however, the South African Coloured equation had a slightly different male-female agreement. Two female and three males were misclassified for the South African Coloured equation whilst the other three equations only had one female and four males misclassified. No individuals were deemed to be of unknown sex.

The results of applying the sectioned Phenice OLR equation onto the St. Owens sample can be seen in Appendix A.2 Table 6. Each equation achieved an overall agreement greater than 90%. The South African Coloured equation correctly sexed all the females in the sample with only two male misclassifications creating an overall agreement of 95.45%. Three other equations achieved the same overall agreement of 95.45% (South African White, South African, and Summary Sex), however they all had one female and one male sexed incorrectly. The remaining two equations, Spitalfields and South African Black, had one additional female sexed differently creating an overall percentage agreement of 93.18%.

The Chumash sample increased to 35 individuals with one female being included. All six equations resulted in the same overall percentage agreement of 82.86%. Each equation had the same male-female misclassification of four females and two males (Appendix A.2 Table 7). No individuals were deemed ambiguous.

Each equation for the Andaman sample correctly sexed each male but incorrectly identified three females with two being deemed ambiguous. This created an overall agreement of 88.89%.

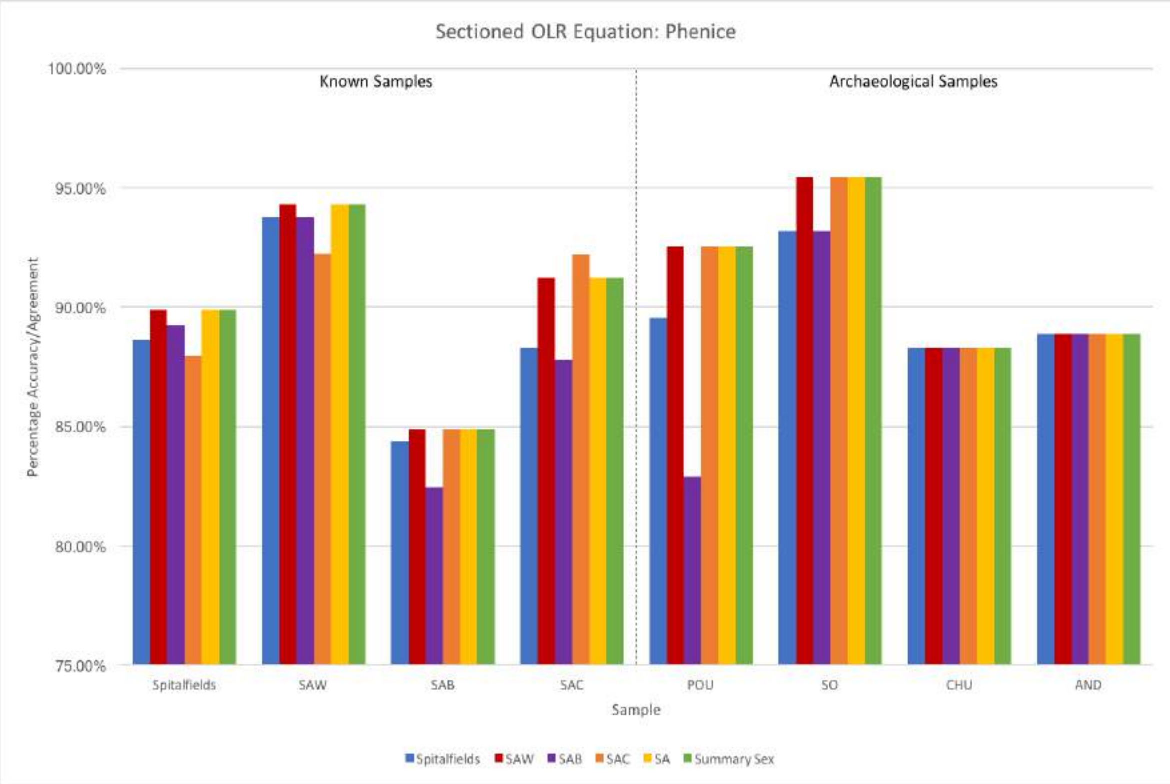


Figure 37: Bar chart showing the results of each Sectioned Phenice OLR equation for the eight samples tested.

7.3.2 Sectioned Posterior Pelvis OLR

After sectioning the original OLR equations that included just the Greater Sciatic Notch and Preauricular Sulcus variables, they were applied on the four known samples (Appendix B.2 Tables 1 – 4). Figure 33 (Appendix B.2 Table 1) summarises the results for the Spitalfields sample which has been increased to 214 individuals. The South African Coloured equation performed the best with an overall accuracy of 88.32% with 14 females and 11 males being misclassified. One of the male misclassifications was deemed ambiguous. The South African Black equation obtained the lowest accuracy of 82.71%. More females were misclassified than males by a wide margin. A total of 30 females were incorrectly sexed compared against the seven males (including one male being classed as ambiguous). The remaining four equations, including the two clustered equations, all gave the same accuracy (83.18%) and male-female classification (Female: 73.15% and Male: 93.40%). As seen by the other equations as well, one male was deemed unknown.

Appendix B.2 Table 2 describes the results for the South African White sample. From all of the equations, two females and one male were calculated to be ambiguous. From there, the South African Coloured equation achieved the greatest accuracy 88.08% with 16 females and seven males being misclassified. The South African Black equation proved to be the least affective for correct sex identification with an accuracy of 80.83%. This resulted from, albeit only four male misclassifications, a total of 33 females being incorrectly sexed. The remaining equations all achieved the same accuracy of 89.53%.

Five of the six equations for the South African Black sample all achieved the same overall agreement of 81.95% (Figure 33). However, the South African Black equation differed from the rest with a different male-female classification ratio. The South African Coloured equation had the only different overall accuracy which was slightly lower than the other at 80.49%. All the equations however classified one female and three males as being unknown.

Four of the six equations for the South African Coloured samples obtained and accuracy of 79.02% with 43 females being the only specimens that were misclassified (One female was determined to be ambiguous). These equations were the Christ Church, Spitalfields, South African White, South African, and Summary Sex equations. The remaining two equations resulted in the highest and lowest accuracies. The South African Coloured equation obtained an overall accuracy of 87.80% with a much higher classification rate for females than the previous four equations (an additional 12 females being correctly sexed = 79.61%). With this being said, four males were misclassified resulting in the loss of the 100% success rate as seen by the other four OLR equations. The South African Black equation obtained the lowest accuracy of 71.22% with less than half of the females being correctly sexed (42.72%). However, all male individuals were correctly sexed. As mentioned before, one female in all of the six analyses was identified as being ambiguous.

Moving from the known age and sex samples, the archaeological sample of Medieval Poulton was assessed using the sectioned OLR equation for variables found at the posterior portion of the pelvis (Figure 33). An increased sample size to 125 individuals was found with the new selection criteria (Females: 63, Males: 62). All but one equation achieved the same overall agreement and agreement rates for both males and females. The two clustered equations, Christ Church, Spitalfields, South African White and Black all achieved an overall agreement of 84.80% which included a higher percentage of females being sexed correctly than males (88.89% and 80.65% respectively). Only the South African Coloured equation resulted in a different agreement accuracy. Overall, the equation performed better with 87.20% of individuals being sexed correctly. Four additional males were correctly sexed increasing its accuracy to 87.10%, however female accuracy did drop by one additional female being sexed differently (87.30%). No individuals in this sample were deemed ambiguous.

The St. Owens Medieval sample increased from 30 individuals to 98 with 44 adult females and 48 adult males being included in the analysis. All equation except the South African Coloured equation obtained the same accuracy for males and females being 75.00% for females and 88.89% for males. This created an overall agreement accuracy of 82.65%. The only difference found when applying the South African Coloured equation onto the St. Owens sample was that two additional females were correctly sexed increasing the female accuracy to 79.55%, which in turn increased the overall agreement to 84.65%. One female and two males were deemed as ambiguous from these six analyses and therefore classed as incorrect sex (Appendix B.2 Table 6).

The Chumash sample was increased to its maximum sample size of 38 with 19 females and 19 males. Similarly to the St. Owens sample, all but one equation achieved the same agreement accuracies. All but the South African Black equation achieved an overall 75.68% with only three females and six males being misclassified. The South African Black equation produced a much lower agreement accuracy of 70.27% which included an additional two males being misclassified. No individuals were deemed to be ambiguous.

The Andamanese sample (N=27) achieved accuracies much higher than those obtained by the Chumash sample as stated previously (Figure 33). Five out of the six equations achieved and accuracy of 85.19% with only three females and one male being sexed differently (76.92% and 92.86% correct agreement, respectively). The only equation to differ was the South African Coloured equation which achieved a slighter higher accuracy of 88.89% (Appendix B.2 Table 8). This higher accuracy was due to one additional female being correctly identified, in turn increasing the female percentage agreement to 84.62%.

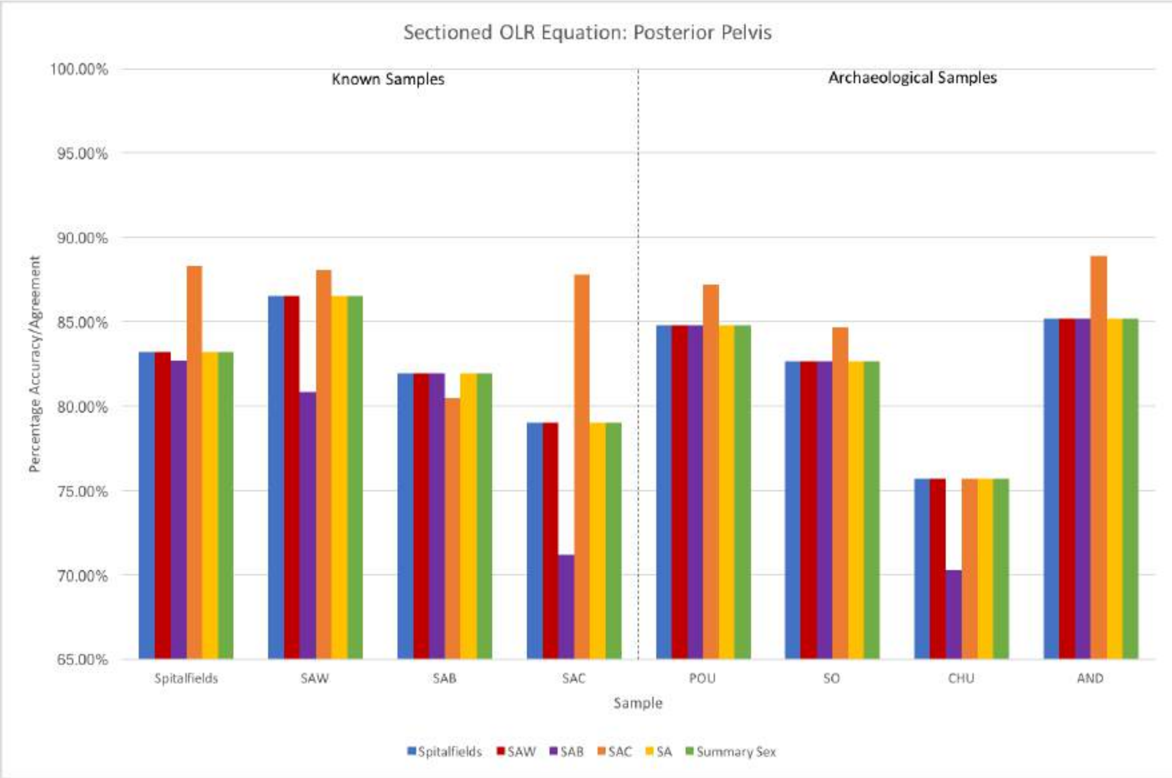


Figure 38: Bar chart showing the results of each Sectioned Posterior Pelvis OLR equation for the eight samples tested.

7.3.3 Sectioned Anterior Pelvis OLR

The final sectioned OLR equation included two variables found on the anterior portion of the pelvis; the subpubic angle and the shape of the pubic body with the results presented in Figure 34 (Appendix C.2 Tables 1 – 8). The first sample to apply the Anterior Pelvis OLR equations was the Christ Church, Spitalfields with an overall sample size of 162 individuals (Females: 81, Males: 81). The three South African equations (White, Black, and Coloured) obtained the highest overall accuracy of 95.68% with only one female and six males being misclassified. The remaining three equations (Christ Church, Spitalfields, South African, and Summary Sex) had a slightly lower accuracy yet still achieving over 95%. The latter three had only one female and seven males misclassified resulting in an overall accuracy of 95.06%. Each equation did however classify one male as ambiguous.

For the South African White sample the Christ Church, Spitalfields equation performed the best with an overall accuracy of 93.78% with two females and seven males being misclassified (94.90% and 92.63% accuracies respectively). The South African Black equation achieved the lowest accuracy of 90.67% with 90.81% of females and 90.53% of males being correctly sexed. The remaining four equations achieved the same overall accuracy (92.22%) and accuracies for females and males (91.84% and 92.63%). With this being said, two males were classified as ambiguous in all six analyses and were deemed as misclassifications.

Three of the six equations obtained an accuracy of 87.80% when applying the sectioned Anterior Pelvis equation on the South African Black sample. The Christ Church, Spitalfields, South African, and Summary Sex equations all correctly identified 82.47% of females and 92.59% of males in the sample. The South African White and Coloured equations obtained the lowest accuracy of 87.32% with one additional female being misclassified. The highest accuracy was obtained by the South African Black equation with an overall accuracy of 89.76% which included 86.60% of females and 92.59% of males being sexed correctly. These analyses did classify two females and one male as unknown however.

All but one equation differed when applying the equations onto the South African Coloured sample. The South African Black equation obtained the highest accuracy of 89.76% with ten female and nine male misclassifications (90.29% and 89.22% accuracies respectively). The remaining five correctly identified one additional male raising its accuracy to 90.20% but misclassified two additional females lowering female accuracy to 88.35%. These five equations all had an overall accuracy of 89.27%. All six equations resulted in one female and six males being classed as being ambiguous.

Moving from the known age and sex samples, the first archaeological sample to have the sectioned Anterior Pelvis OLR equations was the medieval Poulton sample with an overall sample size of 76 (Females: 37, Males: 39). All of the six equations obtained the same male agreement accuracy of 87.18% with no males being classified as ambiguous. The only difference was in the female identification. Four of the equations (Christ Church, Spitalfields, South African Coloured, clustered South African, and Summary Sex) all had five female misclassifications which resulted in an overall percentage agreement of 86.84% for the Poulton sample. The South African White equation had one additional female misclassification, lowering the overall accuracy to 85.53 whilst the South African Black had one additional female classified as female, increasing the overall accuracy to 88.16%. All six equations did deem one female as unknown.

The St. Owens samples increased its overall sample size to 46 with 24 females and 22 males. After applying the six Anterior Pelvis equations, four produced the highest agreement accuracies (95.65%) (Figure 34). The Christ Church, Spitalfields, South African Black, clustered South African, and the Summary Sex equations resulted in only two males being misclassified with all females being sexed in accordance with previous records. The remaining two equations (South African White and Coloured) differed only slightly with one female being misclassified which lowered the overall agreement accuracy to 93.48%. No individuals for this sample were classified as being ambiguous.

The Chumash sample (N=37) achieved the lowest agreement accuracies of any of the eight samples analysed. Five of the six equations achieved an overall agreement of 81.08% which included five females and two males being misclassified (72.22% and 89.47% accuracies respectively). The only equation to differ was the South African Black equation which performed better than the other five with an additional two females being correctly identified raising the overall agreement accuracy to 86.49%.

The last sample to be analysed was the Andaman sample with an overall sample size of 27 (Female: 13, Male: 14). Each of the six equations correctly identified all of the males in the sample. Only the female accuracy differed where two females were misclassified using all but the South African Black equation, where they resulted in an overall accuracy of 92.59%. The South African Black equation misclassified one additional female lowering the accuracy to 88.89%. No individuals were classified as ambiguous.

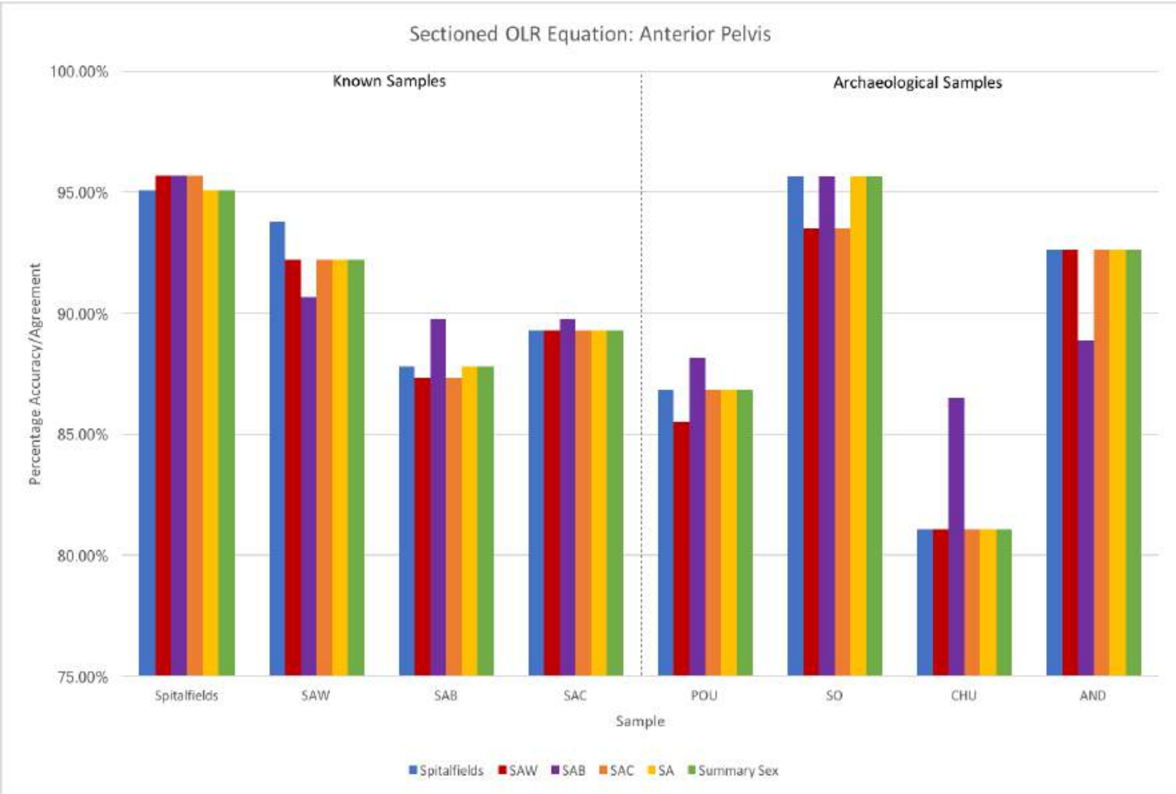


Figure 39: Bar chart showing the results of each Sectioned Anterior Pelvis OLR equations for the eight samples tested.

7.4 New OLR Equations

The previous section of this chapter looked at the possible route of sectioning previously developed OLR equations to help include individuals that were fragmented. To see if sectioning the equations worked to its highest capacity, new OLR equations were developed and applied onto each sample which will later be compared to the its sectioned counterparts. The same three equations were recreated, the Phenice equation (Ventral Arc, Subpubic Concavity, and medial aspect of the Ischiopubic Ramus variables), the Posterior Pelvis equation (Preauricular Sulcus and Greater Sciatic Notch variables), and the Anterior Pelvis equation (shape of the Pubic Body and Subpubic Angle variables). Each of the new OLR equations can be found in Appendices D – F along with their corresponding Wald Test and ROC Curve. Each equation was first applied to the four known age and sex samples then onto the four archaeological samples (Figures 35 – 37).

7.4.1 New Phenice OLR

When applying the six new Phenice OLR equations onto the 18th/19th Century British sample of Spitalfields, we see that none of the six produce accuracies reaching 90%. The best performing equation was the one modelled on itself with an accuracy of 89.87% with females have a slightly higher accuracy (90.91%) than males (88.89%). The South African Black and clustered South African equations obtained the second highest accuracy of 89.24%. These two equations identified two additional females raising the female accuracy to 93.51% but misclassified an additional three males lowering their accuracy to 85.19%. The Summary Sex and South African White equations achieved an accuracy of 88.61% which included six females and 12 males being incorrectly sexed (92.21% and 85.19% accuracy respectively). The lowest accuracy for this sample was achieved by applying the South African Coloured equation resulting in an overall accuracy of 87.34% with more females being correctly sexed than males (90.91% and 83.95%).

All six equations when applied onto the South African White sample achieved accuracies greater than 90%. Four of the six (South African, Summary Sex, South African White and Black equations) obtained the same overall accuracy along with accuracies for males and females (Appendix D.4 Table 2). Females were correctly identified with a percentage 91.84% (eight misclassifications) and males were correctly sexed with a percentage accuracy of 97.89% (two misclassifications) resulting in an overall accuracy of 94.82%. The Christ Church, Spitalfields equation achieved an accuracy of 93.26% with only one male misclassification (98.85% accuracy) but a greater number of females being misclassified resulting in an 87.76% accuracy for females. The lowest accuracy was obtained by the South African Coloured equation with an accuracy of 92.75% with only eight females and six males being incorrectly sexed.

Similar to the results from the South African White sample, four of the six equations obtained the same accuracy for the South African Black sample however they differed by approximately 10%.

The Summary Sex, South African, South African Black and White equations all resulted in an overall accuracy of 84.88% (Figure 35), however the latter equation had a different female and male accuracy to the other three. The South African White equation correctly identified 74 females (76.29%) and 100 males (92.59%) whilst the other three identified one less female and one additional male (75.26% and 93.52% respectively). The South African Coloured equation achieved an accuracy of 82.93% and the Christ Church, Spitalfields equation achieving the lowest accuracy for this sample with 81.95%.

The best performing equation for the South African Coloured sample was its sample specific one. This equation misclassified 13 females and only one male resulting in an overall accuracy of 93.17%. The two cluster equations and the South African Black equation obtained the same accuracies for both males and females with an overall accuracy of 91.22%. The South African White equation was the only other equation that obtained an accuracy greater than 90% with an overall accuracy of 90.73%. This included an accuracy of 83.50% for females and 98.04% accuracy for males. The lowest accuracy was achieved by the Christ Church, Spitalfields equation which incorrectly sexed 21 females and two males creating an overall accuracy of 88.78%.

The Poulton sample increased to 67 individuals by using the new selection criteria of just the three Phenice variables with 34 females and 33 males. Similar to the results achieved by the South African Coloured equation, all but one achieved accuracies greater than 90%. The Christ Church, Spitalfields equation had an overall agreement accuracy of 89.55% and misclassified four females and three males. The South African Coloured equation surpassed the 90% marker with an overall agreement of 91.04% where one less female was sexed as a male. The clustered South African, Summary Sex, and South African Black equations all obtained the same accuracy of 92.54% with only two females and three males being sexed differently. The South African White equation had the highest percentage agreement with 94.03% which had only one female misclassification along with the three male misclassifications.

The urban Medieval St. Owens sample increased to 44 individuals with 24 females and 20 males. The Christ Church, Spitalfields, South African White and Black equations obtained the same accuracy of 93.18%. Only two females and one male being sexed differently for the former two equations and one female and two males for the latter. The remaining three equations all achieved the same overall accuracy of 95.45%, however the female-male accuracies differed slightly. The South African Coloured equation only sexed two males differently whilst the South African and Summary Sex equations sexed one male and one female differently. No individuals were deemed as being ambiguous.

All six equations using the original Phenice variables achieved the same accuracies and the same female-male accuracies for the Chumash sample. Four females and two males were sexed differently than the museum records resulting in an overall accuracy of 82.86%. No individuals from these analyses were deemed to be ambiguous.

The six equations, when applied onto the Andaman sample, all obtained the same overall accuracy and female-male accuracy. No males were sexed differently from the museum records and only three females were sexed as males creating an overall accuracy of 88.89%. No individuals were classed as being ambiguous.

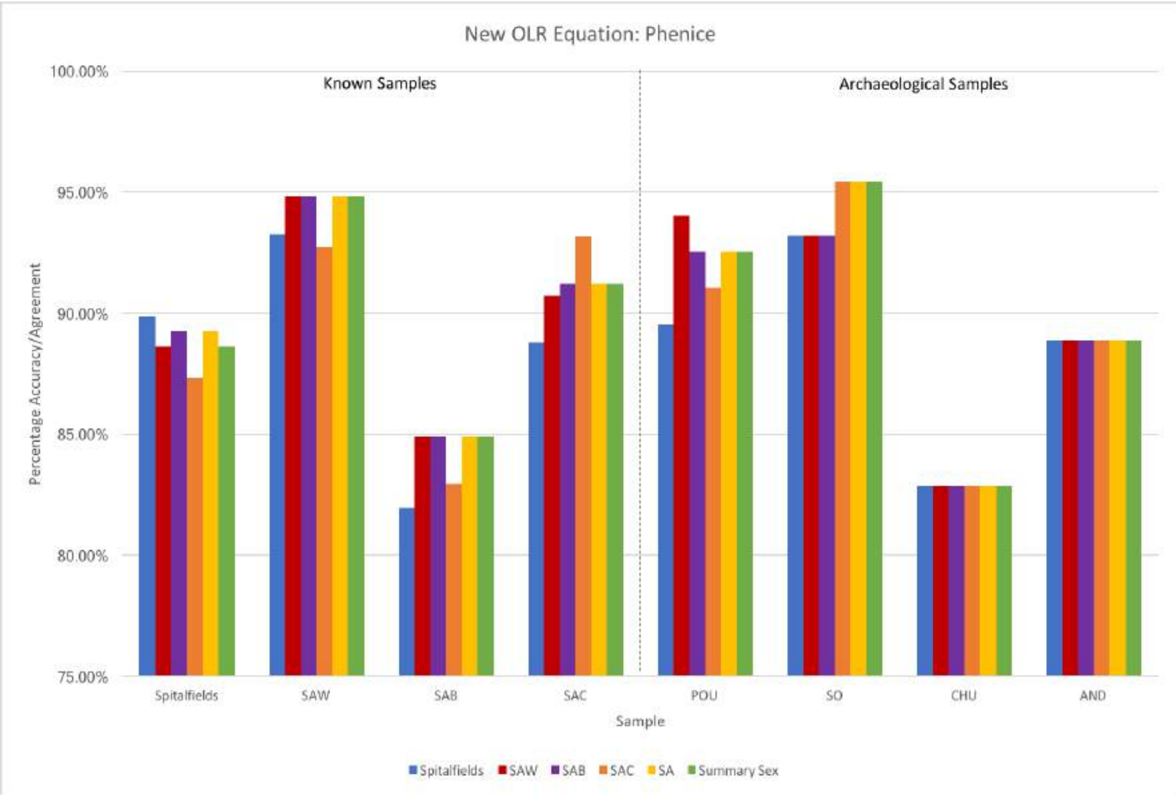


Figure 40:Bar chart showing the results of each New Phenice OLR equation for the eight samples tested.

7.4.2 New Posterior Pelvis OLR

Three of the six new Posterior Pelvis OLR equations achieved the same accuracy for the Spitalfields sample (N=214). The South African White and Black, and the clustered South African equations all obtained the lower accuracy of 83.18% with a much lower proportion of females being correctly sexed than males (73.15% and 93.40% respectively). The remaining three equations achieved an increase in accuracy of 5% with a total of 88.32% of individuals being sexed correctly (Figure 36). An increase of females being correctly sexed grew to 87.15% however, male classification was reduced to 89.62%. One male was classified an ambiguous when using all six equations.

The South African White and clustered South African equations achieved the lowest accuracy for the South African White sample with a total of 86.53% (77.55% for females and 95.79% for males). The remaining four equations achieved only slightly higher accuracies of 88.08% (Christ Church, Spitalfields, South African Black, and Summary Sex equations) and 88.60% (South African Coloured equation). From all six analyses, two females and one male was classed as being ambiguous (Appendix E.4 Table 2).

The South African Black sample achieved the highest accuracy of 81.95% by using the South African White and/or the South African equation with 94.44% of males being correctly sexed yet only 68.04% of females being sexed accurately. All six equations however had the same accuracy when concerning female individuals in the sample with 33 misclassifications. Male accuracy did differ. The South African Black equation had one additional male misclassification than previously resulting in an overall accuracy of 81.46%. The remaining three equations had the same accuracy of 80.49% where there were nine male misclassifications. All six analyses did classify one female and three males as ambiguous.

The final known age and sex sample, South African Coloured, gained its highest accuracy (87.80%) when applying the Christ Church, Spitalfields, South African Coloured, and Summary Sex equations with 79.61% of females and 96.08% of males being correctly sexed. The South African White and Black equations had the lowest overall accuracy of 79.02% (Figure 36) but interestingly achieved 100% accuracy for sexing male individuals. Females were sexed with an accuracy of 58.25%. The South African Coloured equation achieved an accuracy slightly higher than the previous two with an overall accuracy of 80.49%. Unlike the previous two equations, three males were misclassified resulting in a male accuracy of 97.06% with female accuracy similar to the first two equations of 79.61%. All six equations did misclassify one female however.

After applying the new Posterior Pelvis OLR equation onto the rural medieval Poulton sample, three equations obtained the highest accuracy of 87.20% (Figure 36). The Summary Sex, South African Coloured, and Christ Church, Spitalfields equation all correctly identified 87.30% of females and 87.10% of males in the sample. The remaining three equations (South African White and Black, and the clustered South African equation) had a decrease of nearly 3% to 84.80%. This lower percentage was due to one additional female and four additional males being sexed differently (Appendix D.4 Table 5).

The urban medieval St. Owens sample obtained its highest accuracy of 84.65% when applying the Christ Church, Spitalfields, South African Coloured and Summary Sex equations, similar to the Poulton sample. These thee equations sexed 79.55% of females and 88.89% of males when compared to the sex assigned via seriation. The clustered South African and South African White equations sexed 75.00% of females and 88.89% of males the same as the seriation based sexing which gave an overall agreement of 82.65%. The lowest agreement accuracy was achieved by the

South African Black equation which had one addition male misclassification lowering the accuracy to 81.63%. One female and two males were classified as being ambiguous.

All six equations that were applied onto the Chumash sample achieved the same overall accuracy of 75.68% (Figure 36). Three females and seven males were sexed differently from the museum records resulting in an agreement accuracy of 84.21% for females and 63.16% for males. No individuals were classified as being ambiguous (Appendix D.4 Table 7).

The last sample to have the six Posterior Pelvis OLR equations applied to was the Andaman sample. The lowest accuracy was achieved by the clustered South African equation and the South African Black and White equations. They resulted in an accuracy of 85.19% with 76.92% of females and 92.86% of males being sexed in accordance with the museum records. The remaining three equations obtained the highest accuracy of 88.89%. The only difference between the former and latter three equations was that one additional female was sexed correctly raising the female accuracy to 84.62%. No individuals were classed as being ambiguous for the Andaman sample.

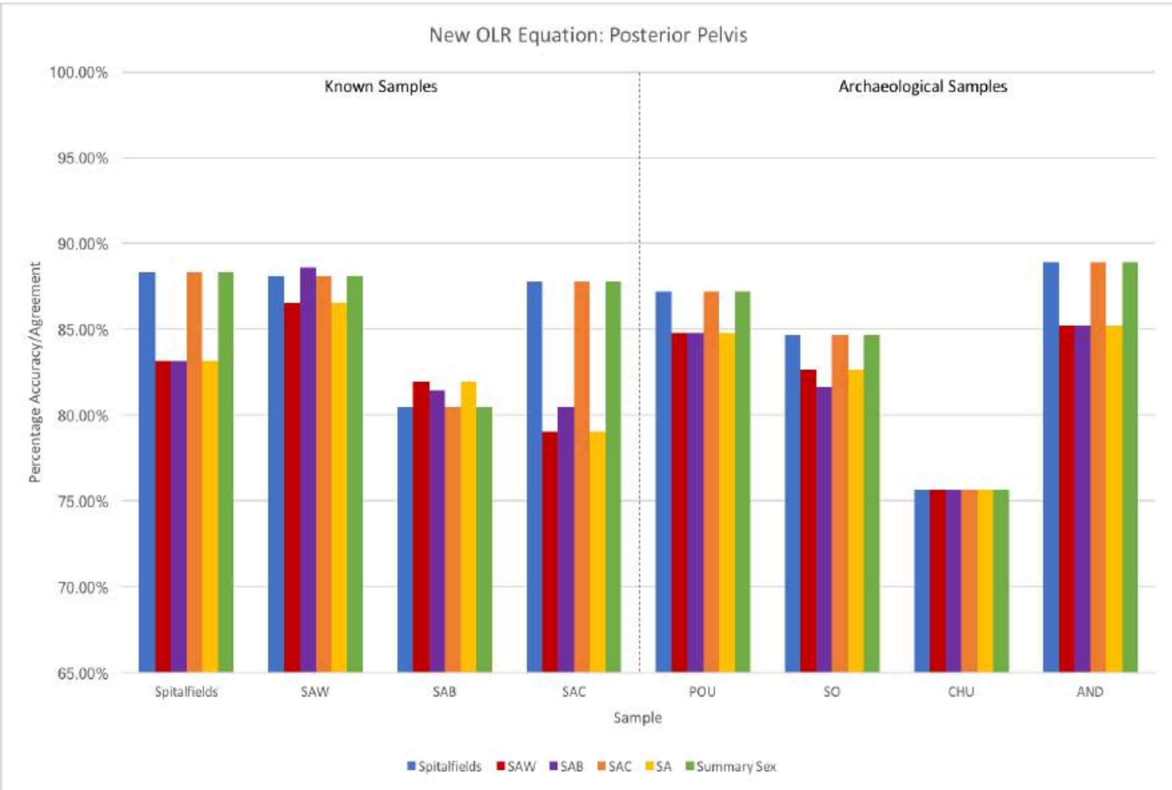


Figure 41:Bar chart showing the results of each New Posterior Pelvis OLR equation for the eight samples tested.

7.4.3 New Anterior Pelvis OLR

The first sample to have the new Anterior Pelvis equations applied was the Christ Church, Spitalfields sample (N=162). The highest accuracy for this sample was achieved by the Spitalfields equation with an overall accuracy of 96.91% with only one female and four males being misclassified (98.77% and 95.06% accuracies respectively). The three South African equations (White, Black, and Coloured) along with the clustered Summary Sex equation had one additional male misclassified as female, lowering male accuracy to 93.83% and overall to 96.30%. The clustered South African equation obtained the lowest accuracy of 95.68% with a further additional male misclassification. No individuals from this sample was deemed ambiguous using these six equations.

Both the South African White and the clustered South African equations achieved the lowest accuracy for the South African White sample with 93.78% (Figure 37). This included five females and seven males being incorrectly sexed (94.40% and 92.63% accuracies). The South African Black equation correctly identified two additional males increasing the overall accuracy to 94.84%. The remaining three equations obtained the highest accuracy of 95.34% with 94.40% of females and 95.79% of males being correctly identified. No individuals in the South African sample was classified as being ambiguous.

The South African Black equation obtained the lowest accuracy when applied onto its own sample with an accuracy of 87.80% which included 81.44% of females and 93.52% of males being sexed correctly. A slight increase in overall accuracy was made when applying either the Christ Church, Spitalfields, South African Coloured, or Summary Sex equations (Figure 37). They achieved an accuracy of 88.29% with an additional female raising the female accuracy to 86.60%. The remaining two equations, the clustered South African and South African White, achieved the highest accuracy of 90.24%. This was due to an increase of female classification up to 86.70%. No individuals in this sample was classified as ambiguous.

The South African White and the clustered South African equations resulted in the highest accuracy for the South African Coloured sample (N=205) with a female accuracy of 90.29% and a male accuracy of 95.10%. This created an overall accuracy of 92.68% (Figure 37; Appendix F.4 Table 2). The remaining four equations obtained a slightly lower overall accuracy of 92.20%. This was due to two females being further misclassified as male, lowering their accuracy to 88.35%. However, male accuracy did rise to 96.08% by correctly identifying one additional male. No individuals were found to be of unknown sex.

The archaeological sample from Poulton (N=76) obtained its lowest accuracy after applying the South African Black and the South African Coloured equations. They achieved an overall agreement of 86.84%. However, they did not have the same female-male accuracies. The South

African Black equation had an agreement of 83.78% for females and 89.74% for males whilst the South African Coloured equation had one less female and one more male in agreement with the sex based seriation (86.49% and 87.18% respectively). The South African White and the clustered South African equations obtained a female agreement of 89.19% and a male agreement of 87.18% resulting in the overall agreement of 88.16%. The remaining two equations (Christ Church, Spitalfields and Summary Sex) also obtained an overall accuracy of 88.16% However, their female agreement accuracy was lower at 86.49% and male agreement accuracy was higher at 89.74%. No individuals were deemed as being ambiguous.

All six equations, apart from the South African Black equation, achieved the same agreement accuracy for the urban medieval St. Owens sample (N=76). These five equations correctly identified each female in the same and only had two males sexed as females creating an overall agreement of 95.65%. The South African Black equation had the same male accuracy of 90.91% but sadly incorrectly sexed one female lowering the accuracy to 95.83% and lowering the overall agreement to 93.48%.

The Chumash sample achieved its highest accuracy when the clustered South African and South African White equations were applied (Figure 37). They both achieved an accuracy of 86.49% with only three females and two males being misclassified. The remaining four equations all achieved a lower accuracy of 81.08% because two more females were misclassified as males. No individuals in the Chumash sample were deemed ambiguous.

The last sample to be tested was the Andaman sample. The highest performing equation was that of the South African Black equation. This equation correctly identified all males in the sample and only misclassified two females resulting in an overall accuracy of 92.59%. The remaining five equations had one additional female being sexed as a male which lowered the female accuracy to 76.92% which in turn lowered the overall accuracy to 88.89%. No individuals were classified as ambiguous.

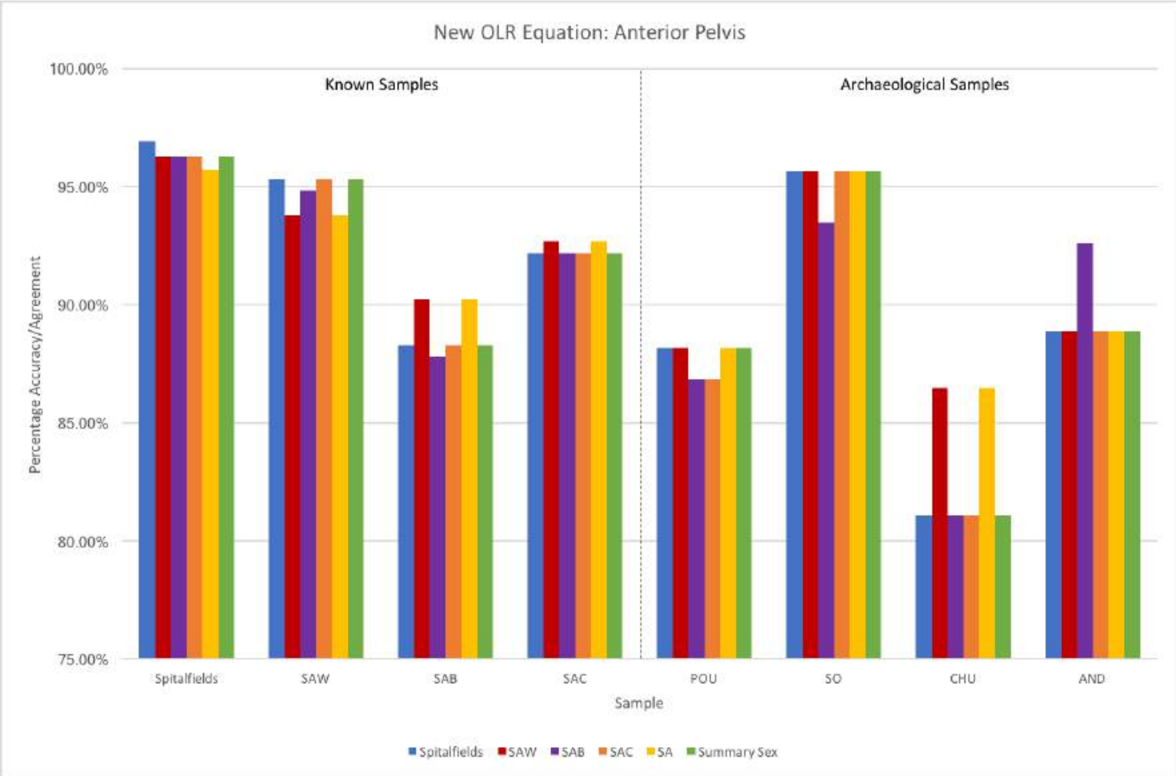


Figure 42:Bar chart showing the results of each New Anterior Pelvis OLR equation for the eight samples tested.

7.5 Summary

This chapter aimed to explore three possible routes in how to analyse samples that have fragmented material. Firstly, we looked at substituting all missing scores with the individual's median value to create a full data set. After this, we looked at using the previously created OLR equations (Results Chapter 1) and sectioning them to help increase the sample size. Three sets of sectioned OLR equations were created being the Phenice OLR Equation (Ventral Arc, Subpubic Concavity, and the medial aspect of the Ischiopubic Ramus variables), the Posterior Pelvis OLR Equation (Greater Sciatic Notch and Preauricular Sulcus variables), and the Anterior Pelvis OLR Equation (shape of the Pubic Body and the Subpubic Angle variables). Lastly, to assess the feasibility of using sectioned OLR equations, new OLR equations were created to mirror the sectioned equations.

From these simple comparisons, it is evident that using the new versions of the OLR equations are a better option in terms of higher accuracy for the eight samples provided. However, only if it is absolutely necessary then sectioning an already created OLR equation will give similar results, albeit a slightly lower sexing accuracy and is not recommended.

From the creation of the three new equations, the group of variables that performed with the highest overall accuracy were those used for the Anterior Pelvis equations (both sectioned and new) and the worst performers were the Posterior Pelvis equations. This can be explained by the *os pubis* having a better discriminatory power/ higher amount of sexual dimorphism than the posterior portion of the *os coxae*.

8. Discussion

The main aim of the PhD research was to provide a plausible universal equation for sex estimation using eight morphoscopic traits from the human pelvis. To do this, eight human skeletal samples from different time periods and different locations around the world were examined. Using these samples, three main objectives were created and were explored in Chapters 5, 6, and 7:

- Create population specific multivariate equations for sex estimation for the four known age and sex samples and test the idea of population specificity (Chapter 5).
- Apply each multivariate equation to four archaeological samples to see if accuracies differ from previous sex estimation results (Chapter 6).
- Investigate the use of using the median score to replace missing data and to explore the use of smaller equations to help assess fragmented skeletal remains (Chapter 7).

This piece of work produces several cross-validated equations that can be used specifically for their sample population and would be able to aid in forensic human identification (South African Samples) (Chapters 5, 7, and Appendices A – F). This research also has produced three new definitions for three morphoscopic sexing traits that previously did not have details (descriptions and/or images) on how to use them for sex determination (Chapter 4) with associated intra/inter observer analysis (Chapters 5 and 6). The final product of this study was the investigation of fragmented remains and how to overcome this issue when using multivariate equations. This investigation found that using the median value was a good substitution for missing values but should only be used as a last resort.

8.1 Study Results

8.1.1 Chapter 5 Results I

Four known age and sex samples were used to investigate the sexing accuracy of eight morphoscopic traits found on the human pelvis. Firstly, accuracies for each morphoscopic trait was explored which resulted in the shape of the pubic body having the greatest accuracy for three samples (Christ Church, Spitalfields: 90.42%; South African White: 88.60%; South African Black: 86.83%) (Tables 4, 5, and 6) and the subpubic angle having the greatest accuracy for the South African Coloured sample (82.93%) (Table 7). For the first null hypothesis ($1H_0$) (Chapter 5) to be accepted each OLR equation would have to have a lower or equal accuracy than the single morphoscopic analysis. The OLR equations greatly increased accuracy for all four populations regarding the single morphoscopic traits with the greatest accuracy (Table 13). An increase of 5.96% was observed for the Christ Church, Spitalfields sample, 6.74% increase for South African

White, 3.41% increase for South African Black, and 11.22% increase for South African Coloured. With these results the first null hypothesis has been rejected:

$1H_1$: The results from the OLR will provide a greater accuracy than that of a single morphoscopic trait for sex estimation.

The second hypothesis in Chapter 5 stated that there would be a difference in sexing accuracy when other OLR equations were applied onto a sample. When the three South African equations were applied onto the Christ Church, Spitalfields sample only the South African Black equation gave a different accuracy which was 0.73% lower than the original accuracy (Table 12 and Table 13). For the South African White sample, only the Christ Church, Spitalfields equation recreated the same accuracy from the original OLR equation (Table 12 and Table 14). Both the South African Black and Coloured equations resulted in different accuracies. All three other OLR equations for the South African Black and Coloured samples resulted in different accuracies (Tables 15 and 16). With different accuracies being observed when other OLR equations are being applied onto the samples, we reject the second null hypothesis:

$2H_1$: There is a difference in sexing accuracy when a different OLR equation is applied onto a sample.

The final hypothesis in Chapter 5 was to create and assess the use of a universal equation to see if it would have solved the issue of population specifity seen with the four OLR equations. If we take the lowest accuracies from each of the samples (Christ Church, Spitalfields: 95.65%, South African White: 94.82%, South African Black: 88.87%, and South African Coloured: 90.24%) (Tables 13 – 16) we get a mean accuracy of 92.37%. From the last analysis, the OLR equations from both the South African sample and "Summary Sex" sample, accuracies of 92.54% and 92.98% (respectively) were calculated (Table 18). This slight improvement of 0.17% (South African OLR) and 0.61% (Summary Sex OLR) shows that they have reduced the classification rate seen by applying different OLR equations onto a sample, barely. With this being the case, the third null hypothesis rejected:

$3H_1$: The creation of a "universal" OLR sexing equation will reduce the misclassification rate seen when applying different OLR sexing equations on populations they were not created from.

8.1.2 Chapter 6 Results II

In Chapter 6, four archaeological skeletal samples were used to evaluate the possible use of pre-made OLR equations on known age and sex skeletal samples. First, the accuracies of each morphoscopic trait were explored, resulting in the subpubic angle having the greatest accuracy for the Poulton and Chumash samples (84.11% and 86.49%) (Tables 19 and 21) as well as the shape of the pubic body for the St. Owens and Andaman samples (83.33% and 92.60%) (Table 20 and 22). To reject the null hypothesis stating that the OLR equations will not yield greater accuracies compared to single morphoscopic traits, each OLR equation would have to result in an equal to, or lower accuracy to those listed above. For Poulton, St. Owens, and Andaman samples, the null hypothesis can be rejected with accuracies ranging from 92.59%-100.00% (Tables 24, 25, and 27).

H_1: The results from the OLR will provide a greater accuracy than that of a single morphoscopic trait for sex estimation.

The Chumash sample however achieved a lower accuracy with all the OLR equations which was 4.14% lower than using just subpubic angle (Table 26). Caution must be taken when viewing the results from the Chumash sample due to issues with the records being noted. This will be fully discussed in Chapter 8.

H_0: The results from the OLR will not differ from the sexing accuracy from a single morphoscopic trait for sex estimation.

8.1.3 Chapter 7 Results III

Chapter 7 aimed to explore three possible routes in how to analyse samples that have fragmented material. Firstly, we looked at substituting all missing scores with the individual's median value to create a full data set. After this, we looked at using the previously created OLR equations (Results Chapter 1) and sectioning them to help increase the sample size. Three sets of sectioned OLR equations were created being the Phenice OLR Equation (Ventral Arc, Subpubic Concavity, and the medial aspect of the Ischiopubic Ramus variables), the Posterior Pelvis OLR Equation (Greater Sciatic Notch and Preauricular Sulcus variables), and the Anterior Pelvis OLR Equation (shape of the Pubic Body and the Subpubic Angle variables). Lastly, to assess the feasibility of using sectioned OLR equations, new OLR equations were created to mirror the sectioned equations.

Four of the eight samples had their missing values substituted with the individuals median score with good success. The best performing OLR equation for all four samples was the South African Coloured equation. If this equation did not perform on par to others, it was better by 2-3% when compared to the other five OLR equations. The original accuracies for the Christ Church, Spitalfields sample (found in Chapter 5) ranged between 95.38% (obtained from the South African Coloured equation) and 97.10% (achieved by the clustered South African and Summary Sex equations). After analysing the Spitalfields sample, that had missing values replaced, lower accuracies were found ranging between 90.00% and 91.11%. The Poulton sample had 85 individuals which were fragmented. After substituting missing values, accuracies ranged between 84.71% (Christ Church, Spitalfields equation) to 88.24% (South African Coloured equation). This is considerably lower than the original analysis on the 47 individuals who had a full data set which range between 93.62% and 95.74%. The same can be said for the St. Owens sample. The substitution of missing values had a much lower agreement accuracy than individuals with a complete data set. There was over a 10% difference in sexing between the two as the highest agreement for a full data set was 100% whilst for those with a substituted missing value it was 88.24%. The Chumash sample had a slightly different result but can be explained by the sample size. Hence comparing the percentage agreement will not be considered but the frequency of misclassifications will be. The difference between the two analyses were four female misclassifications and two male misclassifications, which can be described as not being a large number of misclassifications to begin with. The reason why the missing value substitution has been labelled as a good success is because we were inferring what the missing data could have been following the sexual dimorphism already present in the variables that were present. A dip in accuracy is to be expected when this is the case. Because differences were found however, the first null hypothesis for this chapter has been rejected.

$1H_1$: A difference in accuracy will be found when using median values for missing variables for fragmented individuals and the original accuracies from the previous chapter.

The next possible route was to section the OLR equations, found in the Chapter 5, into three equations that would help increase sample size. When applied onto the eight samples the clustered South African and Summary Sex equations obtained the highest accuracies for the sectioned Phenice OLR equations. The only case where accuracy dipped was when they were applied onto the South African Coloured sample where its own equation performed better with an increase of 0.98%. When comparing the results from the sectioned Phenice OLR to the original analyses in Chapter 5 there is a considerable drop in accuracy for the four known age and sex samples. Original accuracies for the Spitalfields sample predominantly stayed around the 95.00% marker whilst for the sectioned Phenice OLR analysis, it was much lower (~88.00%). The South African White sample had an accuracy drop of 2-3% across the six equations. A drop in 5-8% was found between the original results in Chapter 5 and the Phenice OLR equations for the South African Black sample. The South African Coloured sample averaged a similar accuracy but had a greater range in percentages than the original, this time ranging from 87.80% to 92.20%. For the Poulton sample, a decrease in approximately 4% was seen between the results found in Chapter 6 and this analysis. The largest difference was found when using the South African Black equations as there was over a 10% difference in agreement with the sex estimated by seriation. A slight decrease is agreement was found for the St. Owens samples. With an increase of sample from 30 to 44 a maximum of three individuals were misclassified from all six sectioned Phenice OLR's. The Andaman sample had a decrease in accuracy ranging from 4 – 8% with the greatest difference from all but the South African Black equation. The Chumash sample however had the only increase in accuracy of nearly 6% across all six equations.

The next sectioned OLR equation was the Posterior Pelvis (Figure 33; Appendix B). None of the six equations achieved accuracies reaching 90% for any of the eight samples tested. Even though this is the case, the clustered equations (South African and Summary Sex) were the equations that on average performed the best when only using the Preauricular Sulcus and Greater Sciatic Notch variables. The sample most affected by using these equations was the South African Coloured sample. The greatest difference was nearly 20% which was achieved by applying the South African Black equations. A close second to the South African Coloured sample was the St. Owens sample which had a difference in accuracy of approximately 18%. The majority of the other six samples had differences around 10%. The sectioned Posterior Pelvis OLR equations were all less effective than the original OLR and the sectioned Phenice OLR equations.

The final sectioned OLR equation was from using the shape of the Pubic Body and the Subpubic Angle variables. As a general statement, these equations had a much better performance than the other two sectioned OLR equations (Phenice and Posterior Pelvis). The South African Black Anterior Pelvis OLR equation, on average, performed the best for the eight samples. For the Spitalfields sample, accuracy was very close to the original accuracy of 96% with an accuracy as high as 95.68%. The only negative is that this top accuracy was achieved by the three South African equations (White, Black, and Coloured). A decrease in accuracy between 2-5% was seen for the South African White sample as the results stayed within the low 90's when using the sectioned Anterior Pelvis OLR equations. Accuracies did not drop by much when applied onto the South African Black sample. A maximum of 1.5% difference was found between Spitalfields equations when comparing the original to the sectioned OLR equation. The greatest percentage difference to be found for the South African Coloured sample was from the equation it was derived from which was nearly 5%. The South African Black equation however had the least amount of difference from the sectioned to the original being 0.48%. No equation for this sample achieved an accuracy reaching 90%. Each of the four archaeological samples all resulted in lower percentage agreements to either sex estimated by seriation or the museum records. However, they were much closer to the original results found in Chapter 6 than the results from the sectioned Posterior Pelvis OLR equations.

In terms of the second hypothesis for Chapter 7, it is clear that the null hypothesis has been rejected:

$2H_1$: A difference in accuracy will be found when using sectioned OLR Equations to that of the original accuracies from the previous chapter.

The last route that was explored was to create new, shorter OLR equations that mirrored the sectioned equations and to test their accuracy on the six samples. The best performing equation overall was the clustered South African equation even though a lower accuracy was observed when applied to the South African Coloured and medieval Poulton samples. The first set of equations to be analysed was the new Phenice OLR equations (Appendix D). Similar to what was seen in the sectioned Phenice OLR for the Spitalfields sample, a decrease in accuracy of at least 5% was evident as none of the six equations managed to achieve an overall accuracy of 90%. All results were found in the higher 80's. For the South African White sample, accuracy only decreased by 2% making it very similar to the original OLR results. The South African Black sample on the hand had a much larger difference in accuracy by 4-5% with accuracies as low as 81.95% for the new Phenice OLR equations. However, saying this, the greatest difference seen was between the two Spitalfields equations with a percentage difference of 8%. The South African Coloured sample had a dip in accuracy very similar to the South African White sample, around 2% for the majority of equations, however one equation did surpass this by a difference of 4% which was seen by the two Spitalfields equations. With an increased sample size for the Poulton sample, accuracy did decrease by 3-4% on average, but the majority of the equations resulted in a 90+% agreement (bar the Christ Church, Spitalfields Phenice OLR equation). However, an increase of accuracy was found when comparing the results from the original South African White equation to the new Phenice OLR equation, which was by 1%. A small difference was found when comparing the results from the original results for the St. Owens analysis by having 2-3 individuals being misclassified, lowering its 100% agreement to 93.18% - 95.45%. However, with an increase of 14 individuals, this is only a small difference. There was a difference of less than 1% when comparing the Chumash results but the same cannot be said for the Andaman sample. A maximum difference of 8% was noted between the two sets of analyses dropping its percentage agreement from the mid 90's to the high 80's.

The new Posterior Pelvis OLR analyses (Appendix E) for the Christ Church, Spitalfields sample had very similar accuracies to those found by using the new Phenice OLR equations. No equation reached 90% accuracy. The smallest difference between original and new equations was 11% which was by the samples own specific equations. The greatest difference was seen by the South African Black equations where a difference of 13% can be noted. The South African White sample did not fare any better as again, no equation managed to reach 90% with an average of 7% difference between the original results and the new Posterior Pelvis OLR results. Even more so for the South African Black sample having an average of 9% decrease in sexing accuracy across all six equations. The South African Coloured sample's results for the new OLR equation ranged by 10% from the high 70's to high 80's which is unlike the original results which stayed within the low/mid 90% mark. With this, a substantial difference of 13% can be seen (between the South African White equations) and with the smallest difference of 10% (between the South African Black equations). All of the four archaeological samples exhibited a similar 10% reduction in agreement with either the museum records (Chumash and Andaman samples) or the sex estimated from seriation (Poulton and St. Owens samples).

The last set of new OLR equations were created from using the shape of the Pubic Body and the Subpubic Angle variables (Appendix F). The Summary Sex and Spitalfields equations were the two equations that performed the best overall for each of the eight samples. The Christ Church, Spitalfields sample achieved accuracies greater than 95% for all six equations, with the highest performing equation attaining 96.91% which is higher than the original OLR equation. All six equations achieved accuracies greater than 90% for the South African White sample which mirrors what was originally obtained from the original OLR equations in Chapter 5. The greatest difference found for the South African White sample was 1.5% which was from the equations originally created for this sample. The same can be seen for the South African Black and Coloured samples. The results from the new Anterior Pelvis OLR equations did not greatly differ from the original OLR equations. Both the Poulton and Andaman samples had large difference in percentage agreement with an approximate 10% decline when compared to the original application found in Chapter 6. The Chumash sample on average had a very similar accuracy to the original OLR equation apart from two equations. The clustered South African and the South African White Anterior Pelvis OLR's obtained higher accuracies of nearly 4%. The St. Owens sample had a slight decline in percentage agreement by 5% which included 3 – 5 individuals being sexed differently than the sex based on its seriation.

In terms of the third and final hypothesis for Chapter 7, the null hypothesis has been rejected:

$3H_1$: A difference in accuracy will be found when using the new 'smaller' OLR Equations to that of the original accuracies from the previous chapter.

One thing to consider between the last two possible routes is if it is a valid option, in terms of accuracy, to section a previously constructed equation or to create a brand new equation that fit within the parameters needed. Did accuracies differ greatly between the two methods? As an overall view, when considering the Phenice OLR equations, the newly created equations performed better for each sample. However, some minor differences between the two same sampled equations can be seen. For instance, the Spitalfields sample concerning its own specific equation ranged in accuracy of around 1-2%, having a better accuracy for the new OLR equation. The largest difference in accuracy was seen when applying the South African Black equations on the medieval Poulton sample. The sectioned equation obtained an accuracy of 82.90% whilst the new OLR achieved an accuracy of 92.54%. Similar to the Phenice OLR equations, the new Posterior Pelvis OLR equations had a better accuracy for all six equations for the eight samples. From the sectioned OLR equations, it is clear that the South African Coloured equation was more effective for all samples bar the South African Black sample. For the new Posterior Pelvis OLR equations, primarily the Spitalfields and the Summary Sex equations had such an increase in accuracy to match the accuracies set by the fore runner – South African Coloured equation. Saying this however, even though sample size did increase the most when using the Posterior Pelvis OLR equations, they were some of the lowest accuracies reported for both the sectioned and the new versions of the equation. The final comparison between the Anterior Pelvis OLR equations show that, again the new OLR versions of the Anterior Pelvis equations performed better for all samples bar the Andaman sample. This is more evident in the four known and age sex samples is that the new OLR analysis, more equations were obtaining similar and higher accuracies. This is also seen for the archaeological samples but not to the same degree, especially concerning the Andaman sample where the reverse was seen.

From these simple comparisons, it is evident that using the new versions of the OLR equations are a better option in terms of higher accuracy for the eight samples provided. However, only if it is absolutely necessary then sectioning an already created OLR equation will give similar results, albeit a slightly lower sexing accuracy and is not recommended.

From the creation of the three new equations, the group of variables that performed with the highest overall accuracy were those used for the Anterior Pelvis equations (both sectioned and new) and the worst performers were the Posterior Pelvis equations. This can be explained by the *os pubis* having a better discriminatory power/ higher amount of sexual dimorphism than the posterior portion of the *os coxae*.

8.2 Morphoscopic Trait Analysis

Chapters 5 and 6 investigated the use of OLR generated sexing equations and how they compared against simple morphological analyses. From all eight samples, the majority of the OLR equations outperformed the single trait analysis. Overall, the results from the single trait analysis are what were expected.

Two traits were consistently more effective at estimating sex than the other six. The subpubic angle and the shape of the pubic body were found to have the highest accuracies across all eight samples. The anterior portion of the human pelvis (especially the os pubis) has been described as being the most sexually dimorphic structure (Klales et al. 2012), hence it is not surprising that the subpubic angle and pubic body shape are the best at identifying sex. This can be because the anterior pelvis needs to develop a certain way to help facilitate birth; a wider subpubic anglemakes for easier transitions between rotations the neonate's head has to make in navigating the birth canal. This highlights the importance of using traits that are more sexually dimorphic in sex estimation.

The obturator foramen, on the other hand, was found on the opposite side of the scale, being one of the worst performing morphoscopic traits for sex estimation. When comparing the morphoscopic results to Rogers & Saunders (1994) study, both similarities and differences are present. They found that the obturator foramen was the second most effective sex indicator for their Canadian sample with accuracies as high as 93.8%. There are a number of reasons why this may be the case. Firstly, the individuals from Canada may exhibit the extreme morphologies for males and females for the obturator foramen. Secondly, they use descriptions made by St. Hoyme (1984), where no visual diagrams can be found, therefore trying to recreate their study regarding the obturator foramen is impossible. If the descriptions from European Meeting for Anthropologists (1980) were the same as St. Hoyme (1984), then caution should be taken as there are no visual diagrams for how to score this trait and the only explanation for each score is a maximum of three words. Fourth, because a new scoring method was created for this study, different criteria were used which could account for some of the discrepancies found between this study and Rogers & Saunders' (1994). This study is more in agreement with the interpretation made by St. Hoyme (1984) that the foramen has little value when considering its use for sex estimation. Accuracies did differ between populations in this study but averaged an overall accuracy of 66% which is similar to Arsuaga & Carretero (1994) (63.5% correct classification). Comparing the results to Beirry et al. (2013), significant differences in accuracy rates are present. However, this can be explained by the use of Fourier analysis and Discriminant Function Analysis to categorise the oval (male) and triangular (female) expressions rather than relying on the naked eye, like in Beirry's study. Ridgeway et al. (2008) concluded after analysing 96 female pelves that the considerable amount of shape variation seen in obturator foramen could be due to an individual's height, noting that the greater the foramen area is, the taller the individual was. Ridgeway and colleagues (2008) found no significant difference was found between shape and a person's race. Because an individual's height was not considered, as it would have been largely estimated unlike the known height in Ridgeway's study, their conclusions could not be tested. However, a large amount of variation was seen not only between samples, but within samples as well which coincide with Ridgeway et al. (2008) findings of their not being a link between shape and an individual's race. The obturator foramen seems to be "hit and miss" when applying it to sex determination (this study being a miss). With this in mind, the obturator foramen should not be used to determine sex on its own as it is a poor discriminator. Arguments can be made to include it into a multiproxy assessment however, with other morphoscopic traits that perform much better.

The preauricular sulcus averaged an accuracy of 70% between all eight populations with a range of >30% (Poulton = 84.92%; Chumash = 52.63%). Rogers & Saunders (1994) yielded an accuracy as high as 91% but only scored it as a presence/absence after St. Holmes (1984) descriptions.

Houghton (1974) found that no males had shown the 'GP' groove, but the majority did show some bony modification, which was also found in this study as some males scored "0". The greater sciatic notch has been stated to have a higher accuracy for females than males, as more females show the hyperfeminine description by Walker (2005) which was later supported by Gómez-Valdéz et al. (2012). Walker (2005) analysed the British sample of St. Brides as part of his research to which all females, bar one, was correctly sexed as either a score of "1" or "2". For the Spitalfields sample, a similar pattern can be seen. The vast majority of females showed the hyper feminine morphology with only a few being seen as indeterminate. On the other side, males were spread out across the scores, with the majority scoring intermediate through to hyper-masculine. This pattern mirrors what was seen by Walker (2005). The same can be said for the South African White and St. Owens samples. Interestingly, the opposite occurred for the South African Black sample, where more variation was seen in the females than the males. Walker (2005) and Gómez-Valdéz et al. (2012) achieved accuracies greater than 80%, unlike this study which averaged 74.23% across eight different human populations. This can be explained by the greater diversity of populations analysed for this study and also larger sample sizes. Walker (2005) analysed two American populations (White and Black) from the 20th Century and an 18th/19th Century British population whilst Gómez-Valdéz et al. (2012) only used a contemporary Mexican population. Considering the lower accuracy, the greater sciatic notch identified females at a greater rate than males giving further support to Walker (2005) and Gómez-Valdéz et al. (2012) results.

After observing the three original Phenice (1969) traits using Klales et al. (2012) description, the ventral arc achieved the lowest accuracy. Overall, the ventral arc had an accuracy of 73% across the eight populations which is considerably lower than those accuracies found in other studies. Roger & Saunders (1994) found the ventral arc to correctly identify sex with an 86.9% accuracy. Arsuaga & Carretero (1994) found a similar accuracy of 81.7%. However, these accuracies were a result of using a 3-point grading system "Male, Ambiguous, Female" which does not consider sexual variation of the trait, unlike the Klales et al. (2012) descriptions. They found that when applying a logistic regression onto the ventral arc, accuracies as high as 88.5% could be achieved. The Phenice (1969) trait that performed second best was the subpubic concavity. With an averaged accuracy of 79.52%, it performed best for the Andaman and South African White populations (88.89% and 83.94% respectively). Arsuaga & Carretero (1994) achieved an overall accuracy for 88.6% which is similar to that of Klales et al. (2012) (86.6%) and the Andaman population from this research. Sadly, the same accuracy was not recreated for the Victorian/Georgian British population of Christ Church, Spitalfields resulting in an accuracy as low as 70%. This difference in accuracy could be a result of population differences in sexual dimorphism. The medial aspect of the ischiopubic ramus had the highest accuracy out of the three Phenice (1969) traits (80.09%). Arsuaga & Carretero (1994) found that this trait had the lowest accuracy out of three with 76.7% and the same for Klales et al. (2012) (75.8%). Even with this opposite result, accuracies between the studies are only minor, with a difference of <4%.

As far as single morphoscopic traits are concerned, populational differences and temporal differences are evident. This is definitely the case for the obturator foramen, greater sciatic notch and preauricular sulcus; not because of their lower accuracies, but due to their effectiveness fluctuating across the eight samples studied. The Phenice (1969) traits show slight population differences, more so for the ventral arc rather than the subpubic concavity and medial aspect of the ischiopubic ramus. This is because the latter two have a more stable accuracy across the samples. There are minor differences in the shape of the pubic body and subpubic angle traits when sexing across the populations. These two traits would seem to be the most "stable" and most likely to be universal for sexing human remains. This can be explained by the female pelvis having to accommodate for both bipedalism and parturition (Tague 1992). Listi (2010) found that the "variation" seen was more related to differences in the interpretation of each trait rather than morphological dissimilarities between the populations she studied. With the results from the morphoscopic traits in Chapters 5 and 6, the concept of population specifity is strengthened especially concerning single trait analysis.

Using single traits for sex estimation do have strengths being that they can be used on fragmented specimens without risking the application of a multivariate analysis. However, using single traits for analysis is unwise and one should always utilise multiple traits where possible (Bruzek 2002).

With the main aspect of the morphoscopic analysis being used as a baseline for the multivariate analyses it was hypothesised that they would not yield a greater accuracy when compared against the more robust analysis. All eight morphoscopic traits, including the subpubic angle, did not perform as well as the least effective equation (South African Black) as seen in Chapters 5 and 6. The only exception to this was the Chumash sample. Reasons as to why this may be the case is discussed below. This supports the claim that multivariate approaches to identification are superior to univariate analyses (Liebenberg 2014).

8.3 Multivariate Analysis

As stated previously, the multivariate analyses outperformed the single trait analysis in regard to all the eight samples.

Firstly, when observing the results from the four known age and sex samples, all equations achieved accuracies greater than 90% after a leave-one-out cross validation procedure. How well do these results compare to other sex estimation techniques using morphoscopic traits? Walker (2008) found that his most effective logistic regression formula, that used all five cranial traits, correctly classified a total of 87.8% of the modern skulls in his sample. This is much lower than the least effective population specific logistic regression created here, where its accuracy was 90.24% (South African Black equation). Garvin et al. (2014) applied the five cranial traits described by Walker (2008) but instead of using logistic regressions to calculate accuracies, they instead used canonical discriminant function analysis. With this controversial technique, they managed to achieve accuracies that ranged from 81.0% to 86.1% when collections were pooled to help create larger ancestry samples. Even with a more discriminatory statistical approach, accuracies did not reach the same heights as the four population specific equations, and the two clustered equations seen here. Krüger et al. (2015) used OLR to analyse the five Walker (2008) traits on modern-day White and Black South Africans. They found that the most effective White South African equation had an accuracy of 93%, however, when using the pelvic traits in this study, an increase of 2.34% was achieved on the same sample population. When Krüger et al. (2015) created their Black South African equation, accuracy was significantly lower, which mirrors the results seen in this research where the South African Black equation resulted in an accuracy of 90.24%. This lower classification has been attributed to South Africans showing decreased levels of sexual dimorphism compared to other populations (mainly North Americans) (L'Abbé et al. 2013, Krüger et al. 2015). When looking at the results from the archaeological samples, accuracies ranged from 100% (St. Owens sample) to 82.35% (Chumash sample). Walker (2008) also applied his logistic regressions onto a Native American archaeological skeletal collection from Santa Barbara, California. He managed to achieve an accuracy of 78% after comparing the results to pelvic morphology and museum records. The collection from Santa Barbara is also from the Chumash Native American tribe however their collection spans over several time periods rather than the just the Early Period (Walker 2006). With that being said, pelvic morphoscopic features being applied to OLR was more successful than the cranial traits by 5.35%. The fact that all of the equations from this piece of research yielded higher accuracies than those from cranial analyses re-enforces the idea that the pelvis is more sexually dimorphic than the skull. This is most likely due to the selection pressures the human pelvis is under, rather than the plastic nature of the skull (Spradley & Jantz 2011).

Moving from skull morphology, how did the results from the OLR analysis compare with osteometric analysis of the pelvis? Patriquin et al. (2004) found that the most effective pelvic measurement in their sample of 400 South Africans was the ischial length for South African Whites (averaged accuracy of 86%), and the acetabulum diameter for South African Blacks (averages accuracy of 84%). When all nine of their pelvic measurements were used in the DFA, accuracies of 95.5% and 94% were found. These high accuracies are similar to the results from this study for the four known age and sex samples. Accuracies for the South African White samples are practically the same, the only difference being a 0.16% lower accuracy for the morphological analysis. The biggest difference was the South African Black sample and equation. As mentioned previously, in this study, a 90.24% accuracy was found which is considerably lower than the 94% that was achieved by Patriquin et al. (2004). This difference could be due to not including morphoscopic features that highlight the acetabulum, which exhibits large differences between black males and females from South Africa. The results from this study also marry well with the results found when applying six pelvic measurements to a modern-day Greek sample where their accuracy was 93.5% (cross-validated) (Steyn & Iscan 2008). Normally, multivariate statistics focus on using traditional linear measurements or geometric morphometric data especially in regard to sex classification. Gómez-Valdéz et al. (2011) found that when applying Step-Wise Discriminant Analysis (S-WDA) on nine coxal measurements, accuracy can range between 99.1% to 79.2% depending on a number of variables used. They found that with a maximum of four measurements would give them the highest accuracy. Yet, for these measurements to be taken, complete bones are needed which can be a rarity in both forensic and archaeological cases (Waldron 1987, Stojanowski et al. 2002).

Klales et al. (2012) study on redefining the Phenice method (Phenice 1969) created a cross validated accuracy of 94.5% when experienced observers were collecting data. Klales et al. (2012) results sit firmly within the range of accuracies found in this research; which further validates that their technique is something that can be used universally. Sadly, the results from this study fall short of the 100% accuracy the recalibrated Klales et al. (2012) equation had for the Mexican sample (Gómez-Valdés et al. 2016). Kenyhercz (2012) applied Klales et al. (2012) trait descriptions to South African Whites and Blacks and found that accuracy of sex estimation increased when applying multivariate statistics rather than just the morphoscopic scores which can be seen in Chapter 5. Stull et al. (2013) used a wide array of multivariate analyses for sex estimation using both cranial and pelvic morphoscopic traits. They found that the os coxa was the most reliable even when compared against cranial traits and applying logistic regressions to the scores was overall the better analytical approach. Stull et al. (2013) achieved an accuracy of 99.2% using logistic regressions. Alongside the logistic regressions, they also analysed the pelvic traits using two different DFA's; LDA and FDA (Flexible Discriminant Analysis). These DFA analyses gave a cross validated accuracy of 96.5%. The OLR results in this study range between 90.24% (South African Black) to 96.38% (Christ Church, Spitalfields). When the South African samples were pooled to create the South African equation, accuracy was 92.54%, which is still considerably lower than Stull et al. (2013) accuracy of 99.2%. The difference in accuracies could relate to Stull et al. (2013) having a much smaller sample size than this study and pooling all of three populations together without analysing them separately first. Also, even though they state they had three samples, only having one South African Coloured individual does not equal a sample representative of a population. The other thing to take into consideration with this, is that Stull et al. (2013) used statistical techniques that are used for continuous variables rather than ordinal variables (e.g. DFA). This may have had an effect on the results they had calculated.

To further test Klales' (2012) ordinal logistic regression equation for sex determination, Klales et al. (2016) applied her technique to a worldwide population which included: U.S. White, Black and, "Other" (Native American, Hispanic, and Asian), Thai, and South African Whites and Blacks. From their original equation, accuracies ranged between 75.9%-93.3%. When they recalibrated the equation to each specific population and to include a "global" sample, accuracies increased dramatically to 97.9%-99.9%. Comparing their results to this study's "Summary Sex" equation, a difference of 4.92% - 6.92% was found. They concluded that "sexual dimorphism of the pelvis exceeds population differences", which coincide with the results from this research. Klales et al. (2016) also state that their original equation shows a classification bias towards females. This can be explained by the fact that when scoring the morphoscopic features, the observers are automatically looking for the female morphology rather than the male. For instance, the presence of a ventral arc is a female trait, alongside the concavity of the subpubic ramus. With this being said, results found in Chapter 5 show that classification was biased towards males for the majority of equations which contradicts this viewpoint.

Steyn & Patriquin (2009) observed seven metric measurements on the os coxa in an attempt to understand the cost of using population specific equations on vastly different populations. After creating three Step-Wise Discriminant functions to create the population specific equations, they then created an overall equation pooling all three samples. Their results show that when pooling populations and applying multivariate statistics have similar results to this study, however, the notion of population specificity comes into question. This is because accuracies did not significantly differ between the populations being analysed separately or being analysed as a whole. However, Steyn & Patriquin (2009) did not apply one of the population specific equation solely onto another. This leads to question, would the accuracies have differed if this was done? From the results found in Chapters 5, 6, and 7, applying the different OLR equations onto other populations saw that accuracies did not differ significantly. This is also the case when applying pooled population equations onto other samples as well. They say one of the potential reasons why this is the case is due to having a large sample size. A larger sample size "smooth's out" the small differences that are apparent in relatively differing populations making population specifics (which focuses on these small dissimilarities) null and void. However, the revised Phenice method when applied onto a similar skeletal sample had accuracies that dropped to 86.2% (Klales et al. 2012). When the OLR equations from Chapter 5 were applied onto other samples, the largest discrepancy in accuracies was when the South African Black equation was applied to the South African Coloured sample. This led to a decrease in accuracy of nearly 4%, which is nearly half the size of the near 8% that was recorded by Klales et al. (2012). Therefore, having larger sample sizes does, to some extent, "smooth out" the small differences between populations, but not to the same level that Steyn & Patriquin (2009) were stating.

Klales (2016) reviewed the possibility of secular differences in pelvic trait expressions using the Hamann-Todd Collection and Bass Donated Collection. She noted that secular differences were apparent for all of the traits for females and the subpubic concavity and ventral arc for males. However, considering that significant differences were found for how the traits were expressed, the descriptions set by Klales et al. (2012) are still applicable to archaeological and modern populations. Differences in trait expression were noted but were not fully tested to the same degree as Klales (2016).

8.4 Dealing with Fragmentation

8.4.1 Median Score

Fragmented human remains give rise to many issues, especially when creating a biological profile. Unless the remains can be reconstructed (which can create reliability problems), certain analyses have to be omitted. This research highlighted a potential solution for dealing with this issue when performing morphoscopic analysis. Because data were scored in a similar fashion to a Lickert scale, the median was the measure of central tendency (Manikandan, 2011). As can be seen from Chapters 5, 6, and 7, there were slight differences in the accuracies between the equations that only used complete remains and those that had the addition of fragmented remains with median values substituting missing scores. These slight differences show that medians can be used because it follows the amount of sexual dimorphism already observed in the os coxae. As discussed by Kenyhercz et al. (2016), the median value was the best at replacing missing scores when a large amount of data was already missing. This can definitely be said for the two British Medieval sample of Poulton and St. Owens being highly fragmented. The technique of using the median score is very easy to apply, rather than the weighted means approach that Kjellström (2004) applied to the Sigtuna sample, and also does not dampen the amount of sexual dimorphism that is present. The idea of using a neutral score to help account for this is an odd approach.

8.4.2 Smaller Equations

To try and include more remains that were fragmented, and without using an averaged data point to make a larger equation usable, smaller equations were investigated. Three equations were chosen, all with varying success. The "Posterior Pelvis" equations were the least successful out of the three with a large decline in overall accuracies for all eight samples. This result is not surprising because both the greater sciatic notch and the preauricular sulcus were not great performers at accurately estimating sex (see chapters 5 and 6 for original results for the morphoscopic traits); however, accuracies were slightly higher when combining the two together and applying them to OLR. As far as results go, when the posterior pelvis criteria were applied to the Chumash sample, accuracies mirrored those found by Walker (2008) when using cranial traits for sex estimation. A positive for these equations was that more individuals were included, especially considering how fragmented the Poulton and St. Owens Medieval samples were. This is because the posterior portion of the pelvis is more robust and tends to have a better survival rate than other areas of the pelvis (and skeleton). However, these sets of equations should only be used as a last resort because better equations are available.

The “Anterior Pelvis” equations were better suited to estimating sex. For all samples except the Chumash sample, accuracies were greater than 85%. Indeed the “Anterior Pelvis” equations were the most successful (both sectioned and new) in sex estimation than any of the other equations. This can be explained by the os pubis having a higher degree of sexual dimorphism than the posterior portion of the os coxae. This is idea was highlighted by Bondioli et al. (2006) who estimated sex for a one million-year-old pubic bone fragment. They state that the pubic bone held enough sufficient information for an accurate sex assessment. That said, Rogers & Saunders (1994) found that both the subpubic angle and shape of the pubic body were ranked ninth and sixth (respectively) out of 17 pelvic traits for sex estimation, being outperformed by the sacrum shape and obturator foramen(!). This difference in the traits of the anterior pelvis being less effective can be explained by how they were scored. This study used a five-point ordinal scale, whilst Rogers & Saunders (1994) only scored using a male expression and a female expression. This meant that the amount of sexual variation was not considered.

The sectioned and new Phenice equations can be more readily compared to research by Klales et al. (2012), Kenyhercz et al. (2012), Stull et al. (2013), Klales et al. (2016), and Gómez-Valdés et al. (2016). As the same descriptions were used for scoring each of the three traits and two of the same populations being studied, accuracies should be similar. The biggest difference in accuracy was for the South African Black sample. The accuracy of this sample was much lower in this study than those found by Kenyhercz et al. (2012) and Stull et al (2013). A reason as to why such a difference could have occurred would be due to different collections being used. In this study, the South African Black sample was created by pooling individuals from the Pretoria Bone Collection and the Raymond A. Dart Collection, whilst Kenyhercz et al. (2012) and Stull et al (2013) only used individuals from the Pretoria Bone Collection. This may sound like a small difference, however, individuals that were accessioned into the Raymond A. Dart Collection before 1959 have been known to not have the correct information due to the collection room being flooded. The same can be said for the South African White sample, however not as an extreme difference in accuracies were noted. The Phenice equations (both sectioned and new) did not perform as well the “Anterior Pelvis” equation which can be seen as being surprising. This is because the original

Phenice method (Phenice 1969) and Klales et al. (2012) reinterpretation of the technique hold such a strong opinion for being one of the best morphoscopic techniques to use. But this does not mean that the Phenice method has had its criticism over the decades (e.g. Lovell 1989).

Even though the use of smaller equations is better adapted for fragmented remains, their loss in accuracy is their shortfall. This has been recorded in not just for morphoscopic analysis (this study) but also for metric analysis on the pelvis (Patriquin et al. 2004; Steyn & Iscan 2008; Steyn & Patriquin 2009).

8.5 Limitations

The main limitation that this research has is the use of morphoscopic sexing traits and scoring them along an ordinal scale. The main risk of conducting this type of research is the high possibility of inter and intra observer error, and the general need for an experienced researcher who can consistently score accurately. To try and overcome both the issues of inter/intra observer error and level of experience, using morphoscopic traits with good detailed descriptions and illustrations and/or images that complement each other were used. However, this was only available for the greater sciatic notch (Walker 2005) and the three Phenice traits: ventral arc, subpubic concavity, and the medial aspect of the ischiopubic ramus (Phenice 1969, Klales et al. 2012). Sadly, this does not mean that they are the traits that have the highest accuracy, but all yield moderately high accuracies (subpubic angle and shape of the pubic body being the most accurate). This availability for these four traits is because of their popularity with practitioners. For the subpubic angle, obturator foramen and shape of the pubic body, descriptions and illustrations were expanded for this research and are found in Figures 21, 22, and 23. Images and descriptions for the preauricular sulcus were not expanded for this research but standard scoring was used (Buikstra & Ubelaker 1994). When looking at previous literature on the issue of using morphoscopic traits with observer error, Vance (2007) found no significant difference between her intra observer analysis. However, no descriptions of how she scored the traits are present so performing an inter observer error analysis is impossible. Lavallo (2013) found that there were no differences between scoring the same five traits when testing if there were any intra observer differences. Blanchard (2010) assessed the three Phenice traits (Phenice 1969), greater sciatic notch (Walker 2005) and the preauricular sulcus (Buikstra & Ubelaker 1994) and achieved substantial agreement levels when testing for intra observer differences and moderate agreement results for her inter observer analysis. Comparing her intra/inter observer results to that of this study, no main differences can be seen apart from the preauricular sulcus trait. She found that it was the trait with very low repeatability between observers whilst this study found the complete opposite. The contrast between results in the observer error can be explained by experience. For Blanchard (2010), her second observer was an anthropologist without any prior experience with scoring the preauricular sulcus using this technique; whilst in this study, the second observer had prior experience using the scoring method described by Buikstra & Ubelaker (1994).

Another limitation for this project relates to the Native American Chumash collection, currently held at the Natural History Museum (London). Considering that the remains are approximately 4,000 years old and they are well preserved, issue lies with the museum records. There are many inconsistencies between what is stated in the museum records and what is actually present for that individual. The sample size for this population would have been greater but several individuals were omitted. This was because these skeletons were described as being subadults (<18 years old) even though the iliac crest and femoral head were fully fused. With the skeletons being classed however as subadult, sex was not estimated when the remains were originally curated. The age at death given for these individuals ranged from 12 months to 12 years, despite the skeletons being fully mature adults. Alongside the issue that some individuals were aged incorrectly, problems were found with the original inventory. Many specimens that stated a partial or complete os coxae being present were found not to have any fragments present at all. For this sample's analysis, estimated sex from the multivariate statistics was compared against the sex noted in the museum records. There could be a case that the sex stated in these records could be wrong, so caution should be taken with the results regarding the Chumash population. This could explain the lower accuracy seen in the Chumash population when compared against the other seven populations. What this limitation highlights are the need for a more thorough and reliable data collection for skeletal assemblages/archaeological excavations. Handling old skeletal assemblages with potentially wrong data makes trying to understand past population problematic.

From the Chumash population, another limitation surrounds the two medieval populations; the urban St. Owens (N=100) and the rural Poulton (N=132) collections. The main limitation with these was that sex was estimated using seriation meaning sex was not known. Sadly, due to poor preservation of the Poulton skeletons, ancient DNA (aDNA) could not be extracted (Town 2015) and sex could not be determined for a subsample of the remains. Ancient DNA has not been attempted for the St. Owens collection to date.

The final limitation that relates to populations, is the lack of an Asian sample. Even though a prehistoric Native American sample and a small pygmy sample from the Andaman Islands have been analysed, it is by no means a substitute for this geographic group. One reason why data were not collected and analysed for this geographic group was due to a lack of funding for the project, which allowed only limited travel and consequently a reliance on remains curated in the UK.

9. Conclusions and further study

This piece of PhD research created several sex estimation equations using eight morphoscopic traits on the human pelvis. For those equations that were population specific, accuracies ranged from 90.24% to 96.38%. From these four known and sex samples, an investigation of population specificity was conducted. It was found that after applying the different equations to the other three samples, that accuracy did not alter greatly. The greatest accuracy difference was seen in the South African Black equation that was applied to the South African Coloured sample. A difference of 3.91% was found. As well as separate population specific equation, a country wide equation was calculated that combined three South African samples together. This South Africa equation had an overall accuracy of 92.54% (cross validated). When exploring the possibility of a universal equation, an overall accuracy of 92.98% was calculated after a leave-one-out cross validated procedure was undertaken. This equation was called "Summary Sex".

This piece of work produced the following:

- Several cross-validated equations that can be used specifically for their sample population and would be able to aid in forensic human identification (South African Samples) (Chapters 5, 7, and Appendices A – F).
- Two cross-validates equations that can be used over a wider geographic and temporal span which can be used to aid in forensic human identification (Chapters 5, 7, and Appendices A – F).
- Three new definitions for three morphoscopic sexing traits that previously did not have details (descriptions and/or images) on how to use them for sex determination (Chapter 4) with associated intra/inter observer analysis (Chapters 5 and 6).
- An investigation into fragmented remains and how to overcome this issue when using multivariate equations.

However, for this technique to be classed as truly universal more research is required. Because there was a lack of an Asian population, obtaining data from a known skeletal collection would be needed. After expanding and incorporating the Asian data into the "World" analysis updating the "Summary Sex" equation would be needed. The addition of an American sample would also be beneficial; samples from collections such as the Hamann-Todd human skeletal collection (Cleveland Museum of Natural History) and the Robert J. Terry Anatomical Skeletal Collection (Smithsonian, National Museum of Natural History). The additional Asian and American populations would also count as further investigation into a universal equation. This research also has the potential to include other areas of the human skeleton such as the skull. Would the skull show any populational differences when applying OLR to the Buikstra & Ubelaker (1994) and Walker (2008) morphoscopic scoring methods? Would a mix of cranial and pelvic traits yield a higher accuracy similarly to what was observed by Meindl *et al*. (1985)? Further analysis of secular changes of pelvic traits could be explored, especially concerning the three South African samples.

With this increased sample, further testing of using the median score as a method for missing data imputation could be investigated. Alongside the median score being used, other techniques such as Hotdesk and k Nearest Neighbour imputation would be explored as possible avenues for missing data. With the idea of population specificity, would a particular technique be more beneficial for a certain population, or would the same missing data imputation technique work for all? This would also help with the validation of the smaller equations created in Chapter 7 Results III. Would these more robust equations be better adapted for archaeological samples?

Considering that only visual analysis had been conducted, it would be possible to expand this project using geometric morphometric (GM) analysis. Would it be possible to map the relationships between morphoscopic traits using GM analysis? This would be answered by collecting landmark and semi landmark data firstly from the anterior portion of pelvis then posterior to which statistical methods like General Procrustes Analysis, Thin Plate Splines, Principal Component Analysis, and K-means Cluster Analysis would be utilised to assess the relationships seen between certain traits. Folding in the GM analysis with the visual techniques could provide a better understanding of pelvic morphology, especially concerning obstetrics. This type of data can be collected from several sources like radiography (Computed Tomography images), 2D/3D photogrammetry, 3D laser surface scans, and/or using a 3D Microscribe to digitise the landmarks from the physical specimen.

From only analysing modern anatomical Homo sapiens, expanding the methodology and technique (including the Geometric Morphometric Analysis) to analyse hominin remains (i.e. *Ardipithecus ramidus, Australopithecus aferensis,* and *Homo neanderthalensis*) and other extinct/extant primates (i.e. *Proconsul africanus, Pan troglodytes*, and *Pongo abelii*) will be necessary. A preliminary study has already been conducted (Rennie *et al*. 2015) but needs to be expanded. This research could help to indicate when certain morphoscopic traits started to become sexually dimorphic and what lead those traits to become sexually dimorphic.

Appendices

Appendix A. Sectioned Phenice OLR

Appendix A.1 Sectioned OLR Equation: Phenice

$$Christ\ Church, Spitalfields\ \ Equation = (0.962 * Ventral\ Arc\ Score) + (0.931 * Subpubic\ Concavity\ Score) + (0.986 * Ischiopubic\ Ramus\ Score) + \mathbf{0.000}$$

$$SA: White\ Equation = (0.980 * Ventral\ Arc\ Score) + (1.147 * Subpubic\ Concavity\ Score) + (1.449 * Ischiopubic\ Ramus\ Score) + \mathbf{0.000}$$

$$SA: Black\ Equation = (1.702 * Ventral\ Arc\ Score) + (0.750 * Subpubic\ Concavity\ Score) + (1.378 * Ischiopubic\ Ramus\ Score) + \mathbf{0.000}$$

$$SA: Coloured\ Equation = (0.828 * Ventral\ Arc\ Score) + (1.468 * Subpubic\ Concavity\ Score) + (1.016 * Ischiopubic\ Ramus\ Score) + \mathbf{0.000}$$

$$South\ African\ Equation = (0.985 * Ventral\ Arc\ Score) + (1.053 * Subpubic\ Concavity\ Score) + (1.171 * Ischiopubic\ Ramus\ Score) + \mathbf{0.000}$$

$$Summary\ Sex\ Equation = (0.942 * Ventral\ Arc\ Score) + (1.018 * Subpubic\ Concavity\ Score) + (1.196 * Ischiopubic\ Ramus\ Score) + \mathbf{0.000}$$

Appendix A.2 Tabulated Results

Appendix A.2 Table 2: Classification results for the Christ Church, Spitalfields sample (N=158).

OLR Equation	Correct Females	Correct Males	Overall Agreement (%)
Christ Church, Spitalfields	71/77 (92.21%)	69/81 (85.19%)	88.61%
South African White	73/77 (94.81%)	69/81 (85.19%)	89.87%
South African Black	72/77 (93.51%)	69/81 (85.19%)	89.24%
South African Coloured	72/77 (93.51%)	69/81 (85.19%)	87.97%
South African	73/77 (94.81%)	69/81 (85.19%)	89.87%
Summary Sex	73/77 (94.81%)	69/81 (85.19%)	89.87%

Appendix A.2 Table 3: Classification results for the South African White sample (N=193).

OLR Equation	Correct Females	Correct Males	Overall Agreement (%)
Christ Church, Spitalfields	89/98 (90.81%)	92/95 (96.84%)	93.78%
South African White	90/98 (91.84%)	92/95 (96.84%)	94.30%
South African Black	89/98 (90.81%)	92/95 (96.84%)	93.78%
South African Coloured	90/98 (91.84%)	88/95 (92.63%)	92.23%
South African	90/98 (91.84%)	92/95 (96.84%)	94.30%
Summary Sex	90/98 (91.84%)	92/95 (96.84%)	94.30%

Appendix A.2 Table 4: Classification results for the South African Black sample (N=205).

OLR Equation	Correct Females	Correct Males	Overall Agreement (%)
Christ Church, Spitalfields	73/97 (75.26%)*	100/108 (92.59%)	84.39%
South African White	75/97 (76.53%)*	99/108 (91.67%)	84.88%
South African Black	70/97 (72.16%)*	99/108 (91.67%)	82.44%
South African Coloured	75/97 (76.53%)*	99/108 (91.67%)	84.88%
South African	75/97 (76.53%)*	99/108 (91.67%)	84.88%
Summary Sex	75/97 (76.53%)*	99/108 (91.67%)	84.88%

*One female was classified as unknown.

Appendix A.2 Table 5: Classification results for the South African Coloured sample (N=205).

OLR Equation	Correct Females	Correct Males	Overall Agreement (%)
Christ Church, Spitalfields	83/103(80.58%)	98/102 (96.08%)	88.29%
South African White	89/103(87.25%)	98/102 (96.08%)	91.22%
South African Black	81/103(78.64%)	99/102 (97.06%)	87.80%
South African Coloured	90/103(88.24%)	99/102 (97.06%)	92.20%
South African	89/103(87.25%)	98/102 (96.08%)	91.22%
Summary Sex	89/103(87.25%)	98/102 (96.08%)	91.22%

Appendix A.2 Table 6: Classification results for the Poulton sample (N=67).

OLR Equation	Correct Females	Correct Males	Overall Agreement (%)
Christ Church, Spitalfields	32/34(94.12%)	28/33 (84.85%)	89.55%
South African White	33/34(97.06%)	29/33 (85.29%)	92.54%
South African Black	31/34(91.18%)	24/33 (70.58%)	82.90%
South African Coloured	32/34(94.12%)	30/33 (88.24%)	92.54%
South African	33/34(97.06%)	29/33 (96.08%)	92.54%
Summary Sex	33/34(97.06%)	29/33 (96.08%)	92.54%

Appendix A.2 Table 7: Classification results for the St. Owens sample (N=44).

OLR Equation	Correct Females	Correct Males	Overall Agreement (%)
Christ Church, Spitalfields	22/24 (91.67%)	19/20 (95.00%)	93.18%
South African White	23/24 (95.83%)	19/20 (95.00%)	95.45%
South African Black	22/24 (91.67%)	19/20 (95.00%)	93.18%
South African Coloured	24/24 (100.00%)	18/20 (90.00%)	95.45%
South African	23/24 (95.83%)	19/20 (95.00%)	95.45%
Summary Sex	23/24 (95.83%)	19/20 (95.00%)	95.45%

Appendix A.2 Table 8: Classification results for the Chumash sample (N=35).

OLR Equation	Correct Females	Correct Males	Overall Agreement (%)
Christ Church, Spitalfields	13/17 (76.47%)	16/18 (88.89%)	88.29%
South African White	13/17 (76.47%)	16/18 (88.89%)	88.29%
South African Black	13/17 (76.47%)	16/18 (88.89%)	88.29%
South African Coloured	13/17 (76.47%)	16/18 (88.89%)	88.29%
South African	13/17 (76.47%)	16/18 (88.89%)	88.29%
Summary Sex	13/17 (76.47%)	16/18 (88.89%)	88.29%

Appendix A.2 Table 9: Classification results for the Andaman sample (N=27).

OLR Equation	Correct Females	Correct Males	Overall Agreement (%)
Christ Church, Spitalfields	10/13 (76.92%)*	14/14 (100.00%)	88.89%
South African White	10/13 (76.92%)*	14/14 (100.00%)	88.89%
South African Black	10/13 (76.92%)*	14/14 (100.00%)	88.89%
South African Coloured	10/13 (76.92%)*	14/14 (100.00%)	88.89%
South African	10/13 (76.92%)*	14/14 (100.00%)	88.89%
Summary Sex	10/13 (76.92%)*	14/14 (100.00%)	88.89%

*Two females were classified as unknown.

Appendix B. Sectioned Posterior Pelvis OLR

Appendix B.1 Sectioned OLR Equation: Posterior Pelvis

$$Christ\ Church, Spitalfields\ \ Equation = (1.130 * Greater\ Sciatic\ Notch\ Score) + (1.163 * Preauricular\ Sulcus\ Score) + \mathbf{0.000}$$

$$SA{:}\, White\ Equation = (1.547 * Greater\ Sciatic\ Notch\ Score) + (1.819 * Preauricular\ Sulcus\ Score) + \mathbf{0.000}$$

$$SA{:}\, Black\ Equation = (0.993 * Greater\ Sciatic\ Notch\ Score) + (3.014 * Preauricular\ Sulcus\ Score) + \mathbf{0.000}$$

$$SA{:}\, Coloured\ Equation = (1.713 * Greater\ Sciatic\ Notch\ Score) + (1.482 * Preauricular\ Sulcus\ Score) + \mathbf{0.000}$$

$$South\ African\ Equation = (1.142 * Greater\ Sciatic\ Notch\ Score) + (2.137 * Preauricular\ Sulcus\ Score) + \mathbf{0.000}$$

$$Summary\ Sex\ Equation = (1.425 * Greater\ Sciatic\ Notch\ Score) + (1.755 * Preauricular\ Sulcus\ Score) + \mathbf{0.000}$$

Appendix B.2 Tabulated Results

Appendix B.2 Table 1: Classification results for the Christ Church, Spitalfields sample (N=214).

OLR Equation	Correct Females	Correct Males	Overall Agreement (%)
Christ Church, Spitalfields	79/108 (73.15%)	99/106 (93.40%)*	83.18%
South African White	79/108 (73.15%)	99/106 (93.40%)*	83.18%
South African Black	78/108 (72.22%)	99/106 (93.40%)*	82.71%
South African Coloured	94/108 (87.04%)	95/106 (89.62%)*	88.32%
South African	79/108 (73.15%)	99/106 (93.40%)*	83.18%
Summary Sex	79/108 (73.15%)	99/106 (93.40%)*	83.18%

*One male was classified as unknown.

Appendix B.2 Table 2: Classification results for the South African White sample (N=193).

OLR Equation	Correct Females	Correct Males	Overall Agreement (%)
Christ Church, Spitalfields	76/98 (79.17%)*	91/95 (95.79%)*	86.53%
South African White	76/98 (79.17%)*	91/95 (95.79%)*	86.53%
South African Black	65/98 (66.33%)*	91/95 (95.79%)*	80.83%
South African Coloured	82/98 (83.67%)*	88/95 (92.63%)*	88.08%
South African	76/98 (79.17%)*	91/95 (95.79%)*	86.53%
Summary Sex	76/98 (79.17%)*	91/95 (95.79%)*	86.53%

*Two females and one male were classified as unknown

Appendix B.2 Table 3: Classification results for the South African Black sample (N=205).

OLR Equation	Correct Females	Correct Males	Overall Agreement (%)
Christ Church, Spitalfields	66/97 (68.04%)*	102/108 (94.44%)*	81.95%
South African White	66/97 (68.04%)*	102/108 (94.44%)*	81.95%
South African Black	64/97 (65.98%)*	104/108 (96.30%)*	81.95%
South African Coloured	66/97 (68.04%)*	99/108 (91.67%)*	80.49%
South African	66/97 (68.04%)*	102/108 (94.44%)*	81.95%
Summary Sex	66/97 (68.04%)*	102/108 (94.44%)*	81.95%

*One female and three males were classified as unknown.

Appendix B.2 Table 4: Classification results for the South African Coloured sample (N=205).

OLR Equation	Correct Females	Correct Males	Overall Agreement (%)
Christ Church, Spitalfields	60/103 (58.25%)*	102/102 (100.00%)	79.02%
South African White	60/103 (58.25%)*	102/102 (100.00%)	79.02%
South African Black	44/103 (42.72%)*	102/102 (100.00%)	71.22%
South African Coloured	82/103 (79.61%)*	98/102 (96.08%)	87.80%
South African	60/103 (58.25%)*	102/102 (100.00%)	79.02%
Summary Sex	60/103 (58.25%)*	102/102 (100.00%)	79.02%

*One female was classified as unknown

Appendix B.2 Table 5: Classification results for the Poulton sample (N=125).

OLR Equation	Correct Females	Correct Males	Overall Agreement (%)
Christ Church, Spitalfields	56/63 (88.89%)	50/62 (80.65%)	84.80%
South African White	56/63 (88.89%)	50/62 (80.65%)	84.80%
South African Black	56/63 (88.89%)	50/62 (80.65%)	84.80%
South African Coloured	55/63 (87.30%)	54/62 (87.10%)	87.20%
South African	56/63 (88.89%)	50/62 (80.65%)	84.80%
Summary Sex	56/63 (88.89%)	50/62 (80.65%)	84.80%

Appendix B.2 Table 6: Classification results for the St. Owens sample (N=98).

OLR Equation	Correct Females	Correct Males	Overall Agreement (%)
Christ Church, Spitalfields	33/44 (75.00%)*	48/54 (88.89%)*	82.65%
South African White	33/44 (75.00%)*	48/54 (88.89%)*	82.65%
South African Black	33/44 (75.00%)*	48/54 (88.89%)*	82.65%
South African Coloured	35/44 (79.55%)*	48/54 (88.89%)*	84.69%
South African	33/44 (75.00%)*	48/54 (88.89%)*	82.65%
Summary Sex	33/44 (75.00%)*	48/54 (88.89%)*	82.65%

*One female and two males were classified as unknown.

Appendix B.2 Table 7: Classification results for the Chumash sample (N=38).

OLR Equation	Correct Females	Correct Males	Overall Agreement (%)
Christ Church, Spitalfields	16/19 (84.21%)	12/19 (63.16%)	75.68%
South African White	16/19 (84.21%)	12/19 (63.16%)	75.68%
South African Black	16/19 (84.21%)	10/19 (52.63%)	70.27%
South African Coloured	16/19 (84.21%)	12/19 (63.16%)	75.68%
South African	16/19 (84.21%)	12/19 (63.16%)	75.68%
Summary Sex	16/19 (84.21%)	12/19 (63.16%)	75.68%

Appendix B.2 Table 8: Classification results for the Andaman sample (N=27).

OLR Equation	Correct Females	Correct Males	Overall Agreement (%)
Christ Church, Spitalfields	10/13 (76.92%)	13/14 (92.86%)	85.19%
South African White	10/13 (76.92%)	13/14 (92.86%)	85.19%
South African Black	10/13 (76.92%)	13/14 (92.86%)	85.19%
South African Coloured	11/13 (84.62%)	13/14 (92.86%)	88.89%
South African	10/13 (76.92%)	13/14 (92.86%)	85.19%
Summary Sex	10/13 (76.92%)	13/14 (92.86%)	85.19%

Appendix C. Sectioned Anterior Pelvis OLR

Appendix C.1 Sectioned OLR Equation: Anterior Pelvis

$$Christ\ Church, Spitalfields\ \ Equation = (3.255 * Pubic\ Body\ Shape\ Score) + (2.464 * Subpubic\ Angle\ Score) + \mathbf{0.000}$$

$$SA: White\ Equation = (2.141 * Pubic\ Body\ Shape\ Score) + (0.998 * Subpubic\ Angle\ Score) + \mathbf{0.000}$$

$$SA: Black\ Equation = (1.453 * Pubic\ Body\ Shape\ Score) + (2.211 * Subpubic\ Angle\ Score) + \mathbf{0.000}$$

$$SA: Coloured\ Equation = (4.181 * Pubic\ Body\ Shape\ Score) + (1.134 * Subpubic\ Angle\ Score) + \mathbf{0.000}$$

$$South\ African\ Equation = (1.867 * Pubic\ Body\ Shape\ Score) + (1.551 * Subpubic\ Angle\ Score) + \mathbf{0.000}$$

$$Summary\ Sex\ Equation = (2.127 * Pubic\ Body\ Shape\ Score) + (1.521 * Subpubic\ Angle\ Score) + \mathbf{0.000}$$

Appendix C.2 Tabulated Results

Appendix C.2 Table 1: Classification results for the Christ Church, Spitalfields sample (N=262).

OLR Equation	Correct Females	Correct Males	Overall Agreement (%)
Christ Church, Spitalfields	80/81 (98.77%)	74/81 (91.36%)*	95.06%
South African White	80/81 (98.77%)	75/81 (92.59%)*	95.68%
South African Black	80/81 (98.77%)	75/81 (92.59%)*	95.68%
South African Coloured	80/81 (98.77%)	75/81 (92.59%)*	95.68%
South African	80/81 (98.77%)	74/81 (91.36%)*	95.06%
Summary Sex	80/81 (98.77%)	74/81 (91.36%)*	95.06%

*One male was classified as unknown.

Appendix C.2 Table 2: Classification results for the South Africa White sample (N=193).

OLR Equation	Correct Females	Correct Males	Overall Agreement (%)
Christ Church, Spitalfields	93/98 (94.90%)	88/95 (92.63%)*	93.78%
South African White	90/98 (91.84%)	88/95 (92.63%)*	92.22%
South African Black	89/98 (90.81%)	86/95 (90.53%)*	90.67%
South African Coloured	90/98 (91.84%)	88/95 (92.63%)*	92.22%
South African	90/98 (91.84%)	88/95 (92.63%)*	92.22%
Summary Sex	90/98 (91.84%)	88/95 (92.63%)*	92.22%

*Two males were classified as unknown.

Appendix C.2 Table 3: Classification results for the South African Black sample (N=205).

OLR Equation	Correct Females	Correct Males	Overall Agreement (%)
Christ Church, Spitalfields	80/97 (82.47%)*	100/108 (92.59%)*	87.80%
South African White	79/97 (81.44%)*	100/108 (92.59%)*	87.32%
South African Black	84/97 (86.60%)*	100/108 (92.59%)*	89.76%
South African Coloured	79/97 (81.44%)*	100/108 (92.59%)*	87.32%
South African	80/97 (82.47%)*	100/108 (92.59%)*	87.80%
Summary Sex	80/97 (82.47%)*	100/108 (92.59%)*	87.80%

*Two females and one male were classified as unknown.

Appendix C.2 Table 4: Classification results for the South African Coloured sample (N=205).

OLR Equation	Correct Females	Correct Males	Overall Agreement (%)
Christ Church, Spitalfields	91/103(88.35%)*	92/102 (90.20%)*	89.27%
South African White	91/103(88.35%)*	92/102 (90.20%)*	89.27%
South African Black	93/103(90.29%)*	91/102 (89.22%)*	89.76%
South African Coloured	91/103(88.35%)*	92/102 (90.20%)*	89.27%
South African	91/103(88.35%)*	92/102 (90.20%)*	89.27%
Summary Sex	91/103(88.35%)*	92/102 (90.20%)*	89.27%

*One female and six males were classified as unknown.

Appendix C.2 Table 5: Classification results for the Poulton sample (N=76).

OLR Equation	Correct Females	Correct Males	Overall Agreement (%)
Christ Church, Spitalfields	32/37 (86.49%)*	34/39 (87.18%)	86.84%
South African White	31/37 (83.78%)*	34/39 (87.18%)	85.53%
South African Black	33/37 (89.19%)*	34/39 (87.18%)	88.16%
South African Coloured	32/37 (86.49%)*	34/39 (87.18%)	86.84%
South African	32/37 (86.49%)*	34/39 (87.18%)	86.84%
Summary Sex	32/37 (86.49%)*	34/39 (87.18%)	86.84%

*One female was classified as unknown.

Appendix C.2 Table 6: Classification results for the St. Owens sample (N=46).

OLR Equation	Correct Females	Correct Males	Overall Agreement (%)
Christ Church, Spitalfields	24/24 (100.00%)	20/22 (90.91%)	95.65%
South African White	23/24 (95.83%)	20/22 (90.91%)	93.48%
South African Black	24/24 (100.00%)	20/22 (90.91%)	95.65%
South African Coloured	23/24 (95.83%)	20/22 (90.91%)	93.48%
South African	24/24 (100.00%)	20/22 (90.91%)	95.65%
Summary Sex	24/24 (100.00%)	20/22 (90.91%)	95.65%

Appendix C.2 Table 7: Classification results for the Chumash sample (N=37).

OLR Equation	Correct Females	Correct Males	Overall Agreement (%)
Christ Church, Spitalfields	13/18 (72.22%)	17/19 (89.47%)	81.08%
South African White	13/18 (72.22%)	17/19 (89.47%)	81.08%
South African Black	15/18 (83.33%)	17/19 (89.47%)	86.49%
South African Coloured	13/18 (72.22%)	17/19 (89.47%)	81.08%
South African	13/18 (72.22%)	17/19 (89.47%)	81.08%
Summary Sex	13/18 (72.22%)	17/19 (89.47%)	81.08%

Appendix C.2 Table 8: Classification results for the Andaman sample (N=27).

OLR Equation	Correct Females	Correct Males	Overall Agreement (%)
Christ Church, Spitalfields	11/13 (84.62%)	14/14 (100.00%)	92.59%
South African White	11/13 (84.62%)	14/14 (100.00%)	92.59%
South African Black	10/13 (76.92%)	14/14 (100.00%)	88.89%
South African Coloured	11/13 (84.62%)	14/14 (100.00%)	92.59%
South African	11/13 (84.62%)	14/14 (100.00%)	92.59%
Summary Sex	11/13 (84.62%)	14/14 (100.00%)	92.59%

Appendix D. New Phenice OLR

Appendix D.1 New OLR Equations: Phenice

$$Christ\ Church, Spitalfields\ \ Equation = (1.789 * Ventral\ Arc\ Score) + (1.969 * Subpubic\ Concavity\ Score) + (2.290 * Ischiopubic\ Ramus\ Score) + \mathbf{0.000}$$

$$SA: White\ Equation = (1.758 * Ventral\ Arc\ Score) + (1.860 * Subpubic\ Concavity\ Score) + (2.250 * Ischiopubic\ Ramus\ Score) + \mathbf{0.000}$$

$$SA: Black\ Equation = (1.349 * Ventral\ Arc\ Score) + (1.898 * Subpubic\ Concavity\ Score) + (1.828 * Ischiopubic\ Ramus\ Score) + \mathbf{0.000}$$

$$SA: Coloured\ Equation = (1.385 * Ventral\ Arc\ Score) + (3.113 * Subpubic\ Concavity\ Score) + (1.636 * Ischiopubic\ Ramus\ Score) + \mathbf{0.000}$$

$$South\ African\ Equation = (1.501 * Ventral\ Arc\ Score) + (2.087 * Subpubic\ Concavity\ Score) + (1.963 * Ischiopubic\ Ramus\ Score) + \mathbf{0.000}$$

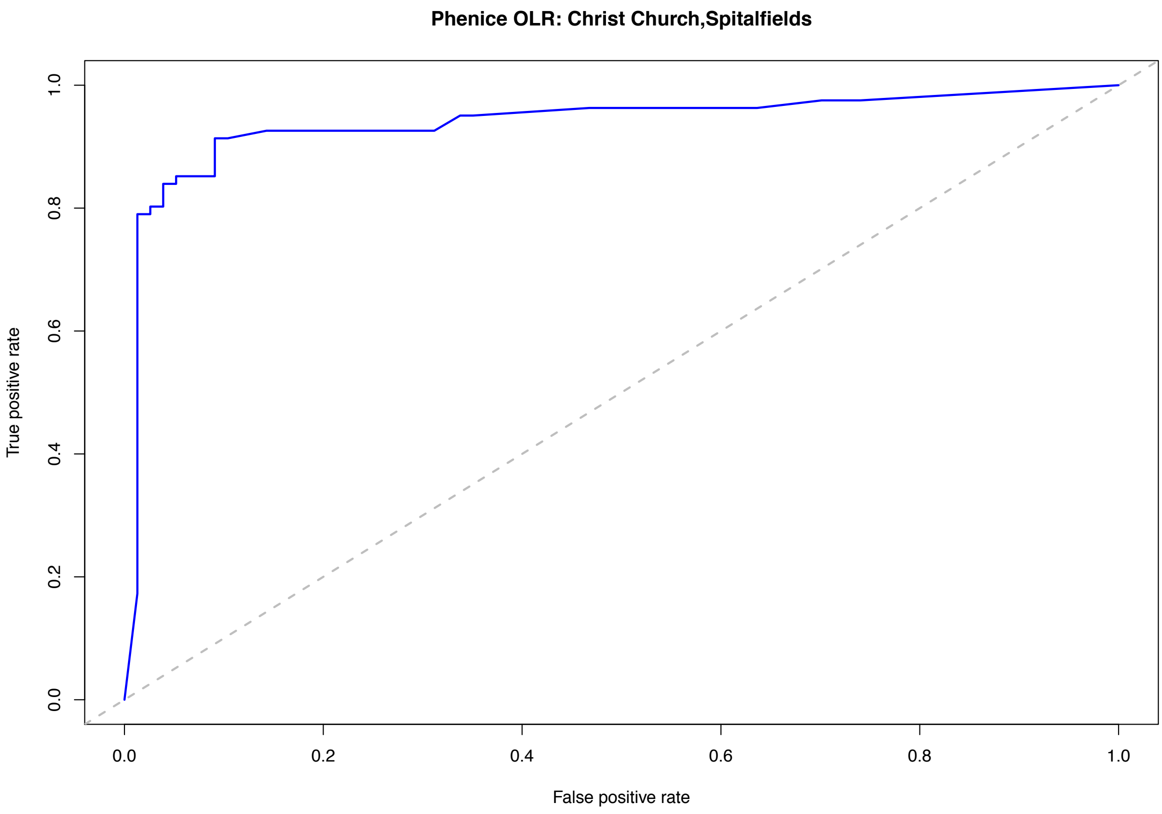
$$Summary\ Sex\ Equation = (1.440 * Ventral\ Arc\ Score) + (1.962 * Subpubic\ Concavity\ Score) + (2.153 * Ischiopubic\ Ramus\ Score) + \mathbf{0.000}$$

Appendix D.2 Wald Test: New Phenice OLR

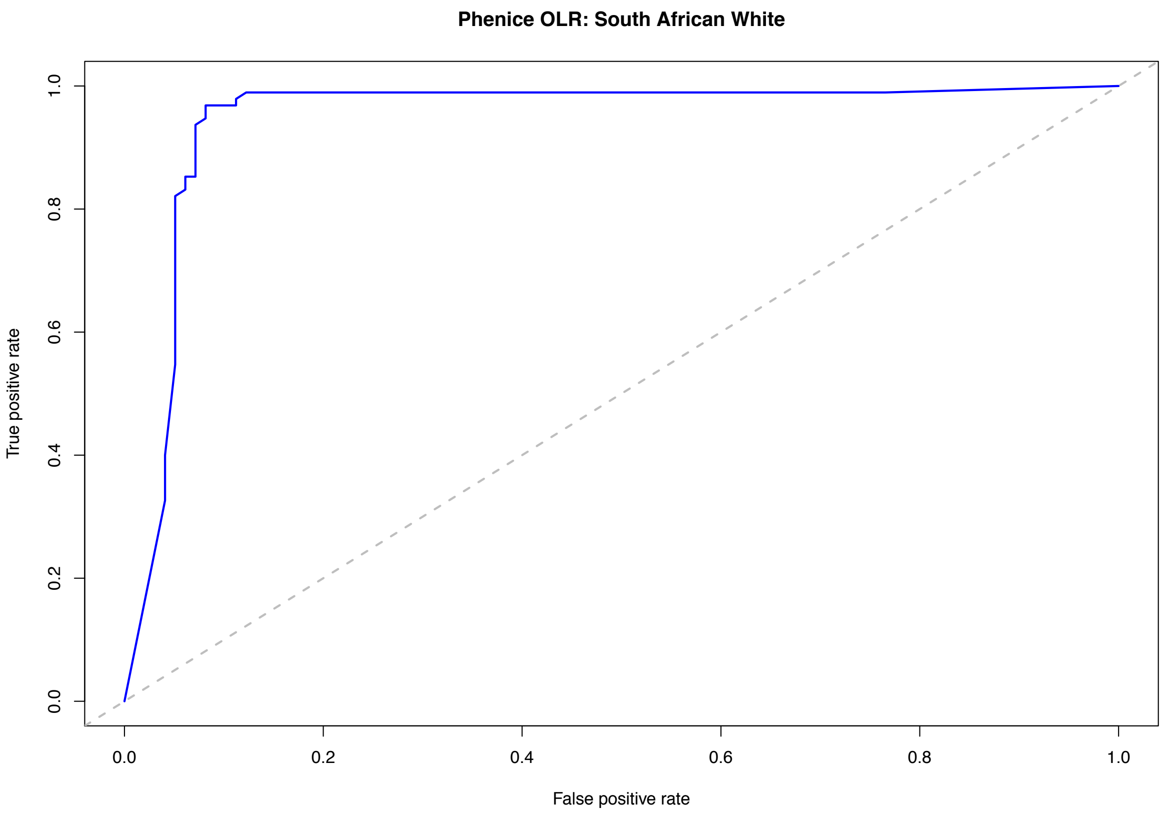
Appendix D.2 Table 1: Results from the Wald Test.

Sample	x^2	df	p Value
Christ Church, Spitalfields	49.5	3	<0.001
South African White	60.9	3	<0.001
South African Black	68.8	3	<0.001
South African Coloured	67.1	3	<0.001
South African	201.1	3	<0.001
Summary Sex	253.3	3	<0.001

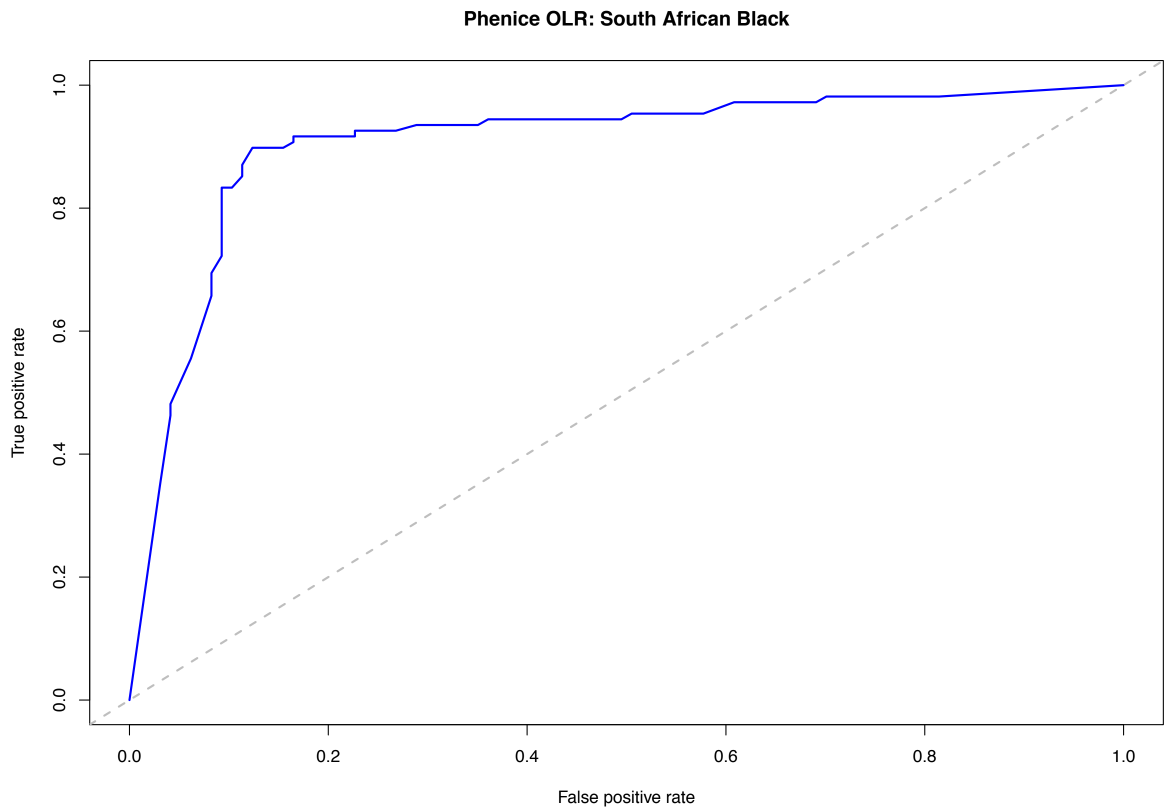
Appendix D.3 ROC Curves: New Phenice OLR



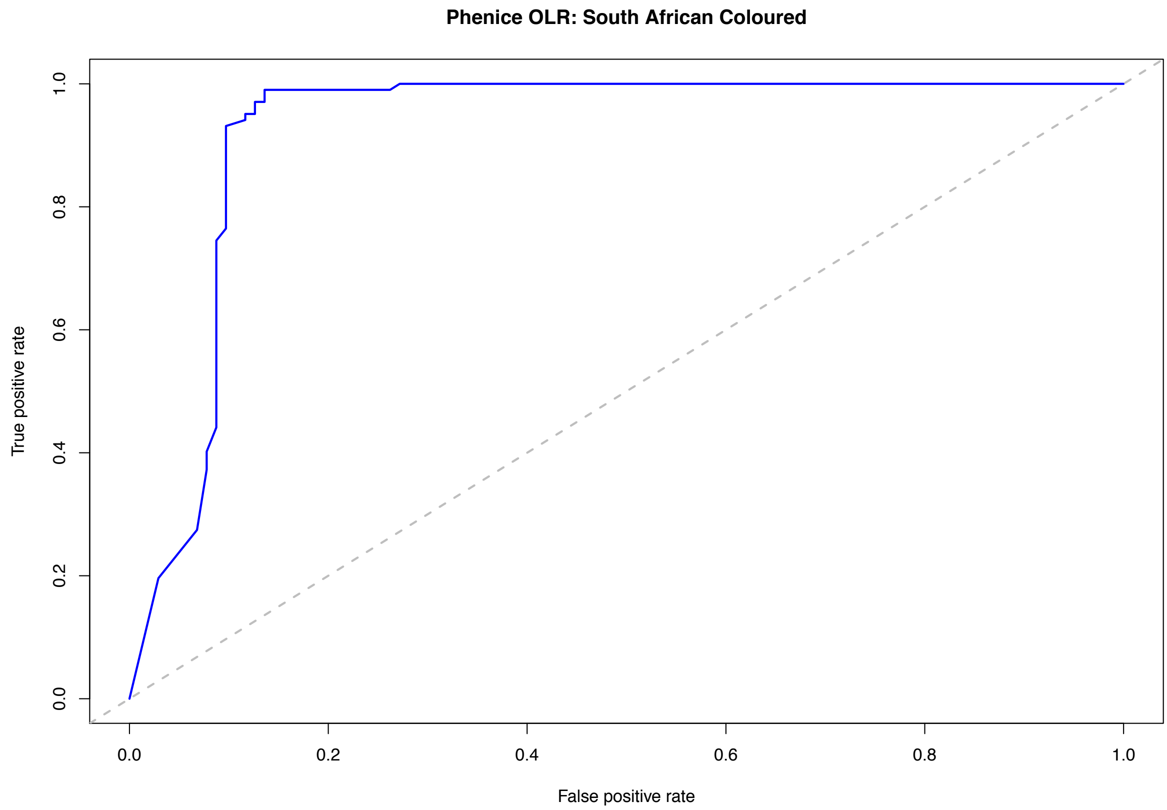
Appendix D.3 Figure 1: ROC curve for Christ Church, Spitalfields New Phenice OLR equation.



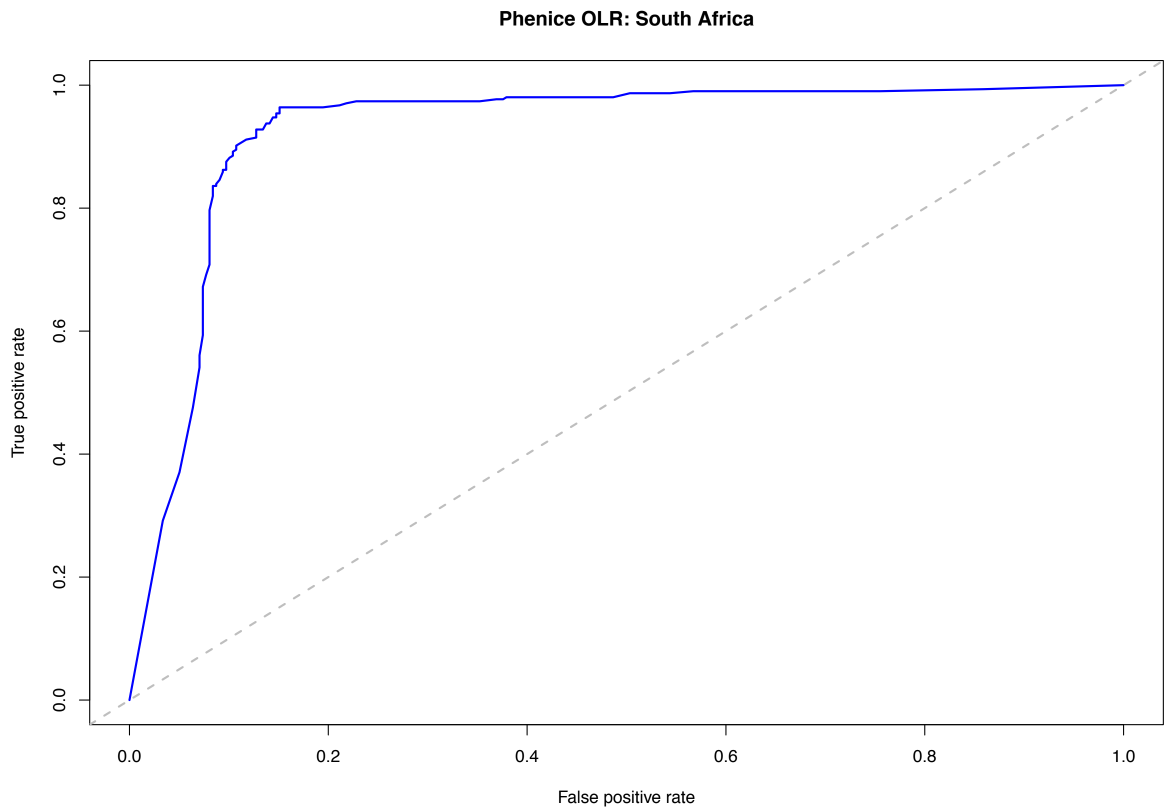
Appendix D.3 Figure 2: ROC curve for South African White New Phenice OLR equation.



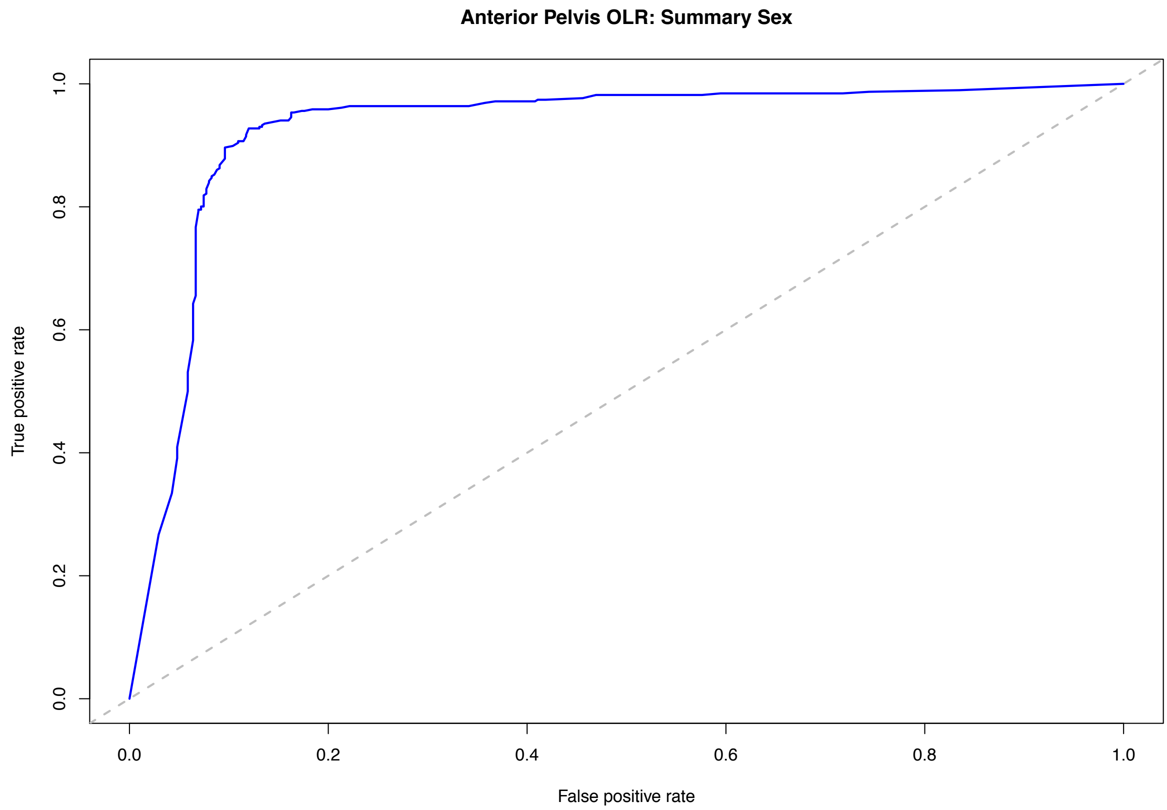
Appendix D.3 Figure 3: ROC curve for South African Black New Phenice OLR equation.



Appendix D.3 Figure 4: ROC curve for South African Coloured New Phenice OLR equation.



Appendix D.3 Figure 5: ROC curve for South African New Phenice OLR equation.



Appendix D.3 Figure 6: ROC curve for Summary Sex New Phenice OLR equation.

Appendix D.4 Tabulated Results: New Phenice OLR

Appendix D.4 Table 1: Classification for the Christ Church, Spitalfields sample (N=158).

OLR Equation	Correct Females	Correct Males	Overall Agreement (%)
Christ Church, Spitalfields	70/77 (90.91%)	72/81 (88.89%)	89.87%
South African White	71/77 (92.21%)	69/81 (85.19%)	88.61%
South African Black	72/77 (93.51%)	69/81 (85.19%)	89.24%
South African Coloured	70/77 (90.91%)	68/81 (83.95%)	87.34%
South African	72/77 (93.51%)	69/81 (85.19%)	89.24%
Summary Sex	71/77 (92.21%)	69/81 (85.19%)	88.61%

Appendix D.4 Table 2: Classification for the South African White sample (N=193).

OLR Equation	Correct Females	Correct Males	Overall Agreement (%)
Christ Church, Spitalfields	86/98 (87.76%)	94/95 (98.95%)	93.26%
South African White	90/98 (91.84%)	93/95 (97.89%)	94.82%
South African Black	90/98 (91.84%)	93/95 (97.89%)	94.82%
South African Coloured	90/98 (91.84%)	89/95 (93.68%)	92.75%
South African	90/98 (91.84%)	93/95 (97.89%)	94.82%
Summary Sex	90/98 (91.84%)	93/95 (97.89%)	94.82%

Appendix D.4 Table 3: Classification for the South African Black sample (N=205).

OLR Equation	Correct Females	Correct Males	Overall Agreement (%)
Christ Church, Spitalfields	67/97 (69.07%)	101/108 (93.52%)	81.95%
South African White	73/97 (75.26%)	101/108 (93.52%)	84.88%
South African Black	74/97 (76.29%)	100/108 (92.59%)	84.88%
South African Coloured	71/97 (73.20%)	99/108 (91.67%)	82.93%
South African	74/97 (76.29%)	100/108 (92.59%)	84.88%
Summary Sex	74/97 (76.29%)	100/108 (92.59%)	84.88%

Appendix D.4 Table 4: Classification for the South African Coloured sample (N=205).

OLR Equation	Correct Females	Correct Males	Overall Agreement (%)
Christ Church, Spitalfields	82/103 (79.61%)	100/102 (98.04%)	88.78%
South African White	86/103 (83.50%)	100/102 (98.04%)	90.73%
South African Black	87/103 (84.47%)	100/102 (98.04%)	91.22%
South African Coloured	90/103 (87.38%)	101/102 (99.02%)	93.17%
South African	87/103 (84.47%)	100/102 (98.04%)	91.22%
Summary Sex	87/103 (84.47%)	100/102 (98.04%)	91.22%

Appendix D.4 Table 5: Classification for the Poulton sample (N=67).

OLR Equation	Correct Females	Correct Males	Overall Agreement (%)
Christ Church, Spitalfields	30/34(88.24%)	30/33 (90.91%)	89.55%
South African White	33/34(97.06%)	30/33 (90.91%)	94.03%
South African Black	32/34(94.12%)	30/33 (90.91%)	92.54%
South African Coloured	31/34(91.18%)	30/33 (90.91%)	91.04%
South African	32/34(94.12%)	30/33 (90.91%)	92.54%
Summary Sex	32/34(94.12%)	30/33 (90.91%)	92.54%

Appendix D.4 Table 6: Classification for the St. Owens sample (N=44).

OLR Equation	Correct Females	Correct Males	Overall Agreement (%)
Christ Church, Spitalfields	22/24 (91.67%)	19/20 (95.00%)	93.18%
South African White	22/24 (91.67%)	19/20 (95.00%)	93.18%
South African Black	23/24 (95.83%)	18/20 (90.00%)	93.18%
South African Coloured	24/24 (100.00%)	18/20 (90.00%)	95.45%
South African	23/24 (95.83%)	19/20 (95.00%)	95.45%
Summary Sex	23/24 (95.83%)	19/20 (95.00%)	95.45%

Appendix D.4 Table 7: Classification for the Chumash sample (N=35).

OLR Equation	Correct Females	Correct Males	Overall Agreement (%)
Christ Church, Spitalfields	13/17 (76.47%)	16/18 (88.89%)	82.86%
South African White	13/17 (76.47%)	16/18 (88.89%)	82.86%
South African Black	13/17 (76.47%)	16/18 (88.89%)	82.86%
South African Coloured	13/17 (76.47%)	16/18 (88.89%)	82.86%
South African	13/17 (76.47%)	16/18 (88.89%)	82.86%
Summary Sex	13/17 (76.47%)	16/18 (88.89%)	82.86%

Appendix D.4 Table 8: Classification for the Andaman sample (N=27).

OLR Equation	Correct Females	Correct Males	Overall Agreement (%)
Christ Church, Spitalfields	10/13 (76.92%)	14/14 (100.00%)	88.89%
South African White	10/13 (76.92%)	14/14 (100.00%)	88.89%
South African Black	10/13 (76.92%)	14/14 (100.00%)	88.89%
South African Coloured	10/13 (76.92%)	14/14 (100.00%)	88.89%
South African	10/13 (76.92%)	14/14 (100.00%)	88.89%
Summary Sex	10/13 (76.92%)	14/14 (100.00%)	88.89%

Appendix E. New Posterior Pelvis OLR

Appendix E.1 New OLR Equations: Posterior Pelvis

$$Christ\ Church, Spitalfields\ \ Equation = (3.545 * Greater\ Sciatic\ Notch\ Score) + (2.298 * Preauricular\ Sulcus\ Score) + \mathbf{0.000}$$

$$SA: White\ Equation = (2.352 * Greater\ Sciatic\ Notch\ Score) + (3.745 * Preauricular\ Sulcus\ Score) + \mathbf{0.000}$$

$$SA: Black\ Equation = (2.324 * Greater\ Sciatic\ Notch\ Score) + (2.936 * Preauricular\ Sulcus\ Score) + \mathbf{0.000}$$

$$SA: Coloured\ Equation = (3.313 * Greater\ Sciatic\ Notch\ Score) + (2.188 * Preauricular\ Sulcus\ Score) + \mathbf{0.000}$$

$$South\ African\ Equation = (12.674 * Greater\ Sciatic\ Notch\ Score) + (2.837 * Preauricular\ Sulcus\ Score) + \mathbf{0.000}$$

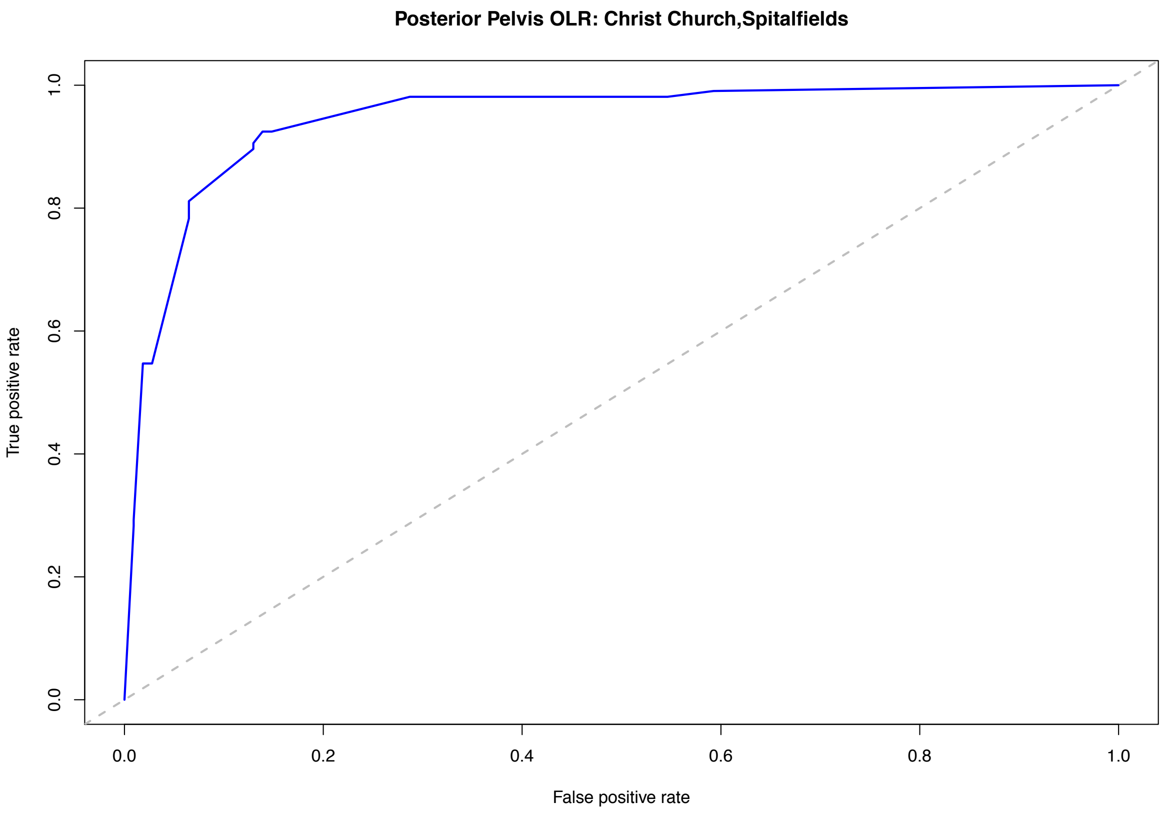
$$Summary\ Sex\ Equation = (2.680 * Greater\ Sciatic\ Notch\ Score) + (2.674 * Preauricular\ Sulcus\ Score) + \mathbf{0.000}$$

Appendix E.2 Wald Test: New Posterior Pelvis OLR

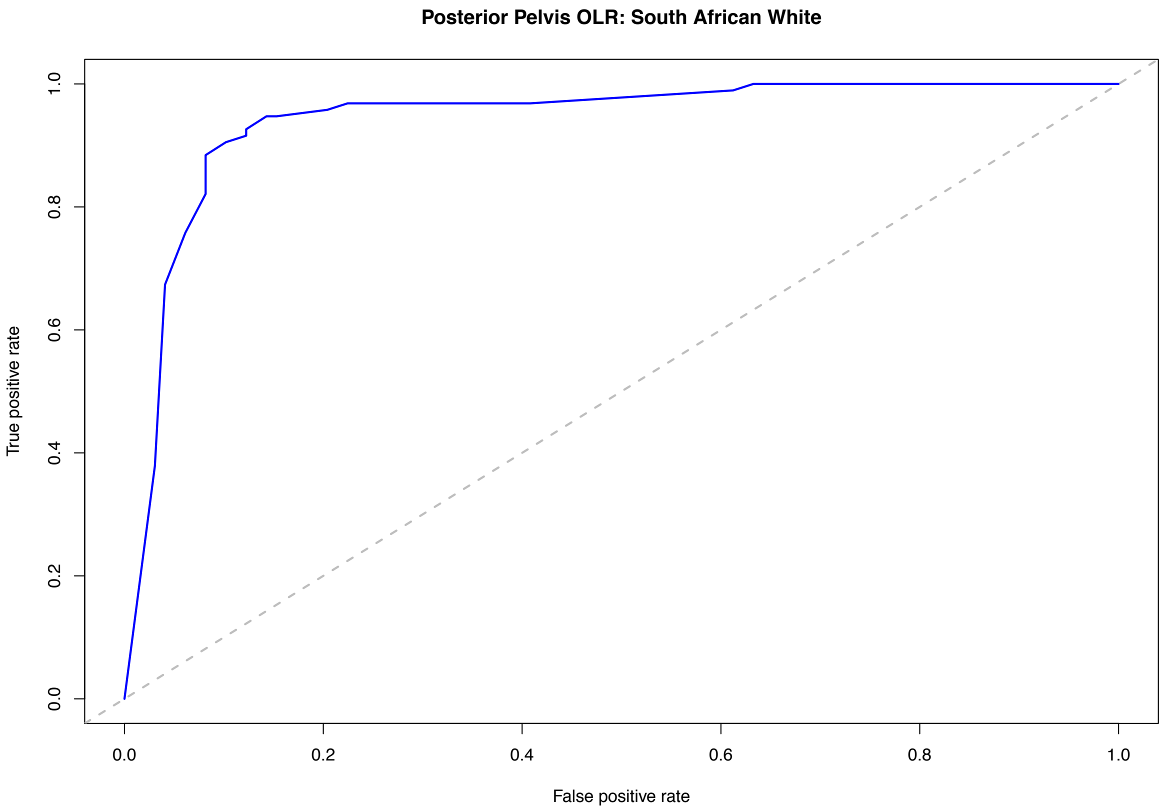
Appendix E.2 Table 1: Results from the Wald test.

Sample	x^2	df	p Value
Christ Church, Spitalfields	60.5	2	<0.001
South African White	57.9	2	<0.001
South African Black	64	2	<0.001
South African Coloured	65	2	<0.001
South African	184.8	2	<0.001
Summary Sex	245	2	<0.001

Appendix E.3 ROC Curves: New Posterior Pelvis OLR

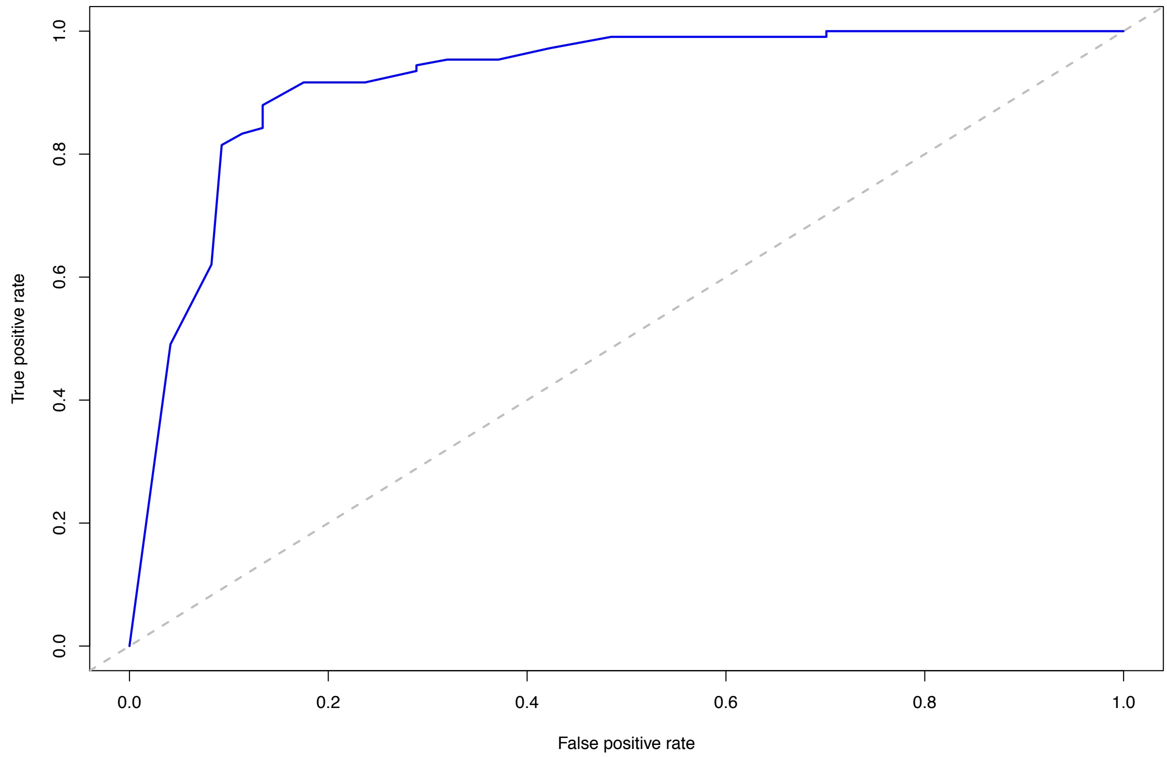


Appendix E.3 Figure 1: ROC curve for Christ Church, Spitalfields New Posterior Pelvis OLR equation.



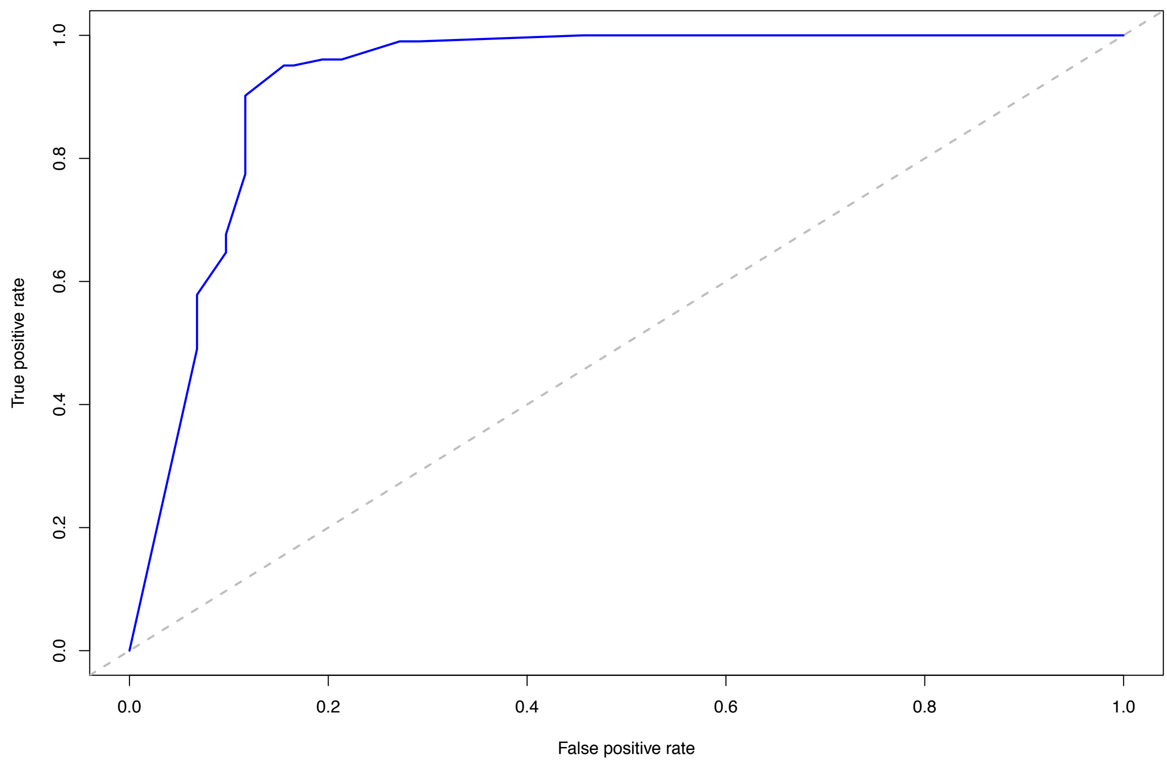
Appendix E.3 Figure 2: ROC curve for South African White New Posterior Pelvis OLR equation.

Posterior Pelvis OLR: South African Black

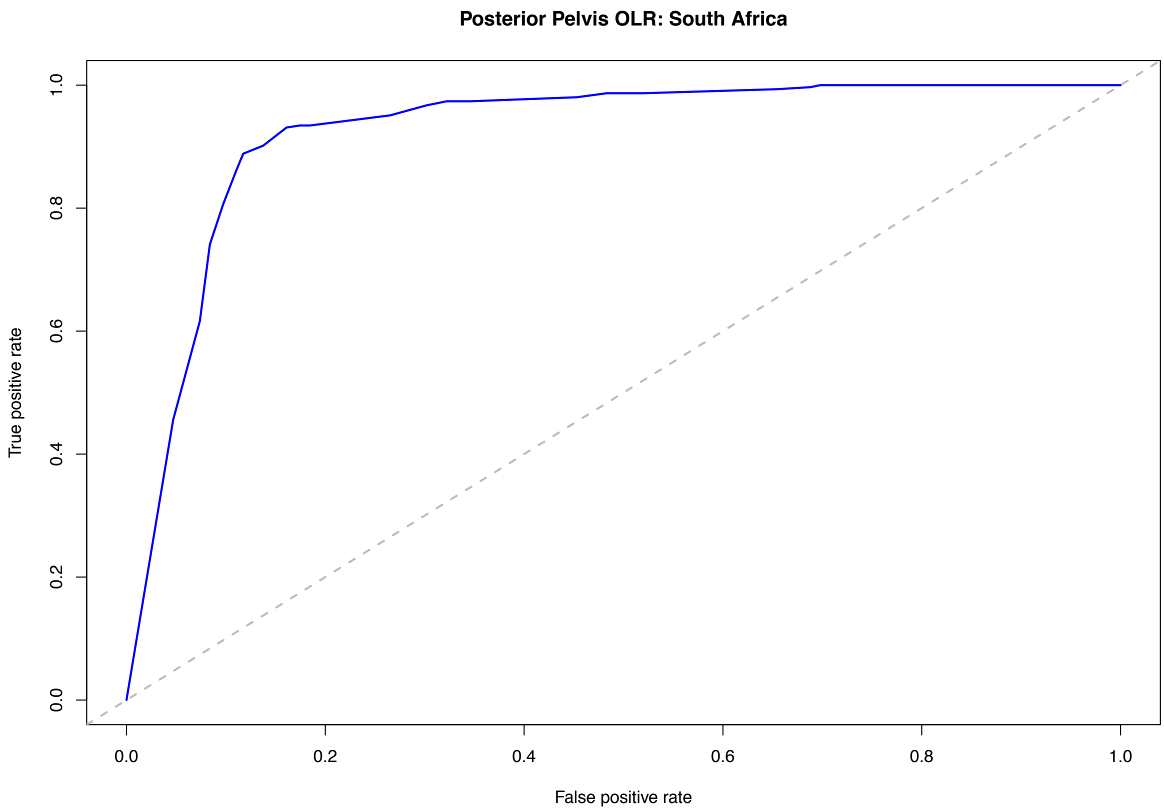


Appendix E.3 Figure 3: ROC curve for South African Black New Posterior Pelvis OLR equation.

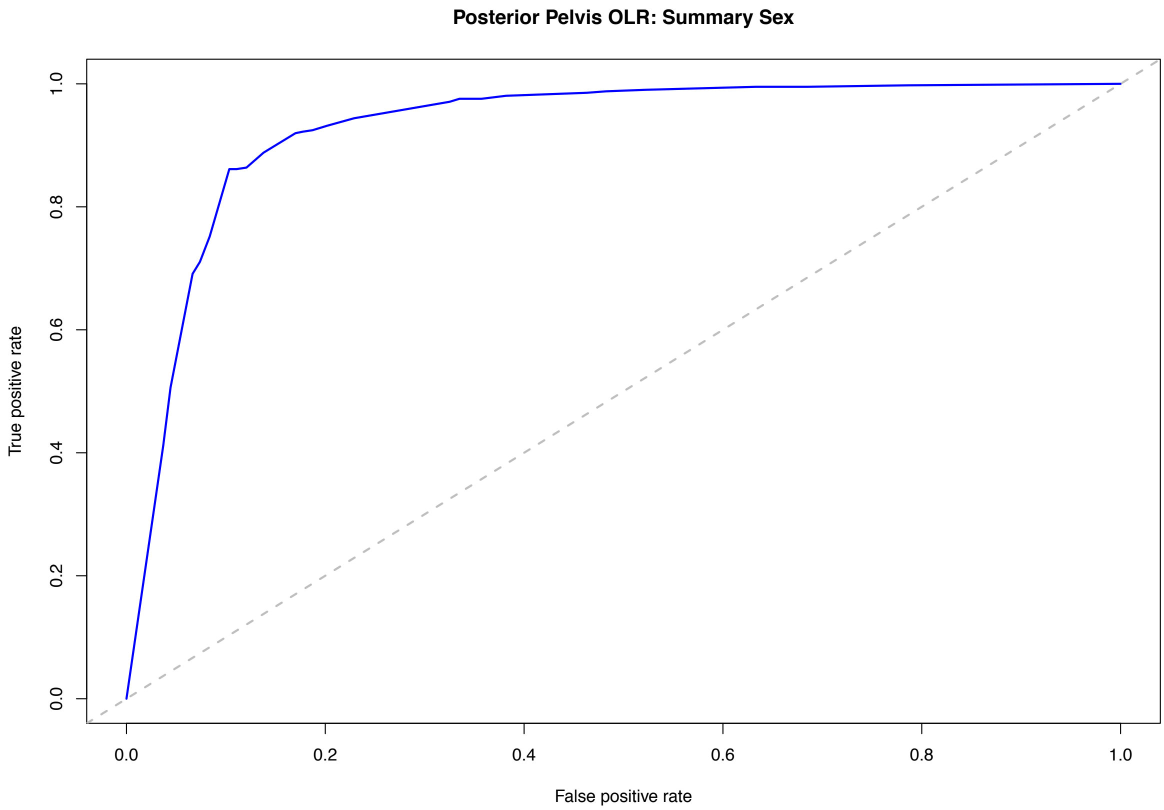
Posterior Pelvis OLR: South African Coloured



Appendix E.3 Figure 4: ROC curve for South African Coloured New Posterior Pelvis OLR equation.



Appendix E.3 Figure 5: ROC curve for South African New Posterior Pelvis OLR equation.



Appendix E.3 Figure 6: ROC curve for Summary Sex New Posterior Pelvis OLR equation.

Appendix E.4 Tabulated Results: New Posterior Pelvis OLR

Appendix E.4 Table 1: Classification results for the Christ Church, Spitalfields sample (N=214).

OLR Equation	Correct Females	Correct Males	Overall Agreement (%)
Christ Church, Spitalfields	94/108 (87.04%)	95/106 (89.62%)*	88.32%
South African White	79/108 (73.15%)	99/106 (93.40%)*	83.18%
South African Black	79/108 (73.15%)	99/106 (93.40%)*	83.18%
South African Coloured	94/108 (87.04%)	95/106 (89.62%)*	88.32%
South African	79/108 (73.15%)	99/106 (93.40%)*	83.18%
Summary Sex	94/108 (87.04%)	95/106 (89.62%)*	88.32%

*One male was classified as unknown.

Appendix E.4 Table 2: Classification results for the South African White sample (N=193).

OLR Equation	Correct Females	Correct Males	Overall Agreement (%)
Christ Church, Spitalfields	82/98 (83.67%)*	88/95 (92.63%)*	88.08%
South African White	76/98 (77.55%)*	91/95 (95.79%)*	86.53%
South African Black	81/98 (82.65%)*	90/95 (94.74%)*	88.60%
South African Coloured	82/98 (83.67%)*	88/95 (92.63%)*	88.08%
South African	76/98 (77.55%)*	91/95 (95.79%)*	86.53%
Summary Sex	82/98 (79.17%)*	88/95 (92.63%)*	88.08%

*Two females and one male were classified as unknown

Appendix E.4 Table 3: Classification results for the South African Black sample (N=205).

OLR Equation	Correct Females	Correct Males	Overall Agreement (%)
Christ Church, Spitalfields	66/97 (68.04%)*	99/108 (91.67%)*	80.49%
South African White	66/97 (68.04%)*	102/108 (94.44%)*	81.95%
South African Black	66/97 (68.04%)*	101/108 (93.52%)*	81.46%
South African Coloured	66/97 (68.04%)*	99/108 (91.67%)*	80.49%
South African	66/97 (68.04%)*	102/108 (94.44%)*	81.95%
Summary Sex	66/97 (68.04%)*	99/108 (91.67%)*	80.49%

*One female and three males were classified as unknown.

Appendix E.4 Table 4: Classification results for the South African Coloured sample (N=205).

OLR Equation	Correct Females	Correct Males	Overall Agreement (%)
Christ Church, Spitalfields	82/103 (79.61%)*	98/102 (96.08%)	87.80%
South African White	60/103 (58.25%)*	102/102 (100.00%)	79.02%
South African Black	66/103 (64.08%)*	99/102 (97.06%)	80.49%
South African Coloured	82/103 (79.61%)*	98/102 (96.08%)	87.80%
South African	60/103 (58.25%)*	102/102 (100.00%)	79.02%
Summary Sex	82/103 (79.61%)*	98/102 (96.08%)	87.80%

*One female was classified as unknown

Appendix E.4 Table 5: Classification results for the Poulton sample (N=125).

OLR Equation	Correct Females	Correct Males	Overall Agreement (%)
Christ Church, Spitalfields	55/63 (87.30%)	54/62 (87.10%)	87.20%
South African White	56/63 (88.89%)	50/62 (80.65%)	84.80%
South African Black	56/63 (88.89%)	50/62 (80.65%)	84.80%
South African Coloured	55/63 (87.30%)	54/62 (87.10%)	87.20%
South African	56/63 (88.89%)	50/62 (80.65%)	84.80%
Summary Sex	55/63 (87.30%)	54/62 (87.10%)	87.20%

Appendix E.4 Table 6: Classification results for the St. Owens sample (N=98).

OLR Equation	Correct Females	Correct Males	Overall Agreement (%)
Christ Church, Spitalfields	35/44 (79.55%)*	48/54 (88.89%)*	84.69%
South African White	33/44 (75.00%)*	48/54 (88.89%)*	82.65%
South African Black	33/44 (75.00%)*	47/54 (87.04%)*	81.63%
South African Coloured	35/44 (79.55%)*	48/54 (88.89%)*	84.69%
South African	33/44 (75.00%)*	48/54 (88.89%)*	82.65%
Summary Sex	35/44 (79.55%)*	48/54 (88.89%)*	84.69%

*One female and two males were classified as unknown.

Appendix E.4 Table 7: Classification results for the Chumash sample (N=38).

OLR Equation	Correct Females	Correct Males	Overall Agreement (%)
Christ Church, Spitalfields	16/19 (84.21%)	12/19 (63.16%)	75.68%
South African White	16/19 (84.21%)	12/19 (63.16%)	75.68%
South African Black	16/19 (84.21%)	12/19 (63.16%)	75.68%
South African Coloured	16/19 (84.21%)	12/19 (63.16%)	75.68%
South African	16/19 (84.21%)	12/19 (63.16%)	75.68%
Summary Sex	16/19 (84.21%)	12/19 (63.16%)	75.68%

Appendix E.4 Table 8: Classification results for the Andaman sample (N=27).

OLR Equation	Correct Females	Correct Males	Overall Agreement (%)
Christ Church, Spitalfields	11/13 (84.62%)	13/14 (92.86%)	88.89%
South African White	10/13 (76.92%)	13/14 (92.86%)	85.19%
South African Black	10/13 (76.92%)	13/14 (92.86%)	85.19%
South African Coloured	11/13 (84.62%)	13/14 (92.86%)	88.89%
South African	10/13 (76.92%)	13/14 (92.86%)	85.19%
Summary Sex	11/13 (84.62%)	13/14 (92.86%)	88.89%

Appendix F. New Anterior Pelvis OLR

Appendix F.1 New OLR Equations: Anterior Pelvis

$$Christ\ Church, Spitalfields\ \ Equation = (3.400 * Pubic\ Body\ Shape\ Score) + (2.813 * Subpubic\ Angle\ Score) + \mathbf{0.000}$$

$$SA: White\ Equation = (3.825 * Pubic\ Body\ Shape\ Score) + (1.681 * Subpubic\ Angle\ Score) + \mathbf{0.000}$$

$$SA: Black\ Equation = (1.367 * Pubic\ Body\ Shape\ Score) + (3.202 * Subpubic\ Angle\ Score) + \mathbf{0.000}$$

$$SA: Coloured\ Equation = (3.187 * Pubic\ Body\ Shape\ Score) + (2.601 * Subpubic\ Angle\ Score) + \mathbf{0.000}$$

$$South\ African\ Equation = (1.367 * Pubic\ Body\ Shape\ Score) + (3.202 * Subpubic\ Angle\ Score) + \mathbf{0.000}$$

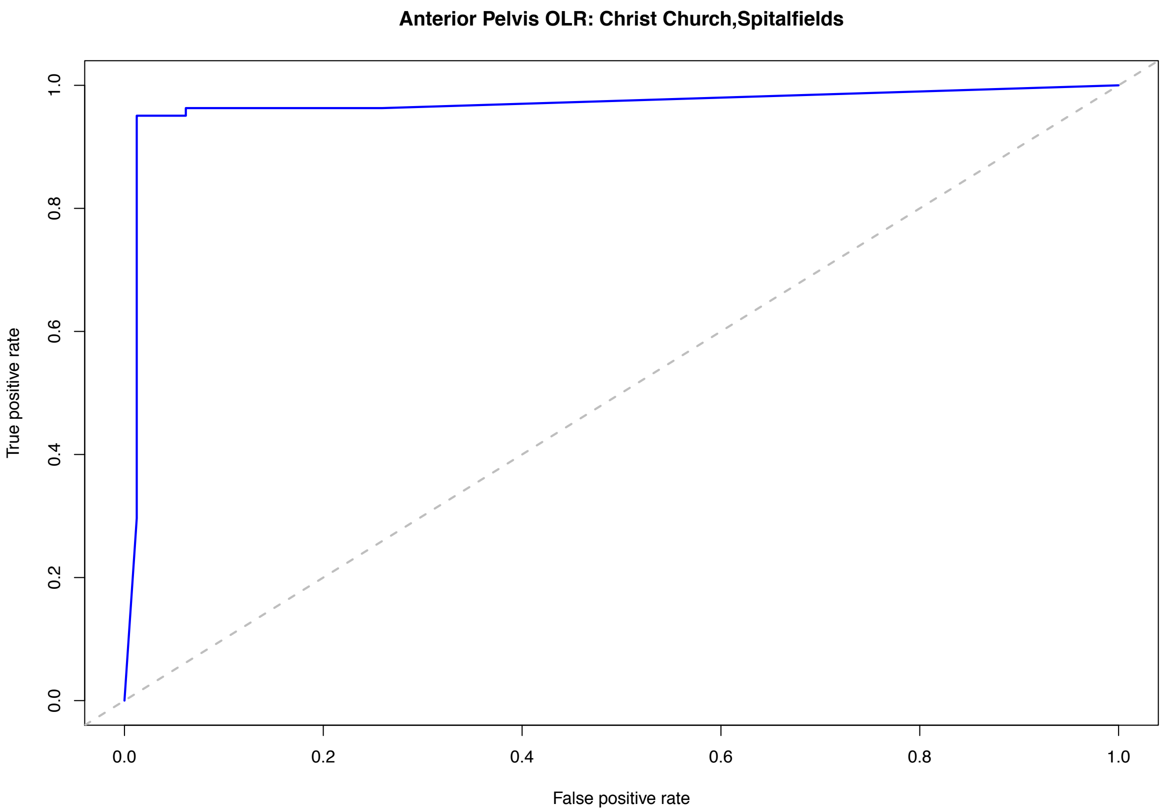
$$Summary\ Sex\ Equation = (2.508 * Pubic\ Body\ Shape\ Score) + (2.408 * Subpubic\ Angle\ Score) + \mathbf{0.000}$$

Appendix F.2 Wald Test: New Anterior Pelvis OLR

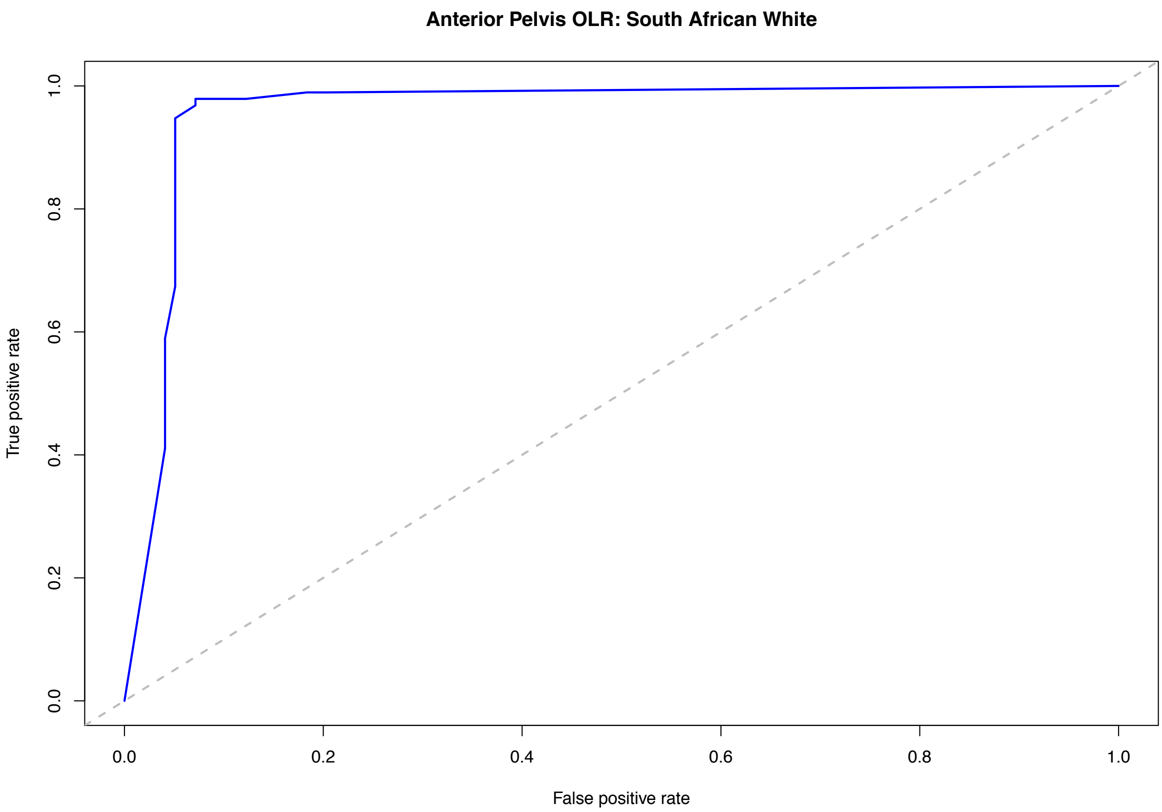
Appendix F.2 Table 1: Results from the Wald Test.

Sample	x^2	df	p Value
Christ Church, Spitalfields	45	2	<0.001
South African White	63.8	2	<0.001
South African Black	73.3	2	<0.001
South African Coloured	62.8	2	<0.001
South African	202.9	2	<0.001
Summary Sex	255.9	2	<0.001

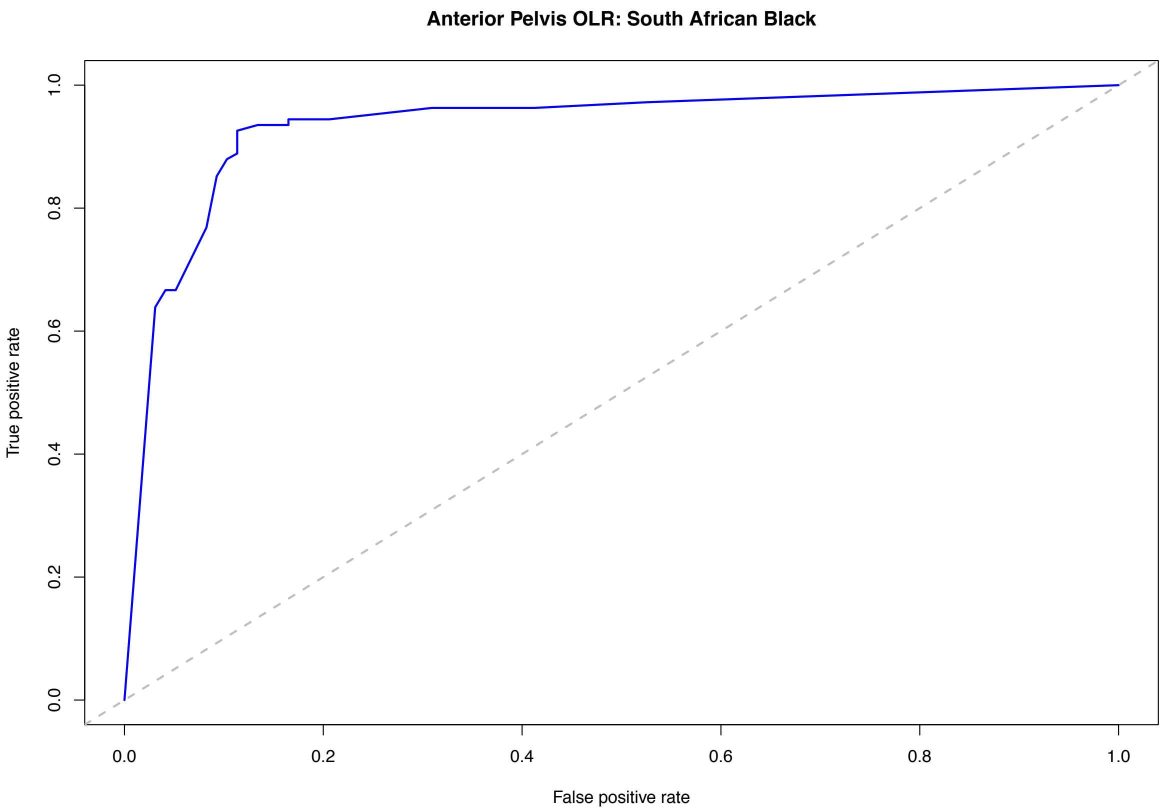
Appendix F.3 ROC Curves: New Anterior Pelvis OLR



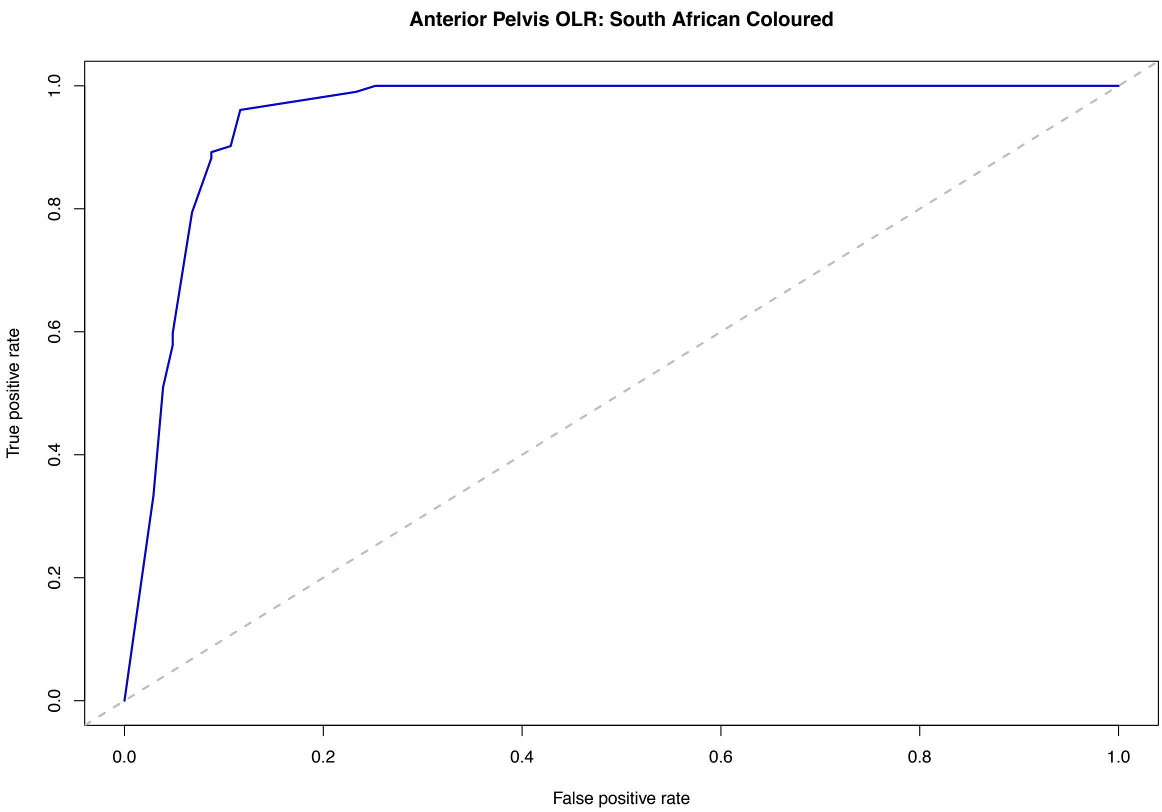
Appendix F.3 Figure 1: ROC curve for Christ Church, Spitalfields New Anterior Pelvis OLR equation.



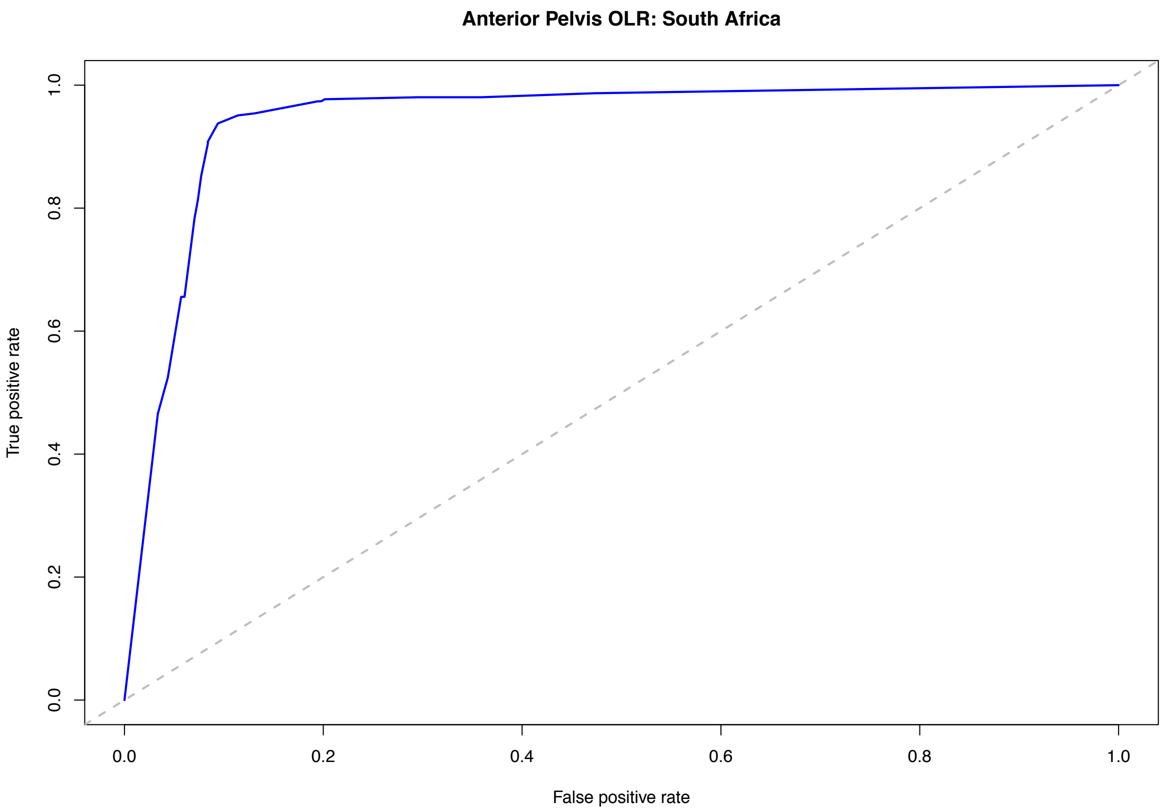
Appendix F.3 Figure 2: ROC curve for South African White New Anterior Pelvis OLR equation.



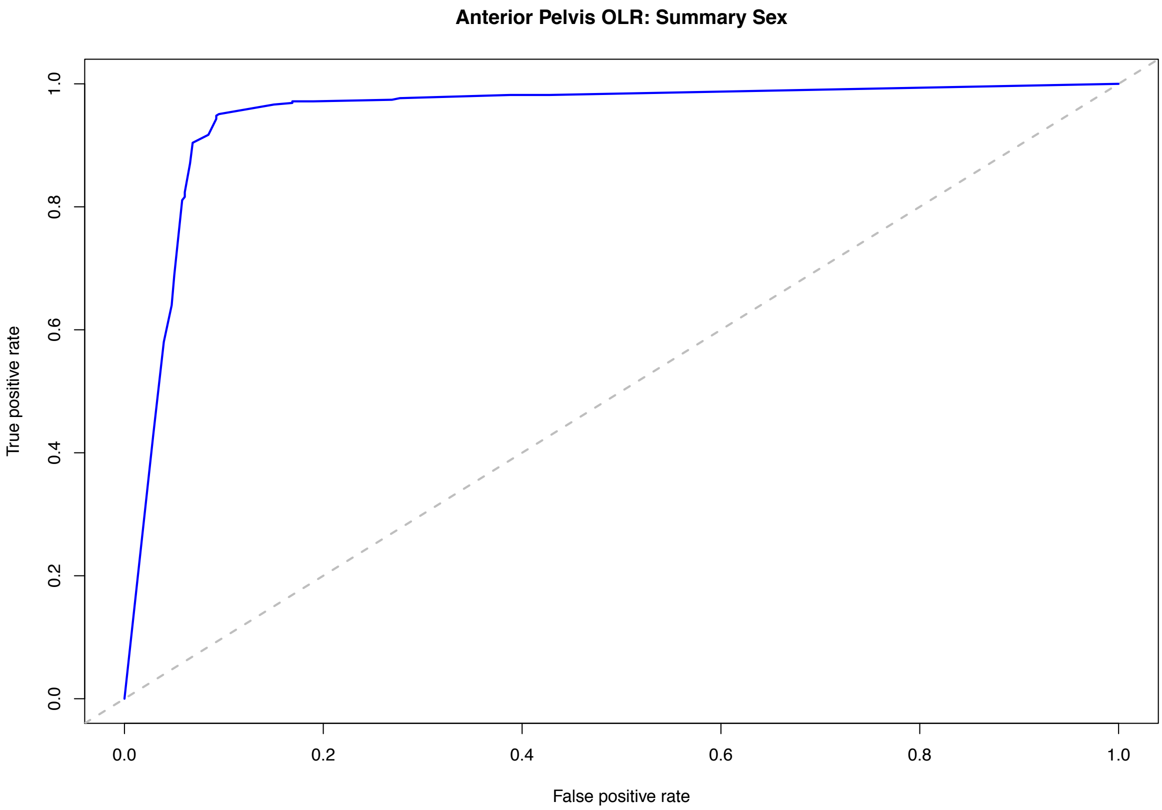
Appendix F.3 Figure 3: ROC curve for South African Black New Anterior Pelvis OLR equation.



Appendix F.3 Figure 4: ROC curve for South African Coloured New Anterior Pelvis OLR equation.



Appendix F.3 Figure 5: ROC curve for South African New Anterior Pelvis OLR equation.



Appendix F.3 Figure 6: ROC curve for Summary Sex New Anterior Pelvis OLR equation.

Appendix F.4 Tabulated Results: New Anterior Pelvis OLR

Appendix F.4 Table 1: Classification results for the Christ Church, Spitalfields sample (N=162).

OLR Equation	Correct Females	Correct Males	Overall Agreement (%)
Christ Church, Spitalfields	80/81 (98.77%)	77/81 (95.06%)	96.91%
South African White	80/81 (98.77%)	76/81 (93.83%)	96.30%
South African Black	80/81 (98.77%)	76/81 (93.83%)	96.30%
South African Coloured	80/81 (98.77%)	76/81 (93.83%)	96.30%
South African	80/81 (98.77%)	75/81 (92.59%)	95.68%
Summary Sex	80/81 (98.77%)	76/81 (93.83%)	96.30%

Appendix F.4 Table 2: Classification results for the South African White sample (N=193).

OLR Equation	Correct Females	Correct Males	Overall Agreement (%)
Christ Church, Spitalfields	93/98 (94.90%)	91/95 (95.79%)	95.34%
South African White	93/98 (94.90%)	88/95 (92.63%)	93.78%
South African Black	93/98 (94.90%)	90/95 (94.74%)	94.84%
South African Coloured	93/98 (94.90%)	91/95 (95.79%)	95.34%
South African	93/98 (94.90%)	88/95 (92.63%)	93.78%
Summary Sex	93/98 (94.90%)	91/95 (95.79%)	95.34%

Appendix F.4 Table 3: Classification results for the South African Black sample (N=205).

OLR Equation	Correct Females	Correct Males	Overall Agreement (%)
Christ Church, Spitalfields	80/97 (82.47%)	101/108 (93.52%)	88.29%
South African White	84/97 (86.60%)	101/108 (93.52%)	90.24%
South African Black	79/97 (81.44%)	101/108 (93.52%)	87.80%
South African Coloured	80/97 (82.47%)	101/108 (93.52%)	88.29%
South African	84/97 (86.60%)	101/108 (93.52%)	90.24%
Summary Sex	80/97 (82.47%)	101/108 (93.52%)	88.29%

Appendix F.4 Table 4: Classification results for the South African Coloured sample (N=205).

OLR Equation	Correct Females	Correct Males	Overall Agreement (%)
Christ Church, Spitalfields	91/103(88.35%)	98/102 (96.08%)	92.20%
South African White	93/103(90.29%)	97/102 (95.10%)	92.68%
South African Black	91/103(88.35%)	98/102 (96.08%)	92.20%
South African Coloured	91/103(88.35%)	98/102 (96.08%)	92.20%
South African	93/103(90.29%)	97/102 (95.10%)	92.68%
Summary Sex	91/103(88.35%)	98/102 (96.08%)	92.20%

Appendix F.4 Table 5: Classification results for the Poulton sample (N=76).

OLR Equation	Correct Females	Correct Males	Overall Agreement (%)
Christ Church, Spitalfields	32/37 (86.49%)	35/39 (89.74%)	88.16%
South African White	33/37 (89.19%)	34/39 (87.18%)	88.16%
South African Black	31/37 (83.78%)	35/39 (89.74%)	86.84%
South African Coloured	32/37 (86.49%)	34/39 (87.18%)	86.84%
South African	33/37 (89.19%)	34/39 (87.18%)	88.16%
Summary Sex	32/37 (86.49%)	35/39 (89.74%)	88.16%

Appendix F.4 Table 6: Classification results for the St. Owens sample (N=46).

OLR Equation	Correct Females	Correct Males	Overall Agreement (%)
Christ Church, Spitalfields	24/24 (100.00%)	20/22 (90.91%)	95.65%
South African White	24/24 (100.00%)	20/22 (90.91%)	95.65%
South African Black	23/24 (95.83%)	20/22 (90.91%)	93.48%
South African Coloured	24/24 (100.00%)	20/22 (90.91%)	95.65%
South African	24/24 (100.00%)	20/22 (90.91%)	95.65%
Summary Sex	24/24 (100.00%)	20/22 (90.91%)	95.65%

Appendix F.4 Table 7: Classification results for the Chumash sample (N=37).

OLR Equation	Correct Females	Correct Males	Overall Agreement (%)
Christ Church, Spitalfields	13/18 (72.22%)	17/19 (89.47%)	81.08%
South African White	15/18 (83.33%)	17/19 (89.47%)	86.49%
South African Black	13/18 (72.22%)	17/19 (89.47%)	81.08%
South African Coloured	13/18 (72.22%)	17/19 (89.47%)	81.08%
South African	15/18 (83.33%)	17/19 (89.47%)	86.49%
Summary Sex	13/18 (72.22%)	17/19 (89.47%)	81.08%

Appendix F.4 Table 8: Classification results for the Andaman sample (N=27).

OLR Equation	Correct Females	Correct Males	Overall Agreement (%)
Christ Church, Spitalfields	10/13 (76.92%)	14/14 (100.00%)	88.89%
South African White	10/13 (76.92%)	14/14 (100.00%)	88.89%
South African Black	11/13 (84.62%)	14/14 (100.00%)	92.59%
South African Coloured	10/13 (76.92%)	14/14 (100.00%)	88.89%
South African	10/13 (76.92%)	14/14 (100.00%)	88.89%
Summary Sex	10/13 (76.92%)	14/14 (100.00%)	88.89%

Bibliography

Abd-El-Hameed SY, Mohamed AA, Thabet HZ. 2009. Determination of subpubic angle in Egyptian population. Mansoura Journal of Forensic Medicine and Clinical Toxicology 17:41-53.

Ackermann JF. 1788. Dissertatio inauguralis anatomica de discrimine sexuum, Maguntiae (Mainz): Alef.

Acsádi G, Nemeskéri J. 1970. History of human life span and mortality. Budapest: Akadémiai Kiadó.

Adhikari M. 2005. Contending approaches to coloured identity and the history of the coloured people of South Africa. History Compass 3:1-6.

Albanese J, Osley SE, Tuck A. 2011. Do century-specific equations provide better estimates for stature? A test of the 19-20th century boundary for the stature estimation feature in Fordisc.3.0. Forensic Science International 219:286.e1-286.e3.

Albanese J, Osley SE, Tuck A. 2016. Do group-specific equations provide the best estimates of stature. Forensic Science International 261:154-158.

Albanese J, Saunders SR. 2006. Is it possible to escape racial typology in forensic identification? In: Schmitt A, Cunha E, Pinheiro J, editors. Forensic Anthropology and Medicine: Complementary sciences from recovery to cause of death. Humana Press. p.281- 315.

Allen JA. 1877. The influence of physical conditions in the genesis of species. Radic Rev 1:108- 140.

AlQahtani SJ, Liversidge HM, Hector MP. 2010. Atlas of tooth development and eruption. American Journal of Physical Anthropology:481-90

AlQahtani SJ. 2008. Atlas of tooth development and eruption. Barts and the London School of Medicine and Dentistry. London, Queen Mary University of London. MClinDent.

Anderson BE. 1990. Ventral arch of the Os Pubis: Anatomical and developmental considerations. American Journal of Physical Anthropology 83:449-458.

Arnold JE. 2001. *The Origins of a Pacific Coast Chiefdom: The Chumash of the Channel Islands*. Utah: The University of Utah Press.

Arsuaga JL, Carretero JM. 1994. Multivariate analysis of sexual dimorphism of the hip bone in a modern human population and in early hominins. American Journal of Physical Anthropology 93:241-257.

Asala SA, Bidmos MA, Dayal MR. 2004. Discriminant function sexing of fragmentary femur of South African blacks. Forensic Science International 145:25-29.

Asala SA. 2001. Sex determination from the head of the femur of South African whites and blacks. Forensic Science International 117:15-22.

Asala SA. 2002. The efficiency of the demarking pint of the femoral head as a sex determining parameter. Forensic Science International 127:114-118.

Atkin M, Garrod AP. 1990. Archaeology in Gloucester 1989. Transaction of the Bristol and Gloucestershire Archaeological Society. 108:185-192.

Bacon AM. 2000. Principal components analysis of distal humeral shape in Pliocene to recent African hominids: the contribution of geometric morphometrics. American Journal of Physical Anthropology 111:479-487.

Barnard A. 1992. Hunters and Herders of South Africa: A comparative ethnology of the Khoisan peoples. Cambridge University Press.

Bass WM. 1987. Human Osteology: A laboratory and field manual, 3rd ed. Columbia, MO: Missouri Archaeological Society.

Bedford ME, Russell KF, Lovejoy CO, Meindl RS, Simpson SW, Stuart-MacadamPL (1993) Test of the multifactorial aging method using skeletons with known ages-at-death from the Grant collection. American Journal of Physical Anthropology 91:287-297.

Benazzi S, Maestri C, Parisini S, Vecchi F, Gruppioni G. 2008. Sex assessment from the acetabular rim by means of image analysis. Forensic Science International 180:7–9.

Berg GE. 2001. Four Chins and a Funeral or Racial Affinity as Determined by Mandibular Morphology. Paper presented at the 53rd Annual Meeting of the American Academy of Forensic Sciences, Seattle, Washington.

Berg GE. 2006. Discriminant Function Analysis as Applied to Mandibular Morphology to Assess Population Affinity. Paper presented at the 58th Annual Meeting of the American Academy of Forensic Sciences, Seattle, Washington.

Berg GE. 2008. Biological affinity and sex determination using morphometric and morphoscopic variables from the human mandible. PhD Thesis, Department of Anthropology, University of Tennessee, Knoxville, TN.

Berge C. 1984. Multivariate analysis of the pelvis for hominids and other extant primates: Implications for the locomotion and systematics of the different species of Australopithecines. Journal of Human Evolution 13:555-562.

Bergmann C. 1847. About the conditions of the thermal economy of animals to their size. Göttingen Stud 3:595-708.

Bidmos MA. 2006. Metrical and non-metrical assessment of population affinity from the calcaneus. Forensic Science International 159:6-13.

Bierry G, Le Minor JM, Schmittbuhl M. 2010. Oval in males and triangular in females? A quantitative evaluation of sexual dimorphism in the human obturator foramen. American Journal of Physical Anthropology 141:626-631.

Blanchard BK. 2010. A study of the accuracy and reliability of sex estimation methods of the human pelvis. California State University, USA.

Boas F. 1931. Race & Progress. Science 1905:1-8.

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

Brace CL. 1995. Region does not mean "race": Reality versus convention in Forensic Anthropology. Journal of Forensic Sciences 40:171-175.

Brooks S, Suchey JM. 1990. Skeletal age determination based on the os pubis: a comparisonof the Acsádi-Nemeskéri and Suchey-Brooks methods. Journal of Human Evolution 5:227-238.

Brooks ST. 1955. Skeletal age at death: Reliability of cranial and pubic age indicators. American Journal of Physical Anthropology 13:567-597.

Brown JL. 2006. Morphological variation of the proximal femur in selected skeletal remains. M.A. Thesis. Wichita State University.

Bruzek J. 2002. A method for visual determination of sex, using the human hip bone. American Journal of Physical Anthropology 117(2):157-168.

Buck TJ, Vidarsdottir US. 2004. A Proposed Method for the Identification of Race in Sub-adult skeletons: A Geometric Morphometric Analysis of Mandibular Morphology. Journal of Forensic Sciences 49:1159-1164.

Buckberry JL, Chamberlain AT. 2002. Age estimation from the auricular surface of the ilium: a revised method. American Journal of Physical Anthropology 119:231-239.

Budinoff LC, Tague RG. 1990. Anatomical and developmental bases for the ventral arc of the human pubis. American Journal of Physical Anthropology 82:73-79.

Buikstra JE, Meilke JH. 1985. Demography, diet, and health. In: RI Gilbert Jr, and HJ Meilke eds: The analysis of prehistoric diets. New York: Academic Press, 359-422.

Buikstra JE, Ubelaker DH. 1994. Standards for data collection from human skeletal remains: proceedings of a seminar at the Field Museum of Natural History, organized by Jonathan Hass. Fayetteville: Arkansas Archeological Survey.

Burrel CL, Carpenter RJ. 2013. Analysis of Human Skeletal Material from the Poulton Research Project: 1995-2013.

Burrell CL, Davenport CAL, Carpenter RJ, Ohman JC. 2013. Biological age estimation of subadult human skeletal remains: Comparison of dental development with the humerus, femur and pars basilaris. British Association for Biological Anthropology and Osteoarchaeology Conference – York.

Burrell CL, Dove ER, Emery MM, Ohman JC. 2015. A review of biological age from dental development and long bone diaphyseal growth in subadult human remains. [ABSTRACT] American Journal of Physical Anthropology S60:65-334.

Burrell CL. 2017. Skeletal variation as a possible reflection of relatedness within three medieval British populations. PhD Thesis. Liverpool John Moores University.

Byers S, Akoshima K, Curran B. 1989. Determination of adult stature from metatarsal length. American Journal of Physical Anthropology 79:275-279.

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

Campbell DM, Hall MH, Barker DJ, Cross J, Shiell AW, Godfrey KM. 1996. Diet in pregnancy and the offspring's blood pressure 40 years later. British Journal of Obstetrics and Gynaecology 103:273-280.

Canavan SM. 2014. "Arrowman" An analysis of a possible medieval cold case from Poulton, Chesire. M.Phil. Thesis. Anthropology Department, Liverpool John Moores University: U.K.

Cardoso H. 2008. Epiphyseal union at the innominate and lower limb in a modern Portuguese skeletal sample, and age estimation in adolescent and young adult male and female skeletons. American Journal of Physical Anthropology 13: 161–170.

Cardoso HFV, Saunders SR. 2008. Two arch criteria of the ilium for sex determination of immature skeletal remains: A test of their accuracy and an assessment of intra- and inter-observer error. Forensic Science International 178:24-29.

Chapman T, Lefevre P, Semal P, Moiseev F, Sholukha V, Louryan S, Rooze M, Van Sint Jan S. 2014. Sex determination using the probabilistic sex diagnosis (DSP: Diagnose sexuelle probabiliste) tool in a virtual environment. Forensic Science International 234:189.e1-189.e8.

Coqueugniot H, Weaver T. 2007. Infracranial maturation in the skeletal collection from Coimbra, Portugal: New aging standards for epiphyseal union. American Journal of Physical Anthropology 134: 424–437.

Cox M. 1989. An evaluation of the significance of "scars of parturition" in the Christ church Spitalfields sample. Ph.D. Thesis. University College London: U.K.

Daniell C. 1998. *Death and Burial in Medieval England, 1066-1550*. London and New York: Routledge.

Davenport CAL. 2017. Combining forensic anthropological and geological approaches to investigate the preservation of human remains in British archaeological populations and their effects on palaeodemography. PhD Thesis. Liverpool John Moores University.

Dayal MR, Kegley ADT, Štrkalj G, Bidmos MA, Kuykendall KL. 2009. The History and Composition of the Raymond A. Dart Collection of Human Skeletons at the Universoty of the Witwatersrand, Johannesburg, South Africa. American Journal of Physical Anthropology 140:324-335.

De Fillippo C, Bostoen K, Stoneking M, Pakendorf B. 2012. Bringing together linguistic and genetic evidence to test the Bantu expansion. Proc R Soc B 279:3256-3263.

De Wit E, Delport W, Rugamika CE, Meintjes A, Möller M, van Helden PD, Seoghie C, Hoal EG. 2010. Genome-wide analysis of the structure if the South African Coloured Population in the Western Cape. Human Genetics 128:145-153.

Decker SJ, Davy-Jow SL, Ford JM, Hilbelink DR. 2011. Virtual Determination of Sex: Metric and Nonmetric Traits of the Adult Pelvis from 3D Computed Tomography Models. Journal of Forensic Sciences 56:1107–1114.

Delphy C. 1993. Rethinking sex and gender. Women's Studies International Forum 16:1-9.

Derry DE. 1911. The significance of the sulcus preauricularis. Anat Anz 39:13-20.

DeSilva JM, Lesnik JJ. 2006. Chimpanzee neonatal brain size: Implications for brain growth in Homo erectus. Journal of Human Evolution 51:207-212.

DeSilva JM. 2011. A shift toward birthing relatively large infants early in human evolution. Proceedings of the National Academy of Sciences 108:1022-1027.

Dixit SG, Kakar S, Agarwal S, Choudhry R. 2007. Sexing of human hip bones of Indian origin by discriminant function analysis. Journal of Forensic and Legal Medicine 14:429-437.

Dove ER. 2017. Stressful times: An investigation into the indicators of childhood health between an urban and rural population dating from the 12th to 17th centuries. PhD Thesis. Liverpool John Moores University.

Dray S, Dufour AB. 2007. The ade4 package: implementing the duality diagram for ecologists. Journal of Statistical Software 22:1-20.

Dudzik B, Jantz, RL. 2016. Misclassifications of Hispanics using Fordisc 3.1: Comparing cranial morphology in Asian and Hispanic populations. Journal of Forensic Sciences 61:1311-1318.

Dudzik B, Langley NR. 2015. Estimating age from the pubic symphysis: A new component-based system. Forensic Science International 257:98-105.

Durić M, Rakočević, Donić D. 2005. The reliability of sex determination of skeletons from forensic context in the Balkans. Forensic Science International 147:159-164.

Edgar HJH, Hunley KL. 2009. Race reconciled: How biological anthropologists view human variation. American Journal of Physical Anthropology 139:1-4.

Emery MM. 1996. The Archaeology of an Ecclesiastical Landscape: Chester Archaeological Society.

Emery MM. 2000. The Poulton Chronicles. Williamsburg, Virginia: Poulton Archaeology Press. 60 p.

Fairgrieve, S. I. and Oost, T. S. (1995), On a test of the multifactorial aging method by Bedford *et al*. (1993). American Journal of Physical Anthropology 97: 83–85.

Falys CG, Schutkowski H, Weston DA. 2006. Auricular surface aging: Worse than expected? A test of the revised method on a documented historic skeletal assemblage. American Journal of Physical Anthropology 130:508-513.

Fazekas IG, Kósa F. 1978. Forensic Fetal Osteology. Budapest: Akadémiai Kiadó.

Feldesman MR, Fountain RL. 1996. "Race" Specificity and the Femur/Stature Ratio. American Journal of Physical Anthropology 100: 207-224.

Fischer B, Mitteroecker P. 2015. Covariation between human pelvis shape, stature, and head size alleviates the obstetric dilemma. Proceedings of the National Academy of Sciences 112:5655-5660.

Franklin D., Cardini, A., Oxnard CE. 2010. A geometric morphometric approach to the quantification of population variation in sub-Saharan African crania. American Journal of Human Biology 22:23-35.

Franklin, D., Cardini, A., Oxnard CE. 2007. Geometric morphometric study of population variation in indigenous southern African crania. American Journal of Human Biology 19:20-33.

Fullbrook-Leggat. 1945. Medieval Gloucester: *I' Transactions of the Bristol and Gloucestershire Archaeological Society* 66:6-9.

Fully G. 1956. Une nouvelle méthode de détermination de la taille. Ann Med Legale 35:266–273.

Gamble J, Blackburn A, Hoppa RD. Congruence of methods for determination of sex using real, virtual and 3-D printed specimens. Proceedings of the 39th Conference in Computer Applications and Quantitative Methods in Archaeology, Beijing, China.

Gilbert BM, McKern TW. 1973. A method of aging the female os pubis. American Journal of Physical Anthropology 38:31-38.

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

Giles E, Elliot O. 1963. Sex determination by discriminant function analysis of crania. American Journal of Physical Anthropology 21:53-58.

Giles E. 1964. Sex determination by discriminant function analysis of the mandible. American Journal of Physical Anthropology 22:129-135.

Giles E. 1966. Statistical techniques for sex and race determination. Some comments in defence. American Journal of Physical Anthropology 25:85-86.

Giles E. 1968. Table I. Age effect in sexing crania by multivariate discriminant function analysis. Proceedings VIIIe. International Congress of Anthropological and Ethnological Sciences, Tokyo 1, pp. 59-61.

Glanville EV. 1967. Sexual dimorphism in the pelvic bones of the Telem, a medieval negro population from the Mali Republic. Konikl Nederl Akad Wetesch Amsterdam 70:368-377.

Goldberg CF. 1993. The application of stable carbon and nitrogen isotope anlaysis to human dietary reconstruction in prehistoric southern California. Ph.D. Thesis. University of California, Los Angeles.

Gómez-Valdés JA, Menéndez Garmendia A, García-Barzola L, Sánchez-Mejorada G, Karam C, Baraybar JP, Klales A. 2017. Recalibration of the Klales *et al*. (2012) method of sexing the human innominate for Mexican populations. American Journal of Physical Anthropology 162:600–604.

Gómez-Valdés JA, Quinto-Sánchez M, Germandia AM, Velemínska J, Sánchez-Mejorada G, Bruzek 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 JA, Ramírez GR, Molgado SB, Sánchez-Mejorada G. 2011. Discriminant function analysis for sex assessment in pelvis girdle bones: Sample from the contemporary Mexican population. Journal of Forensic Sciences 56:297-301.

Gómez-Valdés JA, Ramírez GT, Molgado SB, Sain-Leu PH, Caballero JLC, Sánchez-Mejorada G. 2011. Discriminant function analysis for sex assessment in pelvic girdle bones: Sample from the contemporary mexican population. Journal of Forensic Sciences 56:297–301.

Gonzalez PN, Bernal V, Perez SI. 2009. Geometric morphometric approach to sex estimation of human pelvis. Forensic Science International 189:68–74.

Greef JM. 2007. Deconstructing Jaco: Genetic heritage of an Afrikaner. Ann Hum Genet 71:674-688.

Grumbach MM, Conte FA. 1992. Disorders of sex differentiation. In Wilson JD and Foster DW (eds) Williams textbook of endocrinology, Sanuders, Philidelphia:853-951.

Hall M, Morris A. 1983. Race and Iron Age human skeletal remains from southern Africa: An assessment. Social Dynam 9:29-36.

Hanihari K, Suzuki T. 1978. Estimation of age from pubic symphysis by means of multiple regression analysis. American Journal of Physical of Anthropology 48:233-240.

Harma A, Karakas HM. 2007. Determination of sex from the femur in Anatolian Caucasians: A digital radiological study. Journal of Forensic and Legal Medicine 14:190-194.

Harris EF, Bailit HL. 1988. A Principal Components Analysis of human odontometrics. American Journal of Physical Anthropology 75:87-99.

Hefner JT, Ousley SD. 2014. Statistical classification methods for estimating ancestry using morphoscopic traits. Journal of Forensic Sciences 59:883–890.

Hefner JT. 2009. Cranial Nonmetric Variation and Estimating Ancestry. Journal of Forensic Science 54:985-995.

Hens S, Belcastro M. 2012. Auricular surface aging: a blind test of the revised method on historic Italians from Sardinia. Forensic Science International 214:209.e1–e5.

Herbert RK. 1990. The sociohistory of clicks in Southern Bantu. Anthropol Linguist 32:295-315.

Holobinko A. 2012. Forensic human identification in the United States and Canada: A review of the law, admissible tehniques, and the legal implications of their application in forensic cases. Forensic Science International 222:394e.1-394e.13.

Hoshi H. 1961. On the preauricular groove in the Japanese pelvis with special reference to the sex difference. Okajimas Folia Anatomica Japonica 37:259-269.

Houghton P. 1974. The relationship of the pre-auricular groove and childbirth alterations at the pubis, ilium and sacrum. American Journal of Physical Anthropology 41:381-389.

Houghton P. 1975. The bony imprint of pregnancy. Bull N Y Acad Med 51:655-661.

Howells, W.W. *Cranial Variation in Man: A Study by Multivariate Analysis of Patterns of Difference among Recent Human Populations*. Papers of the Peabody Museum 67. Peabody Museum, Harvard University, Cambridge, Mass, 1973.

IBM Corp. Released 2012. IBM SPSS Statistics for Macintosh, Version 21.0. Armonk, NY: IBM Corp.

Igbigbi PS, Igbigbi AM. 2003. Determination of sex and race from the subpubic angle in Ugandan subjects. American Journal of Forensic Medicine and Pathology 24:168-172.

Irish JD. 2005. Population continuity vs. discontinuity revisited: Dental affinities among late paleolithic through Christian-Era Nubians. American Journal of Physical Anthropology 128:520–535.

İşcan MY, Steyn M. 1999. Craniometric determination of population affinity in South Africans. International Journal of Legal Medicine 112:91-97.

Jantz RL, Ousley SD. 2005. FORDISC 3.0: Personal Computer Forensic Discriminant Functions. Knoxville: University of Tennessee.

Jantz RL. 1992. Modification of the Trotter and Gleser female stature estimation formulae. Journal of Forensic Sciences 37:1230-1235.

Jit I, Kulkarni M. 1976. Times of appearance and fusion of epiphyses at the medial end of the clavicle. Indian Journal of Medical Research 64: 773–782.

Jit I, Singh B. 1971. A radiological study of the time of fusion of certain epiphyses in Punjabees. Journal of Anatomical Society India 20: 1–27.

Johnson DR, O'Higgins P, Moore WJ, McAndrew TJ. 1989. Determination of race and sex of the human skull by discriminant function analysis of linear and angular dimensions. Forensic Science International 41:41-53.

Johnson JR, Thomas WS, Henry OA, Don PM. 1999. Arlington Springs Revisited. In Proceedings of the 5th California Island Symposium, Santa Barbara Museum of Natural History. DR Brown, HS Chaney, and KL Mitchell: 541-545. Department of Interior Minerals Management Service: Carmarillo, CA.

Jovanonic S, Zivanovic S. 1965. The establishment of the sex by the great sciatic notch. Acta Anat (Basel) 61:101-107

Jungers WL, Larson SG, Harcourt-Smith W, Morwood MJ, Sutikna T, Due Awe R, Djubiantono T. 2009. Descriptions of the lower limb skeleton of *Homo floresiensis*. Journal of Human Evolution 57:538-554.

Kelly G. 2016. A (nearly) complete glossary of gender identities for your next census. *The Telegraph*. [online] Available at: http://www.telegraph.co.uk/men/the-filter/a-nearly-complete-glossary-of-gender-identities-for-your-next-ce/ [Accessed 2 Jul. 2017].

Kennedy KAR, Chiment J, Disotell T, Meyers D. 1984. Principal-components analysis of prehistoric South Asian crania. American Journal of Physical Anthropology 64:105-118.

Kenyhercz MW, Passalacqua NV, Hefner JT. 2016. Missing data imputation methods using morphoscopic traits and their performance in the estimation of ancestry. Proceedings of the 68th Annual Meeting of the American Academy of Forensic Sciences, Las Vegas, NV, USA:143

Kenyhercz MW, Passalacqua NV. 2016. Missing Data Imputation Methods and Their Performance With Biodistance Analyses. In: Biological Distance Analysis. Elsevier. p 181–194.

Kenyhercz MW. 2012. Sex estimation using pubic bone morphology in a modern South African sample: A test of the Klales *et al.* method. [ABSTRACT] American Journal of Physical Anthropology S54:179-180.

Key CA, Aiello LC, Molleson T. 1994. Cranial suture closure and its implications for age estimation. International Journal of Osteoarchaeology 4:194-207.

Key P, Jantz RL. 1981. A multivariate analysis of temporal change in Arikara craniometrics: A methodological approach. American Journal of Physical Anthropology 2:247-259.

Kjellström A. 2004. Evaluations of sex assessment using weighted traits on incomplete skeletal remains. International Journal of Osteoarchaeology 14:360-373.

Klales A, Ousley SD, Vollner JM. 2012. A revised method of sexing the human innominate using Phenice's nonmetric traits and statistical methods. American Journal of Physical Anthropology 149:104-114.

Klales AR, Burns TL. 2017. Adapting and Applying the Phenice (1969) Adult Morphological Sex Estimation Technique to Subadults. Journal of Forensic Sciences 62:747–752.

Klales AR, Kenyhercz M, Stull K, McCormick K, Call S. 2016. Worldwide population variation in pelvic sexual dimorphism. [ABSTRACT] American Journal of Physical Anthropology 159:193.

Klales AR, Kenyhercz MW. 2015. Morphological Assessment of Ancestry using Cranial Macromorphoscopics, Journal of Forensic Sciences 60:13–20.

Klales AR. 2016. Secular change in morphological pelvic traits used for sex estimation. Journal of Forensic Sciences 61:295-301.

Kobayashi K. 1967. Trend in the length of life based on human skeletons from prehistoric to modern times in Japan. Journal of the Faculty of Science, the University of Tokyo, Section V, Anthropology, 3: 107–162.

Kramer MS, Kakuma R. 2003. Energy and protein intake in pregnancy. Cochrane Database Syst Rev 4:CD000032.

Kranioti EF, Vorniotakis N, Galiatsou C, Iscan MY, Michalodimitrakas M. 2009. Sex identification and software development using digital femoral head radiographs. Forensic Science International 189:113.e1-113.e7.

Krüger GC, L'Abbé EN, Stull KE, Kenyhercz MW. 2015. Sexual dimorphism in cranial morphology among modern South Africans. International Journal of Legal Medicine 129:869-875.

Kurki HK. 2007. Protection of obstetric dimensions in a small-bodied human sample. American Journal of Physical Anthropology 133:1152-1165.

Kurki HK. 2011. Pelvic dimorphism in relation to body size and body size dimorphism in humans. Journal of Human Evolution 61:631-643.

Kurki HK. 2013a. Bony pelvic canal size and shape in relation to body proportionality in humans. American Journal of Physical Anthropology 151.88-101.

Kurki HK. 2013b. Skeletal variability in the pelvis and limb skeleton of humans: Does stabilizing selection limit female pelvic variation? American Journal of Human Biology 25:795-802.

L'Abbé EN, Kenyhercz M, Stull KE, Keough N, Nawrocki S. 2013. Application of Fordisc 3.0 to explore differences among crania of North Americans and South African blacks and whites. Journal of Forensic Sciences 58:1579-1583.

L'Abbé EN, Loots M, Meiring JH. 2005. The Pretoria Bone Collection: a modern South African skeletal sample. Journal of Comparative Human Biology HOMO 56:197-205.

L'Abbé EN, van Rooyen C, Nawrocki SP, Becker PJ. 2011. An evaluation of non-metric cranial traits used to estimate ancestry in a South African sample. Forensic Science International 209:195-e1.

Lavallo G. 2013. Variation in non-metric traits of the pelvis between Whites, Blacks, and Hispanics. M.A. Thesis. Texas State University-San Marcos, USA.

Lazorthes G, Lhez A. 1940. Les caractés sexuels de la grande échancrure sciatiqu. Soc Anthropol Paris Bull Mem 10:180-185.

Liebenberg L. 2014. Postcraniometric analysis of ancestry among modern South Africans. MSc Thesis, Department of Anatomy, University of Pretoria.

Listi GA, Bassett HE. 2006. Test of an alternative method for determining sex from the os coxae: Applications for modern Americans. Journal of Forensic Science 51:248-252.

Listi GA. 2010. The impact of racial metric variation in the pelvis in the morphological assessment of sex. Journal of Forensic Sciences 55:1157-1161.

Lockwood CA, Kimbel WH, Lynch JM. 2004. Morphometrics and hominoid phylogeny: Support for a chimpanzee-human clade and differentiation among great ape subspecies. Proceedings of the National Academy of Sciences 101:4356-4360.

Lovejoy CO, Meindl RS, Mensforth RP, Barton TJ. 1985a. Multi- factorial determination of skeletal age at death: a method and blind tests of its accuracy. American Journal of Physical Anthropology 68:1–14.

Lovejoy CO, Meindl RS, Pryzbeck TR, Mensforth RP. 1985b. Chronological metamorphosis of the auricular surface of the ilium: a new method for the determination of adult skeletal age at death. American Journal of Physical Anthropology 68:15–28.

Lovell NC. 1989. Test of Phenice's technique for determining sex from the os pubis. American Journal of Physical Anthropology 79:117-120.

Luo YC. 1995. Sex determination from the pubis by discriminant function analysis. Forensic Science International 74:89-98.

Macaluso PJ. 2011. Sex discrimination from the acetabulum in a twentieth-century skeletal sample from France using digital photogrammetry. HOMO- Journal of Comparative Human Biology 62:44–55.

MacLaughlin SM, Bruce MF. 1986. The sciatic notch/acetabular index as a discriminator of sex in European skeletal remains. Journal of Forensic Science 31:1380-1390.

Maclaughlin SM, Bruce MF. 1990. The accuracy of sex identification in European skeletal remains using the Phenice characters. Journal of Forensic Science 35:1384-1392.

Manikandan S. 2011. Measures of central tendency: mean and mode. Journal of Pharmacology & Pharmacotherapeutics 2:214-214.

Martin R. 1914. Lehrbuch der Anthropologie. Jena: G. Fischer.

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

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

Meadows L, Jantz RL. 1992. Estimation of stature from metacarpal lengths. Journal of Forensic Sciences 37:147-154.

Meindl RS, Lovejoy CO, Mensforth RP, Carlos LD. 1985. Accuracy and direction of error in the sexing of the skeleton: Implications for paleodemography. American Journal of Physical Anthropology 68:79-85.

Meindl RS, Lovejoy CO, Mensforth RP, Walker RA. 1985a. A revised method of age determination using the os pubis, with a review and tests of accuracy of other current methods of pubic symphyseal aging. American Journal of Physical Anthropology 68:29-45.

Meindl RS, Lovejoy CO, Mesnforth RP, Don Carlos L. 1985b. Accuracy and direction of error in sexing of the skeleton: Implications for paleodemography. American Journal of Physical Anthropology 68:79-85.

Meindl RS, Lovejoy CO. 1985. Ectocranial suture closure: A revised method for the determination of skeletal age at death based on the lateral-anterior sutures. American Journal of Physical Anthropology 68:57-66.

Milner GR, Boldsen JL. 2012. Transition Analysis: A validation study with known-age modern American skeletons. American Journal of Physical Anthropology 148:98-110.

Mitteroecker P, Bookstein F. 2011. Linear discrimination, ordination, and the visualization of selection gradients in modern morphometrics. Evolutionary Biology 38:100-114.

Molleson T, Cox M. 1993. The Spitalfields Project, Vol. 2: The Anthropology. The Middling Sort. Council for British Archaeology Research Report 86. York: Council for British Archaeology.

Moraitis K, Zorba E, Eliopoulos C, Fox SC. A test of the revised auricular surface aging method on a modern European population. Journal of Forensic Sciences 59:188-194.

Morgan P. 1978. *Domesday Book. No26: Cheshire*. Chichester: Philimore & Co. Ltd.

Morris A, 2012. Biological Anthropology at the Southern Tip of Africa Carrying European Baggage in an African Context. Current Anthropology 53.

Morrison DG. 1969. On the interpretation of discriminant analysis. Journal of Marketing Research 6:156-163.

Msamati BC, Igbigbi PS, Manda JK. 2005. The subpubic angle in adult indigenous Malawian subjects. East African Medical Journal 82:643-648.

Mulhern DM, Jones EB. 2005. Thest of revised method of age estimation from the auricular surface of the ilium. American Journal of Physical Anthropology 126:61-65.

Murail P, Bruzek J, Houët F, Cunha E. 2005. DSP: a tool for probabilistic sex diagnosis using worldwide variability in hip-bone measurements. Bulletins et Mémoires de la Société d'Anthropologie de Paris 17:167-176.

Murphy AMC. 2005. The femoral head: sex assessment of prehistoric New Zealand Polynesian skeletal remains. Forensic Science International 154:210-213.

Nagesh KR, Kanchan T, Bastia BK. 2007. Sexual dimorphism if acetabulum-pubis index in South-Indian population. Legal Medicine 9:305-308.

Naňko O, Šedý J, Jarolím L. 2007. Sulcus Nervi Dorsalis Penis/Clitoridis: its Reliability as a Character for Sex Determination of Isolated Human Pubic Bones. Prague Medical Report 108:167-176

Novak L, Schultz JJ, McIntyre M. 2012. Determining Sex of the Posterior Ilium from the Robert J. Terry and William M. Bass Collections. Journal of Forensic Sciences 57:1155–1160.

Orr PC. 1962. The Arlington Spring site, Santa Rosa Island, California. American Antiquity 27(3):417-419.

Ousley SD, Jantz RL, Freid D. 2009. Understanding Race and Human Variation: Why Forensic Anthropologists are good at Identifying Race. American Journal of Physical Anthropology 139:68-75.

Parsons FG. 1914. The Characteristics of the English Thigh-Bone. Journal of Anatomy and Physiology 48:238-267.

Parsons FG. 1915. The Characteristics of the English Thigh-Bone Part II: The Difficulty of Sexing. Journal of Anatomy and Physiology 49:345-361.

Patriquin ML, Steyn M, Loth SR. 2002. Metric assessment of race from the pelvis in South Africans. Forensic Science International 127:104-113.

Patriquin ML, Steyn M, Loth SR. 2005. Metric analysis of sex differences in South African black and white pelves. Forensic Science International 147:119-127.

Patterson N, Peterson DC, van der Ross RE, Sudoyo H, Glashoff RH, Marzuki S, Reich D, Hayes VM. 2010. Genetic structure of a unique admixed population: Implications for medical research. Human Molecular Genetics 19:411-419.

Perizonius WRK. 1984. Closing and non-closing sutures in 256 crania of known age and sex from Amsterdam (AD 1883-1909). Journal of Human Evolution 13:201-216.

Peterson DC, Libiger O, Tindall EA, Hardie R, Hannick LI, Glashoff RH, Mukerji M, Indian Genome Variation Consortium, Fernandez P, Haacke W, Schork NJ, Hayes, VM. 2013. Complex patterns of genomic admixture within southern Africa. PLOS Genetics 9:e1003309.

Phenice TW. 1969. A newly developed visual method of sexing the os pubis. American Journal of Physical Anthropology 30:297-302.

Powell JF, Neves WA. 1999. Craniofacial morphology of the first Americans: Pattern and process in the peopling of the New World. American Journal of Physical Anthropology 110:153-188.

Press SJ, Wilson S. 1978. Choosing between Logistic Regression and Discriminant Analysis. Journal of the American Statistical Association 73:699–705.

Purkait R. 2003. Sex determination from femoral head measurements: a new approach Legal Medicine 5:347-350.

R Core Team. 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rawes B, and Hodsdon J. (1990). Glevensis: The Gloucester and District Archaeological Research Group Review. ISSN: 0307-210X.

Raxter MH, Auerbach BM, Ruff CB. 2006. Revision of the Fully technique for estimating statures. American Journal of Physical Anthropology 130:374-384.

Reeder, LA, Rick TC, Erlandson JM. 2008. Forty Years Later: What have we learned about the earliest human occupations of Santa Rosa Island, California. North American Archaeologist 29(1):37-64.

Reeves J, Adams M. 1993. The Spitalfields Project, Vol. 1: The Archaeology. Across the Styx. Council for British Archaeology Research Report 85. York: Council for British Archaeology.

Rennie SR, Clegg M, Ohman JC. 2015. Hips don't lie: A multivariate approach to hominin sex determination from the pelvis. [ABSTRACT] American Journal of Physical Anthropology S60:65-334.

Rhine SS. 1990. Non-metric skull racing. In Skeletal Attribution of Race: Methods for Forensic Anthropology, edited by G. W. Gill and S. Rhine, pp. 9-20. Maxwell Museum Anthropological Papers No. 4, Albuquerque, NM.

Ribot I. 2004. Differentiation of Modern Sub-Saharan African Populations: Craniometric interpretations in relation to geography and history. Bull Mém Soc Anthropol Paris 16:143-65.

Ridgeway BM, Arias BA, Barber DB. 2008. Variation of the obturator foramen and pubic arch of the female bony pelvis. American Journal of Obstetrics and Gynecology 198(5):546.e1-546.e4.

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

Rösing FW, Graw M, Marré, Ritz-Timme S, Rothschild MA, Rötzscher K, Schmeling A, Schröder I, Geserick G. 2007. Recommendations for the forensic diagnosis of sex and age from skeletons. HOMO – Journal of Comparative Human Biology 58:75-89.

Ross R. 1999. A Concise History of South Africa. University Press, Cambridge, UK.

RStudio Team. 2015. RStudio: Integrated Development for R. RStudio, Inc., Boston, MA URL http://www.rstudio.com/.

Sampson C. 2010. Chronology and dynamics of Later Stone Age herfers in the upper Seacrow River valley, South Africa. J Arid Environ 4:842-848.

Sashin D. 1930. A critical anlaysis of the anatomy and the pathological changes of the sacroi-iliac joints. The Journal of Bone and Joint Surgery 12:891-910.

Saunders SR. 1992. Subadult skeletons and growth related studies, in: Saunders SR, Katzenberg MA (eds) Skeletal biology of past peoples: Research Methods, Wiley-Liss, New York:1-20.

Schaefer M, Black S, Scheuer L. 2009. Juvenile osteology: `A laboratory and field manual. – Academic Press, London.

Schaefer M. 2008. A summary of epiphyseal union timings in Bosnian males. International Journal of Osteoarchaeology. DOI: 10.1002/oa.959

Schaeffer JP. 1953. Morris' human anatomy. A complete systematic treatise, 11th ed. New York-Toronto: Blakiston.

Scholl TO, Chen X, Khoo CS, Lenders C. 2004. The dietary glycemic index during pregnancy: influence on infant birth weight, fetal growth, and biomarkers of carbohydrate metabolism. American Journal of Epidemiology 159:467-474.

Schultz R, Mühler M, Mutze S, Schmidt S, Reisinger W, Schmeling A. 2005. Studies on the time frame for ossification of the medial epiphysis of the clavicle as revealed by CT scans. International Journal of Legal Medicine 119: 142–145.

Schunke D. 1938. The anatomy and development of the sacro iliac joint in man. Anatomical Record 72:313–31.

Schutkowski H. 1993. Sex determination of infant and juvenile skeleton: I. Morphognostic features. American Journal of Physical Anthropology 90:199-205.

Šedý J, Naňko O, Špačková J, Jarolím L. 2008. Clinical Implications of a Close Vicinity of Nervus Dorsalis Penis/Clitoridis and Os Pubis. Journal of Sexual Medicine 5:1572-1581.

Siddiqi N. 2013. Comparison of osteometric femoral bone dimensions among the South African of different ethnic groups and South African whites. Egypt J Forensic Sci 3:8-14.

Sidler M, Jackowski C, Dirnhofer R, Vock, Thali M. 2007. Use of multislice computed tomography in disaster victim identification – Advantages and limitations. Forensic Science International 169:118-128.

Simpson SW, Quade J, Levin NE, Butler, Dupont-Nivet G, Everett M, Semaw S. 2008. A female *Homo erectus* pelvis from Gona, Ethiopia. Science 322:1089-1092.

Singer R. 1953. Estimation of age from cranial suture closure: a report on its unreliability. Journal of Forensic Medicine 1:52-59.

Singh S, Potturi BR. 1978. Greater sciatic notch in sex determination. Journal of Anatomy 125:619-624.

Sokal RR, and Rohlf FJ. 2012. Biometry. San Francisco: W.H. Freeman and Company. 960 p.

Spradley MK, Jantz RL. 2011. Sex estimation in forensic anthropology: Skull versus postcranial elements. Journal of Forensic Sciences 56:289-296.

Spring DB, Lovejoy CO, Bender GN, Duerr M. The radiographic preauricular groove: Its non-relationship to past parity. American Journal of Physical Anthropology 79:247-252.

St. Hoyme LE. 1984. Sex differentiation in the posterior pelvis. Collegium Anthropologica 8:139-153.

Steel FLD. 1962. The Sexing of Long Bones, with reference to the St Bride's Series of Identified Skeletons. The Journal of the Royal Anthropological Institute of Great Britain and Ireland 92:212-222.

Stein PL, Rowe BM. 1989. Physical Anthropology. 4th ed. McGraw-Hill Publishing Company. p.177-191.

Steudel K. 1978. A multivariate analysis of the pelvis of early hominids. Journal of Human Evolution 7:583-595.

Stewart TD. 1979. Essentials of Forensic Anthropology: Especially as developed in the United States. Charles C Thomas. p.227-238.

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

Steyn M, Iscan MY. 1999. Osteometric variation in the humerus: Sexual dimorphism in South Africans. Forensic Science International 106:77-85.

Steyn M, Iscan MY. 2008. Metric sex determination from the pelvis in modern Greeks. Forensic Science International 179:86.e1-86.e6.

Steyn M, Patriquin ML. 2009. Osteometric sex determination from the pelvis – Does population specifity matter? Forensic Science International 191:113.e1-113.e5.

Steyn M, Smith JR. 2007. Interpretation if ante-mortem stature estimates in South Africans. Forensic Science International 171:97-102.

Stojanowski CM, Siedmann RM, Doran GH. 2002. Differential skeletal preservation at Windover Pond: causes and consequences. American Journal of Physical Anthropology 119:15-26.

Stull KE, Kenyhercz MW, L'Abbe EN. 2013. Non-metric cranial and pelvic traits as a measure of sexual dimorphism in a modern South African population. [Abstract] American Journal of Physical Anthropology S56:64-301.

Stull KE, L'Abbé EN, Ousley SD. 2014. Using multivariate adaptive regression splines to estimate subadult age from diaphyseal dimensions. American Journal of Physical Anthropology 154:376-386.

Stull KE, L'Abbé EN, Ousley SD. 2017. Subadult sex estimation from diaphyseal dimensions. American Journal of Physical Anthropology 163:64-74.

Stull KE. 2013. An osteometric evaluationof age and sex differences in the long bones of South African children from the Western Cape. PhD Thesis: University of Pretoria, South Africa.

Stynder DD. 2009. Craniometric evidence for South African Later Stone Age herders and hunter-gatherers being a single biological population. Journal of Archaeological Science 36:798-806.

Suchey JM, Brooks ST. 1990. Skeletal age determination based on the os pubis: A comparison of the Acsadi-Nemeskeri and Suchey-Brooks methods. Journal of Human Evolution 5:227-238.

Suchey JM. 1979. Problems in aging of females using the os pubis. American Journal of Physical Anthropology 51:467-470.

Sutherland C. 2015. Biological distance among modern and parental South African groups using discrete traits of the skull. MSc Thesis, Department of Anatomy, University of Pretoria.

Tague RG, Lovejoy CO. 1986. The obstetric pelvis of A. L. 288-1 (Lucy). Journal of Human Evolution 15:237-255

Tague RG. 1988. Bone resorption of the pubis and preauricular area in humans and nonhuman mammals. American Journal of Physical Anthropology 76:251-267.

Tague RG. 1992. Sexual dimorphism in the human bony pelvis, with consideration of the Neandertal pelvis from Kebara Cave, Israel. American Journal of Physical Anthropology 88:1-21.

Tague RG. 2000. Do big females have big pelves? American Journal of Physical Anthropology 112:377-393.

Tague RG. 2005. Big-bodies males help us recognize that females have big pelves. American Journal of Physical Anthropology 127:392-405.

Thompson L. 2013. A History of South Africa, Fourth Edition. Yale University Press, New Haven and London.

Todd TW, Lyon DW. 1924. Endocranial suture closure, its progress and age relationship: Part I adult males of the white stock. American Journal of Physical Anthropology 7:325-384.

Todd TW, Lyon DW. 1925a. Cranial suture closure, its progress and age relationship: Part II ectocranial closure in adult males of white stock. American Journal of Physical Anthropology 8:23- 45.

Todd TW, Lyon DW. 1925b. Cranial suture closure, its progress and age relationship: Part III endocranial closure in adult males of negro stock. American Journal of Physical Anthropology 8:47-71.

Todd TW, Lyon DW. 1925c. Cranial suture closure, its progress and age relationship: Part IV ectocranial closure in adult males of negro stock. American Journal of Physical Anthropology 8:149-168.

Todd TW. 1920. Age changes in the pubic bone: I. The white male pubis. American Journal of Physical Anthropology 3:285-334.

Todd TW. 1921a. Age changes in the pubic bone: 11. The pubis of the male Negro-White hybrid III: the pubis of the white female; IV. The pubis of the female Negro- White hybrid. American Journal of Physical Anthropology 4:l-70.

Todd TW. 1921b. Age changes in the pubic bone: V. Mammalian pubic metamorphosis. American Journal of Physical Anthropology 4:333406.

Todd TW. 1921c. Age changes in the pubic bone: VI. The interpretation of variations in the symphyseal area. American Journal of Physical Anthropology 4:407-424.

Town NJ. 2016. The analysis of preserved and degraded human skeletal material: Understanding relationships between bone and the soil environment. Ph.D. Thesis. Anthropology Department, Liverpool John Moores University, UK.

Trotter M. 1970. Estimation of stature from intact long limb bones, in Stewart, T.D. (ed.), *Personal Identification in Mass Disasters*: National Museum of Natural History, Washington, pp. 71-83.

Turner CAL, Burrell CL, Carpenter RJ, Ohman JC. 2013. Biological age estimation of subadult human skeletal remains: Comparison of dental development with the humerus, femur and *pars basilaris*. Poster Presentation. British Association for Biological Anthropology and Osteoarchaeology, University of York.

Ubelaker DH, De La Paz JS. 2012. Skeletal indicators of pregnancy and parturition: A historical review. Journal of Forensic Science 57:866-872.

Valoriani S. 2018. Cranial remains from the graveyard to the classroom: Restoration, conservation, and analysis of medieval British skeletal collections. PhD Thesis. Liverpool John Moores University.

Van Sickle C. 2014. A new examination of childbirth-related pelvic anatomy in Neandertal females. PhD Thesis: University of Michigan (USA).

Vance VLW. 2007. Age related changes in the post-cranial human skeleton and its implication for the determination of sex. Ph.D. Thesis. Faculty of Health Sciences, University of Pretoria, RSA.

VanPutte C, Regan J, Russo A. (2010). Seeley's Essentials of Anatomy and Physiology 7th ed. New York: McGraw-Hill.

Velemínska J, Krajíček V, Dupej J, Gómes-Valdés JA, Velemínský P, Šefčáková A, Pelikán J, Sánchez-Mejorada G, Brůžek J. 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.

Waldron T. 1987. The relative survival of the human skeleton: implications for palaeopathology. In: Boddington A, Garland AN, Janaway RC, editors. Death, decay and reconstruction: approaches to archaeology and forensic science. Manchester: Manchester University press. P 55-64.

Walker PL, DeNiro MJ. 1986. Stable nitrogen and carbon isotope ratios in bone collagen as indices of prehistoric dietary dependence on marine and terrestrial resources in southern California. American Journal of Physical Anthropology 71:51-61.

Walker PL, Erlandson J. 1986. Dental Evidence for Prehistoric Dietary Change on the Northern Channel Islands. American Antiquity 51(2):375-383.

Walker PL. 2005. Greater sciatic notch morphology: Sex, age, and population differences. American Journal of Physical Anthropology 76:183-188.

Walker PL. 2006. Californian Bioarchaeology, in: Handbook of North American Indians, Volume 3: Environment, Origins, and Population, edited by Ubelaker DH. Washington DC: Smithsonian Institution Press.

Walker PL. 2007. California Bioarchaeology. In: Ubelaker DH ed. *Handbook of North American Indians, Volume 3: Environment, Origins, and Population*. Washington DC: Smithsonian Insitution Press.

Walker PL. 2008. Sexing skulls using discriminant function analysis of visually assessed traits. American Journal of Physical Anthropology 136:39-50.

Walker RA, Lovejoy CO. 1985. Radiographic Changes in the clavicle and proximal femur and their use in the determination of skeletal age at death. American Journal of Physical Anthropology 68:67-78.

Webb POA, Suchey JM. 1985. Epiphysal union of the anterior iliac crest and medial clavicle in a modern sample of American males and females. American Journal of Physical Anthropologists 68:457-466.

Wells JCK, DeSilva JM, Stock JT. 2012. The obstetric dilemma: An ancient game of Russian roulette, or a variable dilemma sensitive to ecology? American Journal of Physical Anthropology 149:40-71.

Wescott DJ. 2015. Sexual Dimorphism in Auricular Surface Projection and Postauricular Sulcus Morphology. Journal of Forensic Sciences 60:679–685.

White TD, Folkens PA. 2000. Human Osteology, 2nd ed. San Diego: Academic Press.

Whittock M. 2009. *A Brief History of Life in the Middle Ages*. London: Robinson.

Wilson LA, MacLeod N, Humphrey LT. 2008. Morphometric criteria for sexing juvenile human skeletons using the ilium. Journal of Forensic Sciences 53:269-278.

Wirth T, Wang X, Linz B, Novick RP, Lum JK, Blaser M, Morelli G, Falush D, Achtman M. 2004. Distinguishing human ethnic groups by means of sequences from Helicobacter pylori: Lessons from Ladakh. Proceedings of the National Academy of Sciences 101:4746–4751.

Wittman A, Wall L. 2007. The Evolutionary Origins of Obstructed Labor: Bipedalism, Encephalization, and the Human Obstetric Dilemma. Obstet Gynecol Surv [Internet] 62:739–748.

Wood ET, Stover DA, Ehret C. Destro-Bisol G, Spedini G, McLeod H, Louie L, Bamshad M, Strassman BI, Soodyall H, Hammer MF. 2005. Contrasting patterns of Y chromosome and mtDNA variation in Arica: Evidence for sex-biased demographic processes. European Journal of Human Genetics 13:1-13.

Workshop of European Anthropologists. 1980. Recommendations for age and sex diagnoses of skeletons. Journal of Human Evolution 9:517-549.

Zaaijer T. 1866. Untersuchungen über die form des beckens javanischer frauen. Natuurk Verh Holl Maatsch Wet Haarlem 24:1-42.

Zammit M, Norris RM. 2013. Sex determination in adults of two otariid species (*Neophoca cinerea* and *Arctophoca australis forsteri*) based on osteology of the pelvis. Marine Mammal Science 29:E515-E5519.

Zani S, Berzieri L. 2008. Measuring Customer Satisfaction Using Ordinal Variables: An application in a Survey on a Contact Center. Statistica Applicata 20:331–351.

Zar JH. 2010. Biostatical Analysis fifth ed. New Jersey: Pearson Education. 944 p.