Comparison of sexually dimorphic patterns in the postcrania of South Africans and North Americans

by

Gabriele Christa Krüger

Submitted in fulfilment of the requirements for the degree
Master of Science
MSc (Anatomy)

in the Faculty of Health Sciences
Department of Anatomy
University of Pretoria

2014

Supervisors
Prof EN L'Abbé
Dr KE Stull

Declaration:

I, Gabriele Christa Krüger, declare that the thesis, which I hereby submit for the degree Master of Science (Anatomy) to the University of Pretoria, is my own work and has not previously been submitted by me for a degree at another university.

_____ _____

Gabriele C Krüger Date

Abstract

Recently, a multivariate approach to sex estimation using the postcrania achieved higher accuracies than the commonly used cranium, recommending postcraniometric sex estimation. In South Africa (SA), sex estimation techniques using long bones exist, but they only offer univariate or bone-by-bone models, which do not acknowledge sexual dimorphism in other elements and, as expected, do not utilize robust statistical analyses. The purpose of this study was to evaluate accuracies of sex estimation in the postcrania of modern South Africans using multivariate statistics and to compare pattern expression of sexual dimorphism in blacks, whites and coloureds.

The study included analysing the skeletons of a total of 360 SA black, white and coloured individuals and the data of 240 black and white North Americans (NA) (equal sex and ancestry). Sympercents expressed sexual dimorphism and were compared in the three SA groups and with the NA individuals. Three classification methods assessed the thirty-nine standard measurements taken from 11 postcranial bones. The creation of different bone models and a variety of multivariate models revealed the potential of a multivariate technique. Comparisons of linear discriminant analysis (LDA), flexible discriminant analysis (FDA) and logistic regression indicated which model provided the greatest discriminatory power between sex and sex-ancestry groups in SA.

Sexual dimorphism was most apparent in SA coloureds, whereas whites were least dimorphic. However, overall NA individuals were more sexually dimorphic. The three classification methods achieved accuracies between 74-91% for bone models when estimating sex, with the clavicle classifying best. Multivariate subsets (different combinations of variables) achieved correct classifications that ranged from 88% to 97% for logistic regression, between

87% and 98% for LDA and between 90% and 97% using FDA. When classifying into sex and ancestry, an all-variable subset (six variables) achieved classification accuracies of 79% and 80% for FDA and LDA, respectively. Overall FDA achieved the best results whereas logistic regression achieved the lowest results for both bone models and multivariate subsets.

Postcranial bones achieve comparable classification accuracies to the pelvis and higher accuracies than metric or morphological techniques using the cranium in South Africa. Large differences in sexual dimorphism between NA and SA warrant the creation of population-specific standards and custom databases for South Africa.

Opsomming

Onlangs het 'n meer-veranderlike benadering tot geslagsberaming met die postkraniale skelet hoër akkuraatheid bereik as die meer algemeen gebruikde skedel. Dus, word postkraniometriese geslagsberaming aanbeveel. In Suid-Afrika (SA) is geslagsberamingstegnieke met die gebruik van langbene beskikbaar, maar maak slegs gebruik van een-veranderlike of been-by-been modelle. Die bogenoemde modelle erken net seksuele dimorfisme in enkele elemente en benut swak statistiese ontledings. Die doel van hierdie studie was om die akkuraatheid van geslagsberaming in die postkraniale skelet van moderne Suid-Afrikaners te evalueer met behulp van meer-veranderlike statistiese tegnieke en om die patroonuitdrukking van seksuele dimorfisme in swart, wit en kleurling individue te vergelyk.

Die studie het die postkraniale skelette van 'n totaal van 360 SA swart, wit en kleurling individue asook die data van 240 swart en wit Noord-Amerikaners (NA) elk met gelyke getalle van seks en afkoms geanaliseer. Vlakke van geslagsdimorfisme was vergelyk tussen die drie SA groepe en die NA individue deur gebruik te maak van simmetriese persentasieverskille, of sympercents. Drie klassifikasie metodes was gebruik om nege-en-dertig standaard metings van elf postkranial bene te assesseer. Die samestelling van verskillende beenmodelle en 'n verskeidenheid meer-veranderlike modelle bewys die potensiaal van 'n meer-veranderlike tegniek. Vergelykings van lineêre diskriminante analise (LDA), aanpasbare diskriminante analise (FDA) en logistiese regressie het aangedui watter model die grootste diskriminerende mag bereik tussen seks en seks-afkoms groepe in Suid-Afrika.

Geslagsdimorfisme was mees opmerkbaar in SA kleurlinge, terwyl wittes die minste dimorfisme getoon het. In die algeheel was NA individue meer dimorfies as SA individue. Die drie klassifikasie metodes het akkuraatheid tussen 74-91% bereik vir die beenmodelle, waar die

klavikel-model die beste klassifikasie akkuraatheid bereik het. Meer-veranderlike metingstelle (verskillende kombinasies van veranderlikes) het korrekte klassifikasies behaal wat wissel van 88% tot 97% vir logistiese regressie, tussen 87% en 98% vir LDA en tussen 90% en 97% vir FDA. Wanneer beide geslag en afkoms geassesseer word, het die alle-veranderlike metingstel (ses veranderlikes) klassifikasie akkuraatheid van 79% en 80% bereik vir die FDA en LDA, onderskeidelik. FDA het in die algeheel die beste resultate bereik, terwyl logistiese regressie die laagste resultate vir beide beenmodelle en meer-veranderlike metingstelle bereik het.

In die Suid-Afrikaanse populasie bereik postkraniale bene klassifikasie akkuraatheid gelykstaande aan die pelvis en hoër akkuraatheid as metriese of morfologiese analise wat gebruik maak van die skedel. Verskille in die vlakke van geslagsdimorfisme tussen NA en SA regverdig die skepping van populasie-spesifieke standaarde en ‘n pasgemaakte databasis vir Suid-Afrika.

Acknowledgements

I may have written this thesis but it would not be what it is today withouth many individuals working in the background. I am grateful to my research supervisors, Professor Ericka L'Abbé and Dr Kyra Stull, for providing me with incredible support, encouragement and guidance, both in the writing of this thesis and the many hours of statistics that preceded it. I would also like to thank Mandi Alblas and Linda Greyling for their assistance with the Kirsten Collection in Stellenbosch. Furthermore, I would like to acknowledge the NRF for the generous funding awarded to me, without which this research would not have been possible. This thesis was enriched significantly ($p<0.05$) through helpful and entertaining discussions with my friend and fellow bone enthusiast, Leandi Liebenberg. Finally, thanks must go to my family. My parents for the love and inspiration they have given me throughout my life, my brothers for always pushing me to do better and my cat for helping me relieve stress through laughter.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Sexual dimorphism or the physical dichotomy between males and females occurs because of sexual selection and sex-specific differences in energetic intake, nutrition, body composition, hormonal influences and genetics (Gray and Wolfe, 1980; Frayer and Wolpoff, 1985; Kimmerle et al., 2008; Cabo et al., 2012; Wells, 2012). Sexual dimorphism can be identified and quantified in the skeleton as a means to estimate the sex of an unknown person, and forms the cornerstone of the biological profile which includes ancestry, age-at-death and stature estimates (Loth and İşcan 2000; Gapert, Black, and Last 2009). Sex estimation is a foundational element of the biological profile, as the accuracies of the other parameters are reliant on the sex, and also the ancestry, of the deceased individual (Barrier and L'Abbé, 2008; Blanchard, 2010). Therefore, the accuracies of the methods, the availability of population-specific standards, as well as the availability and preservation of the bones often dictate the morphological and osteometric methods utilized to assess sex from the human skeleton (Bidmos et al., 2010).

Sexual dimorphism is population-specific and can vary tremendously within and among populations. For example, skeletal samples from a genetically heterogeneous group, such as South African populations, show a reduction in sexual dimorphism when compared to populations that show a common sociodemographic background (Kemkes-Grottenthaler, 2005; Steyn and Patriquin, 2009; Ousley and L'Abbé, 2010; L'Abbé et al., 2013a). On account of variation among populations, previous researchers recognised that origins, admixture, secular changes and environment affect skeletal morphology such that the biological estimations using standards based on one population group are not accurate when used on a sample of different geographical origin (King et al., 1998; Asala, 2001; İşcan, 2005; Ramsthaler et al., 2007; Albanese et al., 2008; Rissech et al., 2013; Macaluso and

Lucena, 2014). As a result, the need exists for the accumulation of both morphologic and osteometric population-specific standards to improve estimates for various population groups.

North American or other international standards were regularly used in the past to assess biological parameters, as standards for South Africa did not exist until quite recently. Sex estimation standards were created from a number of different skeletal elements, namely the cranium, humerus, femur, forearm, hand and foot bones of white and black South Africans (Steyn and İşcan, 1997, 1998, 1999; Asala, 2001; Bidmos and Asala, 2003; Barrier and L'Abbé, 2008; Dayal et al., 2008). However, the research studies are statistically dated as they only apply univariate methods, or a bone-by-bone approach, to estimate sex. With a multivariate approach, including the use of methods such as discriminant function analysis, logistic regression and flexible discriminant analysis, more information can be captured. An increase in the number of variables and skeletal elements can provide much higher classification accuracies than univariate methods (Jantz and Ousley, 2005; Ousley and Jantz, 2013).

Numerous studies used small samples from either the Pretoria Bone Collection or the Raymond A. Dart Collection, both located in the Gauteng Province of South Africa. However, not all socially-defined South African population groups are represented in the Gauteng skeletal collections. As expected, the classification accuracy of the discriminant function formulae derived from these studies were reduced when the formulae were applied to populations not represented in the reference collections (e.g. coloured South Africans) (Robinson and Bidmos, 2009). Coloured South Africans form one of the largest minority groups in South Africa (Statistics South Africa, 2011). However, very few studies have included coloured South Africans in the creation of population-specific standards for estimating biological parameters in forensic casework in South Africa. While advanced statistics are more sophisticated and may provide detailed information, limitations exist (van

Vark and Howells, 1984). The limitations include, if the reference sample is representative of the population to which the statistics will be applied; how well the data fits the requirements of each test; and how accurately the data is collected (i.e. measurement or scoring errors) (Tabachnick and Fidell, 2007).

While a visual analysis of the pelvis – specifically, the pubic bone – is generally used to estimate sex with a high degree of reliability and accuracy, the availability and preservation of this bone tends to be poorer than either the cranium or the long bones (Phenice, 1969; Walker, 1995, 2008; Casado, 2010). Historically, the cranium was believed to be the second-best indicator to the pelvis for estimating sex. However, a recent study demonstrated that when using multiple variables, sex estimation from postcranial elements can produce classification accuracies up to 94%, which is higher than the accuracies produced when the skull is analysed (Spradley and Jantz, 2011).

In an attempt to remove subjectivity and employ multivariate statistical analyses for sex, ancestry and stature estimations many forensic anthropologists in North America regularly use the statistical software program, FORDISC 3.1, in their casework. FORDISC 3.1 or FD3.1 employs discriminant function analysis to estimate sex and ancestry based on any combination of cranial or postcranial measurements available and as a result assist anthropologists in their analyses of unidentified skeletonised cranial and postcranial, even fragmentary, remains (Ousley and Jantz, 2013). FD3.1 also has the ability to perform linear regression to estimate stature from the available measurements (Ousley and Jantz, 2012).

While the program incorporates both modern and historical populations from around the world into the reference sample, the program has less use outside of North America. One issue that limits the application of FD3.1 is the statistical assumption that the unknown belongs to one of the reference populations; currently, not all populations are represented in

the Forensic Databank (Ramsthaler et al., 2007). Cranial data of white, black and coloured South Africans was used to create a custom database for use in FD3.1 so that the same statistical rigor can be applied to South African groups (L'Abbé et al., 2013 a). However, the postcranial skeleton of South Africans has yet to be analysed or added to the custom database to assist in the estimation of sex and ancestry in South Africa. If the cranium is absent in forensic cases in South Africa, North American standards are occasionally applied to assist in the identification process, particularly in estimating sex and stature.

Both North America and South Africa have similar population histories in that both countries have individuals who migrated, either forcibly or willingly, from other countries in Europe and Africa. While North Americans and South Africans share a similar geographical origin in West/Central Africa and Europe, differences exist in the temporal aspects of the migration events. The Bantu migration occurred approximately between 5000 and 3000 years ago and eventually resulted in modern black South Africans, whereas the African diaspora, or forced emigration of West Africans to the Americas occurred much later (mid-15th to mid-19th century) and resulted in modern black North Americans (Spradley, 2006; Tishkoff et al., 2009). Further dissimilarities between the populations occurred due to different admixture frequency rates with other groups (Parra et al., 1998; Salas et al., 2005; Tishkoff et al., 2009). Thus, black and white North Americans and South Africans are distinct populations. Previous craniometric research has shown that South African black males and females are overall smaller and larger, respectively, than their North American counterparts (Ousley and L'Abbé, 2010; L'Abbé et al., 2013 a). Furthermore, the degree of sexual dimorphism was discovered to be lower in South Africans than in North Americans (L'Abbé et al., 2013 a). However, the differences in size, as well as sexual dimorphism, between the groups have yet to be quantified in the postcranial skeleton.

This research explores the pattern expression of postcranial sexual dimorphism in modern South Africans using robust statistical techniques; develops population-specific standards for the estimation of sex from the postcranial elements of three peer-reported South African groups (white, black and coloured); and compares sexual dimorphism between North Americans and South Africans.

Chapter 2: Literature Review

2.1. Bone growth and sexual dimorphism

Bone growth is an intricate process that commences *in utero* and continues through to early adulthood. Genetic, environmental, dietary, socioeconomic, developmental, behavioural, nutritional, metabolic, biochemical, and hormonal factors interact to regulate skeletal growth (Veldhuis, 2005). The skeleton develops through processes of endochondral and intramembranous ossification. Endochondral ossification (e.g. long bones) is the continuous replacement of a cartilaginous model with bone tissue, whereas the direct replacement of mesenchyme cells with osseous tissue refers to intramembranous ossification (e.g. skull vault) (White and Folkens, 2005; Young et al., 2006).

Sex hormones, oestrogen and testosterone, have multiple effects on bone growth. At low concentrations oestrogen promotes the secretion of growth hormone (GH), which has a stimulating effect on chondrocyte growth in the proliferation zone. However, at high concentrations oestrogen induces apoptosis of chondrocytes as well as stimulating osteoblast invasion in the growth plate thereby halting or ending linear growth (Juul, 2001). Other functions of oestrogen includes promotion of epiphyseal maturation, stimulation of endosteal and trabecular bone formation, and increase in tensile bone strength (Veldhuis, 2005). Oestrogens generally limit periosteal bone expansion but stimulate endocortical apposition in females (Callewaert et al., 2010).

Testosterone stimulates osteoblastic activity and inhibits apoptosis of osteoblasts and osteocytes. Furthermore, testosterone promotes cortical (periosteal) bone apposition which explains the larger periosteal diameter seen in boys from mid-puberty onward (Veldhuis, 2005; Callewaert et al., 2010). In males, both androgens and oestrogens work together to stimulate periosteal bone expansion and cortical bone growth during puberty (Callewaert et

al., 2010). Altogether, males build up wider and stronger bones than females and present with cortical bone further away from the neutral axis of the long bone as well as with increased bone mass in certain skeletal regions (femoral neck and lumbar vertebrae) (Bonjour et al., 1991). With the exception of the neonatal surge of testosterone seen in infant boys, further effects of testosterone occur only in later stages of puberty, compared with the earlier occurring effects of oestrogen (Knickmeyer and Baron-Cohen, 2006; Callewaert et al., 2010).

The complete fusion of the epiphyseal plates causes the bones to cease growth, with interstitial growth discontinuing at an earlier age than appositional growth. Generally, early-growing regions of the skeleton have been found to be less dimorphic as compared to later-growing parts (Humphrey, 1998). Therefore, in sub-adult and adult sex estimation from long bones, breadth and circumference are often more discriminating than longitudinal distances such as lengths (Steyn and İşcan, 1997; Stull, 2014). The differential growth rates and tissue sensitivities between males and females create sexual dimorphism in the skeleton that can be observed and calculated (Scheuer and Black, 2004). In general male morphology tends to be more robust and female morphology more gracile.

2.2. Factors affecting the expression of sexual dimorphism

For accurate observation and calculation of sexual dimorphism, additional factors need to be examined that may affect the expression of sex differences within a population, such as age, genetic expression, environmental, nutritional and socio-economics.

2.2.1. The effect of age on sexual dimorphism

The human skeleton changes dynamically with the advancement of age after puberty. The addition of bone in early life influences the manner in which the skeleton changes in the third decade of life and beyond (Vance, 2007). Modifications that occur in bone, besides density loss, include changes to the diameters of the bones such that increased dimensions of

the humerus and tibia of older individuals can be observed (Pfeiffer, 1980). Dimensions may increase due to periosteal bone gain, or more specifically a continual positive periosteal bone balance, after the age of 50 years. Mean periosteal femoral diameter is seen to increase as cortical thickness decreases (Walker, 1995). Similar results occur in humeral dimensions, where, quantitative overlap is obvious between the sexes and causes a decrease in skeletal sexual dimorphism in the elderly population. The enlargement of the cortical area results in net bone gain due to the continuing periosteal apposition (CPA) that exceed endosteal resorption (Vance, 2007). Osteoprogenitor cells within the periosteum of the bone are stimulated to produce more bone matrix in order for bone to maintain strength and integrity and to overcome the problem of decreased bone density (Vance et al., 2010). Generally females either maintain their size or became larger in skeletal dimensions, which may leads to misclassification of females over 50 years of age (Walker, 1995; Vance et al., 2010).

The greatest bone loss and increases in bone dimensions occur between middle and old age and is higher in females, particularly of European descent. Both North America and South African white females have shown decreased bone density, particularly in the femoral neck region, when compared to their black counterparts (Nelson et al., 2004). In white females, the cortical bone tends to decrease, whereas the size of the bone structure increases (Pfeiffer, 1980; Walker, 1995). These changes result in larger measurements in the cranial and post-cranial skeleton in older individuals. This in turn may affect the accuracy of sex estimation from unknown remains (Walker, 1995; Galdames et al., 2008).

2.2.2. Genetic expressions of sexual dimorphism

For the most part males tend to be taller, larger and more robust than their female counterparts (Rogers and Mukherjee, 1992; Loth and İşcan, 2000; Gustafsson and Lindenfors, 2004). However, the degree of sexual dimorphism within a group can be affected by genetics and lifestyle as well as by social status, environmental and nutritional factors

(Cameron, 2002; Barrier, 2007). Genes on the X-chromosome show immense variation in the expression and action between males and females. The hormonal environments of males and females differ considerably and the expression of genes controlling a quantitative trait may be influenced by the sex-environment encountered (Cameron, 2002; Weiss et al., 2006). Considerable variation exists across populations in growth in stature, weight, and other body dimensions, as well as in the rate and timing of maturation. Adult stature is a polygenic trait (controlled by multiple genes) and is highly heritable. The multiple genes that regulate stature occur at different loci and in different chromosomes, which leads to the creation of a wide range of stature variation within a population (Cameron, 2002; Weedon et al., 2008). The heritability of stature in a population tends to be lower in developing countries than first-world countries and may be attributed to the larger effect of environmental factors (increased nutritional and disease stresses) on stature in the less affluent countries (Cameron, 2002). However, in developed countries heritability estimates are usually higher, particularly if a significant phenotypic correlation exists between the parents. Phenotypic correlations, particularly with regard to stature, occur more frequently in European populations than in non-European populations (Cameron, 2002). Additionally, changes in the economy may result in differing growth environments between parents and their offspring, thereby decreasing the estimate of total variation attributable to genes (Cameron, 2002).

2.2.3. Environmental factors, nutrition, health and socio-economic status

Unfavourable living conditions can also decrease the degree of sexual dimorphism in a population (Harris, 1997). Conversely, enhanced environments and dietary conditions will cause an increase in the degree of sexual dimorphism, particularly with regard to stature (Tobias, 1971; Gray and Wolfe, 1980; Henneberg and Van den Berg, 1990; Barrier, 2007). Studies on health in third world countries have shown that sexual dimorphism decreases when the body undergoes repeated stress or poor nutrition (Tobias, 1971; Harris, 1997;

Bogin, 1999). Malnourished populations exhibit a lower degree of sexual dimorphism with adolescent males in a nutritionally deprived environment not reaching their full height and muscle development thus making them appear more gracile. (Tobias, 1971, 1975, 1985; de Beer, 2004; Maat, 2005; Barrier, 2007). Despite the strong genetic control of growth status and the rate of growth, evidence exists that the body can adapt to adverse environmental conditions by slowing down the developmental growth rate, most likely as a coping mechanism to manage with suboptimal physiological and metabolic requirements for development. If the adverse conditions are reversed, a period of rapid growth (catch-up growth) may restore the growth deficit. However, catch-up growth is not always complete and depends on the timing, severity, and duration of the insult (Cameron, 2002).

Further variation in growth sexual dimorphism may be attributed to unbalanced nutrition levels between the sexes in the different populations. Females tend to be less affected by nutritional fluctuations on account of hormonal mechanisms that support nutritional demands of reproduction. Males are not as resistant to environmental stress, such as nutritional deficiencies, and are affected to a greater extent than females (Greulich, 1951; Stini, 1969; Stinson, 1985). Therefore, in populations with decreased nutritional intake, sexual dimorphism occurs to a lesser degree because males have not had the environmental support to reach the maximum growth potential. Evidence for the nutritional hypothesis has traditionally been found in studies on sexual dimorphism of stature (Gray and Wolfe, 1980). The complex interaction between genetics and environmental effects influence bone shape and size; thus diet, activity, and disease are factors that can alter bone morphology (Klepinger, 2001; Ruff et al., 2006). Variation between populations occurs through various ontogenic processes as well as environmental and epigenetic influences, which include climate and activity patterns (Kemkes-Grottenthaler, 2005). Correlations are common between short populations and low economic status, while the tall populations can be

correlated to industrialized environments. Therefore, the evidence of secular trends in stature as well as the degree of similarity for stature in high socioeconomic status groups from different geographical areas shows that environmental factors play a significant role in inter-population variation in growth and development (Cameron, 2002; Barrier, 2007).

Activity patterns and mechanical loading also contribute to influencing bone mass and structural morphology (Pearson, 2000; Compston et al., 2007; Maalouf et al., 2007). With a lack of regular weight-bearing exercises to strengthen bone, bone mass loss will occur in both the axial and appendicular skeleton (Skedros et al., 2004). Therefore, if behavioural practices are different between males and females, the sexual dimorphism will be expressed in the skeleton (Anderson and Trinkaus, 1998; Ruff et al., 2006; Vance, 2007). Population and cultural differences in behavioural practices are common, where males and females have distinct roles in their social surroundings which expose them to differing physiological and biomechanical stressors. Therefore, bone density varies between population groups, with African populations most often exhibiting a higher bone density and a lower risk of osteoporosis than Europeans (Vance et al., 2010). A more physically demanding workload may explain the greater bone densities and decreased continuing periosteal apposition in African populations (Vance et al., 2010). Climate also has clear correlations with physique and skeletal proportions (Pearson, 2000). Limb bone robusticity varies with climate as much as with lifestyle. Habitual activity influences the amount of bone apposition and resorption that occurs at specific sites in the skeleton. However, climate and ecology can also influence culturally determined patterns of activity and thus potentially affect skeletal morphology, although this occurs to a lesser degree in adults as compared to adolescent individuals (Pearson, 2000; Pearson and Lieberman, 2004). Groups from cold climates, such as Europeans, tend to have the most robust diaphysis and proportionally the largest epiphyses.

Conversely, groups from hot areas, including Khoesan, fall to the opposite extreme, possessing relatively slender diaphysis and small joints (Pearson, 2000).

As age, environment, nutrition and socio-economic standings vary within and among populations, understanding these factors may help in understanding the pattern expression of sexual dimorphism present in different geographic populations.

2.3. *North American population*

The North American population is mainly made up of white, black and Hispanic individuals, with the minorities self-classifying as Asians or Native Americans. The white group forms the largest portion of the population with 75.49%, followed by the black group (14.91%) (Tise, 2010; U.S. Census Bureau, 2014).

White North Americans have received main genetic contributions from British groups; however, smaller contingents of other Europeans, such as the Spanish, French, Dutch, Swedes and Germans, have also contributed to the white population (Hrdlička, 1925; Davenport, 1926). Admixture between the various contributing groups created a blended group that became the North American white population (Hrdlička, 1925).

The North American black population, or African-Americans, mainly originated from West/Central African labourers brought to the Americas during the African diaspora (Parra et al., 1998; Salas et al., 2004; Spradley, 2006). When the enslaved Africans arrived in the New World, they were subjected to a new environment and plantation labour (Spradley, 2006). The shifts in living conditions and the exposure to different environments may have been responsible for the changes seen in the current African-American population (Spradley, 2006). A study comparing West Africans from the 18th and 19th centuries, display different craniofacial morphology to modern black Americans (Trevor, 1958; Spradley, 2006). The differences between ancestral and modern groups indicate that the enslaved Africans adapted

to the environment both culturally and biologically (Spradley, 2006). Furthermore, admixture occurred with white North Americans to produce the current black North American population (Salas et al., 2004; Spradley, 2006). Genetic exploration of the ancestry of black North Americans showed substantial contributions from Niger-Kordofanian (~71%), with lower additions from European (~13%), and other African (~8%) populations (Tishkoff et al., 2009). Very low levels of Native American ancestry were also discovered (Tishkoff et al., 2009). Gene flow from Europeans was shown to be sex biased with the male contribution being substantially greater than the female contribution (Parra et al., 1998).

Despite African-Americans sharing a considerable amount of genetic variation with white North Americans, the two ancestral populations remain relatively dissimilar in their cranial morphology (Parra et al., 1998; Kittoe, 2013). Robust statistical evidence has shown that morphological differences persist between African-Americans and European-Americans even though both ancestral groups have undergone secular change (Ousley et al., 2009).

On account of the various migrations, forced and willing, in North America, the modern North American population is heterogeneous. However, even with the variation present in the population, distinct skeletal differences exist between males and females and between North American black and white individuals (Spradley and Jantz, 2011; Kittoe, 2013).

2.4. South African population

Similar to the North American population, South African received genetic contributions from a wide range of parent groups and is home to multiple ancestries (Tishkoff et al., 2009; L'Abbé et al., 2011). As of 2011 the South African population was estimated to consist of 52.98 million people of which the largest proportion is black (79.8%), followed by coloureds (9.0%), whites (8.7%) and Indians/Asians (2.5%) (Statistics South Africa, 2013).

White South Africans are descended largely from colonial immigrants including Dutch, French, British and German groups (Steyn and İşcan, 1998; L'Abbé et al., 2011), with recent research has suggested almost equal contributions from all European groups (Greeff, 2007). Genetic evidence also suggests admixture between Europeans males and their slaves from southern Africa, most likely the Khoesan. A female bias of non-European influence was shown and stems from the fact that emigrants were predominantly male, resulting in a male biased sex ratio of adults (Steyn and İşcan, 1998; Greeff, 2007). Low frequencies of alleles typical to the Khoesan and Bantu-speaking peoples are found white South Africans.

The current black South African population arose mainly from Bantu-speaking groups that migrated from near the Nigerian/Cameroon highlands into South Africa within the past 5000 to 3000 years (Beck, 2000; Tishkoff and Williams, 2002; Tishkoff et al., 2009; May et al., 2013). However, gene flow is evident between the historical Bantu-speaking and indigenous Khoesan groups early in history but genetic and morphological differences exist between the groups (Liebenberg et al., in press; Herbert, 1990; Stynder, 2009). Bantu-speaking males were described as having "imposing height" and the females were "no less well-built" (Thompson, 2001:20). However, Bantu-speakers most probably differentiated in southern Africa from their counterparts in Central Africa under the pressure from both a different environment and genetic influence of the Khoesan-speaking peoples (Hiernaux, 1966).

The South African coloured population is a unique, highly admixed population that shows the highest levels of intercontinental admixture of any global population. Parent contributions include nearly equal high levels of Khoesan, Bantu-speaking, Indian, European ancestries and lower contributions from East Asian ancestry (Tishkoff et al., 2009; de Wit et al., 2010; Patterson et al., 2010; Quintana-Murci et al., 2010). In addition, unbalanced gender-specific contributions of the various population genetic components have been shown,

the most pronounced being the massive maternal contribution of Khoisan peoples (more than 60%) and the almost negligible maternal contribution of Europeans with respect to their paternal counterparts (Quintana-Murci et al., 2010). The gender-biased admixture indicates that the modern South African Coloured population results mainly from the early encounters of European and African males with Khoesan females (Quintana-Murci et al., 2010). The Khoesan ancestors were small people, particularly in stature, as well as delicately built, which is reflected in the current coloured population (Thompson, 2001).

During Apartheid (particularly 1948-1994), black South Africans were severely affected by race discrimination and as a result had extremely low socio-economic standings (Møller, 1998; Thompson, 2001). The coloured population shares a common socio-economic status and a shared culture derived from their incorporation into the lower ranks of Cape colonial society. Both Indians and the coloured groups were also affected by race discrimination, but to a lesser degree than black South Africans (Møller, 1998). As a result the coloured group held an intermediate status in the South African racial hierarchy, distinct from the historically dominant white minority and the numerically greater African population (Adhikari, 2005).

Differences in population histories, origins, as well as varying levels of admixture among may affect the expression of sexual dimorphism within and among populations. The composition and history of the population being investigated are further important factors to understanding the varying levels of sexual dimorphism present in the populations.

2.5. Osteometric vs morphoscopic methods

Sexual dimorphism in the skeleton can be identified and quantified to estimate sex. Both morphoscopic and osteometric analyses can be used to estimate sex from the skeleton; however the two approaches are continuously compared in terms of their ability to accurately

and reliably assess biological indicators of sex in a population (Cheverud et al., 1979; Steyn and İşcan, 1998). Morphoscopic analyses include visually assessing and occasionally scoring specific bony features (Buikstra et al., 1994; Loth and İşcan, 2000). However, this method is more subjective than osteometric analysis and experienced anthropologists may achieve higher accuracies than inexperienced ones (Pickering and Bachman, 2009). Furthermore, poor descriptions of morphoscopic traits often leads to inaccuracies in the scoring and also in the repeatability of the analyses (Novotný et al., 1993; Pretorius et al., 2006).

When applying an osteometric approach, measurements are taken from bone dimensions, usually between two landmarks (Loth and İşcan, 2000; Dayal et al., 2008). While linear measurements are not always able to describe the expression of sexual dimorphism visible in shape features (glabella or subpubic concavity), measurements are able to quantify sexual dimorphism present in size differences (maximum length and epicondylar breadths of long bones) (Loth and İşcan, 2000; Bidmos et al., 2010). Furthermore, osteometric analysis for estimating sex from skeletal remains has been found to be considerably less subjective than the use of morphological traits (Steyn and İşcan, 1998; Ousley and Jantz, 2013; Hefner and Ousley, 2014).

The statistical assessment of scores and measurements can range from utilising single variables to combinations of two or more, including complex multivariate analyses (Milner and Boldsen, 2012). Numerous studies have been done on the postcranial skeleton from South African samples but only focused on individual measurements or one or two bones at a time. Discriminant function analyses of combinations of variables have shown to provide better sex estimates (79%-86%) than individual variables (64%-79%) when analysing the calcaneus (Bidmos and Asala, 2003; Bidmos et al., 2010). Some other studies have attained classification accuracies from between 86% - 91% (femoral and tibial measurements) to 96%

(humeral measurements) (Steyn and İşcan, 1997, 1999). However, the high rates of correct classification were only produced when more than one measurement was combined.

Many different skeletal elements have been shown to be useful in estimating sex in adults (Buikstra et al., 1994). The skull, the pelvis, the humerus and the femur are the most commonly used elements. The adaptation of the female pelvis to its specialized function of child-bearing creates a clear distinction between the sexes and is thus also the best skeletal element to use in sex estimation (Purkait and Chandra, 2004; Milner et al., 2008; Spradley and Jantz, 2011). Sex estimation from the pelvis using the Phenice method yields accuracies of 88% (Ubelaker and Volk, 2002). Modifying the technique by adding line diagrams, a five-point ordinal scale and assessing the scores using logistic regression, not only increases the accuracies to 95%, but also creates a reliable and valid method of sex estimation from the human innominate (Ubelaker and Volk, 2002; Klales et al., 2012). Osteometric assessment of the pelvis yields classification accuracies >90% when using the pelvis as a whole (Patriquin et al., 2005).

While the pubic region of the os coxae expresses higher degrees of sexual dimorphism than the skull, the skull is often better preserved than the pelvis (Cheverud et al., 1979; Gapert et al., 2009; Garvin, 2012). The skull produces correct classification accuracies between 84% and 88% when looking at five non-metric traits (Walker, 2008) and 86% to 91% correct classification when measured (Steyn and İşcan, 1998).

When assessing the sexing potential of different parts of the appendicular skeleton, osteometric methods and statistical analyses are usually preferred (Ousley and Jantz, 2013). The maximum length, head diameter and epicondylar breadth measurements of the humerus have yielded accuracies up to 96% when estimating sex (Steyn and İşcan, 1999). The femur and tibia have also been used extensively, both internationally and in South African studies.

Classification accuracies of South African black and white individuals ranged between 86% and 91% (Steyn and İşcan, 1997; Asala, 2001, 2002; Barrier, 2007). Further osteometric studies performed on South African black and white samples include sex assessment of the radius and ulna, which resulted in correct classifications ranging from 76% to 86% (Barrier and L'Abbé, 2008). Examination of the calcaneus yielded correct classification between 82% and 92% for estimating sex in a South African sample.

Very little research has been done on sexual dimorphism and sex estimation in coloured South Africans, particularly the postcranial skeleton. However, a craniometric analysis of the group has shown differences between males and females, which made sex estimation possible with accuracies of 85% (Keen, 1950).

The metric approach to classification has many advantages and has become more popular with the appearance of FORDISC. Generally, osteometric analyses are overall good predictors of sex and ancestry (Patriquin et al., 2005). Furthermore, in light of the *Daubert* decision (Daubert v. Merrell Dow Pharmaceuticals, Inc., 1993), scientific conclusions need to follow certain guidelines, namely using reliable techniques and methods for which estimated error rates have been calculated and which are generally accepted by the peer community (Klales et al., 2012). While the *Daubert* standards are not yet required by the legal system in South Africa, the *Daubert* decision simply reinforces the value of good science.

Understanding the factors affecting sexual dimorphism as well as the populations being investigated is necessary to interpreting the expression of sexual dimorphism within a population. While research has been conducted on modern South Africans as a means to create population-specific standards for sex estimation, the exploration of the pattern expression of sexual dimorphism in South Africans was limited to individual bones. Furthermore, the information on sexual dimorphism in coloured South Africans is lacking,

particularly in the postcranial skeleton, warranting further assessment. Advanced statistical analyses were used as a means to improve on sex estimation techniques as well as to increase the understanding of patterns of sexual dimorphism not only within the South African population but also between North Americans and South Africans.

Chapter 3: Materials and Methods

3.1. Materials

3.1.1. Sample: modern South Africans

The modern South African sample consisted of 360 peer-reported black, white and coloured individuals with an equal distribution of sex and ancestry. All individuals had recorded demographic information. An individual was excluded from the sample if younger than 18 years. Furthermore, a measurement was not taken if the skeletal element showed pathology, trauma or the bones were articulated. The age distribution was from 18 to 87 years with a mean age of 53 years (Table 3.1).

The samples are housed at the Pretoria Bone Collection (PBC) and the Kirsten Collection at the University of Pretoria and the University of Stellenbosch, respectively. Both institutions actively accession skeletons into their collection, which assures a modern sample. Sampling of the different ancestry groups present in South Africa ensures that the variation present in the entire population is captured.

Table 3.1 – Sample sizes and abbreviations for each ancestry and sex group for modern South Africans.

Collection		*n*	*Mean Age*	*Abbreviation*
Pretoria Bone Collection (PBC)	Black Females	60	47.0	BF
	Black Males	60	47.8	BM
	White Females	60	67.3	WF
	White Males	60	61.9	WM
Kirsten Collection	Coloured Females	60	44.1	CF
	Coloured Males	60	52.6	CM

The PBC, housed in the Department of Anatomy at the University of Pretoria, is one of the most contemporary skeletal samples in South Africa with birth dates ranging from 1863 to 1996. The PBC, which started in 1947, is made up of 1135 complete crania, 816

postcrania and 399 incomplete remains. The entire collection is cadaver-based and consists of unclaimed, albeit known, and donated individuals to the Department of Anatomy, University of Pretoria for medical training and/or research. Most individuals in the collection are suspected to have been of low socioeconomic status and as a result, generally poor health (Keough et al., 2009). The individuals are mainly male (78%), of which 77% are black South Africans. Many of the black males were migrant labourers that entered the cities for work, died and were unclaimed by family members (L'Abbé and Steyn, 2012).

The Kirsten Skeletal Collection at the Division of Anatomy and Histology at Stellenbosch University in South Africa houses the largest sample of coloured individuals in the country. The collection is mainly cadaveric with known demographic information. The Kirsten Collection is comprised of 892 individuals; while the Kirsten Collection contains all populations present in South Africa, the most prevalent group (60.9%) is coloured South Africans. Similar to the composition of the Kirsten collection, the Western Cape (WC) is mainly populated by coloured individuals (48.8%), followed by black (30.1%) and finally white (15.7%). The number of females in the WC slightly exceeds the number of males with 96 males for every 100 females (Statistics South Africa, 2011). However, the number of males in this Kirsten collection marginally exceeds the females (54%) (Alblas, 2014).

3.1.2. Sample: modern North Americans

The postcranial measurements of 240 black and white North Americans, with equal sex and ancestry distribution, were sampled from the Forensic Databank (FDB). The ages of the individuals ranged from 18 to 89 years with a mean age of 41 years, which is slightly younger than the South African sample (Table 3.2).

Table 3.2 – Sample sizes and abbreviations for each ancestry and sex group for modern North Americans.

Collection		*n*	*Mean Age*	*Abbreviation*
Forensic Databank (FDB)	Black Females	60	35.1	NABF
	Black Males	60	36.3	NABM
	White Females	60	47.7	NAWF
	White Males	60	44.1	NAWM

The FDB contains measurements and demographic information on positively identified persons, most of whom were born past 1930 (Jantz and Moore-Jansen, 1988). The FDB data consists primarily of black and white North Americans, though other populations are also represented (Tise, 2010). The major data sources of the FDB primarily come from anatomical collections consisting of late 19th and early 20th century skeletal remains. A large number of individuals originate from the Robert Terry Collection at the Smithsonian Institution in Washington D.C and the Hamann-Todd Collection at the Cleveland Museum of Natural History in Cleveland, Ohio.

3.2. Methods

Thirty-nine measurements were collected on 11 postcranial bones (clavicle, scapula, humerus, radius, ulna, sacrum, innominate, femur, tibia, fibula, calcaneus) with an osteometric board, GPM sliding calipers or GPM spreading calipers. Measurement definitions were taken from *Data Collection Procedures for Forensic Skeletal Material* (Moore-Jansen and Jantz, 1994) (Appendix I) and abbreviations were taken from FORDISC (Jantz and Ousley, 2005) (Table 3.3). All statistical analyses were executed in R (R Core Team, 2013), which is an open-source software program. The packages employed for each statistical analysis will be cited accordingly.

Table 3.3 – Measurement abbreviations (measurement definitions are available in Appendix I).

	Abbreviation	*n*	*Measurement*
clavicle	claxln	259	Maximum length of the clavicle
	claapd	268	Antero-posterior diameter of the clavicle at midshaft
	clavrd	271	Vertical diameter of the clavicle at midshaft
scapula	scapht	307	Height of the scapula
	scapbr	341	Breadth of the scapula
humerus	humxln	353	Maximum length of the humerus
	humebr	257	Epicondylar breadth of the humerus
	humhdd	353	Maximum vertical diameter of the head of the humerus
	hummxd	355	Maximum diameter of the humerus at midshaft
	hummwd	356	Minimum diameter of the humerus at midshaft
radius	radxln	350	Maximum length of the radius
	radapd	353	Antero-posterior diameter of the radius at midshaft
	radtvd	354	Transverse diameter of the radius at midshaft
ulna	ulnxln	347	Maximum length of the ulna
	ulndvd	354	Dorso-volar diameter of the ulna
	ulntvd	357	Transverse diameter of the ulna
	ulnphl	352	Physiological length of the ulna
sacrum	sacaht	195	Anterior height of the sacrum
	sacabr	285	Anterior breadth of the sacrum
	sacs1b	257	Transverse diameter of sacral segment 1
innominate	innoht	339	Height of the innominate
	iliabr	336	Iliac breadth
femur	femxln	352	Maximum length of the femur
	fembln	352	Bicondylar length of the femur
	femebr	350	Epicondylar breadth of the femur
	femhdd	354	Maximum diameter of the head of the femur
	femsap	331	Antero-posterior subtrochanteric diameter of the femur
	femstv	344	Transverse subtrochanteric diameter of the femur
	femmap	328	Antero-posterior diameter of the femur at midshaft
	femmtv	347	Transverse diameter of the femur at midshaft
tibia	tibxln	345	Length of the tibia
	tibpeb	337	Maximum epiphyseal breadth of the proximal tibia
	tibdeb	345	Maximum epiphyseal breadth of the distal tibia
	tibnfx	326	Maximum diameter of the tibia at the nutrient foramen
	tibnft	347	Transverse diameter of the tibia at the nutrient foramen
fibula	fibxln	324	Maximum length of the fibula
	fibmdm	327	Maximum diameter of the fibula at midshaft
calcaneus	calcxl	204	Maximum length of the calcaneus
	calcbr	321	Middle breadth of the calcaneus

3.2.1. Inter- and Intra-observer error

Intra-observer error describes differences between repeated measures by one individual, while an inter-observer error describes differences between single measurements taken by two observers (DiGangi and Moore, 2013; Stull, 2013). As the validity of a measurement depends on its repeatability, the intra- and inter-observer error was estimated in order to evaluate consistency within and between observers for all measurements. The analyses involved re-measuring the postcranial remains of five randomly selected individuals (a total of 172 measurements) by the principle investigator (intra-observer error) and two additional observers (inter-observer errors). The error rates were evaluated using absolute technical error of measurement (TEM) and relative technical error of measurement (% TEM) and visualized with Bland-Altman plots.

TEM represents the typical magnitude of the expected measurement error and examines the standard deviations of the repeated measurement sets (Knapp, 1992; Stull, 2013). For the intra-observer agreement the two sets of measurements taken by the principle investigator were assessed and the TEM for each set of measurements was calculated. To calculate the relative TEM and thereby remove the size component, the absolute TEM was converted to a percentage and divided by the mean of each of the measurement pairs taken (Knapp, 1992; Perini et al., 2005; Jamaiyah et al., 2008; Stull, 2013).

The inter- and intra-observer agreement was visually assessed using Bland-Altman plots. The plots show trends in the agreement between the measurements of any two observers by plotting the means of the repeated measurements along the x-axis and the differences between the measurement pairs on the y-axis (Rothwell, 2000; Harris and Smith, 2009). A 95% confidence interval is illustrated by dashed lines and is based on the standard deviation. A higher agreement between observers is displayed as a random scatter of

differences close to a mean of zero, whereas disagreement is indicated by scatter that extends away from the mean (Bland and Altman, 1999; Stull, 2013)

3.2.2. Statistical analyses: modern South Africans

Exploratory statistics included Pearson correlation coefficient, Student's *t*-test, analysis of variance (ANOVA), multivariate analysis of variance (MANOVA) and sympercents. The correlation coefficient was used to analyse the measurements, whereas *t*-tests, ANOVA, MANOVA and sympercents were used to assess sexual dimorphism within and between ancestry groups as well as sexual dimorphism between males and females when the ancestries were pooled.

Correlations assessed the linear relationship between measurements and provided coefficients to indicate the strength and direction of the interactions (Taylor, 1990). A correlation coefficient provides a quantitative measure of the dependence between two variables and can be positive or negative (Rosner, 2011). The coefficient ranges between 0 and 1, where 0 refers to a non-existent relationship and 1 refers to highly correlated measurements. A positive correlation coefficient indicates that an increase in the first variable corresponds to an increase in the second variable, thus implying a direct relationship between the variables. A negative correlation indicates an inverse relationship whereas one variable increases and the second variable decreases (Taylor, 1990). Correlations are provided in the form of a matrix. Each row and each column within a correlation matrix represent a different variable. The value at the intersection between each row and each column is the relationship between the two variables. A matrix is symmetrical about the main diagonal, which means that only one half is necessary to assess the correlations. The entries on the main diagonal are correlations of each variable with itself and are equal to 1. Each variable in the study was

compared with every other variable and the correlation matrix was illustrated using the corrplot function in the corrplot package (R Core Team, 2013; Wei, 2013).

Multicollinearity may be present when two measurements have a correlation coefficient greater than 0.90 and may increase the likelihood of Type II errors by increasing the standard error and weakening the power of the relationship (Grewal et al., 2004; Tabachnick and Fidell, 2007; Afshartous and Preston, 2011). An increase in the standard error may increase the likelihood of an erroneous acceptance of a false null hypothesis. However, in certain measurement pairs, such as femoral epicondylar breadth and proximal epiphyseal breadth of the tibia, a strong correlation is expected because the elements are articulated. The negative effects of multicollinearity can be counteracted when the sample size is large enough and the independent variables describe a great proportion of the variance in the dependent variable (Grewal et al., 2004). Furthermore, Type I errors may occur with multiple analyses. A Type I error refers to the rejection of a true null hypothesis. A Holm's correction was added to the p-values obtained with the correlation coefficients in order to adjust for the effects of multiple tests and thereby decreasing the likelihood of Type I errors (Holm, 1979).

Student's *t*-test is used to compare the means of two samples; a value of 0.05 was used for alpha (McDonald, 2008). In the current study, male and female measurements were compared within each population (black, white and coloured) as well as for differences between the means of males and females when the populations were pooled. In *t*-tests, assumptions included normal distribution of the variables and equal variances; Levene's tests were utilised to assess the data for homogeneity of variance and were run using the car package (McDonald, 2008; Gastwirth et al., 2009; Fox, 2014).

Analysis of variance (ANOVA) was used to evaluate sex, ancestry, and the interaction of sex and ancestry between measurements of the sex-ancestry groups (i.e., BM, BF, WM, WF, CM, CF) (McDonald 2008). Multivariate analysis of variance (MANOVA), an extension of ANOVA, was used to investigate differences between sex-ancestry groups when assessing combinations of measurements or bone models (Tabachnick and Fidell, 2007; Stull, 2013). The independent variables were combined to create single independent measures (bone models), in order to maximize the between group differences (Mertler and Vannatta, 2002). MANOVA then compared the independent variables (*i.e.*, measurements), to two or more dependent variables. Similar to the ANOVA, the interaction between sex and ancestry was examined for the bone models. Both ANOVA and MANOVA were run using the stats package in R (R Core Team, 2013). The bone models included the clavicle, scapula, humerus, radius, ulna, sacrum, innominate, femur, tibia and fibula. The measurements taken on each separate bone were combined to create the bone model (Table 3.4).

Table 3.4 – Measurement included in the bone models. See definitions in Table 3.3.

	n	*Variables*
clavicle	253	claxln; claapd; clavrd
scapula	294	scapht; scapbr
humerus	343	humxln; humebr; humhdd; hummxd; hummwd
radius	344	radxln; radapd; radtvd
ulna	339	ulnxln; ulndvd; ulntvd; ulnphl
sacrum	160	sacaht; sacabr; sacs1b
innominate	331	innoht; iliabr
femur	288	femxln; fembln; femebr; femhdd; femsap; femstv; femmap; femmtv
tibia	293	tibxln; tibpeb; tibdeb; tibnfx; tibnft
fibula	312	fibxln; fibmdm

Sympercents (s%), or symmetric percentage differences, were used to assess the level of sexual dimorphism within the South African population. The ancestry groups were separated and the males and females of each group were compared. Because the size of the skeletal element (i.e. length versus breadth measurements) influences the mean difference between males and females, the unit of measure was removed and the log of a ratio presented

in terms of s% (s% = (100 $\log_e x_2$) – (100 $\log_e x_1$)) (Cole, 2000). The s% of male and female black, white and coloured South Africans was then compared to determine which group showed the greatest differences between the sexes.

Following the exploratory analyses, numerous univariate and multivariate classification methods were used for sex estimation and included sectioning points, linear discriminant analysis (LDA), flexible discriminant analysis (FDA) and logistic regression. The methods were compared to evaluate the strengths of the measurements and the ability of the methods themselves, to distinguish between males and females.

Univariate sectioning points with individual accuracy rates were created for each measurement that had showed significant differences between the sexes (based on *t*-test results). Pooled ancestry and population specific sectioning points were created. Taking the male mean and female mean and dividing by two calculated the sectioning points. Values above the sectioning point were considered male while values below were considered female. Values equal to the sectioning point were considered indeterminate (Spradley and Jantz, 2011). All individuals were classified based on the sectioning points. The classification accuracies for the sectioning points were calculated by dividing the number of correct group assignments by the number of total individuals.

Multivariate classification methods included linear discriminant analysis (LDA), flexible discriminant analysis (FDA), and logistic regression. The multivariate sex estimation techniques are robust and consider the interaction and relationship among multiple measurements instead of assessing only individual measurements (univariate approach).

Discriminant analysis maximizes differences between two or more populations and is frequently used in anthropological studies involving group classification (Pietrusewsky, 2008). Linear discriminant analysis (LDA) is the most commonly used type of discriminant

analysis and focuses on categorical response variables (i.e. coloured males, white females, etc.) and continuous predictor variables (i.e. measurements). In LDA a factor or weight is calculated for each measurement, which when added together for all variables maximises the mean differences among groups (Ousley and Jantz, 2012). A discriminant function is created using the resulting factors. The function is simply a linear combinations of the independent variables, which separate the groups defined by the categorical data (Dawson-Saunders and Trapp, 2004). The results provide a probability that the unknown belongs to one of the groups in the reference sample (Dawson-Saunders and Trapp, 2004). Assumptions associated with LDA include an adequate sample size and normal distribution of measurements. Discriminant analyses are also sensitive to outliers, which should be removed prior to analysis (Ousley and Jantz, 2012). The lda function is in the MASS package (Venables and Ripley, 2002). Equal prior probabilities were used and results were cross-validated using leave-one-out cross validation (LOOCV).

FDA is similar to LDA, except it converts a classification problem into a regression problem via optimal scoring. FDA obtains nonparametric versions of discriminant analysis by replacing linear regression with a nonparametric regression method (Hastie et al., 1994). For this study, FDA built regression models defined by multivariate adaptive regression splines (MARS) (Milborrow, 2013). MARS uses piece-wise linear segments to describe non-linear relationships. To identify the optimal model a forwards/backwards stepwise procedure is implemented along with automatic variable selection (Leathwick et al., 2005). FDA is considered advantageous to LDA and regression because FDA generates a flexible surface to separate the classes and accuracies have been higher when there are more than two response variables (Hastie et al., 1994). FDA was conducted using the fda function in the mda package (Hastie, 2013)

Logistic regression is used to describe relationships between one or more independent variables (i.e. measurements) and a binary dependent variable (i.e. the sex of an individual). A logistic model is designed in which the independent variables are used to describe the probability that an unknown will classify into a certain group (i.e. male or female) (McDonald, 2008; Kleinbaum and Klein, 2010). The parameters in a logistic regression analysis are chosen to maximize the conditional probability of the sample data. With this, the adjustment of the parameters will continue until the likelihood of the data does not change significantly (Tabachnick and Fidell, 2007; Stull, 2013). While logistic regression is similar to LDA in that the method calculates weights for the measurements according to their ability to correctly classify an unknown individual, logistic regression is generally more flexible and has fewer assumptions (Tabachnick and Fidell, 2007). One disadvantage of the method is its inability to classify unknown individuals into more than two groups. The glm package was used for logistic regression models (R Core Team, 2013).

The multivariate methods were employed to reveal the best combinations of measurements – either bone models or multivariate subsets – that can be used to classify according to sex. As the number of different combinations that can be used to classify individuals into male and female groups is vast, only a few models and subsets will be presented herein.

Each multivariate method utilised a variable selection process to explore which of the variables were useful in each particular subset or bone model. While LDA, FDA and logistic regression each choose different variables as valuable, the starting bone models and multivariate subsets were identical for each method (Table 3.4. and Table 3.5). The numbers of variables in each model also depended on the sample sizes of the measurements that were procured. While a larger number of variables in model creation can provide more information, too many variables may produce statistical ‘noise’ by including redundant

information (Ousley and Jantz, 2012). Further, by increasing the number of variables assessed, the sample sizes for each group become smaller.

Sufficient sample sizes are important to counteract the effects of overfitting the data and are necessary for other requirements of the analyses (Ousley and Jantz, 2012). Therefore, the minimum sample size of each group are suggested to be no smaller than three times the number of measurements used in the model for LDA (Huberty, 1975, 1994). The recommended sample size should produce stable estimates that are 'less subject to sampling variation' while permitting enough measurements for accurate classification (Ousley and Jantz, 2012). However, the necessary sample sizes for logistic regression are somewhat higher; suggested sample sizes are approximately 10 to 15 times the number of predictor variables (Hair et al., 2010). As the sample size for each model and subset is limited by the number of measurements that could be taken from each individual in the sample, the number of variables per subset are lower than variable numbers for both LDA and FDA.

Table 3.5 – Multivariate subsets and associated measurements. Measurement definitions are available in Table 3.3.

	n	*Variables*
shoulder	196	claxln; claapd; clavrd; scapht; scapbr; humxln; humebr; humhdd; hummxd; hummwd
hip	130	femxln; fembln; femebr; femhdd; femsap; femstv; femmap; femmtv; innoht; iliabr; sacaht; sacabr; sacs1b
upper	183	claxln; claapd; clavrd; scapht; scapbr; humxln; humebr; humhdd; hummxd; hummwd; radxln; radapd; radtvd; ulnxln; ulndvd; ulntvd; ulnphl
lower	148	femxln; fembln; femebr; femhdd; femsap; femstv; femmap; femmtv; innoht; iliabr; sacabr; sacs1b; tibxln; tibpeb; tibdeb; tibnfx; tibnft; fibxln; fibmdm
breadths	138	claapd; clavrd; scapbr; humebr; hummxd; hummwd; radapd; radtvd; ulndvd; ulntvd; femebr; femsap; femstv; femmap; femmtv; iliabr; sacabr; tibpeb; tibdeb; tibnfx; tibnft
lengths	106	claxln; scapht; humxln; radxln; ulnxln; ulnphl; sacaht; innoht; femxln; fembln; tibxln; fibxln
proxdist	315	humhdd; humebr; femhdd; femebr; tibpeb; tibdeb
all-variable model	124	claapd; clavrd; scapbr; hummxd; hummwd; ulnxln; ulntvd; ulnphl; sacabr; sacs1b; innoht; iliabr; femxln; femebr; femhdd; femmtv; tibxln; tibnfx; tibnft; calcbr

The bone models and multivariate subsets to be assessed by LDA were subjected to both forward and backward stepwise variable selection, which removes all unnecessary variables. Stepwise selection was conducted using the stepclass function in the klaR package (Röver et al., 2014). The initial model is defined by the provided starting variables (Table 3.4 and Table 3.5); the stepwise procedure excludes variables one-by-one until the improvement in accuracy is less than 5% or if the specified maximum number of variables is reached (Weihs et al., 2005). The removed variables are the least effective in classifying unknowns into their correct group.

Similar to LDA, FDA makes use of both forward pass and backward stepwise selection to choose the variables most useful in each of the models and subsets (Milborrow, 2013; Kuhn, 2014). Again the models and subsets analysed using FDA included all variables of the original subsets (Table 3.4 and Table 3.5). Once the stepwise procedure selected all useful variables for model creation, the varImp function of the caret package (Kuhn, 2014) displays which variables influence the model most and their associated values between zero and 100. The importance of a variable for FDA is based on the measure of the effect that a change to the variable has on the observed population. The effect is quantified and used to rate the use of a variable in predicting group membership (Milborrow, 2013). However, if a variable was not used in model creation the value shown using varImp would be equal to zero.

As the necessary sample sizes for the logistic regression analyses need to be bigger than for either LDA or FDA, the number of variables in the bone models and multivariate subsets are somewhat smaller for logistic regression than for the other two multivariate methods. Additionally, in order for analyses of the bone models and multivariate subsets to be valid only variables selected by the stepwise procedure were included in each model or subset (Table 3.9 and 3.10). For all bone and multivariate subset logistic regression models,

results were cross-validated using LOOCV of the cv.glm function in the boot package (Canty and Ripley, 2014).

While multicollinearity may be reduced using large sample sizes, to ensure that no subsets were affected by multicollinearity, principal component analysis (PCA) was conducted on each of the original bone models and subsets to reduce the number of variables and find the maximum variance from the principal components. The principal function in the psych package was used on each bone model and multivariate subset (Revelle, 2011). PCA has the ability to not only remove multicollinearity, but is also useful in reducing dimensionality (removing unnecessary variables) within the models and subsets (Jolliffe, 2002; Ringnér, 2008). PCA assumes correlation of variables and extracts uncorrelated components (principal components) from the original variables, while retaining as much of the variation as possible from the data set (Jolliffe, 2002; Ringnér, 2008). The PC scores, or linear combinations of the original variables, are ordered with the first few components extracting the most variance present in all the original variables of a subset and the last few PC scores retaining the least variance (Jolliffe, 2002; Tabachnick and Fidell, 2007). PC1 is also recognised to account for gross size in linear measures though shape information is also retained (Jolliffe, 2002; Berner, 2011). PC2 is the linear combination of the residual correlations from the first PC score and is orthogonal to PC1. Each additional component follows a similar pattern by extracting variability from residual correlations and is orthogonal to all previously extracted components. After extraction rotation of the factors was used to simplify interpretation and utility (Tabachnick and Fidell, 2007). Varimax, an orthogonal rotation method, was used to simplify factors by maximising the variances of the loadings of each factor (Jolliffe, 2002). The varimax function was a component of the stats package (R Core Team, 2013). Only PC scores with rotated loadings greater than 0.7 were chosen for the models as to reduce the number of variables in each bone model and multivariate subset. The

loading is the weight each standardized raw measurement is multiplied by to obtain the component scores and indicate the contribution. The PC scores replaced the raw measurements in each model and underwent stepwise selection to remove any additional unnecessary variables.

To compare which sex obtained the highest correct classifications, a sex bias was calculated for each classification method. The sex bias was determined by subtracting the male classification accuracy from the female classification accuracy. Therefore a positive sex bias indicates greater correct classifications for females than for males and vice versa.

After the analyses to estimate sex, both LDA and FDA were further utilised to classify the sample according to both sex and ancestry. However, when classifying according to sex and ancestry, the number of dependant variables increase from two to six. Therefore, the number of independent variables (measurements) that could be used to classify the individuals decreased considerably in order to comply with the suggested sample sizes. Amended multivariate subsets were created for classifying according to sex and ancestry using both LDA and FDA. PCA was conducted on the amended subsets to remove multicollinearity.

3.2.3. Statistical analyses to assess the modern North Americans

In order to investigate differences and similarities between North American and South African black and whites, analysis of variance (ANOVA), Tukey's honestly significant differences (HSD) test, and sympercents (s%) were calculated for the groups. Both ANOVA and Tukey's tests were run using the aov and TukeyHSD functions from the stats package (R Core Team, 2013). ANOVA was utilised as more than two dependent variables were assessed. Males and females were separated and the blacks and whites of each population were compared. Generally, significant results in ANOVA indicate that at least one group is

distinct from the other groups. However, the pattern of differences between the means is not always obvious (Abdi and Williams, 2010). In order to analyse the patterns, the means of two groups are compared at a time – pairwise comparisons – to indicate the relationship between the two groups. Tukey's HSD test is a pairwise comparison technique that makes use of the *q* distribution. The *q* distribution gives the sampling distribution of the largest variance between a set of means from the same population. All pairwise variances are evaluated using the same sampling distribution used for the largest differences (Abdi and Williams, 2010). Means resulting from ANOVA of each measurement for each population were also utilised to assess size differences between the groups. The means for females were separated according to population and the North American black means subtracted from the means of both the South African white and South African black females. Similarly, the means for North American white females were subtracted from the means of both the South African white and South African black females. Significance was noted for each mean comparison. The same procedure was completed for the male groups and again any significant differences between the means of two groups were noted.

Sympercents (s%) were again utilised to examine the sexual dimorphism present in North American black and white individuals. Furthermore, the s% of North American black and white individuals were compared to the s% of South Africans black and white individuals to investigate which group showed the greatest levels of sexual dimorphism and whether the pattern of expression of sexual dimorphism in the postcranial elements differed between North Americans and South Africans.

Further, multivariate sex and ancestry estimation using LDA and FDA were run. The original multivariate subsets (Table 3.4), as used to classify the South African sample, were again utilised as the starting variables. However, the number of dependant variables increased to eight variables, when assessing for sex and ancestry, requiring fewer independent variables to

preserve large enough sample sizes. Again new subsets were created using a limited number of the original variables. LDA and FDA assessed the new multivariate subsets to classify North Americans and South Africans according to sex and ancestry.

Chapter 4: Results

4.1. Inter- and intra-observer error

Absolute technical error of measurement (TEM) and relative technical error of measurement (%TEM) were conducted for each measurement and showed an overall lower intra-observer and inter-observer error rates. The mean intra-observer TEM and %TEM are fairly small at 0.13 mm (ranging from 0.0 mm to 0.82 mm) and 0.24% (ranging from 0.0% to 2.24%), respectively. The mean TEM and %TEM for the first inter-observer error were slightly higher at 0.54 mm and 1.57%, and the ranges notably wider at 0.0 mm to 1.22 mm and 0.0% to 6.99%, respectively. The second inter-observer error had intermediate results with a mean TEM of 0.43 mm (0.0 mm to 0.94 mm) and a mean % TEM of 1.04% (0.0% to 4.65%) (Appendix II). Measurements associated with the ulna had the greatest error. The largest intra-observer %TEM was observed for ulndvd, whereas both inter-observer %TEM were largest for the ulntvd, followed by the ulndvd of the same bone.

The Bland-Altman plot of the intra-observer error revealed overall high agreement with no measurement differences exceeding 2mm or systematic bias (Figure 4.1). Inter-observer plots showed similar results, despite a difference of 3 mm for one measurement in the first inter-observer assessment (Figure 4.2). While the means and range of differences does not vary considerably between the intra- and inter-observer errors, the spread of the differences is larger for both inter-observer errors than for the intra-observer error.

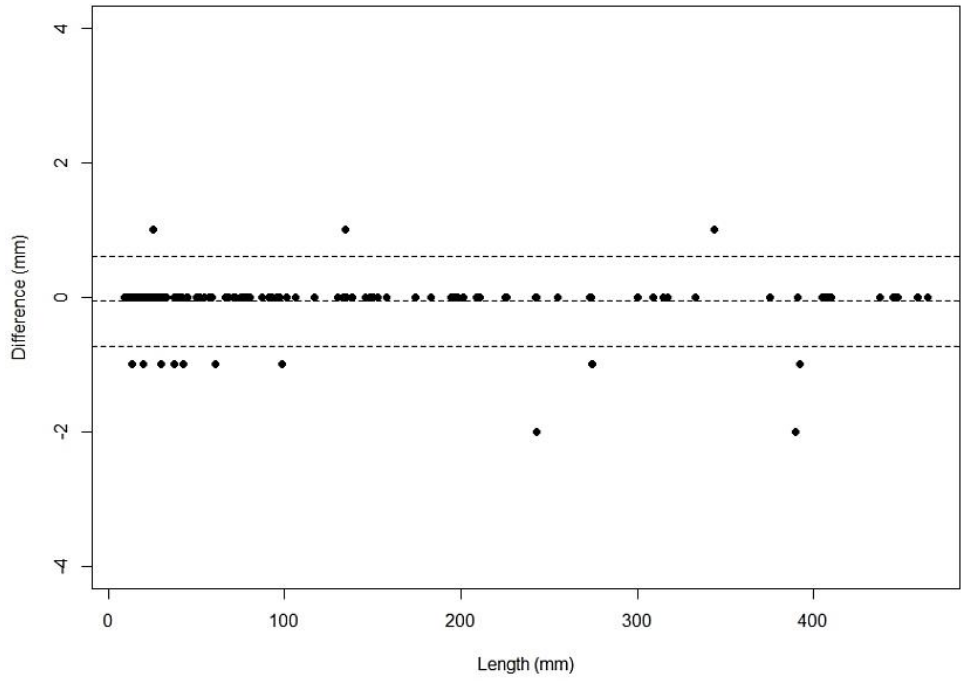

Figure 4.1 – Bland-Altman plot illustrating the intra-observer agreement when the principle investigator re-measured the five randomly selected individuals. The majority of measurements fall within the upper and lower agreement levels (dashed lines), which are based on the standard deviation.

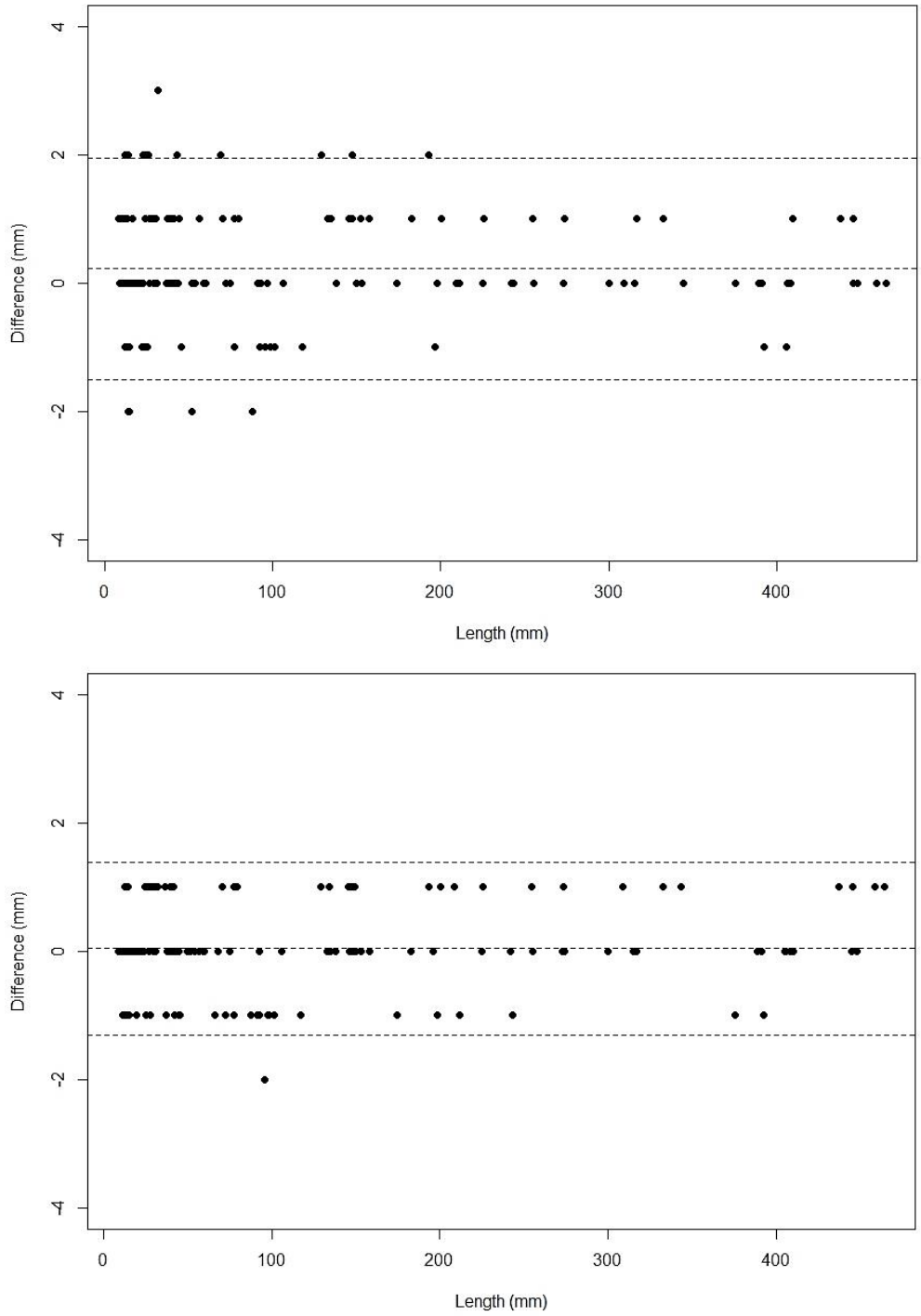

Figure 4.2 – Bland-Altman plots representing the second and third inter-observer errors when both the principle investigator and a second and third observer re-measured the five randomly selected individuals. The majority of measurements fall within the upper and lower agreement levels (dashed lines), which are based on the standard deviations.

4.2. Modern South Africans

4.2.1. Correlations

Inter-variable correlations, with a Holm's adjustment, were conducted on all 39 measurements (Appendix III). The majority of measurements presented with moderate (r = 0.3-0.69) to strong (r > 0.7) positive correlation coefficients. Only a few correlations (7%) were weak (r < 0.30) and an even smaller number (1%) were negative. Strong positive relationships occurred among all long bone lengths such as humxln and femxln (r = 0.91) and articulating elements, such as the femebr and tibpeb (r = 0.92). Furthermore, measurements from an individual bone, such as humxln and humhdd, were also strongly correlated (r = 0.78). Weak correlations were all associated with either the sacrum or pelvis, with the weakest correlation between hummxd and sacabr (r = 0.03). The strongest negative correlation was observed between sacabr and claapd (r = -0.13). In Figure 4.3, inter-variable relationships are visualised with a corrplot (R Core Team, 2013; Wei, 2013).

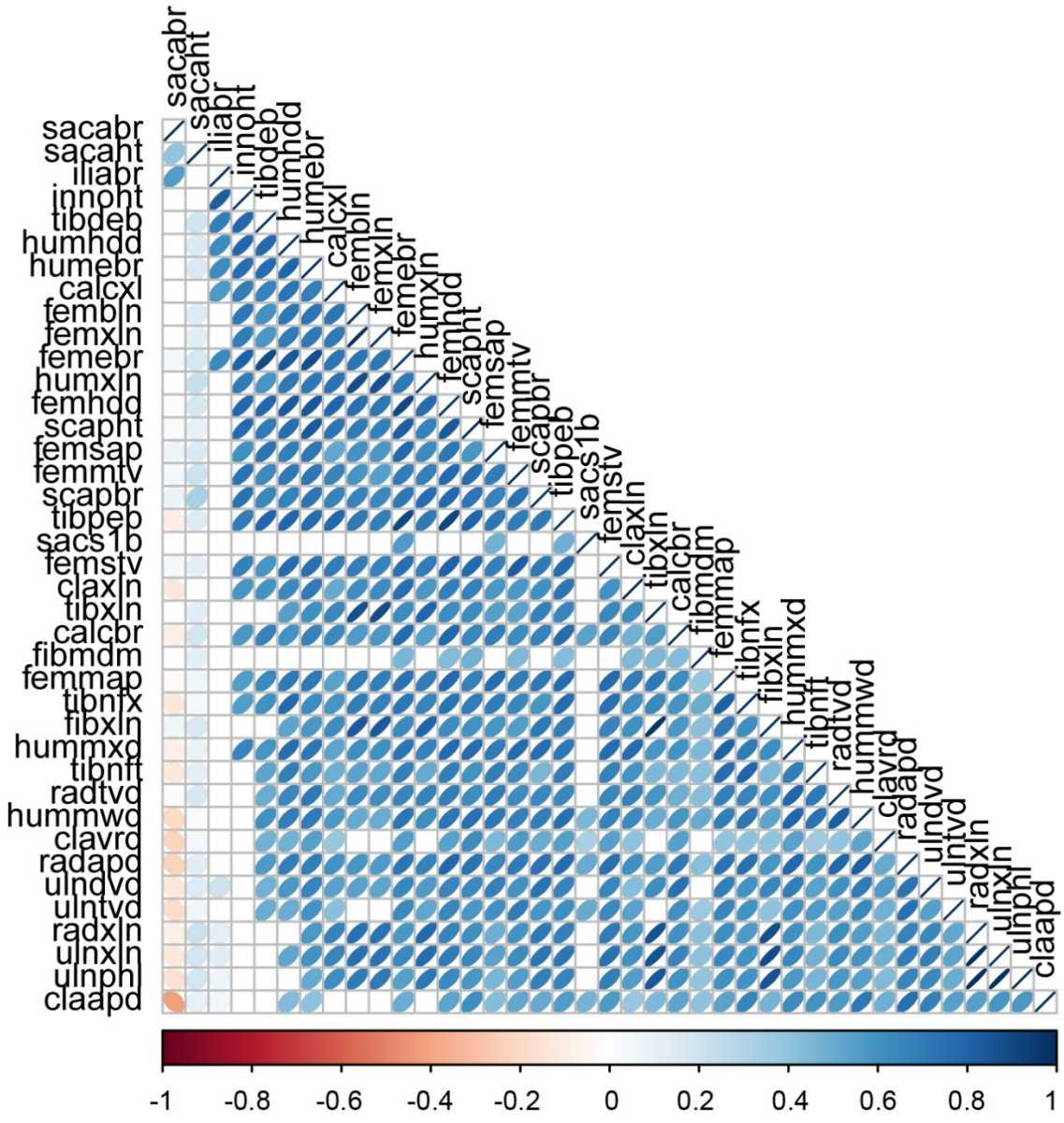

Figure 4.3 – Corrplot (R Core Team, 2013) *illustrating the inter-variable correlations. Ellipses were removed if the correlation, based on a Holm's adjustment, was not significant ($p > 0.05$).*

4.2.2. Student's t-test

A Student's *t*-test revealed significant differences ($p < 0.05$) between the sexes and when ancestry groups were pooled (i.e. all males compared to all females) for all 39 measurements. When the sexes of each ancestral group were compared separately, only black South Africans showed significant sex differences for all measurements. Significant sex differences were observed in white South Africans for all measurements, except sacaht and iliabr ($p=0.4245$ and $p=0.2324$, respectively). Coloured South African males and females were significantly different for all measurements with the exception of sacaht ($p=0.4373$).

4.2.3. ANOVA and MANOVA

Analysis of variance (ANOVA) was used to test for significant differences for the means of both sex and ancestry groups and to assess for any interaction between the two dependent variables. Significant differences between males and females were observed for all measurements. Only claapd, clavrd, ulndvd and calcbr demonstrated no significant differences when explored by ancestry. Significant interactions between sex and ancestry were only present for the iliabr and calcbr (Appendix IV).

Multivariate analysis of variance (MANOVA) assessed significance between the ancestry groups, the sex groups, and evaluated the relationship between sex and ancestry when the measurements were combined into bone models. Significant differences were present between the different ancestry groups and between the two sexes for all bone models (Table 4.1). Significant sex and ancestry interactions were present for the ulna, innominate and femur bone models.

Table 4.1 – Multivariate analysis of variance (MANOVA) results evaluating the statistical significance of sex, ancestry and the interaction between sex and ancestry for the bone models. Bold indicates significance.

	Sex		Ancestry		Sex*Ancestry	
	Pillai	*p-value*	*Pillai*	*p-value*	*Pillai*	*p-value*
clavicle	0.646	**<0.001**	0.148	**<0.001**	0.027	0.34
scapula	0.57	**<0.001**	0.258	**<0.001**	0.003	0.91
humerus	0.64	**<0.001**	0.543	**<0.001**	0.033	0.35
radius	0.652	**<0.001**	0.274	**<0.001**	0.021	0.32
ulna	0.616	**<0.001**	0.314	**<0.001**	0.052	**0.025**
sacrum	0.329	**<0.001**	0.526	**<0.001**	0.027	0.66
innominate	0.43	**<0.001**	0.576	**<0.001**	0.039	**0.012**
femur	0.607	**<0.001**	0.532	**<0.001**	0.094	**0.046**
tibia	0.587	**<0.001**	0.46	**<0.001**	0.055	0.1
fibula	0.346	**<0.001**	0.224	**<0.001**	0.002	0.95
calcaneus	0.478	**<0.001**	0.216	**<0.001**	0.021	0.12

4.2.4. Sexual dimorphism differences in South African groups

Sympercents (s%) were used to compare sexual dimorphism between males and females of each ancestry group. Differences between males and females were observed in the s% calculated for all measurements for all three groups (Table 4.2). While one ancestry group always had a higher s% value compared to the other groups for all measurements, the differences between the groups never exceeded 2.14%. Overall males presented with larger mean values and produced positive sympercents for all measurements except for the sacabr where the females of all three groups were larger than the males. Coloured males and females had the highest degree of sexual dimorphism for the most measurements (51%). In the coloured group, the largest differences occurred between ulndvd (8.90%), ulntvd (8.42%), radapd (8.38%), claapd (8.0%) and hummwd (7.63%). Black males and females presented with the highest levels of sexual dimorphism for 33% of measurements; the highest values were in claapd (9.54%), ulntvd (8.50%), radapd (8.05%), clavrd (7.90%) and ulndvd (6.50 %). White males and females showed the most sexual dimorphism for only 15% of measurements when compared to the other groups (Table 4.4). The largest differences between white males and females were in clavrd (9.53%), radapd (8.39%), hummwd (7.97%), claapd (7.82%) and ulntvd (7.48%). For all three groups the upper limb (clavicle included) was most sexually dimorphic, whereas the pelvis (iliabr and sacrum) and length measurements continuously revealed the least amount of separation between the sexes.

Table 4.2 – Comparison of sympercent differences (s%) between white, black and coloured populations. Negative sympercents indicate females are larger than males and bold indicates the group with the largest sympercent for each measurement.

	White	*Black*	*Coloured*
claxln	4.44	4.01	**4.67**
claapd	7.82	**9.54**	8.00
clavrd	**9.53**	7.90	6.84
scapht	5.13	**5.66**	4.84
scapbr	4.52	**5.51**	5.46
humxln	3.36	3.68	**3.73**
humebr	5.50	4.74	**6.14**
humhdd	4.78	5.79	**6.51**
hummxd	5.74	5.41	**5.90**
hummwd	**7.97**	6.42	7.63
radxln	4.32	4.69	**4.72**
radapd	**8.39**	8.05	8.38
radtvd	**7.13**	6.04	5.98
ulnxln	4.12	4.30	**4.48**
ulndvd	6.76	6.50	**8.90**
ulntvd	7.48	**8.50**	8.42
ulnphl	3.84	**4.63**	4.28
sacaht	0.97	**2.83**	1.08
sacabr	-1.44	-1.49	**-2.34**
sacs1b	2.57	3.39	**3.79**
innoht	3.01	3.57	**3.64**
iliabr	0.34	**2.25**	1.48
femxln	2.88	3.15	**3.36**
fembln	2.91	3.26	**3.39**
femebr	4.13	4.57	**5.03**
femhdd	4.74	5.87	**6.07**
femsap	**4.79**	3.79	4.03
femstv	3.48	4.94	**5.44**
femmap	4.63	**4.83**	4.23
femmtv	4.37	**4.89**	4.63
tibxln	2.99	3.36	**3.49**
tibpeb	4.66	5.01	**5.25**
tibdeb	3.65	4.05	**4.66**
tibnfx	4.17	**5.95**	4.39
tibnft	**5.81**	5.20	5.78
fibxln	2.91	3.56	**3.68**
fibmdm	3.49	**4.28**	3.92
calcxl	3.00	**4.64**	NA
calcbr	3.87	**4.56**	2.12

4.2.5. Univariate sectioning points

Univariate sectioning points were created for all measurements since all were significantly different between the sexes. Individuals were classified as male if the value was greater than the sectioning point and as female if the value was smaller than the sectioning point. Percent correct, when the three ancestry groups were pooled, ranged from 52% to 82%, where the tibpeb and the fibmdm classified best and worst, respectively (Table 4.3). All measurements were able to classify the individuals into one of the two categories and all measurements had accuracies better than chance (i.e. > 50%).

Population-specific univariate sectioning points were also created. The classification accuracies for black males and females ranged from 54% to 90% (Table 4.5). The femhdd measurement provided the highest correct classification and fibmdm the lowest. Sectioning points created for white South Africans included only 37 of the 39 measurements, as sacaht and iliabr were not significantly different between the sexes. Univariate classification using the remaining measurements yielded accuracies between 52% and 86%, with tibpeb classifying best and fibmdm classifying worst (Table 4.3). Population-specific univariate sectioning points for coloured South Africans produced accuracies ranging from 50% to 88%. The tibpeb and calcbr measurements achieved the best and worst classification accuracies, respectively (Table 4.3). Similar to white South Africans, no univariate sectioning point was created for sacaht to classify the coloured males and females.

Table 4.3 – Univariate sectioning points and their associated accuracies when ancestry groups pooled and separated. Bold numbers indicate highest accuracies for each population.

	Pooled Ancestry		**Black**		**White**		**Coloured**	
	Sectioning point	*Accuracy (%)*	*Sectioning point*	*Accuracy (%)*	*Sectioning point*	*Accuracy (%)*	*Sectioning point*	*Accuracy (%)*
claxln	147	77	148	72	152	80	144	81
claapd	10	63	10	61	10	70	12	74
clavrd	12	73	12	76	12	70	10	60
scapht	147	81	145	88	154	84	143	85
scapbr	99	79	98	80	103	78	96	86
humxln	312	73	310	74	325	82	302	80
humebr	59	80	59	80	61	83	57	87
humhdd	43	75	41	83	46	78	41	86
hummxd	21	69	21	71	22	76	21	68
hummwd	17	73	17	73	18	71	17	69
radxln	239	81	244	84	242	85	231	80
radapd	12	75	12	73	12	77	12	73
radtvd	15	64	14	67	15	68	14	61
ulnxln	257	80	263	81	260	82	247	82
ulndvd	15	67	14	65	15	63	14	78
ulntvd	15	73	15	73	15	71	14	77
ulnphl	225	80	231	79	226	82	217	76
sacaht	102	58	101	64	-	-	-	-
sacabr	94	58	90	63	100	57	93	62
sacs1b	45	65	46	61	46	63	43	68
innoht	203	68	197	76	217	79	193	78
iliabr	151	55	146	63	-	-	144	65
femxln	442	74	441	80	456	75	429	79
fembln	439	75	437	79	453	74	426	80
femebr	77	77	76	79	80	80	74	85
femhdd	44	77	43	**90**	45	75	43	83
femsap	27	67	27	67	28	61	26	66
femstv	31	71	30	70	32	66	31	73
femmap	29	67	28	68	29	65	28	66
femmtv	26	67	26	71	28	71	25	69
tibxln	370	73	375	72	378	73	358	76
tibpeb	72	**82**	72	86	75	**86**	70	**88**
tibdeb	47	69	46	68	50	68	45	82
tibnfx	34	70	34	77	35	72	33	70
tibnft	24	68	24	68	25	67	23	73
fibxln	363	75	368	73	370	74	351	79
fibmdm	15	52	15	54	16	52	14	59
calcbr	42	66	41	78	42	70	42	50

4.2.6. Multivariate Sex Estimation

Bone models and multivariate subsets using the original measurements underwent stepwise selection to determine the best variables to classify sex when using LDA (Table 4.4 and 4.5). The stepwise procedure mostly selected a single measurement of each bone model as effective in classifying according to sex; occasionally more than one measurement was used to estimate sex. Most measurements in the multivariate subsets were deemed unnecessary in the stepwise procedure and thus removed for model creation.

Table 4.4 - Bone models for LDA when classifying according to sex only.

	n	*Variables*
clavicle	253	claxln, claapd
scapula	294	scapht
humerus	343	humebr
radius	344	radapd
ulna	339	ulntvd
sacrum	160	sacabr, sacs1b
innominate	331	innoht, iliabr
femur	288	femhdd
tibia	293	tibpeb
fibula	312	fibxln

Table 4.5 - Multivariate subsets for LDA when classifying according to sex only.

	n	*Variables*
shoulder	196	claapd, clavrd, scapbr
hip	130	sacabr, femhdd, femstv
upper	183	humhdd, radapd
lower	148	sacabr, femhdd, tibpeb
breadths	138	scapbr, ulntvd, sacabr
lengths	106	scapht, ulnxln
proxdist	315	humebr, tibpeb
all-variable model	124	femhdd, sacabr

LDA achieved cross-validated accuracies between 75% and 90% for the bone models, where the radius did best and the sacrum and fibula did worst. Females classified better than males for most bone models, except for the ulna and sacrum, where males had the higher classification accuracies (Appendix V). Using LDA and the multivariate subsets, cross-validated correct classifications ranged from 87% to 98% for the proxdist and breadths subsets, respectively (Appendix V). Once again a female sex bias was observed in most subsets, with males only achieving the higher accuracies for the hip and lower bone model subsets.

Cross-validated classification accuracies using the bone models with FDA ranged from 75% to 91% (Appendix VI). The clavicle was best and the fibula worst at classifying the sample into male and female groups. Overall, females classified better than males, but males had higher correct classification for the clavicle, humerus, innominate and femur. The measurements that were used in each bone model and subset were noted according to their importance in the analysis (Table 4.6 and Table 4.7). The measurements that were not incorporated into the model were not included in the tables (Table 4.6 and Table 4.7). Overall classification for the multivariate subsets ranged from 90% to 97% using FDA, where the all-variable model achieved the highest accuracy (Appendix VI). Females obtained the higher correct classifications for all subsets, except the upper and lower models, where males obtained the higher accuracies.

The logistic regression models regularly chose between two and four measurements for each bone model and multivariate subset (Table 4.8 and Table 4.9). Cross-validated classification accuracies for logistic regression ranged between 74% and 91% (Appendix VII). The sacrum achieved the lowest correct classification (74%), followed by the fibula (75%) and innominate (79%). The highest accuracies were associated with the clavicle and radius (91% and 90%, respectively). A female sex bias was observed for all models except

the humerus, sacrum and innominate. When multivariate subsets were used in model creation, cross-validated accuracies ranged from 88% to 97% (Appendix VII). The proxdist subset obtained the lowest classification accuracies, whereas the all-variable model obtained the highest accuracy. Females obtained the higher correct classifications for all subsets, except the upper and lower models, where males obtained the higher accuracies.

All three classification techniques obtained similar cross-validated accuracies using bone models (Table 4.10). However, FDA achieved the overall highest correct classifications with an average of 85% compared to LDA (83%) and logistic regression (84%). The fibula and sacrum achieved the lowest correct classifications for the three methods, whereas the clavicle and radius achieved the highest accuracies. An assessment of the three multivariate classification techniques when using multivariate subsets showed correct cross-validated classifications ranges between 87% and 98% (LDA), 90% and 97% (FDA) and between 88% and 97% (logistic regression) (Table 4.11). Similar to the average correct classification rates obtained using bone models, FDA achieved the highest mean accuracy (95%) followed by LDA (94%). Logistic regression attained the lowest average classification accuracies using multivariate subsets (93%). The all-variables subset produced the highest correct classifications for both FDA and logistic regression, whereas the breadths subset produced the highest accuracies for LDA.

Table 4.6 - Bone models for FDA when classifying according to sex only.

	n	*varImp selected (most to least important)*
clavicle	253	claapd, clavrd, claxln
scapula	294	scapbr, scapht
humerus	343	humebr, hummwd, humhdd
radius	344	radapd, radxln
ulna	339	ulnxln, ulntvd, ulndvd
sacrum	160	sacs1b, sacabr
innominate	331	innoht, iliabr
femur	288	femhdd
tibia	293	tibpeb, tibnfx
fibula	312	fibxln, fibmdm

Table 4.7 – Multivariate subsets for FDA when classifying according to sex.

	n	*varImp selected (most to least)*
shoulder	196	claapd, claxln, clavrd, humebr, hummwd
hip	130	femhdd, sacabr, iliabr, femebr, femmap
upper	183	radapd, claxln, clavrd, claapd, humxln, humhdd
lower	148	femhdd, sacabr, tibnfx, iliabr, femsap
breadths	138	ulntvd, sacabr, radapd, humebr, iliabr, scapbr
lengths	106	scapht, radxln, tibxln, claxln, innoht, fibxln
proxdist	315	humebr, femhdd, tibpeb, tibdeb, humhdd
all-variable model	124	ulntvd, femhdd, sacabr, claapd

Table 4.8 - Bone models for Logistic regression when classifying according to sex only.

	n	*Variables*
clavicle	253	claxln, claapd, clavrd
scapula	294	scapht, scapbr
humerus	343	humebr, humhdd, hummwd
radius	344	radxln, radapd
ulna	339	ulnxln, ulndvd, ulntvd
sacrum	160	sacaht, sacabr, sacs1b
innominate	331	innoht, iliabr
femur	288	femebr, femhdd
tibia	293	tibpeb, tibnfx
fibula	312	fibxln, fibmdm

Table 4.9- Multivariate subsets for logistic regression when classifying according to sex only.

	n	*Variables*
shoulder	196	claxln, claapd, clavrd, humebr
hip	130	femebr, femstv, sacabr
upper	183	humhdd, radapd
lower	148	femhdd, femsap, sacabr
breadths	138	clavrd, ulntvd, iliabr, tibpeb
lengths	106	scapht, innoht, radxln, tibxln
proxdist	315	humebr, femhdd, tibpeb, tibdeb
all-variable model	124	sacabr, femhdd

Table 4.10 - Comparison of multivariate classification methods (LDA, FDA and logistic regression) using bone models to estimate sex. Bold numbers indicate the highest accuracy for each bone model.

	Accuracy (%)		
	LDA	*FDA*	*Logistic regression*
clavicle	88	**91**	**91**
scapula	82	86	83
humerus	84	88	86
radius	**90**	90	90
ulna	86	89	87
sacrum	75	76	74
innominate	78	79	79
femur	86	87	84
tibia	86	86	87
fibula	75	75	75

Table 4.11 - Comparison of robust classification methods (LDA, FDA and logistic regression) using multivariate subsets to estimate sex. Bold indicates the highest accuracy for each method.

	Accuracy (%)		
	LDA	*FDA*	*Logistic regression*
shoulder	91	93	92
hip	96	95	95
upper	95	96	94
lower	97	97	95
breadths	**98**	96	94
lengths	90	93	91
proxdist	87	90	88
all-variable model	97	**97**	**97**

To alleviate any multicollinearity that was present in the bone models and multivariate subsets, PCA was employed for each model and subset. Furthermore, all bone models underwent stepwise selection to eliminate any variables that were not useful in model creation. For all bone models, except the sacrum subset, PC1 explained the largest portion of variance in the sample. Equal portions of variance were explained by PC1 and PC2 in the sacrum bone model. The clavicle model was represented by principal components 1 to 3, which accounted for 100% of the variance. PC1 was most heavily loaded by clavrd, whereas PC2 and PC3 were more heavily loaded by claxln and claapd, respectively. The scapula, innominate and fibula models were reduced to single principal components, which explained 89.3%, 92% and 74.4% of variance, respectively. The PC1 for each of the three models had equal contributions from the two measurements included in the models with equal rotated loadings. For most bone models PC1 was most heavily loaded by maximum lengths. The clavicle was the only long bone where the largest contribution to PC1 was not from a length. For the clavicle, clavrd had the highest loading of the first component, which accounted for the most variance (33.8%) in the model. In the ulna and femur models, PC1 received highest contributions from both maximum and physiological (ulna) or bicondylar (femur) lengths, which explained 47.8% and 28.1% of the model's variance, respectively.

As only component scores with loadings >0.7 were selected for model creation, not all the variance was always explained by the remaining PC scores. All PC scores for the humerus, radius and sacrum models were used and explained 100% of variance in the sample. For the ulna, femur and innominate more than 90% of variation was expressed, with 99.7%, 96.5% and 92%, respectively. The fibula model (74.4%) expressed the least amount of variation, followed by the tibia (86.5%) and scapula (89.3%) models.

Using only the stepwise chosen PC scores in the LDA, FDA, and logistic regression models, classification accuracies were similar to the results when the original measurements

were assessed. Correct classification ranged between 63% and 91% (LDA), 66% and 92% (FDA) and between 63% and 90% for logistic regression (Table 4.12). Accuracies for most of the bone models (clavicle, humerus, radius, ulna, sacrum and femur) increased slightly for at least one method, whereas classification associated with the innominate, tibia and fibula models decreased considerably for all three techniques. For all three methods, the innominate model produced the lowest classification rates and the clavicle and radius produced the highest classification accuracies. Average correct classification was highest for FDA, followed by LDA and lowest for logistic regression. The first principal component was selected for all bone models for both FDA and logistic regression. However, when stepwise chose components for LDA, PC1 was selected for all subsets, except for the humerus subset.

Table 4.12 - Comparison of multivariate classification methods (LDA, FDA and logistic regression) using bone models that underwent PCA. Bold numbers indicate the highest accuracy for each bone model.

	Accuracy (%)		
	LDA	*FDA*	*Logistic Regression*
clavicle	**91**	**92**	**90**
scapula	83	83	83
humerus	86	89	87
radius	90	91	90
ulna	87	87	88
sacrum	74	79	74
innominate	63	66	63
femur	86	88	85
tibia	75	76	74
fibula	72	72	70

For the multivariate subsets, the shoulder, breadths and proxdist subsets did not have length measurements as the largest contribution to PC1. However, unlike the breadths and proxdist subsets, the shoulder model included two length measurements (humxln and claxln). PC1 in the shoulder subset was most heavily loaded by clavrd and accounted for 13.1% of variance. Humxln and claxln contributed more to PC2 and PC4, respectively. PC1 for the hip,

upper, lower and lengths models was most heavily loaded by a variety of maximum length measurements. However, the clavicle maximum length was the only maximum, physiological or bicondylar length that never contributed highly to either PC1 or PC2. When the all-variable model underwent PCA, PC1-PC10 were selected and accounted for 83.1% of variation. PC1 explained the largest proportion of variance (24.7%) and was most heavily loaded by ulnxln, ulnphl, tibxln and femxln. PC3, PC5 and PC8-PC10 had large contributions by midshaft measurements (claapd, clavrd, ulntvd, hummxd and tibnft). PC2 was loaded heavily by the two innominate measurements, whereas PC4, PC6 and PC7 were loaded by measurements associated with joints (calcbr, sacs1b and sacabr, respectively).

Accuracies for the three methods, when using PC scores from the multivariate subsets, ranged from 85% to 92% (LDA) to 82% to 94% (logistic regression) and 90% to 97% (FDA) (Table 4.13). The lengths and lower subsets achieved the highest correct classifications for the three methods, where the lengths subset classified best using LDA and logistic regression and the lower subset classified best using FDA. The lowest classification accuracies was obtained by the upper (LDA and FDA) and hip (logistic regression) subsets. Correct classification using PC scores was substantially lower for almost all subsets using LDA and logistic regression when compared to accuracies obtained from the raw measurements. The only subset for which classification accuracies increased using LDA and logistic regression was the lengths subset. However, when FDA assessed PC scores from multivariate subsets, the accuracies for all subsets, except the shoulder, upper and all-variable models, remained unchanged from the raw measurements.

Table 4.13 - Comparison of multivariate classification methods (LDA, FDA and logistic regression) using multivariate subsets that underwent PCA. Bold numbers indicate the highest accuracy for each bone model.

	Accuracy (%)		
	LDA	*FDA*	*Logistic regression*
shoulder	86	94	90
hip	88	95	82
upper	85	90	84
lower	89	**97**	87
breadths	91	96	84
lengths	**92**	93	**94**
proxdist	86	90	87
all-variable	87	95	85

4.2.7. Multivariate Sex and Ancestry Estimation

Only LDA and FDA were utilised to classify the sample into sex and ancestry groups as logistic regression was unable to classify according to more than two groups. When the sample was classified into sex-ancestry groups, the number of dependent variables increased considerably, which lead to a decrease in classification accuracies. As multivariate subsets classified the sample better according to sex than the bone models, only multivariate subsets were utilised to classify into sex and ancestry. To conserve adequate sample sizes some variables were removed from the original subsets. The subsets once again underwent stepwise selection to select the best variables for model creation (Table 4.14 and Table 4.15).

Table 4.14 - Multivariate subsets for LDA when classifying according to sex and ancestry.

	n	*Variables*
shoulder	198	scapbr, humebr, humhdd, hummwd
hip	228	iliabr, sacabr, femhdd, fembln
upper	224	humhdd, ulnxln, radapd, humxln
lower	226	femmtv, femhdd, femstv, innoht, sacabr, tibnfx
breadths	188	femebr, iliabr, tibpeb, ulntvd, sacabr
lengths	189	innoht, ulnxln, scapht, femxln
proxdist	315	humhdd, tibpeb, tibdeb
all-variable model	226	femhdd, iliabr, ulnxln, ulntvd, calcbr, sacabr

Table 4.15 - Multivariate subsets for FDA when classifying according to sex and ancestry.

	n	*Variables*
shoulder	198	humhdd, claapd, humebr, clavrd
hip	228	iliabr, femhdd, innoht, fembln, sacabr, femebr
upper	224	humhdd, ulnxln, radapd, humxln
lower	226	iliabr, femhdd, femxln, innoht, femstv, femmtv, sacabr, tibnft
breadths	188	femebr, iliabr, tibpeb, hummxd
lengths	189	innoht, ulnxln, scapht
proxdist	315	humhdd, tibpeb, tibdeb, femhdd, humebr
all-variable model	226	femhdd, iliabr, ulnxln, femstv, calcbr, sacabr

LDA classified the sample 53% to 80% correctly (Table 4.16 and Figure 4.4). The all-variable model achieved the highest classification accuracy whereas the shoulder subset achieved the lowest. Overall white females classified the best and coloured males classified the worst. When FDA classified the sample into sex and ancestry using the multivariate subsets, correct classification ranged from 56% to 79% (Table 4.17 and Figure 4.5). The shoulder subset produced the lowest and the all-variable model the highest classification accuracies. Overall, white males classified best and coloured males worst. For the all-variable model overlap was present between black and coloured females and between black and coloured males for both LDA and FDA (Figure 4.4 and Figure 4.5). White males and females, while still overlapping with the other groups, obtained the highest correct classifications. Comparison of the multivariate classification techniques demonstrated that FDA was only slightly better in classifying the sample into sex and ancestry groups with an average accuracy of 66%, compared to the 65% of LDA (Table 4.18).

Table 4.16– Correct classification rates when sex and ancestry were assessed by LDA using multivariate subsets.

	Accuracy (%)						
	Black female	*Black male*	*White female*	*White male*	*Coloured female*	*Coloured male*	*Combined*
shoulder	52	41	64	68	56	37	53
hip	61	52	92	64	65	57	64
upper	59	64	83	70	68	60	67
lower	49	62	84	73	73	70	68
breadths	39	59	88	75	63	50	63
lengths	57	62	70	81	68	48	65
proxdist	46	60	77	70	67	23	56
all-variable	78	78	95	81	79	65	80

Table 4.17- Correct classification rates when sex and ancestry were assessed by FDA using multivariate subsets.

	Accuracy (%)						
	Black female	*Black male*	*White female*	*White male*	*Coloured female*	*Coloured male*	*Combined*
shoulder	44	47	64	67	61	48	56
hip	66	58	83	82	66	36	64
upper	61	71	67	82	74	71	71
lower	71	77	85	82	68	67	75
breadths	41	61	78	84	59	45	62
lengths	49	61	70	80	68	41	60
proxdist	49	57	63	73	60	45	57
all-variable	79	81	94	85	72	63	79

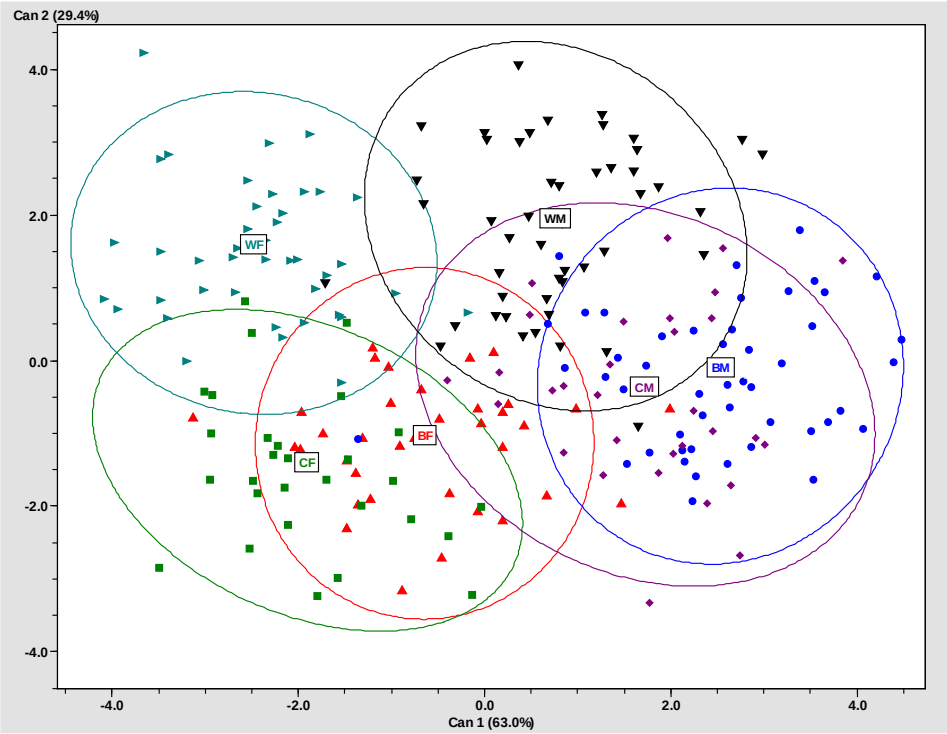

Figure 4.4 – Graph illustrating LDA classification into sex and ancestry groups using the all-variable model.

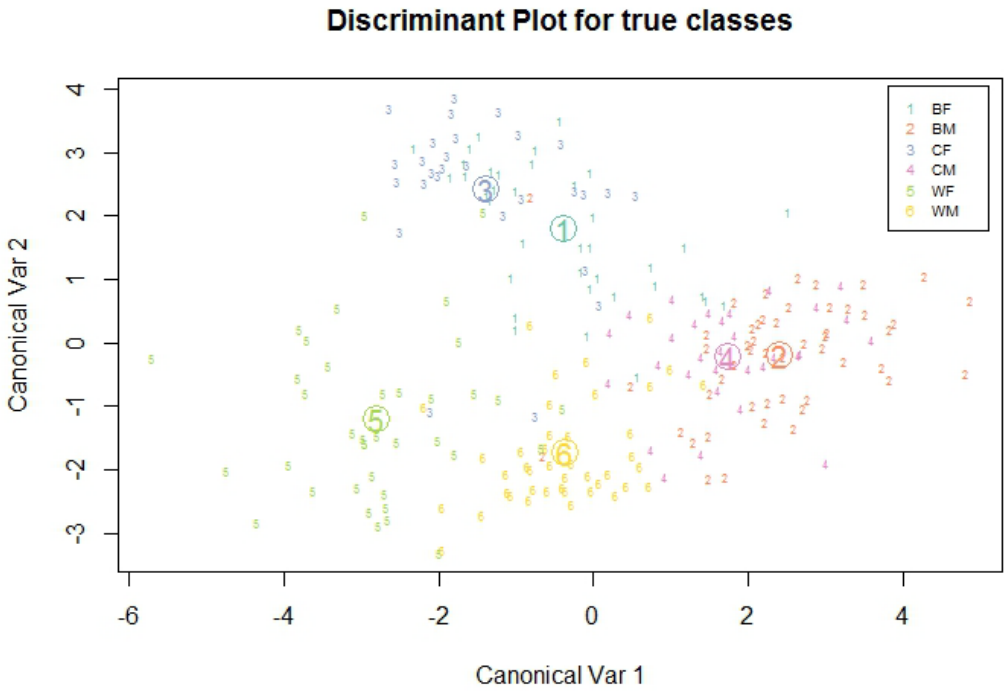

Figure 4.5 – FDA plot illustrating classification of the South African sample into sex and ancestry groups when using the all-variable model.

Table 4.18 - Comparison of LDA and FDA for multivariate subsets

	Accuracy (%)	
	LDA	*FDA*
shoulder	53	56
hip	64	64
upper	67	71
lower	68	75
breadths	63	62
lengths	65	60
proxdist	56	57
all-variable	80	79

Only the three subsets that classified the sample better than 67% (50% better than chance) underwent PCA to remove any multicollinearity between variables within each model. For the upper subset, using LDA, the components obtained from PCA included PC1-PC4 and PC6-PC7. The first PC was highly loaded by ulnxln, ulnphl and humxln and accounted for 25.4% of the total variance. PC2, PC4 and PC7 had highest contributions from the midshaft measurements (clavrd, ulndvd and ulntvd, respectively), whereas PC3 and PC6 were most heavily loaded by humhdd and claxln. Stepwise chose PC1-PC9, with the exception of PC8, for model creation of the lower subset. The first component was heavily loaded by iliabr and innoht, whereas PC2 obtained the largest contribution from femxln, which together with PC1 accounted for 25.7% of variance. After stepwise selection, the all-variable subset comprised of PC1-PC5 and PC7. The first component was highly loaded by ulnxln, which was expected as ulnxln was the only length measurement in the subset. Components two, four and five obtained the largest contributions from calcbr, iliabr, and sacabr, respectively, whereas PC3 and PC7 were most heavily loaded by ulntvd and femmtv.

When the three subsets underwent PCA and stepwise selection for FDA, the upper subset was composed of PC1-PC7. The first component was again most heavily loaded by length measurements (e.g. ulnxln, ulnphl and humxln). Components two, four, five and seven all obtained the largest contributions from midshaft measurements, whereas PC3 and PC6 were most heavily loaded by humhdd and claxln, respectively.

Assessment of the principal components, using LDA and FDA, resulted in slight decreases in correct classifications from when the original measurements were analysed. LDA, using the all-variable model, classified the sample 73% correctly, whereas FDA classified 77% correctly (Table 4.19 and Table 4.20). White females classified best (87%), whereas black females and coloured males classified worst (63% both) when LDA was used. The results for FDA also exhibited the highest classification accuracies for white females (94%) and the lowest accuracies for black females (63%). Overall FDA classified the sample better into sex and ancestry than LDA when PC scores of the upper, lower and all-variable subsets were utilised (Table 4.21).

Table 4.19 – Correct classification rates when sex and ancestry were assessed by LDA using PC scores from multivariate subsets.

	Accuracy (%)						
	Black female	*Black male*	*White female*	*White male*	*Coloured female*	*Coloured male*	*Combined*
upper	47	58	65	67	59	50	58
lower	66	58	84	70	70	70	69
all-variable model	54	74	87	74	82	65	73

Table 4.20 - Correct classification rates when sex and ancestry were assessed by FDA using PC scores from multivariate subsets.

	Accuracy (%)						
	Black female	*Black male*	*White female*	*White male*	*Coloured female*	*Coloured male*	*Combined*
upper	51	65	65	68	71	67	64
lower	63	75	78	73	70	66	70
all-variable model	63	83	94	85	66	70	77

Table 4.21 - Comparison of LDA and FDA using PC scores from multivariate subsets.

	Accuracy (%)	
	LDA	*FDA*
upper	58	64
lower	69	70
all-variable model	73	77

4.3. Modern North Americans

4.3.1. ANOVA and mean comparisons between North Americans and South Africans

ANOVA was utilised to compare the North Americans and their South African counterparts. In Table 4.22, the comparison for all measurements between black and white North Americans and South Africans (NAB, NAW, SAB and SAW, respectively) are shown. The group were separated by sex as a means to examine ancestry differences between the populations. Significant ancestry differences between females were present for all measurements except clavrd and humebr. For the male groups significant differences between ancestry groups were present for most measurements, except two clavicle measurements (claxln, clavrd), radapd, femstv, femmap, tibnfx, tibnft and calcbr.

Closer examinations of the ancestry comparisons were conducted using Tukey's post hoc test. Again males and females were assessed separately to focus on ancestry differences between black and white North Americans and South Africans. The female groups showed overlap among the means for the claxln, clavrd, humebr, radtvd, and calcbr (Appendix VIII). Only humhdd and iliabr showed no overlap among the female ancestry groups. For the sacabr measurements significant differences were present among all groups, except between NAB and SAW. Similarly, innoht means were significantly different for all groups with the exception of NAB and SAB, for which the means were not significantly different. For most measurement means at least three of the female groups showed significant ancestry differences. However, hummxd, sacs1b and fibmdm showed differences between only two of the female groups. Hummxd means were only significantly different between NAW and SAW females. Similarly, fibmdm was only

significantly different between NAW and SAW and sacs1b only between NAW and NAB females.

Tukey's tests of the male ancestry groups revealed that for no measurement were significant differences noted among all four ancestry groups (Appendix VIII). For all measurements the means of at least two groups overlapped. The means of humhdd, sacabr, femxln, fembln and femebr showed overlap between two of the four groups. Humhdd, femxln and fembln were not significantly different between NAW and SAW males, and sacabr and femebr showed overlap between SAW and NAB groups. In contrast, hummxd, radapd and ulntvd means only overlapped between NAB and SAB and the means of radtvd, femhdd and calcxl only overlapped between SAB and NAW. The ulndvd and femmtv means were not significantly different between SAW and NAW, whereas fibmdm was not significantly different between SAW and NAB. For all remaining measurements more than two groups overlapped.

Table 4.22 – ANOVA testing for significant differences ($p < 0.05$) between black and white North American and South African females and males. Bold indicates significance.

		Ancestry (Females)		Ancestry (Males)	
		F-value	*Pr(>F)*	*F*-value	*Pr(>F)*
clavicle	**claxln**	3.12	**0.03**	1.50	0.22
	claapd	7.17	**<0.001**	8.10	**<0.001**
	clavrd	0.09	0.96	1.45	0.2
scapula	**scapht**	11.74	**<0.001**	12.58	**<0.001**
	scapbr	8.34	**<0.001**	5.71	**<0.001**
humerus	**humxln**	15.22	**<0.001**	11.86	**<0.001**
	humebr	2.61	0.05	9.27	**<0.001**
	humhdd	47.69	**<0.001**	41.97	**<0.001**
	hummxd	3.61	**0.01**	4.52	**0.004**
	hummwd	4.43	**0.005**	9.13	**<0.001**
radius	**radxln**	8.04	**<0.001**	10.88	**<0.001**
	radapd	6.17	**<0.001**	2.56	0.06
	radtvd	2.94	**0.03**	5.37	**0.001**
ulna	**ulnxln**	7.68	**<0.001**	10.59	**<0.001**
	ulndvd	15.00	**<0.001**	3.46	**0.02**
	ulntvd	16.93	**<0.001**	4.53	**0.004**
	ulnphl	11.79	**<0.001**	19.35	**<0.001**
sacrum	**sacaht**	5.96	**<0.001**	7.90	**<0.001**
	sacabr	47.13	**<0.001**	49.17	**<0.001**
	sacs1b	3.01	**0.03**	4.41	**0.005**
innominate	**innoht**	54.16	**<0.001**	39.76	**<0.001**
	iliabr	75.38	**<0.001**	28.93	**<0.001**
femur	**femxln**	9.88	**<0.001**	14.00	**<0.001**
	fembln	9.33	**<0.001**	13.12	**<0.001**
	femebr	14.78	**<0.001**	22.46	**<0.001**
	femhdd	18.05	**<0.001**	10.09	**<0.001**
	femsap	4.19	**0.007**	10.26	**<0.001**
	femstv	14.25	**<0.001**	2.56	0.06
	femmap	4.84	**0.003**	2.25	0.08
	femmtv	8.84	**<0.001**	4.69	**0.003**
tibia	**tibxln**	5.02	**0.002**	6.25	**<0.001**
	tibpeb	7.07	**<0.001**	7.61	**<0.001**
	tibdeb	15.50	**<0.001**	11.08	**<0.001**
	tibnfx	8.70	**<0.001**	0.82	0.49
	tibnft	6.91	**<0.001**	2.62	0.05
fibula	**fibxln**	5.78	**<0.001**	4.43	**0.005**
	fibmdm	3.39	**0.02**	5.79	**<0.001**
calcaneus	**calcxl**	14.99	**<0.001**	7.66	**<0.001**
	calcbr	2.68	**0.04**	0.98	0.4

4.3.2. Sexual dimorphism differences in North American groups (sympercents)

Sympercents (s%) were calculated to assess the sexual size dimorphism between North American black and white males and females. North Americans blacks showed a greater degree of sexual dimorphism than white North Americans for the claapd, scapbr, ulndvd, sacabr, sacs1b, iliabr and tibdeb. The five largest differences between North American black males and females were present in the ulndvd, hummwd, claapd, radapd and clavrd. For the remaining 31 measurements, the difference between white males and females was larger than between black males and females. The greatest s% differences between North American whites were displayed in the ulntvd, radapd, clavrd, hummwd and radtvd. Both black and white females had larger mean values for sacral breadth than their male counterparts (Appendix IX). Black females also had a larger sacaht mean compared to black males. The s% differences produced for ulnphl were equal for both groups. The sympercent differences were compared for North American and South African blacks and whites to explore which group presented with the highest levels of sexual dimorphism. When the overall mean for all measurements was calculated for each population, North American whites had the highest mean sympercents (5.6 s%) and South African whites the lowest (4.4 s%) (Table 4.23, Appendix X). North American whites had the largest sympercent difference for most measurements compared to the other three groups. However, North American blacks had the biggest differences between the sexes for scapbr, sacabr, sacs1b, ulndvd and tibpeb. South African blacks were most dimorphic in terms of claapd, iliabr, sacaht, ulnphl and calcxl. South African whites were never more sexually dimorphic than either of the North Americans groups or the South African black group.

Table 4.23 – Sympercent differences between North Americans (NA) and South Africans (SA). Bold indicates largest sympercent difference per measurement.

	NA blacks	*SA blacks*	*NA whites*	*SA whites*
claxln	4.02	4.01	**5.1**	4.44
claapd	8.06	**9.54**	7.04	7.82
clavrd	7.88	7.9	**9.77**	9.53
scapht	5.88	5.66	**6.14**	5.13
scapbr	**5.77**	5.51	5.6	4.52
humxln	3.82	3.68	**4.0**	3.36
humebr	5.65	4.74	**6.87**	5.5
humhdd	5.69	5.79	**6.25**	4.78
hummxd	6.46	5.41	**7.28**	5.74
hummwd	8.26	6.42	**8.98**	7.97
radxln	4.79	4.69	**4.8**	4.32
radapd	8.02	8.05	**10.0**	8.39
radtvd	4.97	6.04	**7.96**	7.13
ulnxln	4.67	4.3	**4.76**	4.12
ulndvd	**11.62**	6.5	7.11	6.76
ulntvd	6.88	8.5	**11.43**	7.48
ulnphl	4.6	**4.63**	4.6	3.84
sacaht	-0.86	**2.83**	1.16	0.97
sacabr	**-1.68**	-1.49	-0.85	-1.44
sacs1b	**5.19**	3.39	4.6	2.57
innoht	3.97	3.57	**4.27**	3.01
iliabr	1.67	**2.25**	1.26	0.34
femxln	3.65	3.15	**3.74**	2.88
fembln	3.8	3.26	**3.86**	2.91
femebr	4.82	4.57	**6.03**	4.13
femhdd	4.95	5.87	**6.21**	4.74
femsap	4.08	3.79	**4.92**	4.79
femstv	4.42	4.94	**6.16**	3.48
femmap	3.63	4.83	5.96	4.63
femmtv	5.23	4.89	**5.95**	4.37
tibxln	3.87	3.36	**4.12**	2.99
tibpeb	5.17	5.01	**5.49**	4.66
tibdeb	**4.96**	4.05	4.81	3.65
tibnfx	5.78	5.95	**7.19**	4.17
tibnft	4.36	5.2	**6.8**	5.81
fibxln	3.5	3.56	**4.36**	2.91
fibmdm	1.97	4.28	**5.82**	3.49
calcxl	3.96	**4.64**	4.53	3.0
calcbr	3.91	4.56	**5.58**	3.87

4.3.3. Multivariate Sex and Ancestry estimation of North Americans and South Africans

North Americans and South Africans were classified according to sex and ancestry using LDA and FDA. Stepwise selection removed unnecessary variables from the multivariate subsets and created smaller models that had adequate sample sizes. Tables 4.24 and 4.25 represent the variables chosen by each method for the different models.

Table 4.24– Multivariate subsets for LDA when classifying North Americans and South Africans according to sex and ancestry (eight groups).

	n	*Variables*
shoulder	334	claapd, humhdd
hip	284	femebr, sacabr, innoht, fembln, iliabr, femsap
upper	320	claapd, humxln, hummwd, humhdd, ulndvd, ulnxln, ulnphl
lower	295	sacabr, femebr, tibpeb, innoht, femxln, iliabr
breadths	281	claapd, ulndvd, femebr, iliabr, sacabr
lengths	262	ulnphl, innoht, ulnxln
proxdist	437	humhdd, femebr, tibpeb
all-variable model	276	hummwd, femebr, sacabr, innoht, ulnxln, ulnphl

Table 4.25 – Multivariate subsets for FDA when classifying North Americans and South Africans according to sex and ancestry (eight groups).

	n	*Variables*
shoulder	334	humhdd, claapd, claxln, clavrd, humxln
hip	284	innoht, sacabr, femxln, femhdd, femebr, femmtv, iliabr
upper	320	humhdd, ulnphl, ulndvd, ulntvd, ulnxln, claapd, radapd, humxln
lower	295	femebr, sacabr, innoht, femxln, femhdd, iliabr, femsap, fibxln, tibpeb
breadths	281	femebr, iliabr, sacabr, ulndvd, ulntvd, tibpeb, hummwd, tibnft
lengths	262	innoht, ulnxln, ulnphl, femxln, scapht, humxln
proxdist	437	humhdd, tibpeb, femebr, humebr, tibdeb
all-variable model	276	ulnphl, innoht, sacabr, ulnxln, iliabr, femebr, scapbr, claapd, calcbr, hummwd, tibnft, ulntvd, femmtv, femxln

Classification of North Americans and South Africans into sex-ancestry groups using LDA yielded accuracies that ranged from 41% to 72% (Appendix XI). The all-variable model achieved the highest correct classifications and the proxdist subset the lowest correct classification. When assessing average correct classification, South African black females and white males classified best, whereas South African black males and North American white females classified worst.

When assessing the results obtained by FDA, classification accuracies ranged from 50% to 86% for FDA (Appendix XI). The all-variable model produced the best classification accuracies and the proxdist subset produced the worst. With FDA North American white males and females classified best, whereas South African white males and females classified worst.

Accuracies for all subsets were higher when FDA was used (Table 4.28). FDA was also better at classifying North Americans, whereas LDA classified South Africans better than FDA. For both methods the proxdist achieved the lowest accuracies, whereas the all-variable model achieved the highest. Using LDA, the all-variable model achieved higher accuracies when South Africans were compared to one another (80%) than when South Africans and North Americans were assessed (72%). However, for FDA the all-variable model achieved better results when North Americans and South Africans were compared (86%) than when the South African groups were compared to one another (79%). The FDA plot illustrating classification of North Americans and South Africans (Figure 4.6) showed grouping of same sex and ancestry but different populations where South Africans black and white males and females showed similarities to their North American counterparts. While the groupings occur, differences among all groups are present for the measurements included in the all-variable subset and are apparent in the high correct classification rates (Table 4.26).

Table 4.26 – Comparison of multivariate classification methods, LDA and FDA, when classifying according to sex and ancestry.

	Accuracy (%)	
	LDA	*FDA*
shoulder	46	52
hip	62	63
upper	64	69
lower	66	71
breadths	59	67
lengths	55	66
proxdist	41	50
all-variable model	72	86

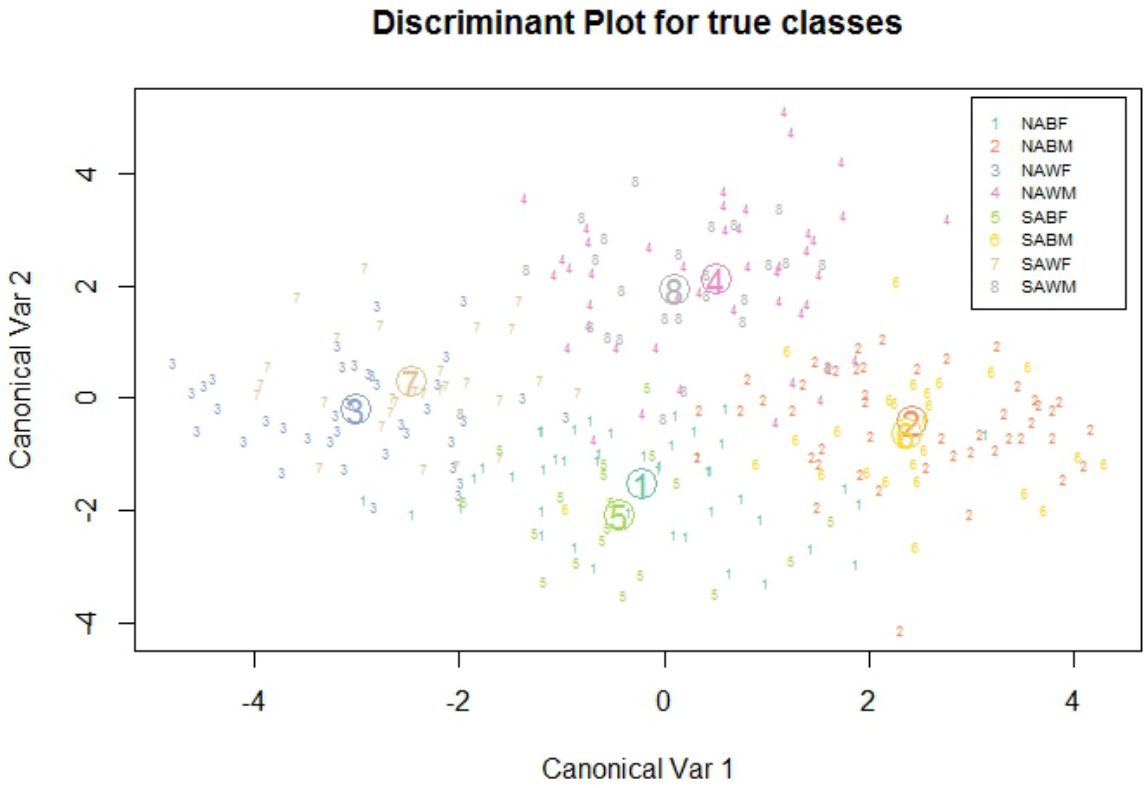

Figure 4.6 – FDA plot illustrating classification of the North American and South African samples into sex and ancestry groups when using the all-variable model.

Chapter 5: Discussion

While sexual dimorphism and sex estimation techniques have been previously researched in South Africa, continuous revisions of current standards are necessary as the composition of populations change and the application of previous methods become outdated. The current research is a comprehensive postcranial analysis that incorporated thorough, novel and robust statistical approaches as a means to improve knowledge of sexual dimorphism with a larger sample, and hence a larger amount of variation, in the South African population; to assess differences between North Americans and South Africans; and to modernize techniques used to estimate sex in South Africa.

5.1. Classification models and statistical methods

Past studies utilised univariate sectioning points to estimate sex and suggested the method was useful in cases where skeletal remains were fragmented (Steyn and İşcan, 1999; Asala, 2001; Bidmos and Asala, 2003). While the current study illustrated that sex estimation is possible using univariate methods, the amount of human variation expressed in a single variable is limited. When incorporating a custom database in FD3.1 the available variables, whether single or multiple measurements, can be used to provide cross-validated accuracies based on data collected from three South African ancestry groups. Another factor to consider is that with an increase in variables more information can be assembled and the classification accuracies may increase (Ousley and Jantz, 2012). Therefore if two or three measurements are available but from different fragmented bones, instead of two or three univariate analyses, FD3.1 can combine the measurements and create a multivariate analysis, which can provide more information on the individual assessed. Bone models (also referred to as the bone-by-bone approach) are also

commonly used in sex estimation (*e.g.*, Steyn and İşcan, 1997, 1999; Asala, 2001; Bidmos and Asala, 2003; Patriquin et al., 2005; Barrier, 2007). Although classification accuracies produced in the current study for a bone model are better than the correct classifications when using univariate methods, the bone (or bone model) assessed may not be the most sexually dimorphic skeletal element and not all measurements of the bone may be useful in estimating sex. Overall, multivariate subsets, combining different variables from different bones, provided the most information and the highest classification accuracies when estimating sex from the postcranial skeleton as they include only the most informative measurements.

For the estimation of sex a variety of statistical methods can be used to assess different models or measurements and depending on the method used classification accuracies may vary. As the names of the discriminant techniques suggest, LDA uses a linear plane to discriminate between males and females, whereas FDA makes use of a flexible surface to classify the sample (Hastie et al., 1994). Therefore the results are not unexpected in that FDA generally achieved higher correct classifications than LDA when classifying according to sex and sex and ancestry. While logistic regression is also defined as a more flexible alternative to LDA (Tabachnick and Fidell, 2007), the method failed to perform better than LDA for classifying the sample according to males and females. Additionally classification using logistic regression is limited to two groups, making classification into sex and ancestry groups impossible.

FDA is considered a valuable method for assessing skeletal data with studies showing superior results to the traditionally utilised LDA (Stull, 2014; Liebenberg, nd). Similarly, FDA obtained the best results in the current study. However, until a program is developed that can employ FDA, the method is difficult to apply in forensic anthropological casework. Furthermore, LDA is the main statistical approach utilised by FD3.1, which simplifies sex and ancestry

estimations. Therefore the results obtained for LDA may be more useful than the outcomes for FDA. However, as the use of FDA for robust statistical analyses becomes more popular in biological anthropology, a program for utilising FDA may be created.

5.2. Sexual dimorphism in black, white and coloured South Africans

Coloured South Africans overall showed the highest degree of sexual dimorphism among the three ancestry groups, but the extent to which sexual dimorphism was greater between coloured males and females than between black males and females and between white males and females, was fairly small. The pattern of sexual dimorphism in the three South African groups showed that overall breadth measurements constantly revealed greater differences between the sexes than bone lengths and are consistent with previous studies. The results can be explained by the different rates of interstitial and appositional growth, where earlier-growing regions (interstitial) tend to be less dimorphic than later-growing areas (appositional) (Steyn and İşcan, 1997; France, 1998; Humphrey, 1998; Spradley and Jantz, 2011; Stull, 2014). Furthermore, the upper limb showed higher levels of sexual dimorphism than the lower limb, although the lower limb was still fairly dimorphic. The sacrum revealed very limited differences between the sexes, which is similar to recent findings in an American population (Spradley and Jantz, 2011). Although a previous osteometric study of the pelvis has shown that sex can be reliably and accurately estimated from pelvic measurements (Patriquin et al., 2005), the pelvic measurements showed fairly low levels of sexual dimorphism in all three ancestry groups in the current study. An explanation may be the interaction between sex and ancestry observed for the innominate. Although the iliac breadth measurement had low levels of sexual dimorphism, this measurement was shown as useful in distinguishing between the ancestry groups (Liebenberg, nd).

The different degrees of sexual dimorphism among the South African ancestry groups may be further explained by separate origins along with historical and modern genetic admixture among the various population groups, which likely interacted to produce the differing expressions in sexual dimorphism in the South African population (Steyn and İşcan, 1998).

While some admixture occurred with Khoesan, the largest genetic contribution to white South Africans is from multiple European groups such as the Dutch, German, British and French (Greeff, 2007; L'Abbé et al., 2011). Because of the many different European groups contributing genes at variable frequencies, within group variation may have had an effect on modern sexual dimorphism. Sexual dimorphism is usually most observable in stature where males tend to be taller than females. However considerable population variation in stature exists and can be seen when the modern Dutch are compared to the other, considerably shorter, European groups. The admixing of the groups has created a wide range of variation in stature for both males and females, which makes stature less sexually dimorphic among white South Africans. Stature has also been highly correlated with long bone measurements and so a decrease in sexual dimorphism in stature means a similar decrease is likely to have occurred in the sexual dimorphism of long bone dimensions.

Coloured South Africans also received genetic contributions from both European and Khoesan groups, similar to white South Africans; however, the levels of influence of each parent group are vastly different. Variation in genetic composition is not only between coloured and white South Africans but also between coloured males and females. Additions from southeast Asians, Indians and Bantu-speakers make the South African coloureds a highly unique and vastly admixed ancestry group (Destro-Bisol et al., 1999; Tishkoff et al., 2009). Due to the sex-specific contributions of the Khoesan, greater similarities occur between modern coloured females and

Khoesan ancestors than between coloured males and the Khoesan groups. Khoesan were described as delicate and small people, particularly with regard to stature (Thompson, 2001; Tishkoff et al., 2009; de Wit et al., 2010; Quintana-Murci et al., 2010). The delicate features were inherited to a greater extent by the coloured females, which may have contributed to the large difference present between the males and females of the group. In contrast the coloured males inherited a greater proportion of genes from their European ancestors when compared to the coloured females (Quintana-Murci et al., 2010).

The black South Africans are descendants of Bantu-speakers that migrated from West Central Africa (possibly Cameroon) to South Africa approximately 5000 to 3000 years ago (Tishkoff et al., 2009; de Filippo et al., 2012). The Bantu-speakers were described as notably tall and well-built; however, the parent group is likely to have differentiated in southern Africa from the ones in Central Africa on account of different environmental pressures, the genetic influence of the Khoesan, and the founder effect. The complex combination of factors led to the intermediate stature of South African blacks (Hiernaux, 1966; Thompson, 2001). As no sex-bias has been mentioned in the literature with regard to the genetic contributions to the South African black population, sexual dimorphism in the group should occur to a higher degree than in the coloured group (Tishkoff et al., 2009). Even though gene flow is evident between the historical Bantu-speaking and indigenous Khoesan, modern black South Africans and the historic Khoesan are morphologically different (Liebenberg et al., in press; Herbert, 1990). Craniometric studies assessing the South African population for ancestry and sex, found that coloureds showed overlap with black and white groups while still classifying into their correct sex groups (L'Abbé et al., 2013 b). Similar findings in the current study showed that coloured males and females misclassified mainly into same sex but different ancestry groups. The Bantu-speaking and

Khoesan influences present in both coloured and black South Africans may be the cause for the overlap between the groups. Similarly, as both white and coloured South Africans have Khoesan and European genetic contributions, some overlap may also occur between coloured and white South Africans.

Age, environment, nutrition and socio-economic, also differed considerably among the South African groups and may have also affected variation in sexual dimorphism. Research has shown that white females are more prone to osteoporotic changes to the skeleton and may as a result become more robust with age, so that bone dimensions appear more similar to those of males. A decrease in bone mass influences an increase in continuing periosteal apposition, which causes an increase in bone dimensions. Bone loss has been associated with a lack of weight-bearing exercise and has been shown as more common in white individuals, specifically females. The mean age of the white females utilised in the study was 68 years, which makes an increase in robusticity a possible explanation as to the decreased level of sexual dimorphism observed between the white South African males and females. As bone loss is not common in black or coloured individuals and on account of the fairly low mean ages for the groups (47 and 44 years, respectively), age-related changes are not useful in explaining patterns of sexual dimorphism in black and coloured groups.

Decreased quality of nutrition and health care may be more likely explanations for the intermediate levels of sexual dimorphism observed between black males and females. Black South Africans were subject to increased environmental stress via intentional discrimination and low socio-economic status, which occurred as a result of political, physical and social segregation during Apartheid. Black South Africans were moved outside of the Republic of South African into separate homelands without access to work or other benefits, including

adequate nutrition. Indians and coloured South Africans were similarly affected by nutritional deficiencies as a result of race discrimination, but this occurred to a lesser degree than among black South Africans. Coloured and Indian groups were considered citizens of South Africa and were allowed to would work within the Republic, but they were also forced to live in their own segregated areas. Overall black groups were the most marginalized and as a result the most impoverished. The lack of quality nutrition may have caused a decrease in sexual dimorphism as the sexes responded differently to unfavourable nutritional conditions. Adolescent males are more likely to be affected by nutritionally deprived environments than their female counterparts and as a result may not reach their full height and muscle development. While the growth deficit may be restored if the stress is removed, the lack of improvement in socioeconomic status and nutrition in the modern black group may have hindered any catch-up growth. As females are not affected to the same extent as males, the degree of separation between the sexes may decrease in the population.

5.3. Comparison of South Africans and North Americans

Exploratory analyses revealed that North American and South African white males were most similar, which resulted in overlap between the two groups. Similarly, considerable overlap was evident between the North American and South African white females, but not to the same extent as the overlap observed among the males. The reason for the similar size in the two population groups may be attributed to their similar origins in multiple European countries and very little genetic influence from other ancestries groups outside of Europe (Davenport, 1926). Furthermore, differences between black and white North Americans can be mainly attributed to genetic influences on the phenotype, rather than socioeconomic standing and environmental

factors. Research has shown that even though some groups have co-existed for hundreds of years, morphologically they have not converged (Sparks and Jantz, 2002).

In contrast to the overlap seen in white North Americans and South Africans, large differences were observed between South African black males and females and their North American counterparts, which can be attributed to different genetic compositions of each of the black populations. While the largest genetic contribution to both populations originated in West/Central Africa, the temporal aspect of the migrations varies considerably. Whereas the Bantu-migration occurred thousands of years ago, the African diaspora occurred merely a few hundred years ago (Spradley, 2006; Tishkoff et al., 2009). As the Bantu-speakers had a longer period of time in which to adapt to new environments and admix with different groups, the resulting modern black South Africans are distinct from their West Central African parental group. In contrast, modern black North Americans have not had the same amount of time in which to distinguish themselves from their West African ancestors, but some differences between the groups have been observed (Spradley, 2006). Furthermore, admixture with different groups over time also added to the substantial differences between North American and South African blacks. North American black males and females received genetic contributions from Europeans and to a lesser extent from Native Americans. However, the European influence was sex-specific as larger contributions were made by white North Americans males than their female counterparts. The sex-bias is most likely due to the more common mating of North American white males with black females when compared to black North American males and white females, which was fairly uncommon. The large differences between black North Americans and South Africans are most likely due to the larger similarities between modern black North Americans and their West African ancestors when compared to modern black South

Africans and their Bantu-speaking ancestors. Furthermore, as only strong and healthy West Africans slaves were chosen and survived the voyage to North America, their larger size and greater strength were passed on to future generations. While North American blacks most likely retained their ancestor's overall large size, the European contributions to the group may also have added to larger measurements which in turn translated to larger differences in the means between North American and South African black individuals.

Comparisons of sexual dimorphism between North Americans and South Africans showed that once again white South Africans were the least dimorphic. The most distinct differences between males and females occurred in the white North Americans. The restricted admixture with other groups as well as the low mean ages of the North American white females (48 years) in the sample most likely resulted in the large observable differences between North American white males and females. Similarly, a recent study observed considerable sexual dimorphism between North American white males and females. Furthermore, black males and females of this population also expressed skeletal differences between the sexes (Spradley and Jantz, 2011).

A craniometric comparison of black North Americans and South Africans indicated a lesser degree of sexual dimorphism in the black South African group (Ousley and L'Abbé, 2010). Similar results were obtained in the current study. In the North American black males, once again, the European admixture as well as the large size inherited from the West African ancestors may have increased the differences between the males and females of the group, thereby creating observable sexual dimorphism. In contrast, while South African black males and females are skeletally different, environmental factors have likely decreased the expression of sexual dimorphism in the group.

5.4. Future recommendations

While this study provides useful information on the pattern of expression of sexual dimorphism and sex estimation methods in South African groups, a few problems have been encountered. As with all studies, increased sample sizes improve validity. However, the lack of young white South Africans in any of the South African skeletal collections makes increased sample sizes without increasing mean ages difficult. The older mean ages of white South Africans makes the sample less comparable to other groups (i.e. white North Americans) and may affect analyses of sexual dimorphism in white South Africans. Therefore, in order to improve on the understanding of the pattern expression of sexual dimorphism, younger white individuals need to be incorporated into the sample. Alternative sources of skeletal material, such as Lodox Statscan or computed tomography radiographic images, will have to be investigated. Another issue was the lack of a user-friendly method to utilise FDA for sex and ancestry estimation. The creation of a program similar to FORDISC could be very useful and improve current sex and ancestry estimation techniques.

Chapter 6: Conclusion

On account of small samples and outdated statistical approaches, a re-evaluation of osteometric techniques with new statistical methods was required. Furthermore, as the South African coloureds form one of the larger minorities in the South Africa, skeletal information on the group were incorporated into modified population-specific standards.

In this study, long bones were shown to reliably distinguish between males and females in three socially-defined South African populations as differences were noted among all sex and ancestry groups. Coloured South Africans demonstrated the greatest degree of sexual dimorphism followed by black and white South Africans, respectively. However, sexual dimorphism was varied between South Africans and North American groups. All South African groups were less dimorphic when compared to their North American counterparts, a situation that may be associated with increased levels of admixture, older mean ages and lower socioeconomic status observed in the South African population. As South Africans showed to be distinct from North Americans population-specific sex estimation standards were created for the South African custom database to be used with FD3.1.

Postcranial bones achieved comparable classification accuracies to morphological analysis of the pelvis and higher accuracies than metric or morphological techniques using the cranium in South Africa (Phenice, 1969; Walker, 2008; Klales et al., 2012; Krüger et al., 2014). Multivariate subsets of postcranial measurements not only captured more information but also produce higher correct classification rates than both bone models and univariate methods. While FDA yielded higher classification accuracies, LDA of multivariate subsets, outperformed logistic regression and obtained better results than previous studies on the South African population.

LDA is the best understood of the three techniques used in the current study and is able to estimate sex with high accuracies from the postcrania. Therefore, until progress can be made regarding the creation of a user-friendly program with which FDA can estimate biological parameters, LDA should be used to assess sex from the skeleton.

References

Abdi H, Williams LJ. 2010. Tukey's Honestly Significant Difference (HSD) test. In: Salkind N, editor. Encyclopedia of Research Design. Thousand Oaks, CA: Sage. p 1–5.

Adhikari M. 2005. Contending approaches to coloured identity and the history of the coloured people of South Africa. Hist Compass 3:1–16.

Afshartous D, Preston RA. 2011. Key result of interaction models with centering. J Stat Educ 19:1–24.

Albanese J, Eklics G, Tuck A. 2008. A Metric Method for Sex Determination Using the Proximal Femur and Fragmentary Hipbone. J Forensic Sci 53:1283–1288.

Alblas A. 2014. Kirsten Skeletal Collection.

Anderson JY, Trinkaus E. 1998. Patterns of sexual, bilateral and interpopulational variation in human femoral neck-shaft angles. J Anat 192:279–285.

Asala SA. 2001. Sex determination from the head of the femur of South African whites and blacks. Forensic Sci Int 117:15–22.

Asala SA. 2002. The efficiency of the demarking point of the femoral head as a sex determining parameter. Forensic Sci Int 127:114–118.

Barrier ILO, L'Abbé EN. 2008. Sex determination from the radius and ulna in a modern South African sample. Forensic Sci Int 179:85.e1–85.e7.

Barrier ILO. 2007. Sex determination from the bones of the forearm in a modern South African sample.

Beck RB. 2000. The History of South Africa. Westport, CT: Greenwood Press.

De Beer H. 2004. Observations on the history of Dutch physical stature from the late-Middle Ages to the present. Econ Hum Biol 2:45–55.

Berner D. 2011. Size correction in biology: how reliable are approaches based on (common) principal component analysis? Oecologia 166:961–971.

Bidmos M, Asala SA. 2003. Discriminant Function Sexing of the Calcaneus of the South African Whites. J Forensic Sci 48:1–6.

Bidmos MA, Gibbon VE, Štrkalj G. 2010. Recent advances in sex identification of human skeletal remains in South Africa. South Afr J Sci 106:1–6.

Blanchard BK. 2010. A study of the accuracy and reliability of sex estimation methods of the human pelvis.

Bland JM, Altman DG. 1999. Measuring agreement in method comparison studies. Stat Methods Med Res 8:135–160.

Bogin B. 1999. Patterns of Human Growth. In: Cambridge Studies in Biological Anthropology. 2nd ed. Cambridge, UK: Cambridge University Press. p 267.

Bonjour J, Theintz G, Buchs B, Slosman D, Rizzoli R. 1991. Critical Years and Stages of Puberty for Spinal and Femoral Bone Mass Accumulation during Adolescence. J Clin Endocrinol Metab 73:555–563.

Buikstra JE, Ubelaker DH, Aftandilian D. 1994. Standards for Data Collection from Human Skeletal Remains. Fayetville, AR: Arkansas Archeological Survey.

Cabo LL, Brewster CP, Luengo Azpiazu J. 2012. Sexual Dimorphism: Interpreting Sex Markers. In: Dirkmaat DC, editor. A Companion to Forensic Anthropology. 1st ed. West Sussex, UK: Wiley-Blackwell. p 248–286.

Callewaert F, Sinnesael M, Gielen E, Boonen S, Vanderschueren D. 2010. Skeletal sexual dimorphism: relative contribution of sex steroids, GH-IGF1, and mechanical loading. J Endocrinol 207:127–134.

Cameron N. 2002. Human Growth and Development. (Cameron N, editor.). San Diego: Academic Press.

Canty A, Ripley B. 2014. boot: bootstrap functions. Available from: http://cran.r-project.org/web/packages/boot/index.html

Casado AM. 2010. Evaluation of Features of the Innominate for Sex Estimation.

Cheverud JM, Buikstra JE, Twichell E. 1979. Relationships between non-metric skeletal traits and cranial size and shape. Am J Phys Anthropol 50:191–198.

Cole TJ. 2000. Sympercents: symmetric percentage differences on the 100 log e scale simplify the presentation of log transformed data. Stat Med 19:3109–3125.

Compston JE, Vedi S, Kaptoge S, Seeman E. 2007. Bone Remodeling Rate and Remodeling Balance Are Not Co-Regulated in Adulthood: Implications for the Use of Activation Frequency as an Index of Remodeling Rate. J Bone Miner Res 22:1031–1036.

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

Davenport CB. 1926. The Old Americans. Science 63:47–48.

Dawson-Saunders B, Trapp RG eds. 2004. Basic and Clinical Biostatistics. 4th ed. New York, NY: McGraw-Hill.

Dayal MR, Spocter MA, Bidmos MA. 2008. An assessment of sex using the skull of black South Africans by discriminant function analysis. HOMO 59:209–221.

Destro-Bisol G, Maviglia R, Caglià A, Boschi I, Spedini G, Pascali V, Clark A, Tishkoff S. 1999. Estimating European admixture in African Americans by using microsatellites and a microsatellite haplotype (CD4/Alu). Hum Genet 104:149–157.

DiGangi EA, Moore MK eds. 2013. Research Methods in Human Skeletal Biology. Waltham, MA: Academic Press.

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

Fox J. 2014. car: Companion to Applied Regression. Available from: http://cran.r-project.org/web/packages/car/index.html

France DL. 1998. Observational and Metric Analysis of Sex in the Skeleton. In: Reichs KJ, editor. Forensic Osteology: Advances in the Identification of Human Remains. 2nd ed. Springfield, IL: Charles C. Thomas. p 163–186.

Frayer DW, Wolpoff MH. 1985. Sexual Dimorphism. Annu Rev Anthr 14:429–473.

Galdames ICS, Matamala DAZ, Smith RL. 2008. Evaluating accuracy and precision in morphologic traits for sexual dimorphism in malnutrition human skull: a comparative study. Int J Morphol 26:876–83.

Gapert R, Black S, Last J. 2009. Sex determination from the occipital condyle: Discriminant function analysis in an Eighteenth and Nineteenth Century British sample. Am J Phys Anthropol 138:384–394.

Garvin HM. 2012. Adult Sex Determination: Methods and Application. In: Dirkmaat DC, editor. A Companion to Forensic Anthropology. 1st ed. West Sussex, UK: Wiley-Blackwell. p 239–247.

Gastwirth JL, Gel YR, Miao W. 2009. The Impact of Levene's Test of Equality of Variances on Statistical Theory and Practice. Stat Sci 24:343–360.

Gray JP, Wolfe LD. 1980. Height and sexual dimorphism of stature among human societies. Am J Phys Anthropol 53:441–456.

Greeff JM. 2007. Deconstructing Jaco: Genetic Heritage of an Afrikaner. Ann Hum Genet 71:674–688.

Greulich WW. 1951. Growth and Development Status of Guamanian School Children in 1947. Am J Phys Anthropol 9:55–70.

Grewal R, Cote JA, Baumgartner H. 2004. Multicollinearity and Measurement Error in Structural Equation Models: Implications for Theory Testing. Mark Sci 23:519–529.

Gustafsson A, Lindenfors P. 2004. Human size evolution: no evolutionary allometric relationship between male and female stature. J Hum Evol 47:253–266.

Hair JF, Black WC, Babin BJ, Anderson RE. 2010. Multivariate Data Analysis. 7th ed. Upper Saddle River, NJ: Prentice Hall.

Harris B. 1997. Growing taller, living longer? Anthropometric history and the future of old age. Ageing Soc 17:491–512.

Harris EF, Smith RN. 2009. Accounting for measurement error: A critical but often overlooked process. Arch Oral Biol 54:S107–S117.

Hastie T, Tibshirani R, Buja A. 1994. Flexible Discriminant Analysis by Optimal Scoring. J Am Stat Assoc 89:1255–1270.

Hastie T. 2013. mda: Mixture and flexible discriminant analysis. Available from: http://cran.r-project.org/web/packages/mda/index.html

Hefner JT, Ousley SD. 2014. Statistical Classification Methods for Estimating Ancestry Using Morphoscopic Traits. J Forensic Sci.

Henneberg M, Van den Berg ER. 1990. Test of socioeconomic causation of secular trend: stature changes among favored and oppressed South Africans are parallel. Am J Phys Anthropol 83:459–465.

Herbert RK. 1990. The Sociohistory of Clicks in Southern Bantu. Anthr Linguist 32:295–315.

Hiernaux J. 1966. Human Biological Diversity in Central Africa. Man 1:287–306.

Holm S. 1979. A Simple Sequentially Rejective Multiple Test Procedure. Scand J Stat 6:65–70.

Hrdlička A. 1925. The Old Americans: A Physiological Profile. New York, NY: Arno Press.

Huberty CJ. 1975. Discriminant Analysis. Rev Educ Res 45:543–598.

Huberty CJ. 1994. Applied discriminant analysis. 1st ed. United States of America: Wiley.

Humphrey LT. 1998. Growth Patterns in the Modern Human Skeleton. Am J Phys Anthropol 105:57–72.

İşcan MY. 2005. Forensic anthropology of sex and body size. Forensic Sci Int 147:107–112.

Jamaiyah H, Geeta A, Safiza MN, Wong NF, Kee CC, Ahmad AZ, Suzana S, Rahmah R, Khor GL, Ruzita AT. 2008. Reliability and Technical Error of Calf Circumference and Mid-half Arm Span Measurements for Nutritional Status Assessment of Elderly Persons in Malaysia. Malays J Nutr 14.

Jantz RL, Moore-Jansen PH. 1988. A Database for Forensic Anthropology: structure content and analysis.

Jantz RL, Ousley SD. 2005. Fordisc 3: computerized forensic discriminant functions. Knoxville, TN: University of Tennessee, Knoxville.

Jolliffe I. 2002. Principal Component Analysis. 2nd ed. New York, NY: Springer-Verlag.

Juul A. 2001. The effects of oestrogens on linear bone growth. Apmis 109:S124–S134.

Keen JA. 1950. A study of the differences between male and female skulls. Am J Phys Anthropol 8:65–80.

Kemkes-Grottenthaler A. 2005. Sex determination by discriminant analysis: an evaluation of the reliability of patella measurements. Forensic Sci Int 147:129–133.

Keough N, L'Abbé EN, Steyn M. 2009. The evaluation of age-related histomorphometric variables in a cadaver sample of lower socioeconomic status: implications for estimating age at death. Forensic Sci Int 191:114.e1–114.e6.

Kimmerle EH, Ross A, Slice D. 2008. Sexual Dimorphism in America: Geometric Morphometric Analysis of the Craniofacial Region. J Forensic Sci 53:54–57.

King CA, İscan MY, Loth SR. 1998. Metric and comparative analysis of sexual dimorphism in the Thai femur. J Forensic Sci 43:954–958.

Kittoe AG. 2013. On Human Biological Diversity: Variation in Sexual Dimorphism of the Skull between African-Americans and European-Americans.

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

Kleinbaum DG, Klein M. 2010. Logistic Regression: A Self-Learning Text. 3rd ed. New York, NY: Springer.

Klepinger LL. 2001. Stature, Maturation Variation and Secualr trends in Forensic Anthropology. J Forensic Sci 46:788–790.

Knapp TR. 1992. Technical Error of Measurement: A Methodological Critique. Am J Phys Anthropol 87:235–236.

Knickmeyer RC, Baron-Cohen S. 2006. Fetal testosterone and sex differences. Early Hum Dev 82:755–760.

Krüger GC, L'Abbé EN, Stull KE, Kenyhercz MW. 2014. Sexual dimorphism in cranial morphology among modern South Africans. Int J Legal Med Accepted.

Kuhn M. 2014. caret: Classification and regression training. Available from: http://cran.r-project.org/web/packages/caret/index.html

L'Abbé EN, Kenyhercz MW, Stull KE, Keough N, Nawrocki S. 2013a. Application of Fordisc 3.0 to Explore Differences Among Crania of North American and South African Blacks and Whites. J Forensic Sci 58:1579–1583.

L'Abbé EN, Kenyhercz MW, Stull KE, Ousley SD. 2013b. Craniometric Assessment of Modern 20th-Century Black, White, and "Colored" South Africans. In: Proceedings of the American Academy of Forensic Sciences. Vol. XIX. Washington, DC: American Academy of Forensic Sciences. p 444.

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 Sci Int 209:195.e1–195.e7.

L'Abbé EN, Steyn M. 2012. The Establishment and Advancement of Forensic Anthropology in South Africa. In: Dirkmaat DC, editor. A Companion to Forensic Anthropology. UK: Wiley-Blackwell. p 626–638.

Leathwick JR, Rowe D, Richardson J, Elith J, Hastie T. 2005. Using multivariate adaptive regression splines to predict the distributions of New Zealand's freshwater diadromous fish. Freshw Biol 50:2034–2052.

Liebenberg L, Stull KE, L'Abbé EN, Botha D. in press. Evaluating the accuracy of cranial indices in ancestry estimation among South African groups. J Forensic Sci.

Liebenberg L. 2015. Postcraniometric assessment of ancestry in among modern South Africans.

Loth SR, İşcan MY. 2000. Sex Determination. Encycl Forensic Sci:252–260.

Maalouf G, Gannagé-Yared MH, Ezzedine J, Larijani B, Badawi S, Rached A, Zakroui L, Masri B, Azar E, Saba E, others. 2007. Middle East and North Africa consensus on osteoporosis. J Musculoskelet Neuronal Interactions 7:131.

Maat GJR. 2005. Two millennia of male stature development and population health and wealth in the Low Countries. Int J Osteoarchaeol 15:276–290.

Macaluso PJ, Lucena J. 2014. Estimation of sex from sternal dimensions derived from chest plate radiographs in contemporary Spaniards. Int J Legal Med 128:389–395.

May A, Hazelhurst S, Li Y, Norris SA, Govind N, Tikly M, Hon C, Johnson KJ, Hartmann N, Staedtler F. 2013. Genetic diversity in black South Africans from Soweto. BMC Genomics 14:644.

McDonald JH. 2008. Handbook of Biological Statistics. Baltimore, MD: Sparky House Publishing.

Mertler CA, Vannatta RA eds. 2002. A Guide to Multivariate Techniques. In: Advanced and Multivariate Statistical Methods. 2nd ed. Los Angeles, CA: Pyrczak Publishing.

Milborrow S. 2013. Notes on the earth package. Available from: http://CRAN.R-project.org/package=earth

Milner GR, Boldsen JL. 2012. Humeral and Femoral Head Diameters in Recent White American Skeletons. J Forensic Sci 57:35–40.

Milner GR, Wood JW, Boldsen JL. 2008. Advances in Paleodemograhy. In: Katzenberg MA, Saunders SR, editors. Biological Anthropology of the Human Skeleton. 2nd ed. Hoboken, New Jersey: Wiley-Liss. p 561–600.

Møller V. 1998. Quality of life in South Africa: Post-apartheid trends. Soc Indic Res 43:27–68.

Moore-Jansen PH, Jantz RL. 1994. Data collection procedures for forensic skeletal material.

Nelson DA, Pettifor JM, Barondess DA, Cody DD, Uusi-Rasi K, Beck TJ. 2004. Comparison of Cross-Sectional Geometry of the Proximal Femur in White and Black Women from Detroit and Johannesburg. J Bone Miner Res 19:560–565.

Novotný V, İşcan MY, Loth SR. 1993. Morphologic and Osteometric Assessment of Age, Sex and Race from the Skull. In: İşcan MY, Helmer RP, editors. Forensic Analysis of the Skull: Craniofacial Analysis, Reconstruction, and Identification. New York, NY: Wiley-Liss. p 71–88.

Ousley SD, Jantz RL, Freid D. 2009. Understanding race and human variation: Why forensic anthropologists are good at identifying race. Am J Phys Anthropol 139:68–76.

Ousley SD, Jantz RL. 2012. Fordisc 3 and Statistical Methods for Estimating Sex and Ancestry. In: Dirkmaat DC, editor. A Companion to Forensic Anthropology. West Sussex, UK: Wiley-Blackwell. p 311–329.

Ousley SD, Jantz RL. 2013. Fordisc 3: Third generation of computer-aided forensic anthropology. Rechtsmedizin 23:97–99.

Ousley SD, L'Abbé EN. 2010. Craniometric Variation in South African and American Blacks. In: Proceedings of the American Academy of Forensic Sciences. Seattle, WA. p 377–378.

Parra EJ, Marcini A, Akey J, Martinson J, Batzer MA, Cooper R, Forrester T, Allison DB, Deka R, Ferrell RE, Shriver MD. 1998. Estimating African American admixture proportions by use of population-specific alleles. Am J Hum Genet 63:1839–1851.

Patriquin ML, Steyn M, Loth SR. 2005. Metric analysis of sex differences in South African black and white pelves. Forensic Sci Int 147:119–127.

Patterson N, Petersen 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. Hum Mol Genet 19:411–419.

Pearson OM, Lieberman DE. 2004. The aging of Wolff's ?law?: Ontogeny and responses to mechanical loading in cortical bone. Am J Phys Anthropol 125:63–99.

Pearson OM. 2000. Activity, Climate, and Postcranial Robusticity: Implications for Modern Human Origins and Scenarios of Adaptive Change. Curr Anthr 41:569–607.

Perini TA, de Oliveira GL, dos Santos Ornellas J, de Oliveira FP. 2005. Technical error of measurement in anthropometry. Rev Bras Med Esporte 11:81–85.

Pfeiffer S. 1980. Age Changes in the External Dimensions of Adult Bone. Am J Phys Anthropol 52:529–532.

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

Pickering R, Bachman D. 2009. The Use of Forensic Anthropology. 2nd ed. Boca Raton, FL: CRC Press.

Pietrusewsky M. 2008. Metric analysis of skeletal remains: methods and applications. In: Katzenberg MA, Saunders SR, editors. Biological Anthropology of the Human Skeleton. 2nd ed. Hoboken, New Jersey: Wiley-Liss. p 375–415.

Pretorius E, Steyn M, Scholtz Y. 2006. Investigation into the usability of geometric morphometric analysis in assessment of sexual dimorphism. Am J Phys Anthropol 129:64–70.

Purkait R, Chandra H. 2004. A study of sexual variation in Indian femur. Forensic Sci Int 146:25–33.

Quintana-Murci L, Harmant C, Quach H, Balanovsky O, Zaporozhchenko V, Bormans C, van Helden PD, Hoal EG, Behar DM. 2010. Strong Maternal Khoisan Contribution to the South African Coloured Population: A Case of Gender-Biased Admixture. Am J Hum Genet 86:611–620.

R Core Team. 2013. R: A language and environment for statistical computing. R Found Stat Comput Vienna Austria [Internet]. Available from: http://www.R-project.org/

Ramsthaler F, Kreutz K, Verhoff MA. 2007. Accuracy of metric sex analysis of skeletal remains using Fordisc® based on a recent skull collection. Int J Legal Med 121:477–482.

Revelle W. 2011. psych: procedures for psychological, psychometric, and personality research. Available from: http://cran.r-project.org/web/packages/psych/index.html

Ringnér M. 2008. What is principal component analysis? Nat Biotechnol 26:303–304.

Rissech C, López-Costas O, Turbón D. 2013. Humeral development from neonatal period to skeletal maturity—application in age and sex assessment. Int J Legal Med 127:201–212.

Robinson MS, Bidmos MA. 2009. The skull and humerus in the determination of sex: Reliability of discriminant function equations. Forensic Sci Int 186:86.e1–86.e5.

Rogers AR, Mukherjee A. 1992. Quantitative Genetics of Sexual Dimorphism in Human Body Size. Evolution 1:226–234.

Rosner B. 2011. Fundamentals of Biostatistics. 7th ed. Boston, MA: Brooks/Cole Cengage Learning.

Rothwell PM. 2000. Analysis of agreement between measurements of continuous variables: general principles and lessons from studies of imaging of carotid stenosis. J Neurol 247:825–834.

Röver C, Raabe N, Luebke K, Ligges U, Szepannek G, Zentgraf M. 2014. klaR: Classification and visualization. Available from: http://cran.r-project.org/web/packages/klaR/index.html

Ruff C, Holt B, Trinkaus E. 2006. Who's afraid of the big bad Wolff?: "Wolff's law" and bone functional adaptation. Am J Phys Anthropol 129:484–498.

Salas A, Carracedo A, Richards M, Macaulay V. 2005. Charting the ancestry of African Americans. Am J Hum Genet 77:676–680.

Salas A, Richards M, Lareu M, Scozzari R, Coppa A, Torroni A, Macaulay V, Carracedo A. 2004. The African diaspora: mitochondrial DNA and the Atlantic slave trade. Am J Hum Genet 74:454–465.

Scheuer L, Black S. 2004. The Juvenile Skeleton. New York, NY: Elsevier Academic Press.

Skedros JG, Hunt KJ, Bloebaum RD. 2004. Relationships of loading history and structural and material characteristics of bone: Development of the mule deer calcaneus. J Morphol 259:281–307.

Sparks CS, Jantz RL. 2002. A reassessment of human cranial plasticity: Boas revisited. Proc Natl Acad Sci 99:14636–14639.

Spradley MK, Jantz RL. 2011. Sex Estimation in Forensic Anthropology: Skull Versus Postcranial Elements. J Forensic Sci 56:289–296.

Spradley MK. 2006. Biological anthropological aspects of the African diaspora: geographic origins, secular trends, and plastic versus genetic influences utilizing craniometric data.

Statistics South Africa. 2011. Census. South Africa. Available from: www.beta2statssa.gov.za

Statistics South Africa. 2013. Mid-year population estimates. South Africa. Available from: www.beta.statssa.gov.za

Steyn M, İşcan MY. 1997. Sex determination from the femur and tibia in South African whites. Forensic Sci Int 90:111–119.

Steyn M, İşcan MY. 1998. Sexual dimorphism in the crania and mandibles of South African whites. Forensic Sci Int 98:9–16.

Steyn M, İşcan MY. 1999. Osteometric variation in the humerus: sexual dimorphism in South Africans. Forensic Sci Int 106:77–85.

Steyn M, Patriquin ML. 2009. Osteometric sex determination from the pelvis—Does population specificity matter? Forensic Sci Int 191:113.e1–113.e5.

Stini WA. 1969. Nutritional stress and growth: sex difference in adaptive response. Am J Phys Anthropol 31:417–426.

Stinson S. 1985. Sex differences in environmental sensitivity during growth and development. Am J Phys Anthropol 28:123–147.

Stull KE. 2013. An Osteometric Evaluation of Age and Sex Differences in the Long Bones of South African Children from the Western Cape.

Stull KE. 2014. An osteometric evaluation of age and sex differences in the long bones of South African children from the Western Cape.

Stynder DD. 2009. Craniometric evidence for South African Later Stone Age herders and hunter–gatherers being a single biological population. J Archaeol Sci 36:798–806.

Tabachnick BG, Fidell LS. 2007. Using multivariate statistics. 5th ed. United States of America: Pearson.

Taylor R. 1990. Interpretation of the correlation coefficient: a basic review. J Diagn Med Sonogr 6:35–39.

Thompson LM. 2001. A History of South Africa. New Haven, CT: Yale University Press.

Tise ML. 2010. A Metric Analysis of the Postcranial Skeleton of Hispanic Individuals to Improve the Estimation of Sex.

Tishkoff SA, Reed FA, Friedlaender FR, Ehret C, Ranciaro A, Froment A, Hirbo JB, Awomoyi AA, Bodo J, Doumbo O. 2009. The genetic structure and history of Africans and African Americans. Science 324:1035–1044.

Tishkoff SA, Williams SM. 2002. Genetic analysis of African populations: human evolution and complex disease. Nat Rev Genet 3:611–621.

Tobias PV. 1971. Growth and Stature in Southern African Populations. In: The human biology of environmental change. Baltyre, Malawi.

Tobias PV. 1975. Stature and secular trend among Southern African Negroes and San (Bushmen). S Afr J Med Sci 40:145–164.

Tobias PV. 1985. The negative secular trend. J Hum Evol 14:347–356.

Trevor JC. 1958. Quantitative traits of the US Negro cranium. The Leech 28:131–138.

U.S. Census Bureau PD. 2014. Annual Estimates of the Resident Population by Sex, Race, and Hispanic Origin for the United States, States, and Counties: April 1, 2010 to July 1, 2013. U.S. Census Bureau, Population Division. Available from: http://factfinder2.census.gov/faces/tableservices/jsf/pages/productview.xhtml?src=bkmk

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

Vance VL, Steyn M, L'Abbé EN, Becker PJ. 2010. A cross-sectional analysis of age related changes in the osteometric dimensions of long bones in modern South Africans of European and African descent. Forensic Sci Int 199:110.e1–110.e9.

Vance VL. 2007. Age related changes in the post-cranial human skeleton and its implication for the determination of sex.

Van Vark GN, Howells WW. 1984. Multivariate Statistical Methods in Physical Anthropology: A Review of Recent Advances and Current Developments. Hingham, MA: Springer.

Veldhuis JD. 2005. Endocrine Control of Body Composition in Infancy, Childhood, and Puberty. Endocr Rev 26:114–146.

Venables W, Ripley B. 2002. Modern Applied Statistics with S. 4th ed. New York, NY: Springer.

Walker PL. 1995. Problems of Preservation and Sexism in Sexing: Some Lessons from Historical Collections for Paleodemographers. In: Saunders SR, Herring A, editors. Grave Reflections: Portraying the Past through Cemetary Studies. Toronto, Canada: Canadian Scholars Press. p 31–47.

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

Weedon MN, Lango H, Lindgren CM, Wallace C, Evans DM, Mangino M, Freathy RM, Perry JRB, Stevens S, Hall AS, Samani NJ, Shields B, others. 2008. Genome-wide association analysis identifies 20 loci that influence adult height. Nat Genet 40:575–583.

Wei T. 2013. corrplot: visualization of a correlation matrix. Available from: http://cran.r-project.org/web/packages/corrplot/index.html

Weihs C, Ligges U, Luebke K, Raabe N. 2005. klaR Analyzing German Business Cycles. In: Baier D, Decker R, Schmidt-Thieme L, editors. Data Analysis and Decision Support. Berlin, Germany: Springer-Verlag. p 335–343.

Weiss LA, Pan L, Abney M, Ober C. 2006. The sex-specific genetic architecture of quantitative traits in humans. Nat Genet 38:218–222.

Wells JCK. 2012. Sexual dimorphism in body composition across human populations: Associations with climate and proxies for short- and long-term energy supply. Am J Hum Biol 24:411–419.

White TD, Folkens eds. 2005. The Human Bone Manual. USA: Elsevier Academic Press.

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

Young B, Lowe JS, Stevens A, Heath JW eds. 2006. Skeletal Tissues. In: Wheater's Functional Histology. London, UK: Churchill Livingston. p 186–206.

Appendix I – Measurement definitions

Maximum length of the clavicle (claxln) - the maximum distance between the most extreme ends of the clavicle (osteometric board).

Antero-posterior diameter of the clavicle at midshaft (claapd) - the antero-posterior (sagittal) distance from the surface of the midshaft (sliding caliper)

Vertical diameter of the clavicle at midshaft (clavrd) - the distance from the cranial to the caudal surface of the midshaft (sliding caliper)

Height of the scapula (scapht) - the direct distance from the most superior point of the superior angle to the most inferior point on the inferior angle (sliding caliper)

Breadth of the scapula (scapbr) - the distance from the midpoint on the dorsal border of the glenoid fossa to midway between the two ridges of the scapular spine on the vertebral border (spreading caliper)

Maximum length of the humerus (humxln) - the direct distance from the most superior point on the head of the humerus to the most inferior point on the trochlea (osteometric board)

Epicondylar breadth of the humerus (humebr) - the distance of the most laterally protruding point on the lateral epicondyle from the corresponding projection of the medial epicondyle (osteometric board or sliding caliper)

Maximum vertical diameter of the head of the humerus (humhdd) - the direct distance between the most superior and inferior points on the border of the articular surface (sliding caliper)

Maximum diameter of the humerus at midshaft (hummxd) - the antero-posterior diameter of the midshaft (sliding caliper)

Minimum diameter of the humerus at midshaft (hummwd) - the distance between the maximum medial and lateral bone surfaces at the midshaft (sliding caliper)

Maximum length of the radius (radxln) - the distance from the most proximally positioned point on the head of the radius to the tip of the styloid process without regard to the long axis of the bone (osteometric board)

Antero-posterior diameter of the radius at midshaft (radapd) - the antero-posterior diameter of the midshaft (sliding caliper)

Transverse diameter of the radius at midshaft (radtvd) - the distance between the maximum medial and lateral bone surfaces at the midshaft (sliding caliper)

Maximum length of the ulna (ulnxln) - the distance between the most superior point on the olecranon and the most inferior point on the styloid process (osteometric board)

Dorso-volar diameter of the ulna (ulndvd) - the maximum diameter of the diaphysis where the crest exhibits the greatest development (sliding caliper)

Transverse diameter of the ulna (ulntvd) - the diameter measured perpendicular to the dorso-volar diameter at the level of greatest crest development (sliding caliper)

Physiological length of the ulna (ulnphl) - the distance between the deepest point on the surface of the coronoid process and the lowest point on the inferior surface of the distal head of the ulna (spreading caliper)

Anterior height of the sacrum (sacaht) - the distance from a point on the promontory in the mid-sagittal plane to a point on the anterior border of the tip of the sacrum (sliding caliper)

Anterior breadth of the sacrum (sacabr) - the maximum transverse breadth of the sacrum at the level of the anterior projection of the auricular surfaces (sliding caliper)

Transverse diameter of sacral segment 1 (sacs1b) - the distance between the two most lateral points on the superior articular surface measured perpendicular to the mid-sagittal plane (sliding caliper)

Height of the innominate (innoht) - the distance from the most superior point on the iliac crest to the most inferior point on the ischial tuberosity (osteometric board or spreading caliper)

Iliac breadth (iliabr) - the distance from the anterior superior iliac spine to the posterior superior iliac spine (spreading caliper)

Maximum length of the femur (femxln) - the distance from the most superior point on the head of the femur to the most inferior point on the distal condyles (osteometric board)

Bicondylar length of the femur (fembln) - the distance from the most superior point on the head of the femur to a plane drawn along the inferior surfaces of the distal condyles (osteometric board)

Epicondylar breadth of the femur (femebr) - the distance between the two most laterally projecting points on the epicondyles (osteometric board)

Maximum diameter of the head of the femur (femhdd) - the maximum diameter of the femur head measured on the border of the articular surface (sliding caliper)

Antero-posterior subtrochanteric diameter of the femur (femsap) - the antero-posterior diameter of the proximal end of the diaphysis measured perpendicular to the transverse diameter at the point of the greatest lateral expansion of the femur below the lesser trochanter (sliding caliper)

Transverse subtrochanteric diameter of the femur (femstv) - the transverse diameter of the proximal portion of the diaphysis at the point of its greatest lateral expansion below the base of the lesser trochanter (sliding caliper)

Antero-posterior diameter of the femur at midshaft (femmap) - the antero-posterior diameter measured approximately at the midpoint of the diaphysis, at the highest elevation of the linea aspera (sliding caliper)

Transverse diameter of the femur at midshaft (femmtv) - the distance between the medial and lateral margins of the femur measured perpendicular to and at the same level as the sagittal diameter (sliding caliper)

Length of the tibia (tibxln) - the distance from the superior articular surface of the lateral condyle of the tibia to the tip of the medial malleolus (osteometric board)

Maximum epiphyseal breadth of the proximal tibia (tibpeb) - the maximum distance between the two most laterally projecting point on the medial and lateral condyles of the proximal epiphysis (osteometric board)

Maximum epiphyseal breadth of the distal tibia (tibdeb) the distance between the most medial point on the medial malleolus and the lateral surface of the distal epiphysis (osteometric board)

Maximum diameter of the tibia at the nutrient foramen (tibnfx) - the distance between the anterior crest and the posterior surface at the level of the nutrient foramen (sliding caliper)

Transverse diameter of the tibia at the nutrient foramen (tibnft) - the straight line distance of the medial margin from the interosseous crest (sliding caliper)

Maximum length of the fibula (fibxln) - the maximum distance between the most superior point on the fibular head and the most inferior point on the lateral malleolus (osteometric board)

Maximum diameter of the fibula at midshaft (fibmdm) - the maximum diameter at the midshaft (sliding caliper)

Maximum length of the calcaneus (calcxl) - the distance between the most posteriorly projecting point on the tuberosity and the most anterior point on the superior margin of the articular facet for the cuboid measured in the sagittal plane and projected onto the underlying surface (sliding caliper)

Middle breadth of the calcaneus (calcbr) - the distance between the most laterally projecting point on the dorsal articular facet and the most medial point on the sustentaculum tali (sliding caliper)

Appendix II – TEM Results for Inter- and Intra-observer Error

Table A2.1 – TEM and %TEM for intra-observer and intra-observer error. Bold indicates the largest errors.

	Intra-observer		Inter-observer 1		Inter-observer 2	
	TEM	*%TEM*	*TEM*	*%TEM*	*TEM*	*%TEM*
claxln	0.00	0.00	0.50	0.34	0.35	0.24
claapd	0.00	0.00	0.50	4.17	0.00	0.00
clavrd	0.00	0.00	0.35	3.67	0.00	0.00
scapht	0.00	0.00	0.71	0.50	0.00	0.00
scapbr	0.35	0.37	0.50	0.52	0.61	0.64
humxln	0.00	0.00	0.45	0.15	0.32	0.10
humebr	0.32	0.56	0.32	0.56	0.00	0.00
humhdd	0.00	0.00	0.55	1.36	0.45	1.10
hummxd	0.00	0.00	0.00	0.00	0.00	0.00
hummwd	0.00	0.00	0.32	2.07	0.32	2.07
radxln	0.00	0.00	0.50	0.21	0.61	0.26
radapd	0.00	0.00	0.45	4.07	0.45	3.99
radtvd	0.00	0.00	0.32	2.34	0.45	3.29
ulnxln	0.35	0.14	0.35	0.14	0.35	0.14
ulndvd	0.32	**2.24**	0.77	5.74	0.55	3.94
ulntvd	0.00	0.00	1.00	**6.99**	0.63	**4.65**
ulnphl	0.71	0.32	0.00	0.00	0.61	0.27
sacaht	0.00	0.00	0.50	0.45	0.50	0.45
sacabr	0.00	0.00	0.87	0.92	**0.94**	0.99
sacs1b	0.00	0.00	0.91	2.11	0.41	0.94
innoht	0.00	0.00	0.84	0.44	0.55	0.29
iliabr	0.32	0.22	0.95	0.66	0.63	0.44
femxln	0.00	0.00	0.45	0.10	0.45	0.10
fembln	0.00	0.00	0.45	0.10	0.45	0.10
femebr	0.00	0.00	0.41	0.55	0.63	0.87
femhdd	0.00	0.00	0.35	0.84	0.32	0.76
femsap	0.32	1.25	1.05	4.23	0.55	2.18
femstv	0.32	1.06	**1.22**	4.22	0.71	2.41
femmap	0.00	0.00	0.45	1.62	0.32	1.15
femmtv	0.00	0.00	0.55	2.27	0.55	2.27
tibxln	0.58	0.15	0.41	0.11	0.58	0.15
tibpeb	0.00	0.00	0.91	1.28	0.50	0.70
tibdeb	0.00	0.00	0.79	1.76	0.35	0.79
tibnfx	0.00	0.00	0.45	1.38	0.32	0.97
tibnft	0.32	1.39	0.45	1.98	0.45	1.96
fibxln	**0.82**	0.22	0.41	0.11	0.58	0.16
fibmdm	0.00	0.00	0.35	2.50	0.00	0.00
calcxl	0.00	0.00	0.71	0.90	0.71	0.90
calcbr	0.50	1.24	0.00	0.00	0.50	1.24

Appendix III – Correlation matrix

Table A3.1 – Coefficients of correlations. Sample sizes ranged from 145 to 357. Correlations that were not significant (p > 0.05) are in bold.

	claxln																			
claxln	1	clavrd																		
clavrd	0.45	1	claapd																	
claapd	0.58	0.52	1	scapht																
scapht	0.77	0.59	0.65	1	scapbr															
scapbr	0.76	0.53	0.6	0.79	1	humxln														
humxln	0.77	0.47	0.54	0.78	0.79	1	humebr													
humebr	0.74	0.61	0.62	0.82	0.77	0.74	1	humhdd												
humhdd	0.7	0.57	0.57	0.79	0.75	0.78	0.79	1	hummxd											
hummxd	0.65	0.62	0.65	0.73	0.69	0.64	0.75	0.7	1	hummwd										
hummwd	0.64	0.62	0.69	0.72	0.69	0.61	0.76	0.71	0.81	1	radxln									
radxln	0.75	0.48	0.57	0.74	0.76	0.85	0.75	0.65	0.62	0.6	1	radapd								
radapd	0.67	0.61	0.71	0.72	0.7	0.66	0.72	0.65	0.77	0.79	0.68	1	radtvd							
radtvd	0.61	0.53	0.55	0.65	0.63	0.62	0.71	0.68	0.77	0.77	0.58	0.73	1	ulnxln						
ulnxln	0.67	0.61	0.71	0.72	0.7	0.82	0.72	0.65	0.77	0.79	0.98	0.67	0.56	1	ulndvd					
ulndvd	0.54	0.57	0.6	0.63	0.62	0.54	0.66	0.57	0.69	0.73	0.57	0.69	0.61	0.58	1	ulntvd				
ulntvd	0.67	0.61	0.71	0.72	0.7	0.66	0.72	0.65	0.77	0.79	0.63	0.75	0.66	0.62	0.64	1	ulnphl			
ulnphl	0.72	0.46	0.54	0.69	0.73	0.81	0.7	0.6	0.58	0.56	0.98	0.65	0.53	0.99	0.55	0.59	1	sacaht		
sacaht	0.35	0.29	0.11	0.35	0.33	0.4	0.31	0.4	0.26	0.25	0.31	0.26	0.35	0.29	0.2	0.25	0.28	1	sacabr	
sacabr	**0.05**	**-0.08**	**-0.13**	0.15	0.15	0.27	**0.06**	0.23	**0.03**	**-0.05**	**0.06**	**-0.07**	**0.1**	**0.07**	**-0.06**	**0.05**	**0.04**	0.34	1	sacs1b
sacs1b	0.48	0.42	0.44	0.51	0.51	0.53	0.56	0.51	0.48	0.5	0.57	0.51	0.48	0.57	0.47	0.55	0.56	0.3	**0.11**	1

Table A3.1 cont'd – Coefficients of correlations. Sample sizes ranged from 145 to 357. Correlations that were not significant (p > 0.05) are in bold.

	claxln	clavrd	claapd	scapht	scapbr	humxln	humebr	humhdd	hummxd	hummwd	radxln	radapd	radtvd	ulnxln	ulndvd	ulntvd	ulnphl	sacaht	sacabr	sacs1b
innoht	0.73	0.46	0.46	0.78	0.79	0.83	0.73	0.83	0.65	0.61	0.66	0.58	0.63	0.65	0.49	0.61	0.61	0.45	0.42	0.54
iliabr	0.5	0.26	0.24	0.63	0.57	0.65	0.53	0.68	0.46	0.4	0.42	0.33	0.47	0.41	0.29	0.41	0.37	0.4	0.54	0.39
femxln	0.75	0.43	0.51	0.74	0.75	0.91	0.74	0.74	0.6	0.58	0.84	0.63	0.59	0.81	0.52	0.62	0.8	0.36	0.26	0.52
fembln	0.75	0.44	0.51	0.74	0.75	0.92	0.74	0.74	0.6	0.58	0.84	0.63	0.59	0.81	0.51	0.62	0.8	0.36	0.26	0.51
femebr	0.74	0.59	0.61	0.81	0.78	0.78	0.84	0.88	0.72	0.74	0.71	0.69	0.69	0.7	0.61	0.7	0.66	0.38	0.13	0.59
femhdd	0.73	0.6	0.62	0.79	0.79	0.79	0.83	0.88	0.72	0.73	0.73	0.72	0.7	0.71	0.63	0.71	0.67	0.38	0.13	0.55
femsap	0.66	0.56	0.55	0.67	0.67	0.69	0.67	0.71	0.71	0.68	0.6	0.64	0.68	0.58	0.55	0.65	0.55	0.34	0.22	0.43
femstv	0.67	0.61	0.71	0.72	0.7	0.66	0.72	0.65	0.77	0.79	0.62	0.69	0.66	0.61	0.62	0.66	0.57	0.21	0.14	0.42
femmap	0.66	0.51	0.56	0.71	0.71	0.66	0.69	0.67	0.76	0.72	0.65	0.67	0.67	0.65	0.62	0.67	0.64	0.31	**0.06**	0.42
femmtv	0.67	0.61	0.71	0.72	0.7	0.66	0.72	0.65	0.77	0.79	0.68	0.69	0.65	0.67	0.57	0.66	0.64	0.32	**0.11**	0.49
tibxln	0.71	0.39	0.47	0.65	0.71	0.84	0.68	0.61	0.56	0.53	0.88	0.59	0.52	0.87	0.48	0.56	0.87	0.32	0.17	0.51
tibpeb	0.74	0.58	0.66	0.8	0.79	0.76	0.84	0.85	0.74	0.75	0.76	0.72	0.7	0.76	0.64	0.71	0.71	0.37	**0.05**	0.57
tibdeb	0.7	0.59	0.54	0.78	0.73	0.77	0.8	0.85	0.69	0.68	0.67	0.64	0.68	0.66	0.53	0.68	0.62	0.36	0.23	0.6
tibnfx	0.64	0.56	0.63	0.71	0.67	0.68	0.68	0.71	0.72	0.72	0.68	0.69	0.65	0.67	0.57	0.66	0.64	0.29	**0.05**	0.5
tibnft	0.62	0.52	0.61	0.66	0.63	0.65	0.67	0.66	0.71	0.72	0.63	0.67	0.71	0.63	0.57	0.65	0.61	0.3	**0.08**	0.49
fibxln	0.73	0.41	0.51	0.69	0.74	0.85	0.71	0.62	0.57	0.52	0.9	0.6	0.54	0.9	0.5	0.56	0.89	0.34	0.19	0.53
fibmdm	0.5	0.37	0.41	0.67	0.61	0.71	0.72	0.7	0.66	0.72	0.65	0.77	0.79	0.5	0.44	0.51	0.48	0.22	0.14	0.41
calcxl	0.65	0.48	0.5	0.77	0.74	0.76	0.74	0.74	0.64	0.63	0.66	0.62	0.59	0.65	0.56	0.61	0.62	0.41	0.26	0.34
calcbr	0.4	0.4	0.44	0.44	0.5	0.43	0.54	0.49	0.5	0.49	0.49	0.48	0.45	0.5	0.44	0.5	0.45	**0.13**	**-0.06**	0.33

Table A3.1 cont'd – Coefficients of correlations. Sample sizes ranged from 145 to 357. Correlations that were not significant (p > 0.05) are in bold.

	innoht	iliabr	femxln	fembln	femebr	femhdd	femsap	femstv	femmap	femmtv	tibxln	tibpeb	tibdeb	tibnfx	tibnft	fibxln	fibmdm	calcxl	calcbr
innoht	1																		
iliabr	0.84	1																	
femxln	0.8	0.62	1																
fembln	0.8	0.62	0.99	1															
femebr	0.82	0.66	0.77	0.77	1														
femhdd	0.8	0.6	0.77	0.78	0.9	1													
femsap	0.76	0.62	0.68	0.67	0.77	0.74	1												
femstv	0.68	0.52	0.62	0.62	0.71	0.75	0.67	1											
femmap	0.66	0.47	0.68	0.68	0.73	0.71	0.7	0.64	1										
femmtv	0.74	0.56	0.66	0.67	0.76	0.74	0.69	0.79	0.63	1									
tibxln	0.69	0.49	0.89	0.89	0.67	0.67	0.58	0.56	0.65	0.63	1								
tibpeb	0.77	0.57	0.75	0.75	0.92	0.88	0.72	0.7	0.72	0.75	0.7	1							
tibdeb	0.83	0.7	0.75	0.76	0.9	0.86	0.76	0.67	0.68	0.72	0.67	0.86	1						
tibnfx	0.68	0.5	0.66	0.66	0.75	0.71	0.71	0.68	0.75	0.69	0.62	0.75	0.71	1					
tibnft	0.65	0.48	0.64	0.64	0.72	0.67	0.69	0.63	0.74	0.65	0.6	0.7	0.69	0.78	1				
fibxln	0.71	0.52	0.9	0.9	0.71	0.69	0.61	0.58	0.66	0.67	0.96	0.74	0.69	0.63	0.61	1			
fibmdm	0.56	0.44	0.47	0.47	0.54	0.5	0.51	0.48	0.54	0.54	0.47	0.57	0.49	0.58	0.59	0.49	1		
calcxl	0.76	0.62	0.71	0.72	0.76	0.77	0.63	0.71	0.59	0.7	0.68	0.79	0.73	0.66	0.55	0.69	0.49	1	
calcbr	0.4	0.24	0.43	0.44	0.57	0.6	0.48	0.54	0.42	0.47	0.45	0.6	0.51	0.49	0.44	0.48	0.32	0.7	1

Appendix IV – ANOVA

Table A4.1 – Analysis of variance (ANOVA) results evaluating the statistical significance of sex, ancestry and the interaction of sex and ancestry for each measurement. Bold indicates significance.

		Sex		Ancestry		Sex*Ancestry	
		F-value	*Pr(>F)*	*F-value*	*Pr(>F)*	*F-value*	*Pr(>F)*
clavicle	**claxln**	196.55	**<0.001**	19.54	**<0.001**	0.28	0.75
	claapd	219.43	**<0.001**	0.73	0.48	1.81	0.17
	clavrd	152.76	**<0.001**	1.73	0.18	0.79	0.46
scapula	**scapht**	305.13	**<0.001**	46.35	**<0.001**	0.76	0.47
	scapbr	334.17	**<0.001**	33.27	**<0.001**	0.72	0.49
humerus	**humxln**	219.71	**<0.001**	63.7	**<0.001**	0.04	0.97
	humebr	434.72	**<0.001**	35.48	**<0.001**	2.02	0.13
	humhdd	442.46	**<0.001**	142.56	**<0.001**	1.41	0.24
	hummxd	234.2	**<0.001**	15.8	**<0.001**	0.2	0.82
	hummwd	302.96	**<0.001**	11.23	**<0.001**	1.71	0.18
radius	**radxln**	319.37	**<0.001**	39.36	**<0.001**	0.23	0.8
	radapd	475.45	**<0.001**	4.36	**0.013**	0.11	0.9
	radtvd	201.97	**<0.001**	22.8	**<0.001**	1.39	0.25
ulna	**ulnxln**	285.9	**<0.001**	41.83	**<0.001**	0.07	0.93
	ulndvd	225.6	**<0.001**	1.28	0.28	2.21	0.11
	ulntvd	347.59	**<0.001**	16.21	**<0.001**	0.21	0.81
	ulnphl	264.77	**<0.001**	39.23	**<0.001**	1.01	0.36
sacrum	**sacaht**	8.11	**0.0049**	11.2	**<0.001**	1.01	0.37
	sacabr	24.9	**<0.001**	61.72	**<0.001**	0.59	0.56
	sacs1b	66.88	**<0.001**	27.79	**<0.001**	0.63	0.53
innominate	**innoht**	198.54	**<0.001**	70.17	**<0.001**	0.11	0.89
	iliabr	27.08	**<0.001**	211.59	**<0.001**	4.29	**0.015**
femur	**femxln**	199.9	**<0.001**	50.43	**<0.001**	0.14	0.87
	fembln	201.37	**<0.001**	50.1	**<0.001**	0.16	0.85
	femebr	387.28	**<0.001**	74.67	**<0.001**	0.47	0.62
	femhdd	445.08	**<0.001**	45.76	**<0.001**	1.37	0.26
	femsap	146.3	**<0.001**	44.85	**<0.001**	1.41	0.25
	femstv	208.55	**<0.001**	20.63	**<0.001**	2.81	0.06
	femmap	175.49	**<0.001**	24.3	**<0.001**	0.46	0.63
	femmtv	200.89	**<0.001**	41.91	**<0.001**	0.11	0.89
tibia	**tibxln**	148.41	**<0.001**	30.92	**<0.001**	0.19	0.83
	tibpeb	439.35	**<0.001**	48.96	**<0.001**	0.12	0.88
	tibdeb	239.23	**<0.001**	87.84	**<0.001**	0.47	0.62
	tibnfx	197.24	**<0.001**	20.02	**<0.001**	2.46	0.87
	tibnft	160.27	**<0.001**	23.48	**<0.001**	0.31	0.74
fibula	**fibxln**	157.83	**<0.001**	35.09	**<0.001**	0.56	0.57
	fibmdm	61.42	**<0.001**	16.26	**<0.001**	0.17	0.85
calcaneus	**calcbr**	122.71	**<0.001**	1.59	0.206	4.78	**0.009**

Appendix V – Classification accuracies and sex bias when assessing sex using LDA

Table A5.1 – Classification accuracies when LDA used bone models (stepwise selected) to estimate sex.

	Accuracy (%)			Sex Bias (%)
	Males	*Females*	*Combined*	
clavicle	86	90	88	4
scapula	79	84	82	5
humerus	84	84	84	0
radius	89	90	90	1
ulna	86	85	86	-1
sacrum	76	74	75	-2
innominate	73	82	78	9
femur	85	88	86	3
tibia	84	88	86	4
fibula	75	75	75	0

Table A5.2 – Classification accuracies when LDA used multivariate subsets (stepwise selected) to estimate sex.

	Accuracy (%)			Sex Bias (%)
	Males	*Females*	*Combined*	
shoulder	87	95	91	8
hip	98	94	96	-4
upper	93	97	95	4
lower	99	95	97	-2
breadths	97	99	98	2
lengths	88	92	90	4
proxdist	84	91	87	7
all-variable model	97	97	97	0

Appendix VI – Classification accuracies and sex bias when assessing sex using FDA

Table A6.1 – Classification accuracies when FDA used bone models to estimate sex.

	Accuracy (%)			Sex Bias (%)
	Males	*Females*	*Combined*	
clavicle	93	90	91	-3
scapula	84	87	86	3
humerus	88	87	88	-1
radius	89	90	90	1
ulna	89	89	89	0
sacrum	74	79	76	5
innominate	80	77	79	-3
femur	87	86	87	-1
tibia	82	90	86	8
fibula	73	77	75	4

Table A6.2 – Classification accuracies when FDA used multivariate subsets to estimate sex.

	Accuracy (%)			Sex Bias (%)
	Males	*Females*	*Combined*	
shoulder	93	93	93	0
hip	94	96	95	2
upper	95	96	96	1
lower	96	97	97	1
breadths	98	95	96	-3
lengths	96	91	93	-5
proxdist	89	90	90	1
all-variable	97	97	97	0

Appendix VII – Classification accuracies and sex bias when assessing sex using logistic regression

Table A7.1 – Classification accuracies when logistic regression used bone models to estimate sex.

	Accuracy (%)			Sex Bias (%)
	Males	*Females*	*Combined*	
clavicle	91	91	91	0
scapula	83	84	83	1
humerus	87	85	86	-2
radius	91	91	90	0
ulna	88	88	87	0
sacrum	76	75	74	-1
innominate	81	77	79	-4
femur	84	87	84	3
tibia	86	88	87	2
fibula	73	76	75	3

Table A7.2 – Classification accuracies when logistic regression used multivariate subsets to estimate sex.

	Accuracy (%)			Sex Bias (%)
	Males	*Females*	*Combined*	
shoulder	93	93	92	0
hip	94	96	95	2
upper	96	94	94	-2
lower	97	95	95	-2
breadths	95	96	94	1
lengths	93	98	91	5
proxdist	89	89	88	0
all-variable	97	97	97	0

Appendix VIII – Mean differences between North Americans and South Africans

Table A8.1 – Mean differences between North American and South African black and white females. Significance differences are noted in bold.

	SAB-NAB difference	p-value	SAB-NAW difference	p-value	SAW-NAW difference	p-value	SAW-NAB difference	p-value
claxln	-2	0.55	1	0.8	4	0.05	1	0.99
claapd	-1.1	**<0.01**	-0.1	1	0.4	0.42	-0.6	0.08
clavrd	-0.1	0.98	0	1	0.1	0.99	0	1
scapht	-6	**<0.01**	-7	**<0.001**	3	0.37	4	0.16
scapbr	-3.8	**<0.01**	-3.3	**0.01**	1.8	0.31	1.3	0.62
humxln	-15	**<0.001**	-8	**<0.01**	8	**0.04**	1	0.99
humebr	-0.3	0.97	-0.1	1	1.4	0.09	1.2	0.2
humhdd	-1.9	**<0.001**	-3.6	**<0.001**	1.2	**0.03**	2.9	**<0.001**
hummxd	-0.6	0.26	0.1	0.99	0.9	**0.04**	0.2	0.91
hummwd	-0.8	**0.03**	0.1	0.97	0.5	**<0.01**	-0.4	0.55
radxln	-8	**<0.01**	3	0.54	2	0.76	-9	**<0.01**
radapd	-0.5	**0.02**	0.2	0.53	0.3	0.32	-0.4	0.07
radtvd	-0.7	0.05	-0.2	0.84	0.4	0.49	-0.1	0.99
ulnxln	-6	0.05	6	0.19	3	0.69	-9	**<0.01**
ulndvd	1.8	**<0.001**	0.8	0.13	1	**0.02**	2	**<0.001**
ulntvd	-1.6	**<0.001**	0.4	0.69	1.1	**<0.01**	-0.9	**0.02**
ulnphl	-11	**<0.001**	1	0.99	-1	0.97	-13	**<0.001**
sacaht	-4.7	0.15	-9.7	**<0.001**	-2	0.89	3	0.67
sacabr	-7.7	**<0.001**	-17.1	**<0.001**	-8	**<0.001**	1.4	0.69
sacs1b	1.1	0.53	-1.1	0.57	-0.5	**0.02**	1.7	0.14
innoht	-1	0.95	-13	**<0.001**	7	**<0.01**	19	**<0.001**
iliabr	-6	**<0.01**	-14	**<0.001**	7	**<0.001**	15	**<0.001**
femxln	-20	**<0.001**	-11	**0.03**	5	0.58	-4	0.71
fembln	-19	**<0.001**	-11	**0.04**	6	0.4	-2	0.94
femebr	-2.4	**<0.01**	-3.2	**<0.001**	1.1	0.39	1.9	**0.02**
femhdd	-2	**<0.001**	-2	**<0.001**	0.8	0.17	0.8	0.14
femsap	-0.4	0.85	-0.5	0.69	1.1	0.09	1.2	**0.047**
femstv	-0.8	0.21	0.2	0.95	2.4	**<0.001**	1.4	**<0.01**
femmap	-1.1	**0.02**	0	1	1.1	0.06	0	1
femmtv	-0.2	0.94	0.3	0.92	1.8	**<0.001**	1.3	**<0.01**
tibxln	-10	**0.03**	4	0.8	8	0.22	-6	0.3
tibpeb	-2.3	**<0.01**	-2.1	**<0.01**	0.5	0.83	0.3	0.94
tibdeb	-0.5	0.76	-2.1	**<0.001**	1.1	0.13	2.7	**<0.001**
tibnfx	-0.6	0.58	0.2	0.98	2.1	**<0.001**	1.3	**0.02**
tibnft	-0.9	0.18	0.9	0.16	1.6	**<0.01**	-0.2	0.97
fibxln	-12	**<0.01**	3	0.9	8	0.25	-7	0.26
fibmdm	0.1	0.99	0.2	0.89	0.9	**0.02**	0.8	0.07
calcxl	-2.3	**0.04**	-4	**<0.001**	1.7	0.26	3.4	**<0.01**
calcbr	-1.1	0.17	0.1	0.99	1	0.23	-0.2	0.97

Table A8.2 – Mean differences between North American and South African black and white males. Significance differences are noted in bold.

	SAB-NAB difference	p-value	SAB-NAW difference	p-value	SAW-NAW difference	p-value	SAW-NAB difference	p-value
claxln	-2	0.61	-2	0.66	2	0.66	2	0.71
claapd	-0.8	0.1	0.7	0.18	0.7	0.17	-0.8	0.07
clavrd	-0.1	0.98	-0.6	0.33	-0.1	0.99	0.4	0.63
scapht	-8	**<0.001**	-10	**<0.001**	-1	0.96	1	0.95
scapbr	-6	**<0.001**	-4	**0.01**	0	1	-2	0.57
humxln	-18	**<0.001**	-12	**<0.01**	3	0.78	-3	0.8
humebr	-1.6	0.07	-3.1	0.07	-0.4	0.9	1.1	0.33
humhdd	-2.3	**<0.001**	-4.5	**<0.001**	0	1	2.2	**<0.001**
hummxd	-1.2	**<0.01**	-0.9	0.05	0.1	0.99	-0.1	0.96
hummwd	-1.8	**<0.001**	-1	**0.03**	0.2	0.93	-0.6	0.32
radxln	-10	**<0.01**	3	0.62	0	1	-13	**<0.001**
radapd	-0.6	**0.04**	-0.3	0.58	-0.1	0.97	-0.4	0.25
radtvd	-0.4	0.56	-1	**0.01**	0.1	0.96	0.7	0.12
ulnxln	-9	**<0.01**	3	0.72	-1	0.99	-13	**<0.001**
ulndvd	0.3	0.7	0.6	0.22	1	**0.01**	0.7	0.13
ulntvd	-1.4	**<0.01**	-0.7	0.25	-0.1	0.97	-0.8	0.11
ulnphl	-11	**<0.001**	1	0.97	-6	0.12	-18	**<0.001**
sacaht	4	0.35	-6	**0.03**	-2	0.79	8	**0.02**
sacabr	-6.5	**<0.001**	-18.2	**<0.001**	-8.2	**<0.001**	3.5	0.11
sacs1b	-0.8	0.82	-2.6	**0.02**	-2.9	**<0.01**	-1.1	0.64
innoht	-3	0.51	-18	**<0.001**	1	0.96	16	**<0.001**
iliabr	-4	0.07	-11	**<0.001**	4	0.13	11	**<0.001**
femxln	-28	**<0.001**	-18	**<0.001**	-4	0.81	-14	**0.01**
fembln	-26	**<0.001**	-18	**<0.001**	-4	0.86	-12	**0.03**
femebr	-3.2	**<0.001**	-6.6	**<0.001**	-2.5	**0.01**	0.9	0.74
femhdd	-1.2	0.07	-2.7	**<0.001**	-0.7	0.53	0.8	0.47
femsap	-0.6	0.55	-1.3	**0.02**	1.1	0.07	1.8	**<0.001**
femstv	-0.5	0.74	-0.7	0.52	0.7	0.49	0.9	0.31
femmap	-0.5	0.73	-0.9	0.2	0.2	0.96	0.6	0.52
femmtv	-0.4	0.64	-0.4	0.72	1	**0.04**	1	0.07
tibxln	-16	**<0.01**	-3	0.89	-2	0.96	-15	**<0.01**
tibpeb	-2.9	**<0.001**	-3.2	**<0.001**	-0.9	0.62	-0.6	0.9
tibdeb	-1.6	0.06	-3.1	**<0.001**	-0.1	0.99	1.4	0.1
tibnfx	-0.5	0.8	-0.9	0.44	-0.2	0.99	0.2	0.99
tibnft	-0.5	0.78	0.1	0.99	1.3	0.06	0.7	0.51
fibxln	-13	**0.01**	-4	0.74	-4	0.69	-13	**<0.01**
fibmdm	0.9	0.07	-0.3	0.78	0.1	0.99	1.3	**<0.01**
calcxl	-1.3	0.57	-4.3	**<0.001**	-1.2	0.63	1.8	0.25
calcbr	-0.5	0.77	-0.9	0.37	-0.7	0.63	-0.3	0.95

Appendix IX – Sympercents in North Americans

Table A9.1 – Comparison of s% of North American black and white males and females

	NA black females	*NA black males*	*s%*	*NA white females*	*NA white males*	*s%*
claxln	143	157	4.02	140	157	**5.10**
claapd	12	14	**8.06**	11	12	7.04
clavrd	9	11	7.88	9	11	**9.77**
scapht	142	162	5.88	142	164	**6.14**
scapbr	96	110	**5.77**	95	108	5.60
humxln	312	341	3.82	305	335	**4.00**
humebr	56	64	5.65	56	65	**6.87**
humhdd	41	46	5.69	42	49	**6.25**
hummxd	20	24	6.46	20	23	**7.28**
hummwd	16	20	8.26	15	19	**8.98**
radxln	239	267	4.79	228	254	**4.80**
radapd	11	13	8.02	10	13	**10.00**
radtvd	14	16	4.97	14	16	**7.96**
ulnxln	256	285	4.67	245	273	**4.76**
ulndvd	12	15	**11.62**	13	15	7.11
ulntvd	15	17	6.88	13	17	**11.43**
ulnphl	229	255	**4.60**	218	242	4.60
sacaht	102	100	-0.86	107	110	**1.16**
sacabr	100	96	**-1.68**	109	107	-0.85
sacs1b	43	49	**5.19**	45	50	4.60
innoht	190	208	3.97	202	223	**4.27**
iliabr	148	154	**1.67**	156	161	1.26
femxln	445	485	3.65	436	475	**3.74**
fembln	440	480	3.80	432	472	**3.86**
femebr	74	83	4.82	75	86	**6.03**
femhdd	42	47	4.95	42	49	**6.21**
femsap	26	28	4.08	26	29	**4.92**
femstv	29	32	4.42	28	32	**6.16**
femmap	28	30	3.63	27	31	**5.96**
femmtv	25	28	5.23	24	28	**5.95**
tibxln	372	406	3.87	357	393	**4.12**
tibpeb	70	79	5.17	70	80	**5.49**
tibdeb	45	50	**4.96**	46	52	4.81
tibnfx	32	37	5.78	31	37	**7.19**
tibnft	24	26	4.36	22	26	**6.80**
fibxln	365	396	3.50	350	387	**4.36**
fibmdm	14	15	1.97	14	16	**5.82**
calcxl	76	84	3.96	78	87	**4.53**
calcbr	40	44	3.91	39	44	**5.58**

Appendix X – Sympercents of North Americans and South Africans

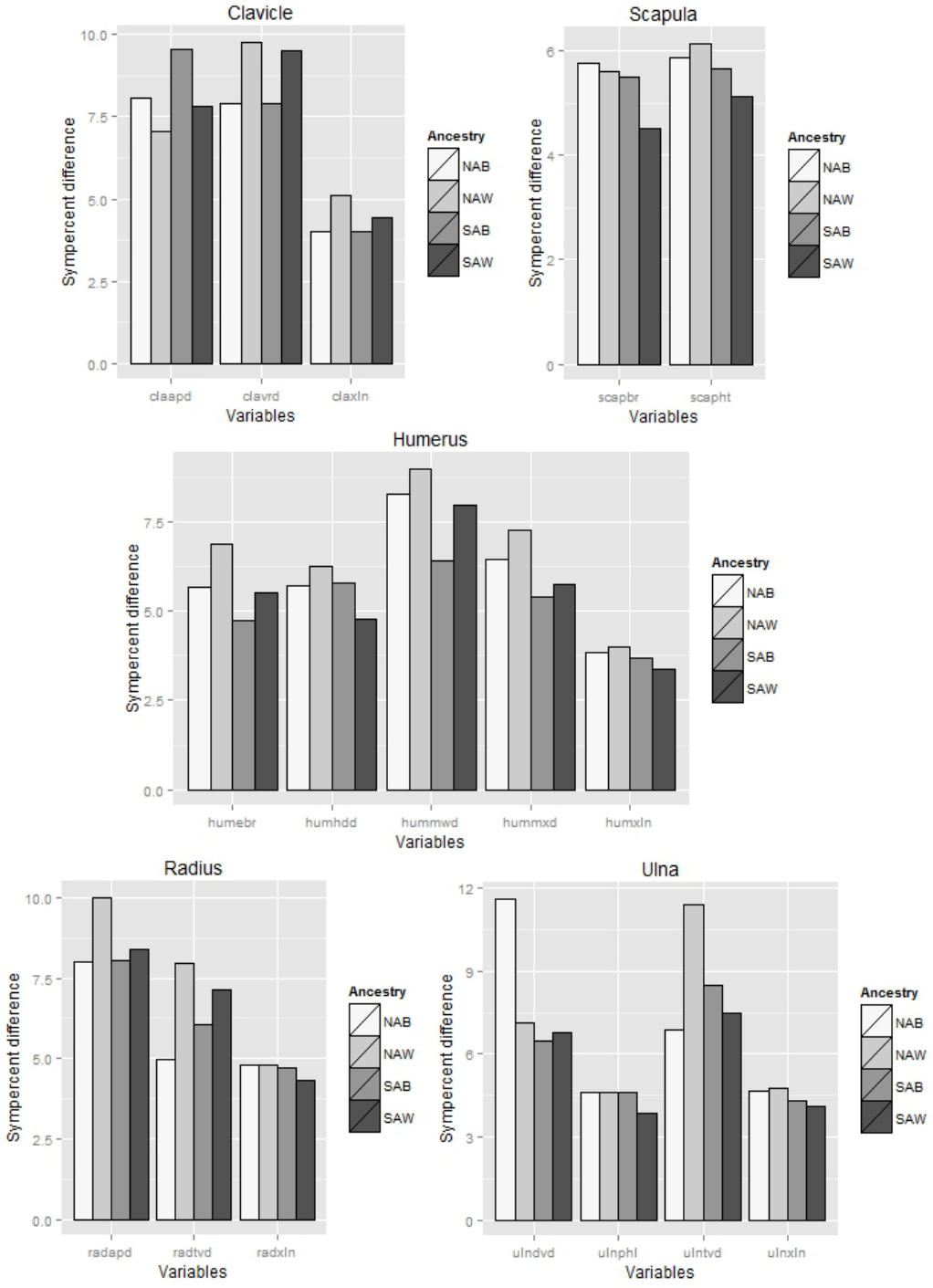

Figure A10.1 – Differences in sexual dimorphism (sympercents) between North American and South African black and white males and females for the clavicle, scapula, humerus, radius and ulna measurements.

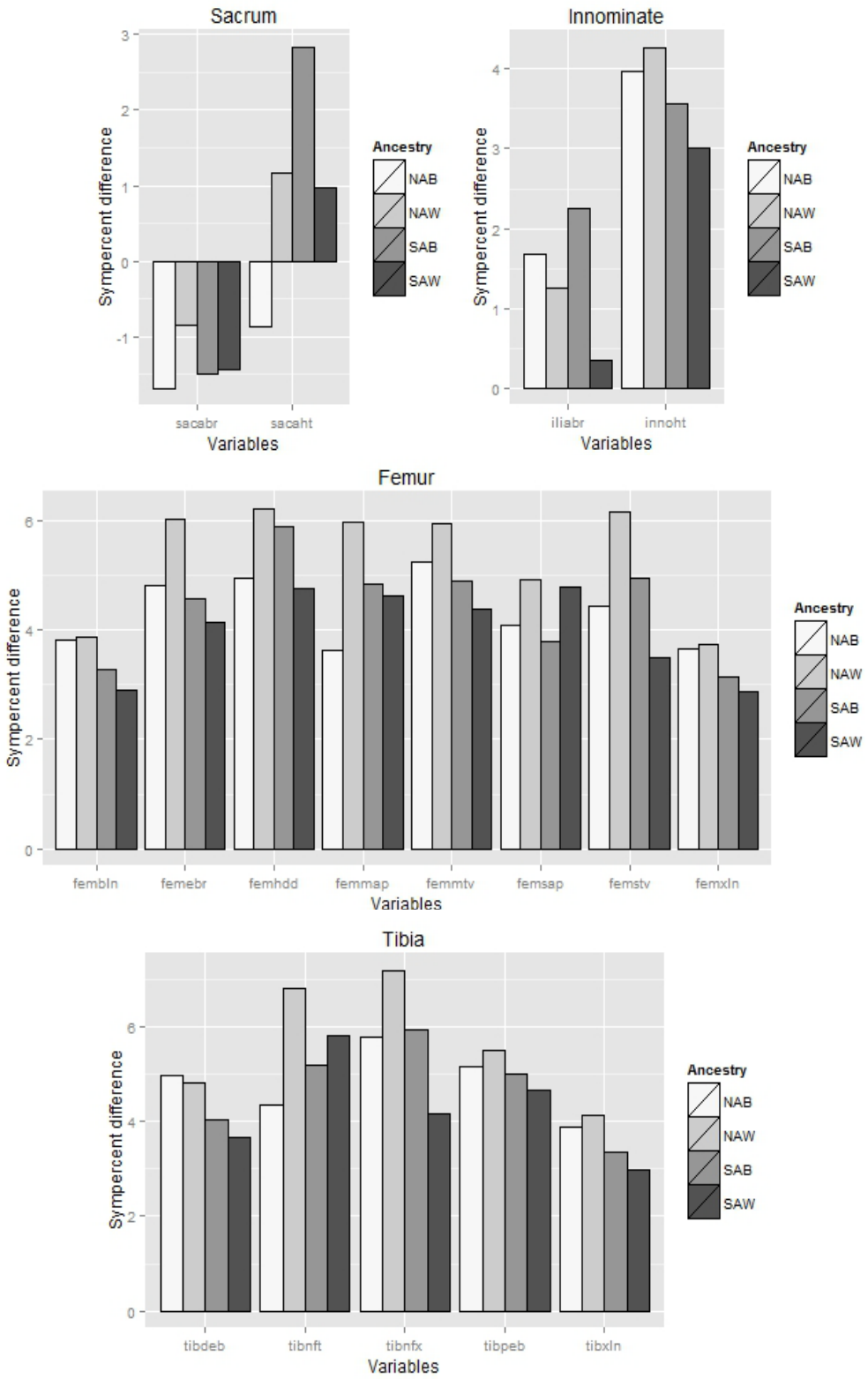

Figure A10.2 – Differences in sexual dimorphism (sympercents) between North American and South African black and white males and females for the sacrum, innominate, femur and tibia measurements.

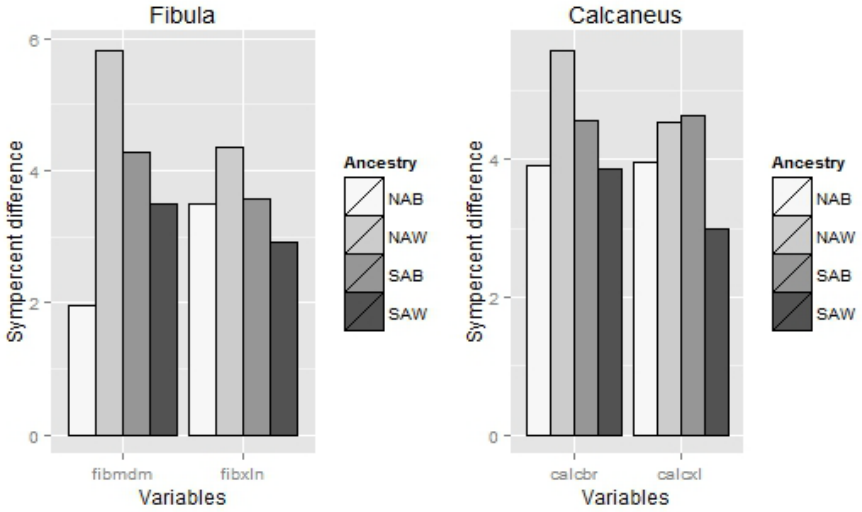

Figure A10.3 – Differences in sexual dimorphism (sympercents) between North American and South African black and white males and females for the fibula and calcaneus measurements.

Appendix XI – Classification accuracies for North Americans and South Africans when classified according to sex and ancestry using LDA and FDA

Table A11.1 – Multivariate subsets for LDA when classifying North Americans and South Africans according to sex and ancestry

	Accuracy (%)								
	NA back female	*NA black male*	*NA white female*	*NA white male*	*SA black female*	*SA black male*	*SA white female*	*SA white male*	*Combined*
shoulder	35	61	46	40	72	36	36	54	46
hip	63	59	64	60	54	66	71	67	62
upper	70	66	49	66	79	61	61	73	64
lower	64	67	62	67	64	74	68	64	66
breadths	58	38	57	69	73	64	65	63	59
lengths	57	54	65	51	68	30	55	75	55
proxdist	28	43	28	56	60	44	45	26	41
all-variable	69	72	72	76	79	59	79	77	72

Table A11.2 – Multivariate subsets for FDA when classifying North Americans and South Africans according to sex and ancestry

	Accuracy (%)								
	SA back female	*SA black male*	*NA black female*	*NA black male*	*NA white female*	*NA white male*	*SA white female*	*SA white male*	*Combined*
shoulder	48	73	75	48	54	36	38	45	52
hip	58	68	91	78	51	63	44	48	63
upper	76	74	75	74	73	46	74	52	69
lower	61	78	83	89	62	71	56	62	71
breadths	80	78	83	82	59	55	53	33	67
lengths	72	74	75	83	61	57	42	42	66
proxdist	44	49	39	62	69	43	48	50	50
all-variable	80	88	93	90	89	85	77	81	86