Postcraniometric analysis of ancestry among modern South Africans

Leandi Liebenberg

Submitted in fulfilment of the requirements for the degree

MSc (Anatomy)

In the School of Medicine, Faculty of Health Sciences, University of Pretoria

2014

DECLARATION

I, Leandi Liebenberg, declare that this thesis is my own work. It is being submitted for the degree of Master of Science in Anatomy at the University of Pretoria. It has not been submitted before for any other degree or examination at this or any other University.

Sign_____

This _____ day of _____ 2014.

ABSTRACT

The primary role of a physical anthropologist is to provide sufficient information to assist in the individualisation of unknown skeletal remains. This is often achieved in establishing a biological profile of the deceased, of which ancestry is an essential aspect. Several successful osteometric and morphological approaches have been developed to facilitate the estimation of ancestry from the cranium. However, the cranium is not always available for analysis, emphasising a need for postcranial alternatives. The postcranial skeleton is frequently labelled as too variable and unreliable to provide an accurate assessment of ancestry. Yet, numerous studies utilise the postcrania for sex and stature estimation, where the a priori knowledge of ancestry results in higher accuracy. Thus, the presence of postcranial differences among populations when investigating other biological parameters inherently demonstrates the potential for the estimation of ancestry. The purpose of this study was to quantify postcranial variation among modern, peer-reported black, white and coloured South Africans. A series of 39 standard measurements were taken from 11 postcranial bones, namely the clavicle, scapula, humerus, radius, ulna, sacrum, pelvis, femur, tibia, fibula and calcaneus. The sample consisted of 360 modern South African individuals (120 black, 120 white, 120 coloured) from the Pretoria Bone and Kirsten Collections housed at the University of Pretoria and the University of Stellenbosch, respectively. Group differences were explored with ANOVA and Tukey's honestly significant difference test (HSD). Group means were used to create univariate sectioning points for each variable indicated as significant with ANOVA. Where two of the three groups had similar mean values, the groups were pooled for the creation of the sectioning points. Multivariate classification models were employed using linear and flexible discriminant analysis (LDA and FDA, respectively). Classification accuracies were compared to evaluate which model yielded the best results.

The results demonstrated variable patterns of group overlap. Black and coloured South Africans displayed similar means for breadth measurements, and black and white South Africans showed similar means for the maximum length of distal limb elements. The majority of group variation is attributed to differences in size and robusticity, where white South Africans are overall larger and more robust than black and coloured South Africans. Accuracies for the univariate sectioning points ranged from 43% to 87%, with iliac breadth performing the best. However, the majority of the univariate sectioning points can only classify individuals into two groups rather than three because of similar group means. Multivariate bone models created using all measurements per bone resulted in accuracies

ranging from 46% to 62% (LDA) and 41% to 66% (FDA). Multivariate subsets consisting of numerous different measurement combinations from several skeletal elements achieved accuracies as high as 85% (LDA) and 87% (FDA).

Ultimately the best results were achieved using combinations of different variables from several skeletal elements. Overall, the multivariate models yielded better results than the univariate approach, as the inclusion of more variables is generally better for maximising group differences. Furthermore, FDA achieved higher accuracies than the more traditional approach of LDA. Despite the significant overlap among the groups, the postcranial skeleton has proven to be proficient in distinguishing the three groups. Thus, even in a heterogeneous population, a multivariate postcraniometric approach can be used to estimate ancestry with high accuracy.

OPSOMMING

Die primêre rol van ‘n fisiese antropoloog is om voldoende inligting te verskaf om die identifikasie van skelette moontlik te maak. Dit word dikwels bereik deur die samestelling van ‘n biologiese profiel, waarvan afkoms ‘n belangrike aspek is. Verskeie suksesvolle osteometriese en morfologiese metodes is tot dusver ontwikkel om die bepaling van afkoms met behulp van die skedel te bewerkstellig. Die skedel is egter nie altyd beskikbaar vir analise nie, en beklemtoon dus ‘n behoefte aan postkraniale alternatiewe. Die postkraniale skelet word dikwels gebrandmerk as te wisselvallig en onbetroubaar om ‘n akkurate assessering van afkoms te voorsien. Tog maak talle studies gebruik van die postkraniale skelet vir die beraming van geslag en liggaamslengte, waar hoër akkuraatheid bereik kan word weens die a priori kennis van afkoms. Die teenwoordigheid van postkraniale verskille tussen populasies wanneer ander biologiese parameters ondersoek word dui inherent op die potensiaal vir die bepaling van afkoms. Die doel van die studie was om postkraniale variasie onder moderne, portuur-gerigte swart, wit en kleurling Suid-Afrikaners te kwantifiseer. ‘n Reeks van 39 standaard metings was geneem van 11 postkraniale bene, naamlik die klavikel, skapula, humerus, radius, ulna, sakrum, pelvis, femur, tibia, fibula en kalkaneus. Die steekproef het bestaan uit 360 moderne Suid-Afrikaanse individue (120 swart, 120 wit, 120 kleurling) van die Pretoria Been en Kirsten versamelings, gehuisves by die Universiteit van Pretoria en Universiteit van Stellenbosch, onderskeidelik. Groepverskille was ondersoek met ANOVA en Tukey se eerlik-beduidende verskil toets. Die groep gemiddeldes was gebruik om een-veranderlike skeidingspunte te skep vir elke veranderlike wat met ANOVA as betekinsvol aangedui was. Waar twee van die drie groepe soortgelyke gemiddeldes gehad het, was die groepe gekombineer vir die skepping van die skeidingspunte. Lineêre diskriminante analise (LDA) en aanpasbare (“flexible”) diskriminante analise (FDA) was toegepas op meer-veranderlike modelle. Die klassifikasie akkuraatheid was vergelyk om te bepaal watter model die beste resultate oplewer.

Die resultate het wisselende patrone van groepoorvleueling getoon. Swart en kleurling Suid-Afrikaners het soortgelyke gemiddeldes vir breedte metings getoon, en swart en wit Suid-Afrikaners het soortgelyke gemiddeldes getoon vir die maksimum lengte van die distale elemente van die ledemate. Die meerderheid van groep variasie word toegeskryf aan verskille in grootte en liggamsbou, waar wit Suid-Afrikaners in die geheel groter as swart en kleurling Suid-Afrikaners was. Die akkuraatheid van die een-veranderlike skeidingspunte het gewissel van 43% tot 87%, met die breedte van die ilium wat die beste resultate opgelewer het. Maar

aangesien groepe gekombineer was weens soortgelyke gemiddeldes, kan meeste van die een-veranderlike skeidingspunte slegs individue in een van twee groepe klassifiseer, in plaas van drie groepe. Meer-veranderlike beenmodelle wat alle metings per been in ag neem het ‘n akkuraatheid bereik wat wissel van 46% tot 62% (LDA) en 41% tot 66% (FDA). Meer-veranderlike metingstelle wat bestaan uit talle verskillende meting kombinasies van verskeie bene het akkuraatheid so hoog as 85% (LDA) en 87% (FDA) bereik.

In die algeheel het die meer-veranderlike modelle beter resultate opgelewer as die een-veranderlike benadering. Gevolglik is die insluiting van meer veranderlikes oor die algemeen beter om groepverskille te identifiseer. Daarbenewens het FDA hoër akkuraatheid as die meer tradisionele benadering van LDA bereik. Ten spyte van aansienlike oorvleueling tussen die groepe was die postkraniale skelet as vaardig bewys met die onderskeiding van die drie groepe. Dus, selfs in ‘n heterogene populasie kan ‘n meer-veranderlike postkraniometriese benadering gebruik word om afkoms met ‘n hoë akkuraatheid te bepaal.

ACKNOWLEDGEMENTS

I was very fortunate to have amazing support while writing this dissertation. First and foremost I would like to thank my supervisors, Prof Ericka L'Abbé and Dr Kyra Stull, for providing advice, knowledge, encouragement and when needed, hugs. Thank you for teaching me everything I know and always pushing me to be better. A special thanks to Kyra for introducing me to the joy that is R. You showed me that doing statistics doesn't have to be a painful experience. I want to offer my gratitude to Mandi Alblas and Linda Greyling at Stellenbosch University for granting access to the Kirsten Collection. I also would like to gratefully acknowledge the funding received from the NRF, without which this thesis would not have been possible. Big thanks to Gabi Krüger for being my partner in crime (Gandi Rules!) and for never judging me when my crazy starts to show. You are like a ninja! And finally, thank you to my family for supporting and putting up with me and my thesis-induced moods. I appreciate it.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

The primary role of a physical anthropologist is to provide sufficient information on unknown skeletal remains to either a forensic pathologist or investigating officer in order to facilitate the identification. This is achieved by establishing a biological profile that includes estimations of age at death, sex, stature and ancestry. Race, ethnicity and ancestry are all terms synonymously used to describe, and ultimately classify, different human groups. The traditional race concept refers to biologically homogeneous groupings as defined by a few phenotypic features (Bophal and Donaldson, 1998; Agyemang et al., 2005; Gravlee, 2009). Although racial studies historically formed a major component of physical anthropology, numerous researchers began to demonstrate the instability of racial types and the intricacies of human variation (Caspari, 2010). While many physical anthropologists acknowledge the existence of morphological variation among humans, it has become common to deny its association with the traditional concept of race (Edgar and Hunley, 2009). The clinal view on race, as established by Livingstone (1962) and Lewontin (1972), argues that human variation is clinally distributed. In other words, the distribution of morphological traits among humans follows a geographic outline, rendering the concept of distinct races inadequate to describe human variation. If all possible groups that vary based on a set of arbitrary traits were considered separate races, there would be hundreds – and maybe thousands – of different races (Ousley et al., 2009). Thus, rather than debating the existence of human variation the focus shifted to whether the variation can be explained taxonomically. With the study of clinal variation the term population replaced race, where the argument for local populations and populational variation has more biological integrity than large, discrete races (Caspari, 2010).

In 2009, the American Journal of Physical Anthropology arranged a symposium in an attempt to reconcile the concept of race and human variation. Edgar and Hunley (2009) summarised their conclusions: the symposium reached the agreement that considerable variation can be observed among individuals within different populations. Furthermore, the noted variation can be apportioned among individuals in different populations and larger population groupings. Finally, within- and between-group variation patterns are largely shaped by numerous extrinsic factors such as culture, language and geography (Edgar and Hunley, 2009). Simply, while significant variation exists within each group, different population groups can be distinguished from one another. Human populations are socially

constructed groups with diverse histories influenced by numerous factors, all of which contributes to human variation (Ousley et al., 2009; L'Abbé et al., 2013a). Thus, ancestry – the term now preferred rather than race – assesses the relation between the socially constructed populations and the morphological variation associated with them which exists as a result of varying origins and population histories, without supporting the existence of biological races.

Some anthropologists argue that ancestry estimates are generally more difficult, less precise and overall less reliable than the estimates of other biological parameters (White and Folkens, 2005). Although estimating ancestry is a difficult task to accomplish, providing answers about population background is as essential to the individualisation of remains as an age or sex estimate. With the development of new methodological approaches and techniques, the estimation of ancestry is becoming much more reliable and valid (Gill, 1990; 2009). The cranium is generally used to facilitate the estimation of ancestry, with mid-facial features, such as the nose, often producing the best results (Gill, 1998; Hefner, 2009). While the cranium has proven to be proficient in discriminating among modern South Africans, yielding accuracies between 73% and 98%, it is not always available for analysis (İşcan and Steyn, 1999; McDowell et al., 2012; L'Abbé et al., 2013a; L'Abbé et al., 2013b; Stull et al., 2014).

Few published studies on the postcranial analysis of ancestry are available. Some of the older studies (e.g. Hrdlička, 1942 and Walensky, 1965) mostly utilise qualitative morphoscopic traits and because of the heuristic approach, lack validity and reliability and ultimately fail to meet the *Daubert* requirements for evidentiary standards in court. More recent studies employ robust multivariate statistics, which present a more objective approach and result in considerably high accuracy rates (Dibennardo and Taylor, 1983; İşcan, 1983; Patriquin et al., 2002; Bidmos, 2006). However, the current application of non-cranial methods for estimating ancestry are rather limited; a recent survey among South African forensic practitioners (n = 30) indicated that 55% of the participants would not even attempt to estimate ancestry if the cranium was not present. Some anthropologists have labelled postcrania as exceedingly variable and unreliable for estimating ancestry (Stewart, 1979; Albanese and Saunders, 2006). Based on the absence of published studies, the postcranial skeleton continues to retain this poor reputation.

Despite an apparent reluctance to utilise postcrania for the estimation of ancestry in South Africa, numerous studies use postcranial elements for both the estimation of sex and the calculation of stature (Steyn and İşcan, 1997; 1999; Asala, 2001; Patriquin et al., 2005; Barrier and L'Abbé, 2008; Lundy and Feldesman, 1987; Steyn and Smith, 2007; Bidmos, 2008; Dayal et al., 2008). Several factors play a role in producing morphological variation in humans including climate, physical activity, socio-economic status (SES) and genetic composition (Pearson, 2000; Ruff et al., 2006; Henneberg et al., 1998; Karsenty and Wagner, 2002). The diverging levels of variability among humans resulting from the complex interaction of the above-mentioned factors prompt studies to emphasise the importance of population-specific standards, as significant differences are consistently demonstrated to exist among populations. Essentially the a priori knowledge of ancestry can increase the accuracy of sex and stature estimates, which inherently indicates the potential to elucidate population differences from the postcranial skeleton specifically for the estimation of ancestry.

Based on the numerous factors influencing postcranial dimensions, varying group origins and the results of previous studies, certain postcranial differences are expected among the South African groups. Specifically, white South Africans are expected to present with greater long bone lengths and epiphyseal breadths along with wider pelves, whereas black South Africans are expected to have slender postcrania with small articular heads, slight muscle markings and narrow pelves (Farally and Moore, 1975; Krogman and İşcan, 1986; Patriquin et al., 2002). Although literature on the postcrania of coloured South Africans are limited (Steyn and Smith, 2007), they are expected to present with postcranial dimensions similar to black South Africans.

The purpose of this study was to explore variation in the postcranial skeleton of modern, peer-reported black, white and coloured South Africans. Additionally, this study aimed to quantify postcranial differences in order to develop country-specific standards that can be used to estimate ancestry in South Africa.

CHAPTER 2: LITERATURE REVIEW

2.1 Race and the concept of ancestry

The validity of the race concept as a means to define human populations has been thoroughly debated for more than a century (Ousley et al., 2009). In past traditional systems of classification, arbitrary traits such as skin colour, facial characteristics, skeletal morphology and even behavioural attributes were used to divide humans into distinct biological groups such as “Caucasoid”, “Negroid” and “Mongoloid” (Stein and Rowe, 1989). Discrete exclusive “types” of racial groups inspired the notion of the existence of ideal race categories. Biological characteristics became a tool with which certain groups ascertained their rights as “superior” human beings, ultimately leading to racist thinking, discrimination and segregation among groups (Brace, 1995). Smay and Armelagos (2000:23) described the typological view, or trait list approach as a “kind of cookbook set of instructions” for identifying race and partitioning groups that did not take into account the complexity of human variation.

Human populations are not simply biological subgroups as any combination of non-biological parameters such as culture, religion, language and geography can be used to define significant groupings of people (L’Abbé et al., 2013a; Edgar and Hunley, 2009). Therefore, the concept of race is a social rather than a biological concept. Edgar (2009) defines the idea of social race, or folk taxonomy, as the categorisations people use in their everyday judgements concerning themselves and the people around them, in other words how people designate or identify themselves.

Recent studies move away from the historical approach of estimating defined racial “types” to assessing inherent differences between populations based on population origins, known as ancestry (Ousley et al., 2009). This emerging view on the concept of human variation focuses on differences and similarities between populations that occur due to broad differences in ancestral geographic origins. Where race focused on arbitrary superficial, morphological characteristics, ancestry focuses on the quantifiable traits that can be correlated with the environment of a certain group’s ancestors, which is then translated into a social labelling system as used in everyday society (Sauer, 1992; Posel, 2001). Ancestry is not a vindication of the race concept or the notion of biological “types” (Sauer, 1992).

Forensic anthropologists are asked to estimate ancestry from skeletal remains to assist in the identification of an unknown person. Social classifications used in South Africa are often

referred to as categories of apartheid, as the designation of black, white and coloured South Africans were established under the 1950 Population Registration Act for the purpose of physically separating groups based on skin colour (Erasmus, 2012; Ellison and de Wet, 1997). However, the social system of classification was created prior to apartheid and has been retained after 1994 for the purpose of "transformation" and redress (Posel, 2001). This begs the question: if race is a social construct, how can anthropologists identify ancestry from skeletal remains? The estimation of ancestry is possible because a relationship exists between skeletal dimensions and self-designated or peer-reported social groupings, which is partly the result of the different origins and population histories of each group (Ousley et al., 2009; Jantz and Ousley, 2005). Past segregation laws and social barriers also played a role in the morphological variation present in a population. Institutional racism associated with apartheid influenced mating patterns, which limited gene flow and produced distinct skeletal differences among social groups (L'Abbé et al., 2013a). While renewed socio-political climate permits gene flow among groups, social behaviour continues to limit it. Therefore, the biological consequences of racism can be used to help explain the range of variation currently observed in the South African population.

2.2 Ancestry estimation

Numerous studies consistently recognise the skull as the best element to estimate ancestry (Gill and Gilbert, 1990; İşcan and Steyn, 1999; Hefner, 2009; L'Abbé et al., 2011; McDowell et al., 2012). To date the assessment of the postcranial skeleton has not been a major focal point in ancestry research as many investigators believe postcrania provide inconsistent or invalid estimations (Stewart, 1979; St. Hoyme and İşcan, 1989; Albanese and Saunders, 2006). The void of research using the postcranial skeleton to estimate ancestry possibly stems from the methodological approach of previous studies. Specifically, the use of outdated statistics and unquantifiable traits.

The majority of previous studies utilised qualitative morphological characteristics from long and short bones to distinguish population differences, such as the anterior curvature of the femur (Walensky, 1965; Gilbert, 1976; Stewart, 1979), variations in the shape of the superior border of the scapula (Hrdlička, 1942) and the number of talar articular facets on the superior aspect of the calcaneus (Bunning and Barnett, 1965; Bidmos, 2006). However, few published validation and/or reliability studies are available for the application of these morphological

characteristics to skeletal remains, which is a precarious situation as the methodologies employed with morphological analyses generally fail to satisfy the *Daubert* standards.

The *Daubert* standards were set during United States Federal legal proceedings (*Daubert* v. Merrell Dow Pharmaceuticals, 1993) and applies to the admissibility of expert witness testimony (Grivas and Komar, 2008). The criteria are in place to ensure quality and consistency in the field of forensic sciences through critical evaluation of the methodology rather than focusing solely on the conclusions or results obtained. Guidelines for meeting the *Daubert* criteria state that expert witness testimony needs to: (1) be testable, (2) be peer-reviewed, (3) have established standards, (4) have known error-rates and (5) be accepted by the scientific community (Grivas and Komar, 2008; Christensen and Crowder, 2009). Essentially, an anthropological assessment can no longer be based on "years of experience" (Dirkmaat et al., 2008). While South African courts do not utilise the *Daubert* guidelines, similar standards are expected (Meintjes-Van Der Walt, 2003).

The heuristic morphoscopic approach has faced significant criticism because of the reliance on the skill and knowledge of the anthropologist and the poor repeatability demonstrated by most validation studies (Hefner, 2009; L'Abbé et al., 2011). Hefner (2009:986) states, "this method is an art, an art that is intuitive, untestable, unempirical and consequently unscientific". The application of morphological characteristics without the inclusion of detailed descriptions and more powerful statistical analyses is bound to produce variable and inaccurate results.

Studies that make use of postcraniometric techniques for estimating ancestry mainly focus on the calculation of indices (Brown, 2006; Krogman and İşcan, 1986; Flower and Garson, 1879; Broca, 1878; Wescott, 2005), such as the platymeric index (PI) of the femur. The PI is calculated by dividing the anterio-posterior subtrochanteric diameter by the transverse subtrochanteric diameter and multiplying the resultant value by 100. The index evaluates the shape of the proximal femur and classifies the variations as more rounded in shape (eurymeric) or having a more flattened cross-section (platymeric) (Brown, 2006). Although indices can be used to obtain more information on the shape of an object, indices are only representative of two dimensions. Thus, an index is not able to capture enough information regarding the range of variation among populations. The application of indices to forensic case work is associated with numerous statistical and methodological problems and the reduction of complex biological data to a single value index may mask important information

and produce misleading results (Green and Chapman, 2011; Holliday and Ruff, 2001; Liebenberg et al., n.d.).

In contrast to morphological approaches and indices, multivariate techniques are statistically robust and are useful in a wide range of problems (Campbell, 1978). Multivariate techniques are capable of maximising group separation and assessing the relationship between variables and how they function when combined, which is required when working with complex biological data. Giles and Elliot (1962) first used multivariate discriminant function analysis as a classification technique in physical anthropology; since this time many studies have employed the methodology to differentiate population groups. However, the majority of studies use the skull. For the limited number of studies that applied multivariate techniques to the postcranial skeleton (Patriquin et al., 2002; Bidmos, 2006; İşcan, 1983; Dibennardo and Taylor, 1983; Uhl et al., 2007), the results were promising with accuracies around 80%, though some studies have shown accuracies as high as 97% using measurements from the pelvis and the femur. Such high accuracies demonstrate the potential utility of multivariate methods to estimate ancestry from the postcrania. However, limitations exist to the published studies, such as the inclusion of limited samples, the use of non-standard measurements, and discriminant function formulae based on limited variables.

With regard to the studies using South African populations, only black and white South Africans were included in the samples (Patriquin et al., 2002; Bidmos, 2006). According to population estimates for 2013, coloured South Africans make up around 9% of the total South African population (www.statssa.gov.za), but within certain provinces the population comprises the majority of the population (48.8%, Western Cape). With the exclusion of coloured South Africans, the reference samples do not effectively represent the South African population and the variation observed therein. Furthermore, because the demographic distribution of groups varies among different regions (e.g. Gauteng vs Western Cape), geography limits where accurate anthropological casework can be conducted if all possible groups are not included in the analysis. Second, novel, non-standard measurements were incorporated into the studies without a presentation of repeatability. Even though the newly developed measurements may reveal significant ancestral differences, some issues with measurement repeatability may exist, which ultimately affects the accuracy and applicability of the results. Lastly, the studies generally present discriminant function formulae with a set number of variables. If an extensive review of variable combinations was not performed, the provided formulae may actually limit the potential of the data. Furthermore, providing

discriminant function formulae requires all of the variables in the formula to be present. If one of the measurements could not be taken – as is often the case with remains exposed in an outdoor context – the formula cannot be used (Ousley and Jantz, 2012). Despite the drawbacks, the above-mentioned studies demonstrate the capability of multivariate models to elucidate ancestry using postcranial elements.

2.3 Inter- and intra-observer agreement and measurement repeatability

While metric data is less subjective and generally more reliable, variation in inter- and intra-observer agreement of measurements can compromise statistical analyses and render results useless. Most errors associated with quantitative data collection result from the transposition of numbers (*i.e.* recording 61 instead of 16), decimal place errors, failure to "zero" digital instruments, and a lack of understanding of the measurement definitions or location of landmarks (Adams and Byrd, 2002; Harris and Smith, 2009). While the other sources of variation can usually be identified and fixed during data cleaning, erroneous values as a result of measurement misinterpretation are not as apparent. Several postcranial measurements frequently identified as problematic include femur subtrochanteric measures (femsap and femstv), ulnar diameters (ulndvd and ulntvd) and sacral breadth (sacabr) (Adams and Byrd, 2002). The problem with the above-mentioned measurements mostly originates from confusion surrounding the exact location of the landmark and orientation of the instrument in relation to the bone rather than unfamiliarity with the measurement descriptions. For instance, uncertainty (and disagreement) exists regarding how close to the lesser trochanter the subtrochanteric measurements should be taken (Adams and Byrd, 2002). Another potential problem is the presence of bony growths such as osteophytes (commonly found on the ulna and calcaneus) or marked muscle attachment sites. Whereas some investigators might choose to include the bony growths/crests, others might exclude them or omit the measurements entirely. Potential sources of error need to be acknowledged when comparing osteometric measures of skeletal material. Although frequently underestimated, quality control is an essential step in the data gathering process so as to minimise measurement errors and prevent misleading results (Goto and Mascie-Taylor, 2007; Bennet, 1986).

2.4 Factors influencing the morphology of postcrania

Several researchers have noted postcraniometric variation among populations (Meadows-Jantz and Jantz, 1999; Todd and Lindala, 1928; Farrally and Moore, 1975; Steyn and Smith,

2007). Discernable differences in the general morphology of skeletal elements develops from both proximate and ultimate influences (Pearson, 2000). Proximate factors that influence the variation of skeletal morphology include immediate, environmental factors that act to alter the morphology within an individual's lifespan. The physical adaptation to environmental changes is often referred to as plasticity (Bogin and Loucky, 1997; Buck et al., 2010). Ultimate causes of variation are factors that produce genetic or epigenetic differences in the manner in which the skeleton grows and remodels and may accumulate over time through differences in selective pressures (*i.e.* genetic constraint) (Pearson, 2000). In essence, greater plasticity results in greater variation among groups, whereas higher levels of genetic constraint are expected to reduce variability. The complex interaction between proximate and ultimate factors is responsible for observed differences among populations.

2.4.1 Proximate factors

2.4.1.1 Sexual dimorphism, stature and population differences

Although beyond the scope of this study, variation in sexual dimorphism and population-specificity should be taken into consideration when interpreting variation associated with ancestral differences. As an ultimate cause of variation, sexual dimorphism is considered to have a strong genetic basis with varying selective pressures being exerted differently on males and females (Lande, 1980). However, sexual dimorphism may also be a proximate influence where dimorphism in body composition is the product of several hormones. Gonadal steroids, or sex hormones, play a role in skeletal homeostasis as they add bone during puberty and maintain skeletal integrity throughout life (Riggs et al., 2002; Callewaert, 2010). For instance, estrogen may repress bone growth by inducing epiphyseal closure, while testosterone stimulates periosteal bone expansion, leading to increased cortical bone growth (Juul, 2001; Callewaert, 2010). Additionally, contrasting strategies among the sexes for the investment of energy during growth and development may also result in population differences (Wells, 2012). Sexual dimorphism in humans is primarily attributed to differences in size, where males are generally larger and more robust than females. Numerous studies utilise the postcranial skeleton for sex estimation and emphasise the importance of population-specific standards because considerable size differences exist, both in absolute dimensions and magnitude, among populations (İşcan and Shihai, 1995; Loth and Henneberg, 1996; Steyn and İşcan, 1997; Steyn and İşcan, 1999; Patriquin et al., 2002; Bidmos and

Asala, 2003; 2005; Barrier and L'Abbé, 2008; Vance et al., 2010; Spradley and Jantz, 2011; Tise et al., 2013).

Research directed at the calculation of stature note proportional differences and variation in long bone length among populations (Gray and Wolfe, 1980; Wolfe and Gray, 1982; Lundy and Feldesman, 1987; Steyn and Smith, 2007; Bidmos, 2008; Dayal et al., 2008). Steyn and Smith (2007) demonstrated that white South Africans – both males and females – are significantly taller (*i.e.* have greater long bone lengths) than black and coloured South Africans, who are similar in stature. In contrast, black North Americans were noted to be similar in stature or even taller than white North Americans (Eveleth and Tanner, 1976; Micklesfield et al., 2007). Thus, stature and sexual dimorphism have been shown to be variable among populations, which is indicative of different genetic compositions. Furthermore, the a priori knowledge of ancestry when estimating sex and stature can increase classification accuracy.

Possible confounding effects of the interaction between sex, stature and ancestry should also be considered. Sexual dimorphism and stature differences are primarily attributed to differences in size. The variation in magnitude of sexually dimorphic patterns between different populations may impede ancestry estimation. For example, when the sexes are pooled, populations exhibit substantial ranges in overall size variation. Thus, individuals may misclassify into the wrong ancestral group as a result of differences in sexual dimorphism. A similar issue was noted with the cranium; L'Abbé et al. (2013a) found that sex misclassified more than ancestry for black and white South Africans (*i.e.* a black male would still be classified as black, but would misclassify as female).

Given the larger range of variation and overlap observed in the postcranial skeleton, a greater risk exists that an individual may misclassify into the wrong ancestral group as a result of gross variance in size due to sex differences. In other words, males of group *a* may classify as females of group *b*, because group *b* has an overall larger size. For instance when compared to black and white North Americans, Hispanic (Tise et al., 2013) and Japanese (İşcan et al., 1998) males are generally smaller and more gracile and thus are more likely to misclassify as females. Based on the stature distribution in South Africa (Steyn and Smith, 2007), robust black males might classify into the white group and gracile black females might classify into the coloured group, ultimately affecting classification accuracies.

2.4.1.2 Socio-economic status (SES)

Studies suggest that optimal environmental conditions favours skeletal maturation, whereas less favourable environments may lead to the retardation of both skeletal growth and maturation (Schmeling et al., 2000; Bogin and Loucky, 1997; Frisancho et al., 1970). Thus, individuals from unfavourable conditions (*i.e.*, black and coloured South Africans) may display stunted growth and as a consequence may be smaller than white South Africans (Zere and McIntyre, 2003; Steyn and Smith, 2007).

A lot has been written about the legacy of apartheid in South Africa, the outcome of which perpetuates unequal living circumstances among groups (Norris et al., 2008). Under apartheid policy, South African groups were subjected to acute socio-economic contrasts ranging from dire Third World conditions to well-off First World conditions (Henneberg and Lavelle, 1999; Cameron, 2003). Following the implemented racial hierarchy, most black South Africans lived in poverty, while white South Africans experienced more comfortable conditions (Posel, 2001; Thompson, 2001). Although similarly affected with racial discrimination as black South Africans, Indian and coloured groups experienced socio-economic conditions ranging between those of black and white South Africans (Møller, 1998; Thompson, 2001). Ultimately black and to a lesser extent coloured children were exposed to numerous factors known to negatively impact skeletal growth and maturation, namely poor nutrition, sanitation and hygiene as well as an increased risk of infection and disease (Vidulich et al., 2006; Norris et al., 2008). The inauguration of the new government in 1994 saw attempts to redress the corollaries associated with apartheid, and although there were shifts in SES for all South Africans, the living standards for the majority of black South Africans were only marginally improved (Møller, 1998).

Because small body-size and short stature are frequently cited as the resultant effects of poor health and low SES (Louw and Henneberg, 1997; Henneberg, 1998; Kurki, 2011), the repercussions of apartheid became widely accepted as the primary cause of the inherent size differences among South Africans. According to Cameron (1992; 2003), black and white South African children might reach similar patterns of growth and development if raised in similar environmental conditions; an outcome that was expected after the abolition of apartheid. However, growth studies performed on post-apartheid children have not observed this trend. Although the post-apartheid cohorts demonstrate improvement in growth and development compared to children born in the 1970's, size differences persist between black

and white children (Cameron, 2003; Richter, 2007; Anholts, 2013). While Cameron (2003) posits that the rate of change in size difference has been slower than expected, additional explanations need to be considered.

Because stature is a complex trait, the possibility exists that individual genetic potential contributes more to growth and overall size than SES and the environment (Henneberg and LaVelle, 1999). Numerous published studies (e.g. Ginsburg et al., 1998 and Livshits et al., 2002) present models exhibiting the substantial heritability of anthropometric traits, thereby demonstrating the influence of genetic control on body size and proportions within individuals and among populations. Stewart (1980) noted that Japanese immigrants to the United States grew to be taller than their counterparts residing in Japan; however they remained below the North American standard for stature. Thus, intrinsic factors restrict the American-born Japanese from reaching the size of black and white North Americans, despite an improved quality of life (Stewart, 1980). Similarly, regardless of improved living conditions/SES among previously disadvantaged groups in South Africa, variation in genetic endowment will lead to variation in body size and stature.

2.4.2 Ultimate factors

2.4.2.1 Climate and body size

Climate is a complex feature of the environment and encompasses temperature, humidity and altitude (Pearson, 2000). Biologists have long recognised strong correlations between skeletal proportions in relation to distance from the equator (Pearson, 2000; Holliday and Ruff, 2001), with the best-known being Bergmann's (1847) and Allen's (1877) ecogeographic rules. The rules state that individuals from colder climates tend to have a larger body size with relatively shorter extremities and more robust diaphyses and joints than their counterparts from warmer regions (Pearson, 2000; Kurki et al., 2008; Plavcan, 2012). Thus, individuals with slender or ectomorphic physiques are typically found in hot, tropical areas, as opposed to the stockier, mesomorphic/endomorphic physiques of arctic populations (Roberts and Bainbridge, 1963). Climatic adaptations represent the cumulative effect of generations of selective pressure on the human physique.

Sub-Saharan Africans are generally grouped into a "tropical, warm-adapted" category (Kurki et al., 2008). However, this category does not necessarily reflect the physiques of the different groups that make up the populations. For instance the robust, large body size of

white South Africans suggest that they are better adapted to colder environments, coinciding with the environments of their European ancestors. Furthermore, Khoesan (indigenous South Africans) present with relatively short limbs, where the upper limb tends to be shorter than the lower limb – in contrast to the long limbs and short trunks of black South Africans (Kurki et al., 2008; Pfeiffer, 2012).

Newman (1975) advocates that any group residing in a given environment long enough for natural selection to affect the group will show certain adaptive characteristics. However, certain populations sharing a common environment, such as South Africans or North Americans, present with significant differences in body shape (Holliday and Falsetti, 1999). Kurki et al. (2008; 2012) suggest that climatic factors are less influential on body proportionality in geographic regions that are not subjected to climatic extremes. Therefore, a lack of thermoregulatory stress may preclude natural selection of a certain body type, resulting in populations retaining the proportions of their ancestors.

2.4.2.2 Genetic components

Although adaptive and secular changes may be responsible for creating differences between populations, the underlying hereditary contributions must also be acknowledged (Holliday and Ruff, 2001). A critical component of skeletal development is the different genes and genetic pathways that control the overall shape and size of a particular skeletal element (Karsenty and Wagner, 2002).

Numerous researchers argue for plasticity of the human body, or how the body responds to the environment (Sparks and Jantz, 2002). Boas (1912) was an advocate of plasticity as demonstrated in his studies on immigrant populations. Studies on the growth and development of the cranium have confirmed that the facial form, especially the nose, has been linked to climatic adaptation (Harvati and Weaver, 2006). However, Sparks and Jantz (2002) argue that even though black and white North Americans have experienced significant change in cranial morphology over the past 150 years, they have not converged to a common morphology as might be expected if environmental plasticity played such a major role. Hence, the dimensions of the cranium preserves “genetic patterns” in humans, despite environmental influences (Sparks and Jantz, 2002; Harvati and Weaver, 2006; Manica et al., 2007). Pollitzer and Anderson (1989:1245) state “genes do not determine destiny but rather they set the stage upon which environment operates”. Therefore, environmental factors such

as climate, SES and physical activity work in conjunction with genetic components to determine the varying physiques of different populations (Pollitzer and Anderson, 1989). The prenatal form of the human body is believed to be largely genetically determined as it is insulated from the world, and it is only after birth that epigenetic and environmental forces have a major impact (Vidarsdottir and O'Higgins, 2003; Holliday and Hilton, 2010). Although the genetic control of the postcranial skeleton is not as well documented as the cranium, the results of previous anthropometric analyses suggest that postcrania may also retain a population history signal (Betti et al., 2012). Subsequently, the population signal accounts for differences among populations located in a similar environment such as the South African population.

CHAPTER 3: MATERIALS AND METHODS

3.1 Materials

3.1.1 Skeletal sample

The skeletal material originated from the Pretoria Bone Collection (University of Pretoria) as well as the Kirsten Collection (University of Stellenbosch). Postcraniometric data were collected from 360 individuals (120 South African black, 120 South African white, 120 South African coloured) with equal sex distribution. Table 3.1 provides the mean age distribution for each group. The skeletal material in the stated collections was from either donated or unclaimed bodies and received by the respective institutions under regulation of the National Health Act 61 (2003). All individuals accessioned into the skeletal collections are of known sex, age at death and ancestry (L'Abbé et al., 2005; Alblas 2014, pers. comm). Ethical approval (s297/2013) was obtained from the Faculty of Health Sciences Research Ethics Committee at the University of Pretoria.

Table 3.1 – Mean age distribution of the skeletal sample.

	Black			White			Coloured		
	Males (*n*=60)	**Females (*n*=60)**	**Pooled (*n*=120)**	**Males (*n*=60)**	**Females (*n*=60)**	**Pooled (*n*=120)**	**Males (*n*=60)**	**Females (*n*=60)**	**Pooled (*n*=120)**
Age	48	47	47	62	67	65	53	44	48

3.1.2 South African groups

South Africa is a country with a vast and rich heritage. Population estimates for 2013 show that black South Africans comprise 79.8% of the population, while coloured and white groups comprise 9.0% and 8.7% of the population, respectively (www.statssa.gov.za). The different groups that make up the diverse South African population hail from varying ancestral and geographic origins. Furthermore, past segregation laws and social barriers have limited gene flow between the groups, resulting in a population with a large amount of variation within and among the groups.

3.1.2.1 Black South Africans

The origin of black South Africans is speculative but is widely accepted as the result of a series of southward migrations by Bantu-speaking inhabitants from western-central Africa (Ribot, 2004; Tishkoff and Williams, 2002; Franklin et al., 2007; Franklin et al., 2010). The

term “Bantu” is a linguistic notion which refers to the general assemblage of several closely-related African languages. These languages belong to the Niger-Kordofanian (Niger-Congo) linguistic phylum and are spoken throughout the majority of western and sub-Saharan Africa (Ribot, 2004). Further divisions among the southern Bantu-speakers have led to numerous subgroups, or tribes, with differential customs and cultural systems (Hall and Morris, 1983; Herbert, 1990). Some of the subgroups that currently inhabit South Africa include the Sotho (Tswana, Pedi and Southern Sotho), Nguni (Xhosa, Zulu, Ndebele and Swazi), Venda and Shangaan-Tsonga (Franklin et al., 2007).

The large-scale migrations from western and central Africa is believed to have instigated a homogenisation process among the southern Bantu-speaking groups as a result of the restructuring of gene flow and a decrease in isolation (De Villiers, 1968; Ribot, 2004). Testing differences among subgroups can be difficult if skeletal remains are not accurately assigned (*i.e.* not self-designated). Some researchers maintain that although southern “Bantu” groups are closely related, differences are noted on the tribal level that can be used to separate the subgroups (Franklin et al., 2007; Siddiqi, 2013). Phenotypic disparities among the subgroups are thought to be the result of intermixture with indigenous Khoesan groups (Franklin et al., 2007).

The literature presents numerous examples citing evidence of contact between the Khoesan and Bantu-speaking groups. The Xhosa and Zulu (Nguni) have adopted several loan words and click consonants commonly associated with Khoesan into their own languages. Oral history among Xhosa and Phuti (a mixed Sotho-Nguni group) addresses the integration of Khoesan groups into their tribes (Herbert, 1990). Wood et al. (2005) and Petersen and co-workers (2013) have identified genetic signatures that indicate intermarriage between the Khoesan and Bantu-speakers; this is especially evident among Xhosa groups. Cranial morphology, specifically vault shape and facial profile, further confirms a historic Khoesan influence amongst some black South African subgroups (Franklin et al., 2010)

3.1.2.2 White South Africans

South Africa experienced two periods of colonisation. During the 17th century many European settlers arrived in the country. The Dutch East Indian Company (“Vereenigde Oost-Indische Compagnie”, or VOC) established the Cape colony which saw the influx of predominantly Dutch settlers, with later additions of British, French and German immigrants (Thompson, 2001; Greeff, 2007). South Africa’s turbulent history with race and the

associated laws and legislatures led to the development of a fairly homogenous white populace, which has been shown to be easily distinguished from other South African groups. However, Greeff (2007) estimates that as much as 7% of the genetic heritage of white South Africans are of non-European descent – possibly a result of gene flow between European settlers and indigenous populations. Additionally, white South Africans differ skeletally from their European and North American counterparts as a result of founder's effect and adaptation to environmental conditions (Steyn and İşcan, 1997; Steyn and İşcan, 1999; Steyn and Smith, 2007).

3.1.2.3 Coloured South Africans

In South Africa, the term "coloured" refers to a phenotypically varied social group of highly diverse social and geographical origins (Adhikari, 2005). This group is extremely heterogeneous, containing a combination of features from several groups (Patterson et al., 2010). Coloured South Africans, also known as the Cape coloureds, consist of individuals of mixed ancestry, descendants of slaves brought to South Africa from countries such as Malaysia and Indonesia as well as the descendants of the historic Khoesan (Adhikari, 2005; Patterson et al., 2010; Petersen et al., 2013).

Khoesan (previously Khoisan) is a collective term derived from the Nama words for 'person' and 'forager' (Schlebusch, 2010). As the name suggests, the Khoesan consist of both the Khoe pastoralists and the San hunter-gatherers (historically referred to as the "Hottenttots" and "Bushmen", respectively) (Hitzeroth, 1972; Barnard, 1992). The groups can be further subdivided: Khoe groups include the Nama, Kede and Hei//om while the !Xu (!Kung), Ju/'hoansi, G//ana, G/wi, G!aokxite and !Xo are included among San groups (Chen et al., 2000; Tishkoff et al., 2007). The Khoesan exclusively occupied southern Africa prior to the expansion of the Bantu-speaking groups (Franklin et al., 2007).

While these two groups are clustered into one social race, some differences in craniofacial morphology and stature exist between the Khoe and San (Rightmire, 1970). The Khoe are generally characterised as more robust in cranial features and larger in stature than the San (Morris, 1987). Biological differences have been attributed to the outcome of the socio-economic influences associated with their different subsistence strategies (herding versus hunting and gathering) between the groups (Stynder, 2009). Marginalised hunters suffering from nutritional stress are expected to be smaller in stature than pastoralists with sufficient stock. The possible selective value of small body size pertaining to food limitation and

mobility in hunter-gatherers should also be considered (Kurki et al., 2008; Pfeiffer, 2012). Nonetheless, the Khoe and San are recognised as different components of a single genetic population (Stynder, 2009). With the inception of the colonial period, the Khoesan were systematically dispossessed of their land and power and some were forced to work as slaves for European settlers. Khoesan numbers were further significantly decreased with a series of smallpox epidemics that struck in the 18th century (Stynder, 2009). Today contemporary Khoesan groups are restricted mainly to the arid Kalahari region of Namibia and Botswana (Petersen et al., 2013).

Another large contributing population to the heritage of modern-day coloured South Africans is the various groups of slaves that were brought to the country to maintain the Cape outpost and to work for the European settlers. The slaves were primarily shipped in by the VOC from the East – including countries such as Indonesia, Malaysia, India and Sri-Lanka – Madagascar and Mozambique as well as some Khoesan groups (Petersen et al., 2013). As the colonists consisted mainly of males, unions between European men and slaves and/or indigenous women were not uncommon (Patterson et al., 2010; de Wit et al., 2010). This is evident in the genotypes of contemporary coloured individuals as sex-biased gene flow, where their ancestral contributions present with strong European paternal signatures, and Khoesan maternal signatures (Patterson et al., 2010; de Wit et al., 2010; Petersen et al., 2013).

3.2 Methods

The skeletal elements utilised include the clavicle, scapula, humerus, radius, ulna, sacrum, pelvis, femur, tibia, fibula and calcaneus. A total of 39 standard measurements (Moore-Jansen et al., 1994) were taken to the nearest whole millimeter using an osteometric board and sliding and spreading calipers. Refer to Table 3.2 for measurement abbreviations and Appendix I for measurement definitions.

When both sides of each skeletal element were present, only the left side was measured. All individuals in the sample were older than 18 years of age and any skeletal material with visible pathology, antemortem trauma or postmortem damage that prevented an accurate measurement was excluded.

Table 3.2 – Measurement abbreviations

Clavicle maximum length	**claxln**	Innominate height	**innoht**
Clavicle sagittal midshaft diameter	**claapd**	Iliac breadth	**iliabr**
Clavicle vertical midshaft diameter	**clavrd**	Femur maximum length	**femxln**
Scapula height	**scapht**	Femur bicondylar length	**fembln**
Scapula breadth	**scapbr**	Femur epicondylar breadth	**femebr**
Humerus maximum length	**humxln**	Femoral head diameter	**femhdd**
Humerus epicondylar breadth	**humebr**	Femur A-P subtrochanteric diameter	**femsap**
Humeral head diameter	**humhdd**	Femur transverse subtrochanteric diameter	**femstv**
Humerus maximum diameter	**hummxd**	Femur A-P midshaft diameter	**femmap**
Humerus minimum diameter	**hummwd**	Femur transverse midshaft diameter	**femmtv**
Radius maximum length	**radxln**	Tibia condylo-malleolar length	**tibxln**
Radius A-P midshaft diameter	**radapd**	Tibia proximal epiphyseal breadth	**tibpeb**
Radius transverse midshaft diameter	**radtvd**	Tibia distal epiphyseal breadth	**tibdeb**
Ulna maximum length	**Ulnxln**	Tibia maximum diameter at nutrient foramen	**tibnfx**
Ulna dorso-volar diameter	**ulndvd**	Tibia minimum diameter at nutrient foramen	**tibnft**
Ulna transverse diameter	**ulntvd**	Fibula maximum length	**fibxln**
Ulna physiological length	**ulnphl**	Fibula maximum diameter	**fibmdm**
Sacrum anterior height	**sacaht**	Calcaneus maximum length	**calcxl**
Sacrum anterior breadth	**sacabr**	Calcaneus middle breadth	**calcbr**
Transverse diameter of S1	**sacs1b**		

3.3 Statistical analysis

All statistical tests were performed using R (version 3.0.1) (R Core Team, 2013). Outliers were detected and removed prior to analysis using univariate boxplots, bivariate scatterplots as well as principal component analysis (PCA). Inter- and intra-observer agreement was assessed to gauge measurement repeatability. Evaluation of measurement error and variability is of the utmost importance when creating standards based on metric data, as measurements need to be both repeatable and valid. Because the current study focused on the quantification of differences among modern South Africans, several exploratory techniques were employed and univariate sectioning points and multivariate classification models were used to test for group differences. The accuracy of the classification methods were compared to evaluate their discriminative power and to determine which method was most useful for the estimation of ancestry. Though exploratory techniques were used to evaluate the effects of sex, the objectives of the current study were not to assess sex differences among South Africans. Thus, the sexes for each group were pooled for all analyses.

3.3.1 Technical error of measurement (TEM) and Bland-Altman plots

Ten individuals were randomly selected from the sample to assess the inter- and intra-observer agreement. The repeatability of the measurements was gauged using two methods. First, the technical error of measurement (TEM) was calculated, which refers to the variability encountered between dimensions with repeated measurement of a specimen (Harris and Smith, 2009). Calculation of TEM provides an accuracy index that expresses error margins through the standard deviation between repeated measurements (Perini et al., 2005; Goto and Mascie-Taylor, 2007). The lower the variability among repeated measurements, the higher the precision of the results (Stomfai et al., 2011). Additionally, absolute TEM was converted to relative TEM. Relative TEM (%TEM), expressed as a percentage, provides an estimate of magnitude of the measurement error relative to the size of the measurement (Goto and Mascie-Taylor, 2007; Stomfai et al., 2011). For instance, a deviation of 2mm from a measurement of 10mm has much larger implications than a 2mm deviation from a measurement of 100mm.

Bland-Altman plots were used to visualise measurement agreement. A Bland-Altman plot is a graphical representation of the comparison of differences between multiple measurements, techniques or observers. In other words, a Bland-Altman plot depicts the TEM and can be

used to identify and reduce errors in measurement (Bland and Altman, 1999; Harris and Smith, 2009). On the plot, the *x*-axis represents the mean size (mm) of the measurements while the *y*-axis represents the difference in a pair of measurements compared between two observers. The upper and lower levels of agreement, or confidence limits, are plotted as horizontal, dotted lines. The limits are generally defined as the mean difference ± 1.96 standard deviation (Hanneman, 2008).

3.3.2 Exploratory analyses

Pearson correlations were conducted on all measurements to evaluate the linear relationship among the variables. As a result of inherent covariation human anthropometric/osteometric traits are often phenotypically dependent on each other and may present with strong correlations (Adjeroh et al., 2010; Uhl, 2014). Any correlation coefficient (*r*) that demonstrates a high degree of association is indicative of multicollinearity (*i.e.* linear dependencies among variables). Thus, multicollinearity is expected when assessing human proportions. While no universal cut-off point is used to indicate high correlations, multicollinearity is suggested to occur with correlations greater than 0.90 (Tabachnick and Fidell, 2007). The presence of multicollinearity in data may produce misleading results; however, variable reduction techniques may be conducted to contend with the effects of multicollinearity (Schroeder, 1990). A Holm's adjustment was performed with the correlations to counteract the effect of multiple comparisons. Holm's adjustment modifies the alpha level to maintain statistical power while controlling for the probability of Type I errors occurring (*i.e.* rejecting the null hypothesis when it is true) (Abdi and Williams, 2010). The correlations were visualised with the corrplot package in R, which helps to interpret the relationships among the variables and to search for multicollinearity using shape and colour (Wei, 2013; Friendly, 2002). Different shades of red indicate negative, or inverse, correlations while varying shades of blue indicate positive correlations. The size of the ellipse demonstrates the magnitude of the correlation between variables; the smaller the ellipse, the stronger the correlation. Additionally, the plot was combined with a significance test; if the correlation coefficient between variables were not significant, the tile was left blank.

An analysis of variance (ANOVA) was performed to simultaneously compare multiple group means per variable without inflating alpha (Samuels and Witmer, 2003). ANOVA is useful for identifying and assessing different sources of variation within a data set (Kachigan, 1991). For the current study the effects of ancestry, sex and the interaction between ancestry and sex

was tested. In addition, Tukey's honestly significant difference (HSD) test was used as a post hoc analysis in conjunction with ANOVA to determine if the means of each measurement for the three groups are significantly different from each other (Fenech, 1979). Simply, where ANOVA shows the significance of each model when all three populations are compared at once, Tukey's HSD identifies significant differences within the model.

Similarly, multivariate analysis of variance (MANOVA) was performed to compare the group means per bone. Essentially a MANOVA model analyses data where more than one dependant variable is explored at a time (*i.e.* all the measurements taken from one skeletal element), whereas ANOVA only looks at the significance of one measurement at a time. Unlike ANOVA, MANOVA is capable of considering the intercorrelations among the dependant variables that could not be addressed with a combination of several univariate analyses (Haase and Ellis, 1987).

3.3.3 Population means and univariate sectioning points

Variables indicated as significant with ANOVA and Tukey's HSD were used to create univariate sectioning points for each measurement. Sectioning points were made by taking the sum of the means for each measurement between two groups and dividing by two to obtain the midpoint of the means (*i.e.* the mean of the means). The midpoint serves as the sectioning point that can be used to assign unknown remains into a group. For example, in order to differentiate between group A and B, sectioning point x is created by adding the means for group A and B, divided by 2. Any measurement smaller than x will classify an unknown into group A, while any value greater than x will classify into group B. Similarly, a second sectioning point is made to distinguish between group B and C. The classification rate was subsequently calculated to establish the accuracy of each variable. Although univariate sectioning points have shown to generally be outperformed by more statistically robust multivariate techniques, this is an easy method to use and can be applied to incomplete, fragmentary or poorly preserved remains (Albanese et al., 2005).

3.3.4 Discriminant Analysis

A multitude of multivariate classification models were employed to identify the most accurate combination of predictor variables. Bone models were created using all of the measurements taken per bone to test the classification accuracy of each skeletal element and multivariate subsets were created using numerous measurement combinations from several

skeletal elements (Table 3.3). Regardless of the ANOVA and Tukey's HSD results, none of the measurements were excluded from the models as the multivariate combination of variables are more proficient in teasing out population differences than univariate approaches. All classification models were assessed using discriminant analysis (DA).

DA is a statistical procedure used to identify relationships between qualitative criteria and quantitative predictor variables (Kachigan, 1991). When applied in anthropology, DA incorporates skeletal observations into a mathematical formula to evaluate biological aspects of unknown individuals (Ousley and Jantz, 2012). DA is particularly useful for summarising overall differences and identifying boundaries between groups while accounting for the effects of individual variation within each group (Albrecht, 1980). Two types of DA were included in the study, namely linear discriminant analysis (LDA) and flexible discriminant analysis (FDA).

Table 3.3 - Variables included in each multivariate subset. See Table 3.2 for measurement abbreviations.

	n	**Variables**
Upper limb	316	humxln, humebr, humhdd, hummxd, humwd, radxln, radapd, radtvd, ulnxln, ulndvd, ulntvd, ulnphl
Upper limb + Shoulder girdle	183	claxln, clavrd, claapd, scapht, scapbr, humxln, humebr, humhdd, hummxd, humwd, radxln, radapd, radtvd, ulnxln, ulndvd, ulntvd, ulnphl
Lower limb	225	femxln, fembln,femebr, femhdd, femsap, femstv, femmap, femmtv, tibxln, tibpeb, tibdeb, tibnfx, tibnft, fibxln, fibmdm
Lower limb + Pelvic girdle	177	femxln, fembln, femebr, femhdd, femsap, femstv, femmap, femmtv, tibxln, tibpeb, tibdeb, tibnfx, tibnft, fibxln, fibmdm, sacabr, innoht, iliabr
Pelvis + Femur	265	femxln, fembln, femebr, femhdd, femsap, femstv, femmap, femmtv, innoht, iliabr
Proximal elements	280	humxln, humebr, humhdd, hummxd, humwd, femxln, fembln,femebr, femhdd, femsap, femstv, femmap, femmtv
Distal elements	247	radxln, radapd, radtvd, ulnxln, ulndvd, ulntvd, ulnphl, tibxln, tibpeb, tibdeb, tibnfx, tibnft, fibxln, fibmdm
All-measurement	140	claapd, scapht, humxln, humhdd, ulnxln, ulnphl, sacabr, iliabr, femstv, tibpeb, calcbr

3.3.4.1 Linear Discriminant Analysis (LDA)

LDA, the most commonly used DA, creates linear boundaries to separate groups. Several key assumptions are associated with the use of LDA, namely multivariate normal data, equal variance-covariance matrices, and a sufficient sample size (Kachigan, 1991; Ousley and Jantz, 2012). As a general rule the sample size of the smallest group in the analysis should exceed the number of predictor variables included in the model (Tabachnick and Fidell, 2007). Following the recommendation of Huberty (1994), the current study used a minimum sample size three times the number of variables included per model (3*m).* A sample of 3*m* yields estimates that are less subject to sampling variation while allowing sufficient measurements required for group separation (Ousley and Jantz, 2012).

Although the inclusion of more variables is generally considered superior, a limit exists to the number of variables that are beneficial in classification. The presence of too many variables may result in redundant information that can cause the classification accuracy to plateau and/or decrease (Tabachnick and Fidell, 2007; Ousley and Jantz, 2012). Therefore, stepwise variable selection was employed. Rather than using an entire set of variables, stepwise selection identifies and removes variables that do not contribute to increased classification (Kachigan, 1991). As a result, the most accuracte discriminant model is created using the smallest number of variables. The reduction of variables is beneficial to maintain a large enough sample size and also assists in avoiding problems with multicolliearity (Tabachnick and Fidell, 2007).

LDA was conducted on all models with the use of equal priors and leave-one-out cross validation (LOOCV), which involves the removal of one individual from the sample, after which a discriminant function is created using all of the remaining individuals. The function then attempts to correctly classify the removed individual. The process is repeated for every individual in the sample to test the reliability of the model and to provide the prediction error estimate (Hastie et al., 2009; Ousley and Jantz, 2012). Essentially LOOCV provides more realistic results with more valid accuracies, as this technique avoids optimistic bias and overfitting of the data (Ousley and Jantz, 2012).

3.3.4.2 Flexible Discriminant Analysis (FDA)

FDA is a low dimensional classifier which combines linear discriminant analysis (LDA) with non-parametric regression in order to generate flexible boundaries between groups (Mallet et

al., 1996). This technique differs from LDA in that a curved or bent surface is utilised rather than a linear hyperplane in order to better separate the classes (Milborrow, 2011). Using the fda function in the mda package (Hastie and Tibshirani, 2013), the FDA is generated by using multivariate adaptive regression splines (MARS), which operates by producing a basis expansion of the derived variables and then performing linear regression in the expanded multivariate space (Hastie et al., 2009). Thus, FDA demonstrates enhanced flexibility when attempting to model variables with non-linear relationships, particularly as the MARS algorithm makes no assumptions about the variables (Butte et al., 2010; Stull et al., 2014). FDA is a useful alternative when the data do not meet the requirements for LDA, and is thought to be superior to LDA when more than two response categories exist (Feldesman, 2002, Hastie, 1994). Comparable to the stepwise procedure, FDA automatically executes variable selection using a backwards elimination feature (Kuhn, 2013). The classification accuracy of FDA was compared to that of the LDA to test its discriminatory power as a tool to estimate ancestry.

CHAPTER 4: RESULTS

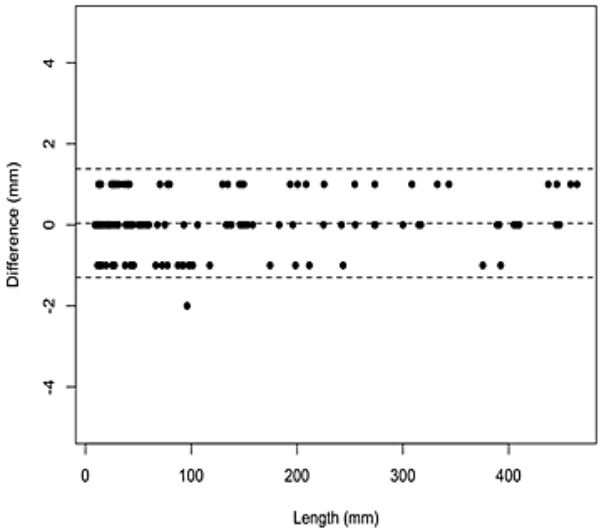
4.1 Measurement repeatability

Technical error of measurement (TEM) was used to demonstrate the degree of measurement repeatability between three observers: the principal investigator and two additional observers. The intra-observer TEM and %TEM ranged between 0 and 0.71mm and 0% and 4.52%, respectively. The inter-observer error was slightly higher, with TEM and %TEM ranging between 0 and 1.10mm and 0% and 5.36% respectively (Table 4.1). The measurements with the highest TEM for both intra- and inter-observer agreement were the midshaft measures of the clavicle, humerus and ulna as well as the subtrochanteric measurements of the femur. Midshaft measurements of the radius also presented higher levels of measurement error for the inter-observer analysis.

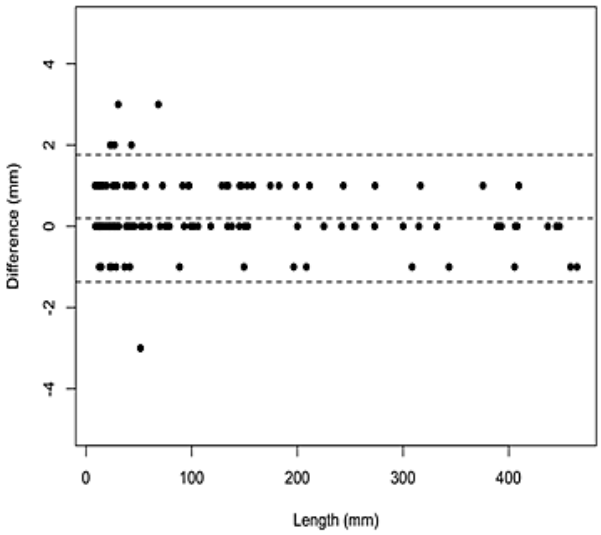
Bland-Altman plots were used to illustrate measurement agreement among the three observers. Sufficient agreement was noted among the three observers, with only a few measurements presenting with more than 2mm difference (Figure 4.1). The intra-observer plot presented with high precision, as the majority of the measurements did not have more than 1mm difference and had an overall narrower distance between the upper and lower agreement levels (Figure 4.2). For both the inter- and intra-observer agreement more inconsistencies were noted in smaller measurements (*i.e.* midshaft measurements) rather than the larger measurements (*i.e.* maximum lengths). Because measurement error was minimal, the inter- and intra-observer agreement was considered satisfactory, and all measurements were retained in the analyses.

Table 4.1 – Technical error of measurement (TEM) and relative technical error of measurement (%TEM) for inter- and intra-observer agreement.

	Intra-observer error		Inter-observer error 1		Inter-observer error 2	
	TEM	**%TEM**	**TEM**	**%TEM**	**TEM**	**%TEM**
claxln	0.00	0.00	0.35	0.24	0.35	0.24
claapd	0.35	3.58	0.35	3.67	0.00	0.00
clavrd	0.00	0.00	0.50	4.17	0.00	0.00
scapht	0.00	0.00	0.71	0.50	0.00	0.00
scapbr	0.35	0.37	0.35	0.37	0.61	0.64
humxln	0.00	0.00	0.55	0.18	0.32	0.10
humebr	0.00	0.00	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	0.32	2.07	0.45	2.90	0.32	2.07
radxln	0.35	0.15	0.35	0.15	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	0.00	0.00	0.35	0.14
ulndvd	0.63	4.52	0.71	5.36	0.55	3.94
ulntvd	0.00	0.00	0.55	3.78	0.63	4.65
ulnphl	0.00	0.00	0.61	0.27	0.61	0.27
sacaht	0.00	0.00	0.00	0.00	0.50	0.45
sacabr	0.71	0.74	0.61	0.65	0.94	0.99
sacs1b	0.00	0.00	1.00	2.33	0.41	0.94
iliabr	0.00	0.00	1.10	0.58	0.55	0.29
innoht	0.32	0.22	0.55	0.38	0.63	0.44
femxln	0.32	0.07	0.45	0.10	0.45	0.10
fembln	0.32	0.07	0.45	0.10	0.45	0.10
femebr	0.00	0.00	0.41	0.54	0.63	0.87
femhdd	0.00	0.00	0.35	0.84	0.32	0.76
femsap	0.00	0.00	0.95	3.87	0.55	2.18
femstv	0.45	1.55	0.77	2.71	0.71	2.41
femmap	0.00	0.00	0.32	1.14	0.32	1.15
femmtv	0.00	0.00	0.00	0.00	0.55	2.27
tibxln	0.00	0.00	0.41	0.11	0.58	0.15
tibpeb	0.00	0.00	0.82	1.14	0.50	0.70
tibdeb	0.00	0.00	1.06	2.34	0.35	0.79
tibnfx	0.00	0.00	0.55	1.70	0.32	0.97
tibnft	0.32	1.38	0.77	3.40	0.45	1.96
fibxln	0.00	0.00	0.41	0.11	0.58	0.16
fibmdm	0.00	0.00	0.32	2.34	0.00	0.00
calcxl	0.00	0.00	0.00	0.00	0.71	0.90
calcbr	0.35	0.87	0.41	0.99	0.50	1.24
Mean	0.13	0.40	0.48	1.42	0.43	1.04
Min	0.00	0.00	0.00	0.00	0.00	0.00
Max	0.71	4.52	1.10	5.36	0.94	4.65



Inter-observer 1



Inter-observer 2

Figure 4.1 –Bland-Altman plots illustrating the inter-observer agreement of measurements compared among three observers.

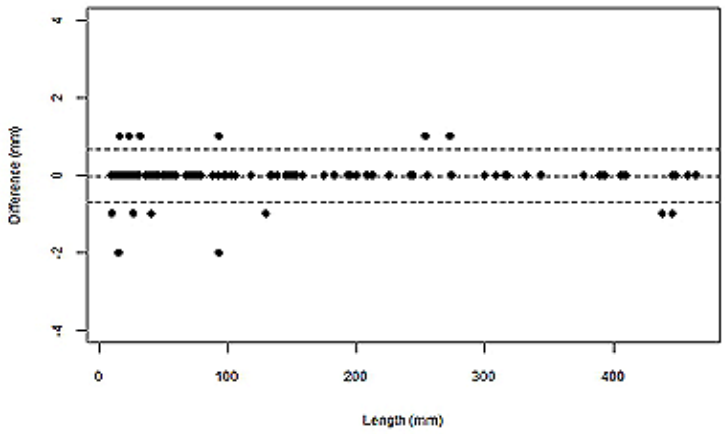


Figure 4.2 – Bland-Altman plot illustrating the intra-observer agreement of measurements.

4.2 Exploratory statistics

4.2.1 Correlations

Correlation coefficients demonstrated the relationships among the variables that ranged from 0.00 to 0.99 (Figure 4.3). Sacabr was the only variable to show negative correlations, with coefficients between -0.331 and -0.017. Substantial positive correlations (r >0.8) were noted between the maximum lengths of the long bones. Strong correlations (0.96-0.98) were also observed between the lengths of paired bones, namely the radius and ulna, and the tibia and fibula. In addition, some measurements associated with skeletal elements that articulate with each other, such as femebr and tibpeb, presented with correlations greater than 0.8.

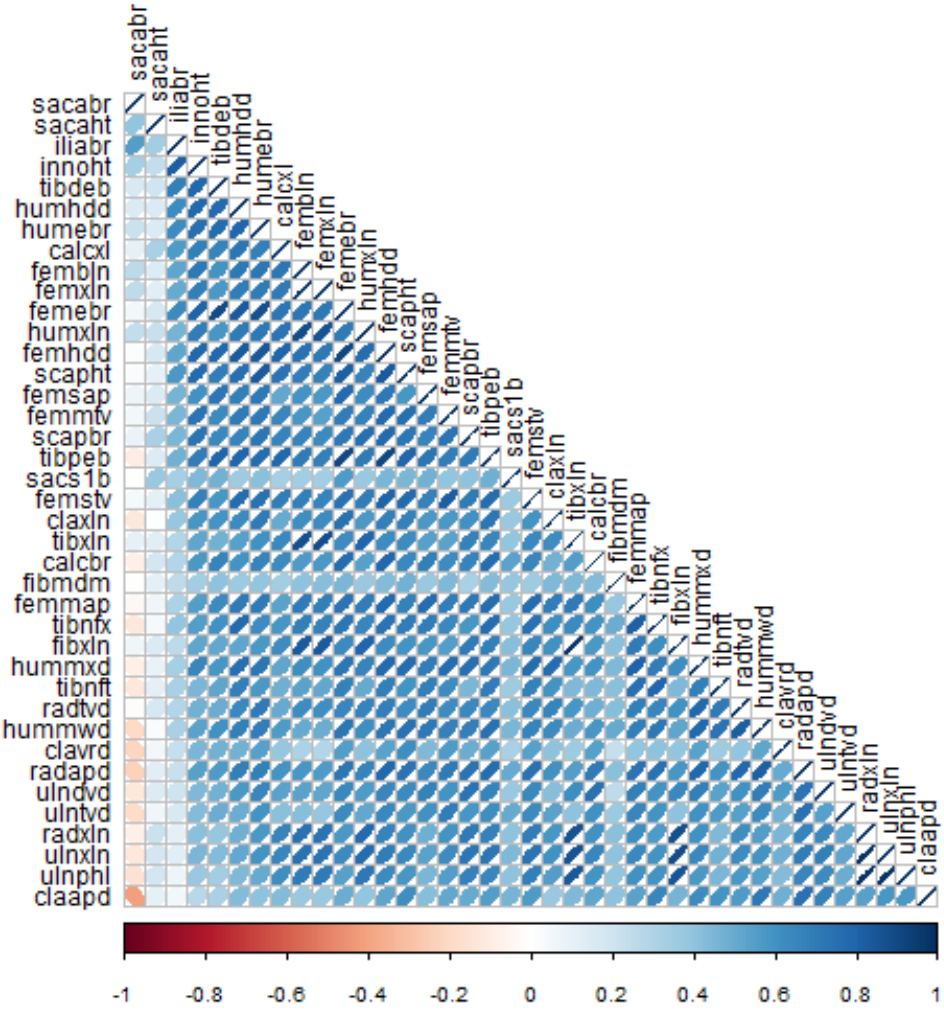


Figure 4.3 – Correlation plot depicting the relationship among the variables. Red indicates negative correlations, blue indicates positive correlations, and the size of the ellipse indicates the magnitude of the correlation.

4.2.2 ANOVA, Tukey's HSD and MANOVA

Table 4.3 presents the desriptive statistics including sample size, mean and standard deviation for each measurement. The ANOVA detected statistically significant differences in ancestry for all variables, except clavrd, claapd, radapd, ulndvd and calcbr. None of the variables besides iliabr, calcxl, and calcbr were found to have significant interaction between ancestry and sex (Appendix II). When further explored with Tukey's HSD, results revealed significant differences between all three groups for humxln, femxln, fembln, femebr, femmap, tibpeb, tibdeb and fibmdm (see Appendix III). As expected, the variables without noteworthy differences in the ANOVA demonstrated considerable overlap among all three groups in the Tukey's HSD. Several variables displayed similar means and subsequent lack of statistical differences for two of the three groups (Table 4.2). Overall, black South Africans were intermediate in size between coloured and white South Africans and demonstrated substantial overlap with the other two groups (Appendix IV). With MANOVA, all bones were found to be significant with regard to ancestry. The ulna, pelvis and femur also demonstrated a significant interaction between ancestry and sex (Appendix V).

Table 4.2 – Break down of group overlap of measurement means based on ANOVA and Tukey's HSD. See Table 3.2 for measurement abbreviations. Abbreviations: B = black; C = coloured; W = white.

No Overlap	All groups	B and C*	B and W*
humxln	clavrd	scapht	claxln
femxln	claapd	scapbr	radxln
fembln	radapd	humebr	ulnxln
femebr	ulndvd	humhdd	ulnphl
femmap	calcbr	hummxd	sacs1b
tibpeb		hummwd	tibxln
tibdeb		radtvd	tibnft
fibmdm		ulntvd	fibxln
		sacaht	
		sacabr	
		innoht	
		iliabr	
		femhdd	
		femsap	
		femstv	
		femmtv	
		tibnfx	

4.3 Sectioning points

Univariate sectioning points were created using the population means that demonstrated significant differences with the ANOVA tests. No sectioning points were formulated for the variables that presented substantial overlap ($p > 0.05$) with the ANOVAs and Tukey's HSD. Additionally – despite significant results with ANOVA – if population means were too similar (e.g. 28 mm and 30 mm) practical sectioning points were not created. Measurements with similar mean values that could not have sectioning points created included radtvd, ulntvd, femsap, femmap, tibpeb, tibdeb and fibmdm. Groups were pooled when overlap was observed between two of the three groups (e.g. black and coloured were similar to each other but both demonstrated significant differences to white). In these instances the sectioning points classify individuals into two groups rather than three (Appendix VI).

Classification accuracies utilising the sectioning points ranged from 43% to 87%, with femxln and fembln performing the worst and iliabr performing the best (Table 4.4). As would be expected, the sectioning points with two groups pooled obtained higher accuracies than the three-way sectioning points (humxln, femxln, fembln, femebr). As the intermediate group, the positive predictive value for the black South Africans were overall fairly low compared to the other groups. Thus, black South Africans were misclassified most often and obtained the lowest correct classifications (13% for femebr). Only iliabr and innoht were capable of obtaining classification accuracies greater than 75%.

Table 4.3 - Summary statistics showing the means and standard deviations for black, white and coloured South Africans for each measurement. See Table 3.2 for measurement abbreviations.

		Black		White		Coloured	
Measurement	***n***	**mean**	**s.d**	**mean**	**s.d**	**mean**	**s.d**
claxln	259	148.2	10.7	151.4	11.8	143.6	11.5
claapd	268	11.9	1.8	12.0	1.7	11.7	1.6
clavrd	271	9.9	1.5	10.1	1.6	9.9	1.5
scapht	307	145.1	12.7	153.4	13.3	141.9	11.3
scapbr	341	98.2	8.9	102.5	8.2	96.2	7.9
humxln	353	309.8	21.3	325.5	20.5	301.8	20.6
humebr	357	58.8	4.6	60.9	5.2	57.3	5.2
humhdd	353	41.2	3.6	45.8	3.6	41.0	3.9
hummxd	355	21.2	2.1	22.1	2.3	20.9	2.2
hummwd	356	16.7	1.9	17.5	2.3	20.9	2.2
radxln	350	244.2	19.5	242.2	16.9	230.1	18.1
radapd	353	11.7	1.5	11.8	1.5	11.5	1.4
radtvd	354	14.4	1.6	15.3	2.0	14.1	1.7
ulnxln	347	262.9	19.8	259.7	17.7	246.8	18.9
ulndvd	354	14.5	1.8	14.7	2.1	14.4	2.0
ulntvd	357	14.6	2.0	15.3	2.0	14.2	1.8
ulnphl	352	230.7	18.1	226.5	15.2	216.4	16.9
sacaht	195	101.15	10.1	106.4	11.3	97.8	7.8
sacabr	285	90.3	6.2	99.6	6.9	92.4	6.4
sacs1b	257	46.3	4.3	46.1	3.0	42.8	3.3
innoht	339	196.9	12.7	216.7	13.1	193.4	13.1
iliabr	336	145.8	8.3	164.0	8.8	143.9	8.3
femxln	352	441.0	25.7	456.6	26.3	429.0	27.2
fembln	352	437.3	26.0	453.5	26.8	425.8	27.0
femebr	350	75.8	5.6	79.9	5.4	73.9	5.6
femhdd	354	43.0	3.6	45.4	3.7	42.6	3.9
femsap	331	26.5	2.1	28.4	2.9	26.1	2.1
femstv	344	30.0	2.7	31.8	2.5	30.8	2.9
femmap	328	28.4	2.6	29.4	2.6	27.4	2.3
femmtv	347	26.0	2.5	27.5	2.2	25.4	2.2
tibxln	345	375.6	24.6	378.0	24.8	357.8	26.8
tibpeb	337	72.1	5.6	74.7	5.5	69.9	5.3
tibdeb	345	46.4	3.4	49.5	3.5	45.0	3.5
tibnfx	326	33.7	3.4	35.1	3.0	33.0	2.8
tibnft	347	24.3	2.9	25.1	2.8	23.1	2.6
fibxln	324	367.9	25.1	371.3	24.3	349.6	24.5
fibmdm	327	15.0	1.8	15.6	1.7	14.4	1.5
calcbr	321	41.5	3.1	41.8	3.1	42.1	3.63

Table 4.4 - Percentage of correct group membership for univariate sectioning points, overall and separated per group. Bold indicates the highest correct group classification.

Variable	Black		Coloured		White		Total	
	N	%	N	%	N	%	N	%
iliabr	114	86	109	**89**	113	85	336	**87**
innoht	116	71	105	74	118	**79**	339	75
sacabr	100	**78**	82	68	103	72	285	73
humhdd	119	**71**	117	68	117	61	353	67
tibxln	114	61	115	64	116	**66**	345	64
radxln	120	63	116	**64**	114	60	350	63
ulnxln	118	**66**	110	63	119	61	347	63
scapht	104	57	98	**71**	105	58	307	62
scapbr	114	63	111	**65**	116	58	341	62
ulnphl	117	**66**	116	61	119	60	352	62
sacaht	80	58	57	**75**	58	55	195	62
humebr	120	58	117	**66**	120	50	357	58
sacs1b	93	54	68	**79**	96	46	257	58
femhdd	119	54	117	56	118	**57**	354	56
claxln	72	40	104	**70**	83	48	259	55
hummxd	119	58	116	**62**	120	47	355	55
fibxln	115	33	105	64	104	**67**	324	54
hummwd	120	**49**	118	46	118	48	356	48
humxln	120	20	116	59	117	**61**	353	46
femebr	120	13	115	**69**	115	53	350	45
femxln	119	15	116	**59**	117	55	352	43
fembln	119	17	116	**59**	117	56	352	43

4.4 Multivariate analyses

4.4.1 Bone models

Multivariate models were created by combining all measurements taken for each bone. Bone models could not be created for the calcaneus because calcaneus length (calcxl) was not available for coloured South Africans. Using LDA, the classification accuracies ranged from 46% to 62%. When stepwise variable selection was applied to the bone models, the accuracies presented with a slight improvement with correct classifications between 47% and 63% (Table 4.5). The bone model using sacral measurements performed the best (Figure 4.4). With the exception of femxln, all maximum lengths were selected as key variables in the stepwise models for the long bones, while midshaft diameters and breadth measures were generally eliminated from the models. With the irregular bones, namely the scapula, pelvis and sacrum the height measures were removed, while the breadth measurements were deemed more discriminative for classification. Overall the coloured South Africans obtained the highest number of correct classifications (66%). The misclassification rates were highly variable, with at least one group per model showing very low positive predictive values (Table 4.6). The pattern of classification accuracies suggests coloured and white South Africans have the greatest degree of between-group variation.

Because FDA automatically executes variable selection, some variables may have been removed from the bone models. FDA produced classification accuracies that varied from 41% with the fibula to 66% with the sacrum (Table 4.5; Figure 4.5). Overall, FDA acquired more correct classifications than LDA for the majority of the models. With the models where LDA outperformed FDA, the difference in accuracy was small. Furthermore, variable selection differed between FDA and LDA. FDA placed slightly less variable importance on maximum bone lengths and more on epiphyseal breadths such as humebr, femebr and tibdeb. Unlike the LDA stepwise variable selection, FDA also included scapht, sacaht and innoht in the bone model functions.

Table 4.5 - Overall classification accuracies (%) for each multivariate bone model and classification method. Bold indicates the highest correct classification.

	n	LDA	Stepwise LDA	FDA
Clavicle	253	46	**47**	46
Scapula	294	46	**47**	42
Humerus	343	57	58	**62**
Radius	344	51	**53**	**53**
Ulna	339	53	53	**54**
Sacrum	160	61	63	**66**
Pelvis	331	**62**	61	61
Femur	288	57	**60**	**60**
Tibia	293	53	56	**57**
Fibula	312	47	**48**	41

Table 4.6 – Overall classification accuracies (%) for the multivariate bone models comparing LDA and FDA separated according to population. Bold indicates the highest correct classification.

Model	Black		White		Coloured	
	LDA	FDA	LDA	FDA	LDA	FDA
Clavicle	18	38	57	43	**59**	**50**
Scapula	23	31	56	**48**	**59**	40
Humerus	49	58	**70**	**70**	51	57
Radius	46	**56**	47	50	**59**	53
Ulna	50	**55**	45	53	**65**	54
Sacrum	54	**74**	**66**	69	65	55
Pelvis	42	50	**86**	**77**	58	53
Femur	45	57	62	**63**	**68**	60
Tibia	47	52	**58**	57	55	**63**
Fibula	31	20	48	39	**65**	**51**

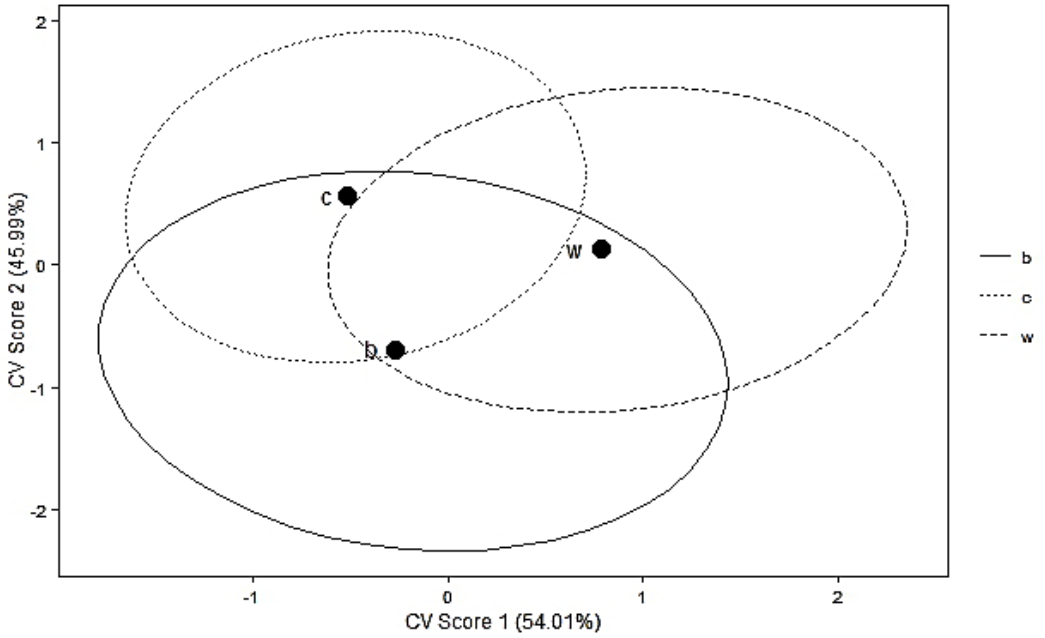


Figure 4.4 – Visualisations of the group means and variation using LDA for the best model, the sacrum with 63% classification accuracy.

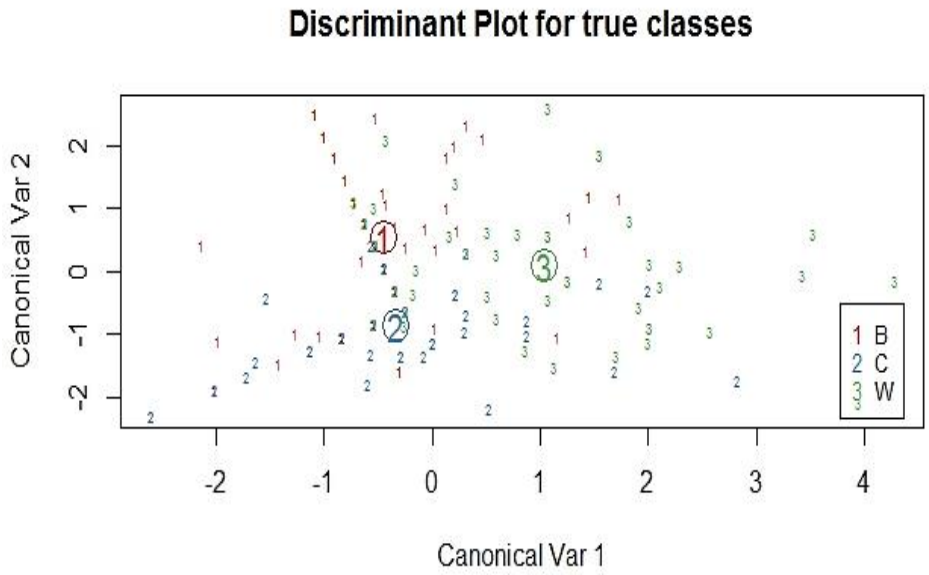


Figure 4.5 – Visualisations of group separation and variation using FDA for the best model, the sacrum, with 66% classification accuracy.

4.4.2 Multivariate subsets

A series of multivariate subsets were created to present accuracies for possible combinations of predictor variables. The classification accuracies ranged from 63% to 85% using LDA, and 62% to 87% with FDA (Table 4.7). While the upper and lower limbs obtained comparable accuracies (70-72% and 68-72%, respectively), when individually combined with the shoulder girdle and pelvic girdle, the lower limb performed better. The higher percent correct observed with the combined lower limb model most likely results from the inclusion of pelvic variables, which were identified as the best univariate classifiers. The proximal elements subset (68-73%) outperformed the distal elements subset (56-62%). With each distal skeletal element, the breadth measures demonstrated significant overlap for black and coloured South Africans and the length variables showed overlap between black and white South Africans. Thus, as a consequence of the variable overlap, the model cannot sufficiently separate the groups, resulting in low classification accuracies. Ultimately the all-measurement model, showcasing a collection of variables from several different bones, obtained the highest accuracy with 85% to 87%. The degree of variation among the three groups is shown in Figures 4.6 and 4.7. The large ellipses indicate higher levels of within-group variation for black and coloured South Africans than the overall smaller ellipse for white South Africans.

For model creation, both LDA and FDA generally retained maximum length variables and epiphyseal breadths when stepwise selection/variable reduction was conducted. The pelvic variables (iliabr and innoht) were retained in all models utilising the pelvis. The two measurements that were repeatedly ranked as key variables based on the variable importance measure used with FDA were iliabr and humhdd. The group classification accuracies shows that white South Africans were most frequently correctly classified. Confusion matrices revealed that black and coloured South Africans tended to misclassify as one another, while white South Africans would misclassify as either coloured or black. With the subsets, white South Africans consistently achieved the highest positive predictive values. As seen with the bone models, FDA demonstrated superior accuracies to LDA for all models, bar the distal elements subset. Overall the subsets yielded moderate to high classification accuracies, with the least accurate model – the distal elements subset (56-62%) – outperforming the majority of the univariate sectioning points as well as the multivariate bone models.

Table 4.7 - Overall classification accuracies (%) for each multivariate subset and classification method. Bold indicates the highest correct classification. Refer to Table 3.3 for subset details.

	n	LDA	stepwise LDA	FDA
Upper limb	316	70	70	**72**
Upper limb + Shoulder girdle	183	68	72	**78**
Lower limb	225	65	68	**72**
Lower limb + Pelvic girdle	177	76	80	**82**
Pelvis + Femur	265	72	73	**75**
Proximal elements	280	65	68	**73**
Distal elements	247	56	**63**	62
All-measurement	140	84	85	**87**

Table 4.8 – Overall classification accuracies (%) for each multivariate subset and classification method, separated according to population. Bold indicates the highest correct classification. Refer to Table 3.3 for subset deatails.

Model	Black		White		Coloured	
	LDA	FDA	LDA	FDA	LDA	FDA
Upper limb	69	72	**77**	**80**	63	64
Upper limb + shoulder girdle	68	76	**76**	**84**	61	73
Lower limb	62	**75**	**72**	**72**	63	68
Lower limb + pelvic girdle	63	87	**81**	**96**	78	79
Pelvis + Femur	63	73	**84**	**89**	73	67
Proximal elements	63	73	**70**	**78**	63	69
Distal elements	52	**64**	**61**	60	57	**64**
All-measurement	78	90	**92**	**94**	83	77

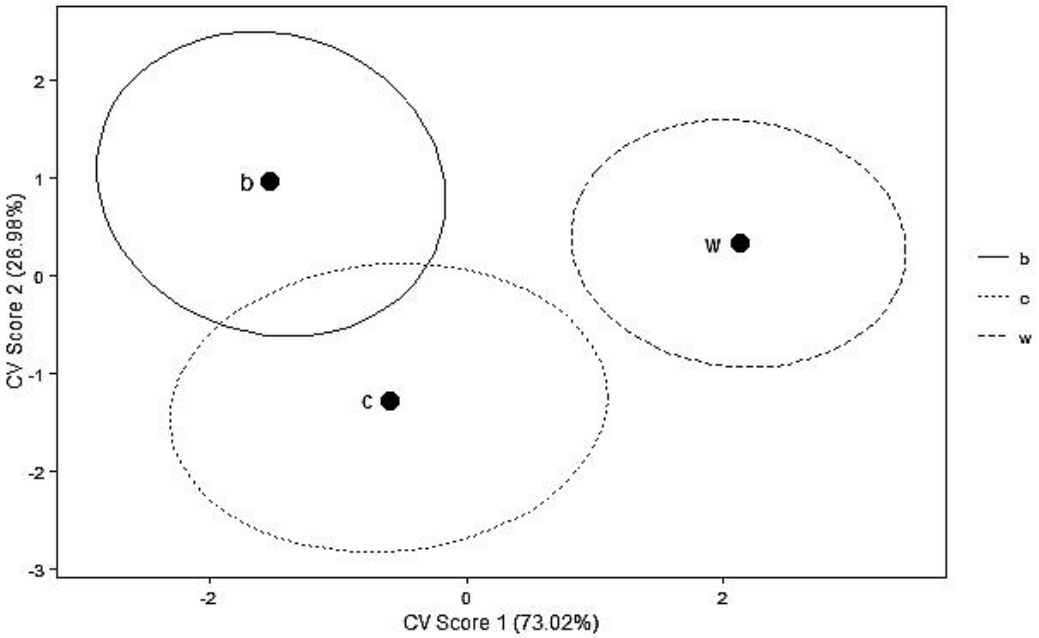


Figure 4.6 – Visualisations of the group means and variation using LDA for the best subset, the all-measurement model, with 85% classification accuracy.

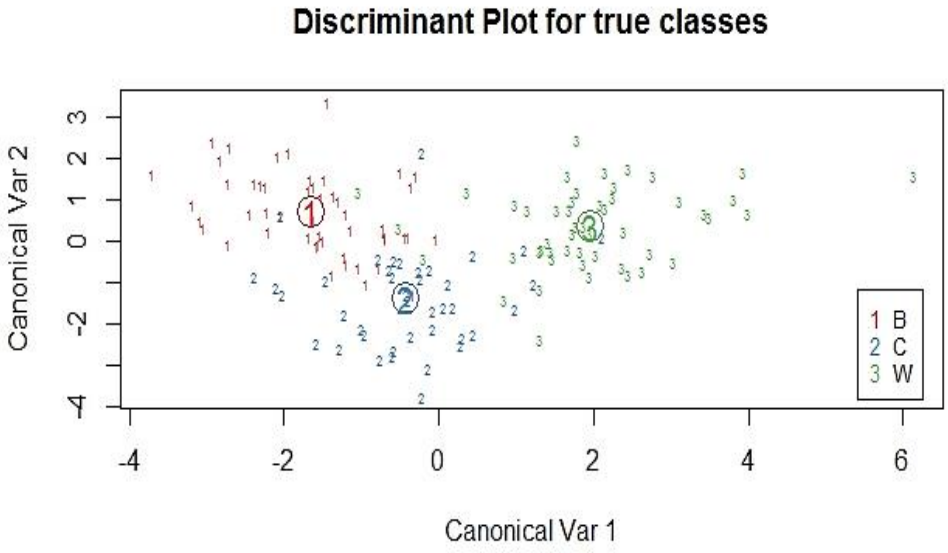


Figure 4.7 – Visualisations of group separation and variation using FDA for the best subset, the all-measurement model with 87% classification accuracy.

CHAPTER 5: DISCUSSION

The current study demonstrated postcraniometric differences among three major social races in South Africa. Differences among the groups can be attributed to varying population origin(s) of each group and to different experiences and selective influences maintained through positive assortative mating within the last hundred years, and particularly during the period of apartheid in South Africa.

Sexual relations and marriage among white and “non-white” groups was banned under the Prohibition of Mixed Marriages Act of 1949 and the Immorality Act of 1950 (Jacobson et al., 2004). Although the laws were abolished in 1991 with the end of apartheid, social constraints continue to limit interracial marriages (Jacobson and Heaton, 2008). Based on the population history, black and coloured South Africans share more morphological similarities, attributable to less restricted gene flow, resulting in increased group overlap. White South Africans present with greater differences because of limited gene flow, different socio-economic conditions and geographical origins that are most dissimilar to that of the other two groups. White South Africans generally presented with less within-group variation, which suggests that the group is fairly homogeneous. Greater levels of variability in the distribution of black and coloured individuals around the respective group centroids indicate more heterogeneity within each group. Furthermore, the considerable size difference between white South Africans and the black and coloured groups resulted in higher levels of between-group variation, making white South Africans easier to distinguish from the other two ancestral groups.

The variable geographic origins of the three populations, their historical circumstances, and the modern socio-political situation in South Africa provides a foundation from which postcraniometric traits were evaluated and deemed useable in a forensic context; classification methods are compared; and the application of this research to forensic anthropology are discussed.

5.1 Best indicators for ancestry estimation

Despite the significant overlap between the groups, which is expected in such a variable population, the postcranial skeleton has proved to be proficient in distinguishing the three groups with high accuracy. The best results were achieved using combinations of different variables from several skeletal elements. For instance, black South Africans have fairly

narrow diaphyses and pelves – similar to coloured South Africans – however, long limbs distinguish them from coloured South Africans, who display shorter limbs. Both of these groups can be distinguished from white South Africans who have a combination of long limbs, robust epiphyseal and diaphyseal dimensions and large pelves.

Craniometric research has shown that coloured South Africans typically demonstrate intermediacy between black and white South Africans which is contrary to results from the postcrania where they are shown to be much smaller in metric dimensions (L'Abbé et al., 2013b). The intermediacy observed with the cranium caused black and white groups to primarily misclassify as coloured, but rarely misclassify as one another. The pattern of misclassification in the postcrania depended largely on the predictor variables included in the analyses. When more epiphyseal and midshaft measurements are included in the model creation, black and coloured South Africans are more likely to misclassify as one another rather than into the white group. In contrast, when the lengths of the distal limb elements are included, black and white South Africans are less likely to misclassify into the coloured group. Even though white South Africans were the most distinct of the three groups, the group displays more overlap in the postcrania with black and coloured South Africans than observed with the cranium.

The postcranial variability observed among modern South Africans is largely related to differences in size. The maximum lengths of the long bones were considered some of the most discriminatory variables in the multivariate models. Long bone length and stature has shown to be highly correlated. The inclusion of length measurements in the multivariate classification models is therefore expected as a strong correlation with stature exists, which in turn demonstrates distinct group differences among South Africans (Steyn and Smith, 2007). Similar to the results of previous research, white South Africans presented with overall larger postcranial dimensions than either black and/or coloured South Africans. Coloured South Africans had the smallest means for all variables, except calcbr for which they presented with the largest values of all three groups. Black South Africans, who were intermediate in size for all variables except the lengths of the distal limb elements, demonstrated substantial overlap with the other groups, subsequently resulting in increased misclassifications.

The pelvis was consistently identified as one of the best indicators for ancestry, where the inclusion of pelvic variables in the models typically resulted in increased classification accuracies. Published literature reports several theories as to what causes variation in the

pelvic girdle (e.g. Angel, 1976; Jordaan, 1976; Pearson, 2000). Recent studies describe the possible effects of neutral evolutionary processes, such as mutation and genetic drift, resulting in population signatures. In other words, where neutral forces are assumed to be the major contributing factors to patterns of phenotypic variation rather than selective forces, quantitative traits (such as postcranial measurements) can be used to elucidate past population history and ancestry (Betti et al., 2012; von Cramon-Taubadel, 2013). The cranium is believed to be largely influenced by neutral forces, which could be the reason for the accuracy with which the cranium can be used to distinguish among population groups (Manica et al., 2007; von Cramon-Taubadel, 2013). Similarly, Betti et al. (2012) argue that pelvic morphology is mainly driven by neutral evolutionary processes, preserving a neutral demographic signal. The effects of population history are evident in the pelvis, thus the pelvis is useful for distinguishing groups with different geographic and social histories, such as socially defined South African groups (Betti et al., 2012). The differences in pelvic morphology resulting from inherent population signatures warrant extensive research regarding the pelvic girdle and evolution of pelvis shape among various populations.

Contrary to the pelvis, long bones have been found to be substantially correlated with climate (Betti et al., 2012). Therefore, as climatic selection has obliterated the population signatures, long bones should – in theory – not be of much use for group separation. Yet, the current study found that the maximum lengths and epiphyseal breadths of long bones, variables of which the dimensions are typically associated with climate, were considered some of the most discriminatory among the South African groups and thus useful for ancestry estimation. The extent to which aspects of postcranial variation are the result of neutral evolutionary forces has not been sufficiently addressed and requires further research (von Cramon-Taubadel and Weaver, 2009).

5.2 Comparison of classification methods

Univariate sectioning points derived from the postcranial skeleton are frequently employed to facilitate the estimation of sex (Steyn and İşcan, 1999; Spradley and Jantz, 2011; Tise et al., 2013); however, univariate analyses are generally not used for the estimation of ancestry. While sectioning points are quite useful to classify an individual into one of the dichotomous categories of sex, univariate measures struggle to capture the wide range of variation and overlap associated with ancestry. The majority of the variables presented with substantial overlap so that two of the three groups were pooled together. For the breadth measures, this

includes black and coloured South Africans while the maximum length measurements of the distal limb segments indicated that black and white South Africans are more similar to each other. By pooling the overlapping, groups the sectioning points can only effectively distinguish between two groups – namely coloured and white South Africans. As black South Africans are intermediate in size, they were always merged with either white or coloured South Africans depending on the variable in question. Thus, as a result of the overlap observed among the groups, the univariate sectioning points are not capable of identifying black South Africans as a group on their own. As black South Africans form the majority of the South African population, a problem arises if the black group cannot be distinguished from the white or coloured groups.

Ultimately, the classification accuracies for the majority of the univariate sectioning points were very low, demonstrating the limited discriminative power of single variables in separating populations. Coloured South Africans yielded the highest number of correct classifications and black South Africans had the lowest accuracies. The two-way classifications (*i.e.* where two groups were pooled due to substantial overlap) yielded better results than the three-way classifications (where ANOVA and Tukey's HSD presented with no group overlap). This decrease in accuracy and discriminative power is not unexpected when attempting to classify an individual into one of three groups as opposed to a classification into two groups. Although, as previously stated, merging groups limits the practical applicability of the sectioning points. Single measurements are only representative of one dimension which cannot provide an accurate estimation of an individual's ancestral background. Essentially the large amount of variation and group overlap associated with ancestry confounds the univariate approach, resulting in increased misclassifications.

While the bone models are superior to the univariate sectioning points as they are capable of discriminating among all three groups rather than just two groups, the observed accuracies are too low to be considered reliable for a classification model. Thus, the specific multivariate combinations of variables utilised in each bone model are not necessarily the best combination for separating the groups. The multivariate subsets achieved higher accuracies than both the univariate and multivariate bone models. The subsets incorporate numerous measurements from more than one skeletal element, producing greater discriminatory combinations. Thus, the inclusion of a larger number of variables was better for maximising group differences, subsequently producing increased classification accuracies (Pietrusewsky,

2008; Ousley and Jantz, 2012). While the bone models used multivariate combinations, only a small number of variables (two to eight measurements) were included in each model. Furthermore, the variables combined for each bone model is not necessarily the best combination of postcranial variables for group separation. The subsets tend to use more measurements, therefore providing more variables from which the optimal combination can be selected for the function to identify group boundaries. For example, when looking at the variable importance and the measurements chosen with stepwise selection, bone lengths and epiphyseal breadths typically carried the most weight. The all-measurement model included several length and epiphyseal variables in the model creation, which is why the model yielded the greatest classification accuracy. Ultimately, the multivariate approach is better than the univariate approach, and multivariate subsets are better at maximising between-group variation than multivariate bone models.

The current study assessed the classification potential of FDA compared to the more traditionally employed LDA. In support of the literature (Hastie, 1994; Stull, 2013), the current study noted that FDA is capable of outperforming LDA. When subjected to FDA the multivariate subsets yielded accuracies between 2% and 6% higher than with LDA. Therefore, FDA is a suitable classification technique for the estimation of ancestry and can be utilised as an alternative to LDA. LDA is arguably the most popular classification technique and until recently was the only method applied for forensic osteometric analysis (Mercer, 2013). A drawback associated with LDA is the numerous assumptions regarding the data that need to be fulfilled. Although LDA can produce satisfactory results if the assumptions are not met, the statistical validity and interpretability of the models may be adversely affected (Ousley and Jantz, 2012). Quadratic discriminant analysis (QDA) is often proposed as an alternative technique when the application of LDA is not possible. However, QDA, with its own parametric requirements, has been shown to produce overfitted data and performs poorly with data not included in the reference sample (Mercer, 2013). The major advantage of FDA is that the function creates less rigid boundaries between groups and involves fewer assumptions (Le Roux and Gardner, 2006). Thus, FDA is a better classifying method than LDA for the estimation of ancestry using the postcrania.

5.3 Application in forensic anthropology

The standards created in the current study follows the Scientific Working Group for Forensic Anthropology (SWGANTH) guidelines for best practice, in that the study was conducted on a

forensically relevant reference sample of sufficient size using appropriate statistics with reported accuracies. Therefore the postcranial data can be employed as an additional technique to corroborate results obtained with the cranium, or used to assess ancestry when the cranium is not available. As the three major groups – black, white and coloured South Africans – were included in the sample, the standards can be applied universally throughout South Africa and is not limited to a specific geographic region. The large number of variables analysed makes the formulation of specific discriminant function formulae impractical. With the use of specialised software becoming increasingly accessible and popular, the data can be incorporated as a custom database in Fordisc (FD3). Using FD3 would be more useful than limited formulae, as the software can create discriminant functions applying numerous combinations of variables on a case-specific basis. Thus, enabling accurate analysis of incomplete and fragmentary remains without having to resort to inferior univariate techniques. Furthermore, FD3 provides numerical results specific to each analysis, such as cross-validated accuracies and posterior and typicality probabilities (Ousley and Jantz, 2013). The numerical results are essential in evaluating the classification performance and validity of the models, which is required to satisfy the *Daubert* criteria. However, FD3 only makes use of LDA. The current results demonstrating the superior performance of FDA warrants the production of additional software to facilitate the application of FDA.

CHAPTER 6: CONCLUSION

The current study is the first to assess the entire postcranial skeleton to estimate ancestry in a modern South African population. Through the use of numerous exploratory, univariate and multivariate approaches the results demonstrated that postcraniometric differences exist among peer-reported black, white and coloured South Africans. The majority of the group variation is attributed to differences in size and overall robusticity: white South Africans are overall larger and more robust than black and coloured South Africans. The heterogeneity and overlap observed in the sample reflects past admixture and intermingled histories despite recent population separation.

When using the group mean-based univariate approach three pelvic measurements, namely iliac breadth, innominate height and anterior breadth of the sacrum yielded the highest accuracies. This suggests that the pelvis is the best univariate indicator of ancestry; however the univariate sectioning points for the pelvis can only be used to distinguish white South Africans from the other groups. As with the univariate approach, the sacrum and pelvis obtained the best results among the multivariate bone models. Although the bone models could distinguish between all three groups, the accuracies were fairly low. Thus, the variable combinations included in the bone models cannot sufficiently separate the groups. By including more variables, the multivariate subsets yielded higher accuracies than both the univariate sectioning points and the bone models.

FDA generally outperformed LDA, especially when using multivariate subsets. Because FDA can create more flexible boundaries this technique tends to be more useful in distinguishing between groups when a considerable amount of variable overlap is present. Overall the highest classification accuracy observed in this study (stepwise-selected subset with 84% to 87%) is comparable to the accuracies noted in previous multivariate studies conducted on postcrania as well as the cranium.

While the range of variation and group overlap should be acknowledged, this study has shown that the postcranial skeleton can be used as a tool to estimate ancestry in South Africa with high accuracy. The results prove the classification potential of the postcranial skeleton. The large amount of measurements analysed in the study makes it impractical to create structured discriminant function formulae for all possible variable combinations. As is, the data can be uploaded as a custom postcranial database in FD3 where LDA models can be

conducted on a case-specific basis. However, the superior results obtained with FDA warrants the development of a similar software program that can utilise this statistical technique.

REFERENCES

Abdi H, Williams LJ. 2010. Principal component analysis. WIREs Comp Stats 2:433-459.

Adams BJ, Byrd JE. 2002. Interobserver variation of selected postcranial skeletal elements. J Forensic Sci 47:1193-1202.

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

Adjeroh D, Cao D, Piccirilli M, Ross A. 2010. Predictability and correlation in human metrology. Information Forensics and Security (WIFS): IEEE International Workshop. p. 1-6.

Agyemang C, Bhopal R, Bruijnzeels M. 2005. Negro, Black, Black African, African Caribbean, African American or what? Labelling African origin populations in the health arena in the 21st Century. J Epidemiol Community Health 59:1014-1018.

Albanese J, Cardoso HF, Saunders SF. 2005. Universal methodology for developing univariate sample-specific sex determination methods: An example using the epicondylar breadth of the humerus. J Archaeol Sci 32:143-152.

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

Albrecht GH. 1980. Multivariate analysis and the study of form, with special reference to canonical variate analysis. Amer Zool 20:679-693.

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

Angel JL. 1976. Colonial to modern skeletal change in the USA. Am J Phys Anthropol 45:723-735.

Anholts A. 2013. Secular trends in the height and weight of South African children aged 6 to 10 years [Honnours thesis]. University of Pretoria.

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

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

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.

Bennett KA, Osborne RH. 1986. Measurement reliability in Anthropometry. Hum Biol 58:751-759.

Betti L, von Cramon-Taubadel N, Lycett SJ. 2012. Human pelvis and long bones reveal differential preservation of ancient population history and migration out of Africa. Hum Biol 84:139-152.

Bergmann C. 1847. Über die Verhältnisse der Wärmeökonomie der Tiere zu ihrer Größe. Göttinger Studien 3:595-708.

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

Bidmos MA. 2008. Stature reconstruction using fragmentary femora in South Africans of European descent. J Forensic Sci 53:1044-1048.

Bidmos MA, Asala SA. 2003. Discriminant function sexing of the calcaneus of the South African whites. J Forensic Sci 48:1213-1218.

Bidmos MA, Asala SA. 2005. Calcaneal measurement in estimation of stature of South African blacks. Am J Phys Anthropol 126:335-342.

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

Boas F. 1912. Changes in the bodily form of descendants of immigrants. Am Anthropol 14:530-562.

Bogin B, Loucky J. 1997. Plasticity, political economy, and physical growth status of Guatemala Maya children living in the United States. Am J Phys Anthropol 102:17–32.

Bophal R, Donaldson L. 1998. White, European, Western, Caucasian or what? Inappropriate labelling in research on race, ethnicity and health. Am J Public Health 88:1303-1307.

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

Broca P. 1878. Sur les indices de largeur de l'omoplate chez l'homme, les singes et dans la série des mammiféres. Bull Soc Anthrop.

Brown JL. 2006. Morphological variation of the proximal femur in selected skeletal remains [dissertation]. Wichita State University.

Buck LT, Stock JT, Foley RA. 2010. Levels of intraspecific variation within the catarrhine skeleton. Int J Primatol 31:779-795.

Bunning PSC, Barnett CH. 1965. A comparison of adult and foetal talocalcaneal articulations. J Anat 99:71-76.

Butte NF, Wong WW, Adolph AL, Puyau MR, Vohra FA, Zakeri IF. 2010. Validation of cross-sectional time series and multivariate adaptive regression splines models for the prediction of energy expenditure in children and adolescents using double labelled water. J Nutr 140:1516-1523.

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, Kgamphe JS, Leschner KF, Farrant PJ. 1992. Urban-rural differences in the growth of South African Black children. Ann Hum Biol 19:23-33.

Cameron N. 2003. Physical growth in a transitional economy: The aftermath of South African apartheid. Econ Hum Biol 1:29-42.

Campbell NA. 1978. Multivariate analysis in biological anthropology: some further considerations. J Hum Evol 7:197-203.

Caspari R. 2010. Deconstructing race: Racial thinking, geographic variation, and implications for biological anthropology. In: Larsen CS, editor. A companion to Biological Anthropology. John Wiley and Sons. p.104-123.

Chen YS, Olckers A, Schurr TG, Kogelnik AM, Huoponen K, Wallace DC. 2000. mtDNA variation in the South African Kung and Khwe – and their genetic relationships to other African populations . Am J Hum Genet 66:1362-1383.

Christensen AM, Crowder CM. 2009. Evidentiary standards for Forensic Anthropology. J Forensic Sci 54:1211-1216.

Dayal MR, Steyn M, Kuykendall KL. 2008. Stature estimation from bones of South African whites. S Afr J Sci 104:124-128.

De Villiers H. 1968. The skull of the South African Negro: A biometrical and morphological study. Witwatersrand University Press, Johannesburg.

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.

DiBennardo R, Taylor TV. 1983. Multiple discriminant function analysis of sex and race in the postcranial skeleton. Am J Phys Anthrop 61:305-314.

Dirkmaat DC, Cabo LL, Ousley SD, Symes SA. 2008. New perspectives in forensic anthropology. Am J Phys Anthropol 137:33-52.

Edgar HJH. 2009. Biohistorical approaches to “race” in the United States: Biological distances among African Americans, European Americans and their ancestors. Am J Phys Anthropol 139:58-67.

Edgar HJH, Hunley KL. 2009. Race reconciled: How biological anthropologists view human variation. Am J Phys Anthropol 139:1-4.

Ellison GT, De Wet T. 1997. The use of “racial” categories in contemporary South African health research. A survey of articles published in the South African Medical Journal between 1992 and 1996. S Afr Med J 87:1671-1679.

Erasmus Z. 2012. Apartheid race categories: Daring to question their continued use. Transformation: Critical Perspectives on Southern Africa 79:1-11.

Eveleth PB, Tanner JM. 1976. Worldwide variation in growth. Cambridge University Press. p.222-240.

Farrally MR, Moore WJ. 1975. Anatomical differences in the femur and tibia between negroids and caucasoids and their effects upon locomotion. Am J Phys Anthrop 43:63-70.

Feldesman MR. 2002. Classification trees as an alternative to linear discriminant analysis. Am J Phys Anthropol 119:257-275.

Fenech AP. Tukey's method of multiple comparison in the randomized blocks model. Am J Statist Assoc 74:881-884.

Flower WH, Garson JG. 1879. On the scapular index as a race character in man. J Anat and Phys 14:13-17.

Franklin D, Freedman L, Milne N, Oxnard CE. 2007. Geometric morphometric study of population variation in indigenous southern African crania. Am J Hum Biol 19:20-33.

Franklin D, Cardini A, Oxnard CE. 2010. A geometric morphometric approach to the quantification of population variation in sub-Saharan African crania. Am J Hum Biol 22:23-35.

Friendly M. 2002. Corrgrams: Exploratory displays for correlation matrices. Am Stat 56:316-324.

Frisancho AR, Garn SM, Ascoli W. 1970. Unequal influence of low dietary intakes on skeletal maturation during childhood and adolescence. Am J Clin Nutr 23:1220–1227.

Gilbert BM. 1976. Anterior femoral curvature: Its probable basis and utility as a criterion of racial assessment. Am J Phys Anthrop 45:601-604.

Giles E, Elliot O. 1962. Race identification from cranial measurements. J Forensic Sci 7:147-157.

Gill GW. 1990. Introduction. In: Gill GW and Rhine S, editors. Skeletal attribution of race: Methods for forensic anthropology. University of New Mexico, Albuquerque. p.vii-xii.

Gill GW. 1990. Craniofacial Criteria in the Skeletal Attribution of Race. In: Reichs KJ, editor. Forensic Osteology: Advances in the identification of Human Remains. Charles C. Thomas, Springfield, IL. p.293-317.

Gill GW. 2009. Assessing ancestry (race) from the skeleton. In: Pickering R and Bachman D, editors. The use of Forensic Anthropology. 2nd ed. CRC Press. p.103-111.

Gill GW, Gilbert BM. 1990. Race identification from the midfacial skeleton: American blacks and whites. In: Gill GW and Rhine S, editors. Skeletal attribution of race: Methods for forensic anthropology. University of New Mexico, Albuquerque. p.47-53

Ginsburg E, Livshits G, Yakovenko K, Kobyliansky E. 1998. Major gene control of human body height, weight and BMI in five ethnically different populations. Ann Hum Genet 62:307-322.

Goto R, Mascie-Taylor CGN. 2007. Precision of measurement as a component of human variation. J Physiol Anthropol 26:256-256.

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

Gravlee CC. 2009. How race becomes biology: Embodiment of social inequality. Am J Phys Anthropol 139:47-57.

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

Green R, Chapman PM. 2011. The problem with indices. Marine Poll Bull 62:1377-1380.

Grivas CR, Komar DA. 2008. *Kumho, Daubert* and the nature of scientific inquiry: Implications for Forensic Anthropology. J Forensic Sci 53:771-776.

Haase RF, Ellis MV. 1987. Multivariate analysis of variance. J Couns Psychol 34:404.

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

Hanneman SK. 2008. Design, analysis and interpretation of method-comparison studies. AACN Adv Crit Care 19:223-234.

Harris EF, Smith RN. 2009. Accounting for measurement error: A critical but overlooked process. Arch Oral Biol 54:107-117.

Harvati K, Weaver TD. 2006. Human cranial anatomy and the differential preservation of population history and climate signatures. Anat Rec Part A 288:1225-1233.

Hastie T, Tibshirani R, Buja A. Flexible Discriminant Analysis by optimal scoring. J Am Statist Assoc 89:1255-1270.

Hastie T, Tibshirani R. 2013. mda: Mixture and flexible discriminant analysis.

Hastie T, Tibshirani R, Friedman J. 2009. The elements of statistical learning: Data mining, inference and prediction. 2nd ed. New York: Springer-Verlag.

Hefner JT. 2009. Cranial Nonmetric Variation and Estimating Ancestry. J Forensic Sci 54:985-995.

Henneberg M, Harrison GA, Brush G. 1998. The small child: Anthropometric and physical performance characterisctics of short-for-age children growing in good and in poor socio-economic conditions. Eur J Clin Nutr 52:286-291.

Henneberg M, LaVelle M. 1999. Varying effects of socio-economic categories on the growth of urban and rural South African “Cape Coloured” boys. Perspect Hum Biol 4:41-49.

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

Hitzeroth, HW. 1972. Fisiese Antropologie van die inheemse mense in Suidelike Afrika. Afrika-Instituut. p. 1-8.

Holliday TW, Falsetti AB. 1999. A new method for discriminating African-American from European-American skeletons using postcranial osteometrics reflective of body shape. J Forensic Sci 44:926-930.

Holliday TW, Hilton CE. 2010. Body proportions of circumpolar peoples as evidenced from skeletal data: Ipiutak and Tigara (Point Hope) versus Kodiak Island Inuit. Am J Phys Anthropol 142:287-302.

Holliday TW, Ruff CB. 2001. Relative variation in human proximal and distal limb segment length. Am J Phys Anthropol 116:26-33.

Hrdlička A. 1942. The adult scapula: Additional observations and measurements. Am J Phys Anthropol 29:363-415.

Huberty C. 1994. Applied discriminant analysis. New York: Wiley and sons.

İşcan MY.1983. Assessment of race from the pelvis. Am J Phys Anthrop 62:205-208.

İşcan MY, Cotton TS. 1985. The effect of age on the determination of race from the pelvis. J Hum Evol 14:275-282.

İşcan MY, Loth SR, King CA, Shihai D, Yoshino M. 1998. Sexual dimorphism in the humerus: A comparative analysis of Chinese, Japanese and Thais. Forensic Sci Int 98:17-29.

İşcan MY, Shihai D. 1995. Sexual dimorphism in the Chinese femur. Forensic Sci Int 74:79-87.

İşcan MY, Steyn M. 1999. Craniometric determination of population affinity in South Africans. Int J Legal Med 112:91-97.

Jacobson CK, Acheampong AY, Heaton TB. 2004. Inter-racial marriages in South Africa. J Comp Fam Stud 35:443-458.

Jacobson CK, Heaton TB. 2008. Comparative patterns of inter-racial marriage: Structural opportunities, third-party factors and temporal change in immigrant societies. J Comp Fam Stud 39:129-148.

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

Jordaan HVF. 1976. Neonatal and maternal cranial form. S Afr Med J 50:2064-2068.

Juul A. 2001. The effects of oestrogens on linear bone growth. Hum Reprod Update 7:124-134.

Kachigan SK. 1991. Multivariate statistical analysis: A conceptual introduction. Radius Press. p. 216-235.1991

Karsenty G, Wagner EF. 2002. Reaching a genetic and molecular understanding of skeletal development. Dev Cell 2:389-406.

Krogman WM, İşcan MY. 1986. The human skeleton in forensic medicine. Charles C Thomas. p.268-301.

Kuhn M. 2013. Caret: Classification and Regression Training.

Kurki HK. 2011. Pelvic dimorphism in relation to body size and body size dimorphism in humans. J Hum Evol 631-643.

Kurki HK, Ginter JK, Stock JT, Pfeiffer S. 2008. Adult proportionality in small-bodied foragers: A test of ecogeographic expectations. Am J Phys Anthropol 136:28-38.

Kurki HK, Pfeiffer S, Stynder DD. 2012. Allometry of head and body size in Holocene foragers of the South African Cape. Am J Phys Anthropol 147:462-471.

L'Abbé EN, Loots M, Meiring JH. 2005. The Pretoria Bone Collection: A modern South African skeletal sample. Homo 56:197-205.

L'Abbé EN, van Rooyen C, Nawrocki SP, Becker PJ. 2011. An evaluation of non-metric cranial traits used to estimate ancestry in a South African sample. Forensic Sci Int 209:195-e1.

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 6:1579-1583.

L'Abbé EN, Kenyhercz M, Stull KE, Ousley S. 2013b. Craniometric Assessment of Modern 20th Century black, white, and coloured South Africans. Proceedings of the 65th Annual Meeting of the American Academy of Forensic Sciences American Academy of Forensic Sciences. Washington, DC.

Lande, R. 1980. Sexual dimorphism, sexual selection and adaptation in polygenic characters. Evolution 34:292-305.

Le Roux NJ, Gardner S. 2006. A biplot based approach to discriminant analysis with categorical variables in the presence of reversals: Theory and methods. S Afr Stat J 40:1-31.

Lewontin RC. 1972. The apportionment of human diversity. Evol Biol 6:381–398.

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

Livingstone F. 1962. On the non-existence of human races. Curr Anthropol 3:279–281.

Livshits G, Roset A, Yakovenko K, Trofimov S, Kobyliansky E. 2002. Genetics of human body size and shape: Body proportions and indices. Ann Hum Biol 29: 271-289.

Loth SR, Henneberg M. 1996. Mandibular ramus flexure: A new morphologic indicator of sexual dimorphism in the human skeleton. Am J Phys Anthropol 99:473-485.

Louw GJ, Henneberg M. 1997. Lack of a secular trend in adult stature in South African males born between 1954 and 1974. Homo 47:54-61.

Lundy, JK, Feldesman, MR. 1987. Revised equations for estimating living stature from the long bones of the South African Negro. S Afr J Sci 83:54 -55.

Mallet YD, Coomans D, De Vel O. 1996. Recent developments in discriminant analysis on high dimensional spectral data. Chemometr Intell Lab 35:157-173.

Manica A, Amos W, Balloux F, Hanihara T. 2007. The effect of ancient population bottlenecks on human phenotypic variation. Nature 448:346-348.

McDowell JL, L'Abbé EN, Kenyhercz MW. 2012. Nasal aperture shape evaluation between black and white South Africans. Forensic Sci Int 222:397.e1-397.e6.

Meadows Jantz L, Jantz RL. 1999. Secular Change in Long Bone Length and Proportion in the United States, 1800–1970. Am J Phys Anthropol 110:57–67.

Meintjies-Van der Walt L. 2003. The proof of the pudding: The presentation and proof of expert evidence in South Africa. J Afr Law 47:88-106.

Mercer DA. 2013. Nonparametric discriminant analysis in forensic ancestry estimation: An assessment of utilized and alternative statistical methods [dissertation]. University of Tennessee.

Micklesfield LK, Norris SA, Nelson DA, Lambert EV, van der Merwe L, Pettifor JM. 2007. Comparisons of body size, composition and whole body bone mass between North American and South African children. J Bone Miner Res 22:1869-1877.

Milborrow S. 2011. Earth: Multivariate Adaptive regression spline models.

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

Moore-Jansen PM, Ousley SD, Jantz RL. 1994. Data collection procedures for forensic skeleton material. The University of Tennessee, Knoxville: Department of Anthropology. p. 70-82.

Morris AG. 1987. The reflection of the collector: San and Khoi skeletons in museum collections. S Afr Archaeol Bull 42:12-22.

Newman RW. 1975. Human adaptation to heat. In: Damon A, editor. Physiological Anthropology. New York : Oxford University Press. p.80-92.

Norris SA, Sheppard ZA, Griffiths PL, Cameron N, Pettifor JM. 2008. Current Socio-Economic measures, and not those measured during infancy, affect bone mass in poor urban South African children. J Bone Miner Res 23:1409-1416.

Ousley SD, Jantz RL. 2012. FORDISC 3 and Statistical Methods for Estimating Sex and Ancestry. In: Dirkmaat DC, editor. A companion to Forensic Anthropology. Blackwell Publishing LTD. p.311-329.

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

Ousley SD, Jantz RL, Freid D. 2009. Understanding Race and Human Variation: Why Forensic Anthropologists are good at Identifying Race. Am J Phys Anthropol 139:68-75.

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 Gen 19:411-419.

Patriquin ML. 2001. A comparative analysis of differences in the pelves of South African blacks and whites [dissertation]. University of Pretoria.

Patriquin ML, Steyn M, Loth SR. 2002. Metric assessment of race from the pelvis in South Africans. Forensic Sci Int 127:104-113.

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.

Pearson OM. 2000. Activity, Climate, and Postcranial Robusticity: Implications for Modern Human Origins and scenarios of Adaptive Change. Curr Anthropol 41:569-607.

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

Petersen DC, Libiger O, Tindall1 EA, Hardie R, Hannick LI, Glashoff RH, Mukerji M, Indian Genome Variation Consortium, Fernandez P, Haacke W, Schork NJ, Hayes VM. 2013. Complex patterns of genomic admixture within southern Africa. PLOS Genet 9:e1003309.

Pfeiffer S. 2012. Conditions for evolution of small adult body size in southern Africa. Curr Anthropol 53:383-394.

Pietrusewsky M. 2008. Metric analysis of skeletal remains: Methods and applications. In: Katzenberg MA, Saunders SR, editors. Biological anthropology of the human skeleton. John Wiley and Sons. p. 487-532.

Plavcan JM. 2012. Body size, size variation and sexual size dimorphism in early Homo. Curr Anthropol 53:409-423.

Pollitzer WS, Anderson JJB. 1989. Ethnic and genetic differences in bone mass: A review with a hereditary VS environmental perspective. Am J Clin Nutr 50:1244-1259.

Posel D. 2001. What's in a name? Racial categorisations under apartheid and their afterlife. Transformation 47:50-74.

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Ribot I. 2004. Differentiation of Modern Sub-Saharan African Populations: Craniometric interpretations in relation to geography and history. Bull Mém Soc Anthropol Paris 16:143-65.

Richter L, Norris S, Pettifor J, Yach D, Cameron N. 2007. Cohort profile: Mandela's children: The 1990 birth to twenty study in South Africa. Int J Epidemiol 36:504-511.

Rightmire GP. 1970. Bushman, Hottentot and South African negro crania studied by distance and discrimination. Am J Phys Anthropol 33:169-196.

Riggs BL, Khosla S, Melton LJ. 2002. Sex steroids and the construction and conservation of the adult skeleton. Endocr Rev 23:279-302.

Roberts DF, Bainbridge DR. 1963. Nilotic physique. Am J Phys Anthropol 21:341-370.

Ruff CB. 2002. Variation in human body size and shape. Annu Rev Anthropol 31:211-232.

Ruff CB, 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.

Samuels ML, Witmer JA. 2003. Statistics for the Life Sciences. Pearson Education, Inc. p. 463-523.

Sauer N. 1992. Forensic anthropology and the concept of race: If races don't exist, why are forensic anthropologists so good at identifying them? Soc Sci Med 34:107-111.

Schlebusch C. 2010. Issues raised by use of ethnic-group names in genome study. Nature 464:487.

Schroeder MA. 1990. Diagnosing and dealing with multicollinearity. Western J Nurs Res 12:175-187.

Schmeling A, Reisinger W, Loreck D, Vendura K, Markus W, Geserick G. 2000. Effects of ethnicity on skeletal maturation: consequences for forensic age estimations. Int J Legal Med 113:253–258.

Siddiqi N. 2013. Comparison of osteometric femoral bone dimensions among the South Africans of different ethnic groups and South African whites. Egypt J Forensic Sci 3:8-14.

Smay D, Armelagos G. 2000. Galileo wept: A critical assessment of the use of race in forensic anthropology. Transform Anthropol 9:19-29.

Sparks CS, Jantz RL. 2002. A reassessment of human cranial plasticity: Boas revisited. PNAS 99:14636-14639.

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

Stein PL, Rowe BM. 1989. Physical Anthropology. 4th ed. McGraw-Hill Publishing Company. p.177-191.

Stewart TD. 1979. Essentials of Forensic Anthropology: Especially as developed in the United States. Charles C Thomas. p.227-238.

Stewart TD. 1980. Responses of the human skeleton to changes in the quality of life. J Forensic Sci 25:912-921.

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. 1999. Osteometric variation in the humerus: Sexual dimorphism in South Africans. Forensic Sci Int 106:77-85.

Steyn M, Smith JR. 2007. Interpretation of ante-mortem stature estimates in South Africans. Forensic Sci Int 171:97-102.

St Hoyme E, İşcan MY. 1989. Determination of sex and race. In: İşcan MY and Kennedy KAR, editors. Reconstruction of life from the skeleton. Alan R. Liss, Inc. p. 53-93.

Stomfai S, Ahrens W, Bammann K, Kovács É, Mårild S, Michels N, Moreno LA, Pohlabeln H, Siani A, Tornaritis M, Veidebaum T, Molnár D. 2011. Intra- and inter-observer reliability in anthropometric measurements in children. Int J Obesity 35:45-51.

Stull KE. 2013. An osteometric evaluation of age and sex differences in the long bones of South African children from the Western Cape [dissertation]. University of Pretoria.

Stull KE, Kenyhercz MW, L'Abbé EN. Ancestry estimation in South Africa using craniometrics and geometric morphometrics. In press: Forensic Sci Int.

Stull KE, L'Abbé EN, Ousley SD. 2014. Using multivariate adaptive regression splines to estimate subadult age from diaphyseal dimensions. Am J Phys Anthropol 154:376-386.

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. Boston: Pearson Education.

Thompson LM. 2001. A history of South Africa. 3rd ed. New Haven: Yale University Press.

Tise ML, Spradley MK, Anderson BE. 2013. Postcranial sex estimation of individuals considered Hispanic. J Forensic Sci 58:9-14.

Tishkoff SA, Gonder MK, Henn BM, Mortensen H, Knight A, Gignoux C, Fernandopulle N, Lema G, Nyambo TB, Ramakrishan U, Reed FA, Mountain JL. 2007. History of click-speaking populations of Africa inferred from mtDNA and Y chromosome genetic variation. Mol Biol Evol 24:2180-2195.

Tishkoff SA, Williams SM. 2002. Genetic analysis of African populations: Human Evolution and complex disease. Nat Rev Genet 3:611-621.

Todd TW, Lindala A. 1928. Dimensions of the body: Whites and American Negroes of both sexes. Am J Phys Anthropol 12:35-119.

Uhl N. 2014. Using multivariate calibration to evaluate hominin brain/body size relationships [dissertation]. University of Illinois at Urbana-Champaign.

Uhl N, Hefner JT, Schultz JJ. 2007. Geometric morphometrics of the scapula: An assessment of ancestry. Proceedings of the 59th Annual Meeting of the American Academy of Forensic Sciences American Academy of Forensic Sciences. San Antonio, TX.

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.

Vidarsdottir US, O'Higgins P. 2003. Developmental variation in the facial skeleton of anatomically modern Homo sapiens. In: Thompson JL, Krovitz GE, Nelson AJ, editors. Patterns of growth and development in the genus homo. Cambridge: Cambridge University Press. p. 114-143.

Vidulich L, Norris SA, Cameron N, Pettifor JM. Differences in bone size and bone mass between black and white 10-year-old South African children. 2006. Osteoporos Int 17:433-440.

von Cramon-Taubadel N. 2013. Evolutionary insights into global patterns of human cranial diversity: Population history, climatic and dietary effects. J Anthropol Sci 91:1-36.

von Cramon-Taubadel N, Weaver TD. 2009. Insights from a quantitative genetic approach to human morphological evolution. Evol Anthropol 18:237-240.

Walensky NG. 1965. A study of the anterior femoral curvature in man. Anat Rec 151:559-570.

Wei T. 2013. Corrplot: Visualization of a correlation matrix.

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.

Wescott DJ. 2005. Population variation in femur subtrochanteric shape. J Forensic Sci 50:1-8.

Wheat AD. 2009. Assessing ancestry through nonmetric traits of the skull: A test of education and experience [dissertation]. Texas State University.

White TD, Folkens PA. 2005. The human bone manual. Elsevier Inc. p. 400-404.

Wolfe LD, Gray JP. 1982. Subsistence practices and human sexual dimorphism of stature. J Hum Evol 11:575-580.

Wood ET, Stover DA, Ehret C, Destro-Bisol G, Spedini G, McLeod H, Louie L, Bamshad M, Strassman BI, Soodyall H, Hammer MF. 2005. Contrasting patterns of Y chromosome and mtDNA variation in Africa: Evidence for sex-biased demographic processes. Eur J Hum Genet 13:1-10.

Zere E, McIntyre D. 2003. Inequities in under-five child malnutrition in South Africa. Int J Equity Health 2:7-17.

APPENDIX I - MEASUREMENT DEFINITIONS

***** All measurement definitions taken from Moore-Jansen et al. (1994).

** For midshaft measurements, determine the midpoint of the diaphysis on the osteometric board and mark with a pencil.

Clavicle maximum length (claxln) – The maximum distance between the most extreme ends of the clavicle (osteometric board).

Clavicle sagittal midshaft diameter (claapd) – The anterio-posterior distance of the midshaft surface (sliding caliper).

Clavicle vertical midshaft diameter (clavrd) – The superior-inferior distance of the midshaft surface (sliding caliper).

Scapula height (scapht) – The direct distance from the most superior point of the superior angle to the most inferior point of the inferior angle (osteometric board).

Scapula breadth (scapbr) – The distance from the midpoint of the dorsal border of the glenoid fossa to a point midway between the two ridges of the scapular spine on the vertebral border (sliding caliper).

Humerus maximum length (humxln) – The direct distance from the most superior point on the humeral head to the most inferior point on the trochlea (osteometric board).

Humerus epicondylar breadth (humebr) – The distance from the most protruding point on the lateral epicondyle to the corresponding projection on the medial epicondyle (osteometric board).

Humeral head diameter (humhdd) – The direct distance between the most superior and inferior points on the border of the articular surface (sliding caliper).

Humerus maximum diameter (hummxd) – The maximum diameter of the midshaft located inferior to the deltoid tuberosity. Take note, the bone should be turned until the maximum is found, which is not necessarily in the anterio-posterior plane (sliding caliper).

Humerus minimum diameter (hummwd) – The minimum diameter of the midshaft located inferior to the deltoid tuberosity (sliding caliper).

Radius maximum length (radxln) – The distance from the most proximally positioned point on the head to the most distal point on the styloid process (osteometric board).

Radius A-P midshaft diameter (radapd) – The diameter of the midshaft in the anterio-posterior plane (slaiding caliper).

Radius transverse midshaft diameter (radtvd) – The diameter of the midshaft in the medio-lateral plane, perpendicular to the A-P diameter (sliding caliper).

Ulna maximum length (ulnxln) – The distance between the most superior point on the olecranon and most inferior point on the styloid process. If there are pronounced osteophytes on the olecranon the measurement should not be taken (osteometric board).

Ulna dorso-volar diameter (ulndvd) – The maximum diameter diameter of the diaphysis where the crests exhibits the greatest development. Take note, this measurement is not necessarily midshaft (sliding caliper).

Ulna transverse diameter (ulntvd) – The diameter taken perpendicular to the dorso-volar diameter at the level of greatest crest development (sliding caliper).

Ulna physiological length (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. Take note, the caliper should be placed on the head, not in the groove between the head and the styloid process (spreading caliper).

Sacrum anterior height (sacaht) – The distance from midpoint on the promontory to a point on the anterior, inferior border of the fifth sacral segment. If the coccyx is fused to the sacrum, do not take the measurement (sliding caliper).

Sacrum anterior breadth (sacabr) – The maximum transverse breadth of the sacrum at the level of the anterior projection of the auricular surfaces (sliding caliper).

Transverse diameter of S1 (sacs1b) – The distance between the most lateral points on the superior articular surface of the first sacral segment. Where lipping is present, the original articular borders can be approximated. If extensive lipping prevents approximation of the borders the measurement should not be taken (sliding caliper).

Innominate height (innoht) – The distance from the most superior point on the iliac crest to the most inferior point on the ischial tuberosity (osteometric board).

Iliac breadth (iliabr) – The distance from the anterior superior iliac spine to the posterior superior iliac spine (osteometric board).

Femur maximum length (femxln) – The distance from the most superior point on the head of the femur to the most inferior point on the condyles. Take note, the bone should be placed parallel to the long axis of the board, with only the medial condyle placed against the vertical endboard (osteometric board).

Femur bicondylar length (fembln) – The distance from the most superior point on the head of the femur to the most inferior point on the condyles. Take note, both condyles should be placed against the vertical endboard. The bone will not be parallel to the long axis of the board (osteometric board).

Femur epicondylar breadth (femebr) – The distance between the most projecting points on the epicondyles (osteometric board).

Femoral head diameter (femhdd) – The maximum diameter of the femoral head measured on the border of the articular surface (sliding caliper).

Femur A-P subtrochanteric diameter (femsap) – The anterio-posterior diameter of the proximal end of the diaphysis measured below the base of the lesser trochanter (sliding caliper).

Femur transverse subtrochanteric diameter (femstv) – The transverse diameter of the proximal end of the diaphysis, taken perpendicular to the A-P diameter below the base of the lesser trochanter (sliding caliper).

Femur A-P midshaft diameter (femmap) – The anterio-posterior diameter taken at the midpoint of the diaphysis at the highest elevation of linea aspera (sliding caliper).

Femur transverse midshaft diameter (femmtv) – The transverse diameter taken perpendicular to the A-P diameter at the midpoint of the diaphysis (sliding caliper)

Tibia condylo-malleolar length (tibxln) – The distance from the superior articular surface of the lateral condyle to the most inferior point of the medial malleolus. Take note, the intercondylar eminence should not be included (osteometric board).

Tibia proximal epiphyseal breadth (tibpeb) – The maximum distance between the most projecting points on the medial and lateral condyles of the proximal epiphysis. Take note, the

bone should be rotated to obtain the maximum, however the articular facet for the fibula should not be included (osteometric board).

Tibia distal epiphyseal breadth (tibdeb) – The distance between the most medial point of the medial malleolus and the lateral surface of the distal epiphysis (osteometric board).

Tibia maximum diameter at nutrient foramen (tibnfx) – The distance between the anterior crest and post surface of the diaphysis at the level of thev nutrient foramen. Take note, the bone should be rotated to obtain the maximum (sliding caliper).

Tibia minimum diameter at nutrient foramen (tibnft) – The transverse diameter of the diaphysis at the level of the nutrient foramen, taken perpendicular to the maximum diameter (sliding caliper).

Fibula maximum length (fibxln) – The maximum distance between the most superior point on the head and the most inferior point on the lateral malleolus (osteometric board).

Fibula maximum diameter (fibmdm) – The maximum diameter of the diaphysis taken at midshaft (sliding caliper).

Calcaneus maximum length (calcxl) – The distance between the most projecting point on the tuberosity and the most anterior point on the superior margin of the cuboidal articular facet. If there is extensive osteophyte development, the measurement should not be taken (sliding caliper).

Calcaneus middle breadth (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 – ANOVA

Table A2 - ANOVA results evaluating the effects of ancestry, sex and the interaction between ancestry and sex for each measurement. Bold indicates significant.

		Ancestry		Sex		Ancestry*Sex	
		F-value	**Pr (< F)**	**F-value**	**Pr (< F)**	**F-value**	**Pr (< F)**
Clavicle	claxln	19.54	**<0.001**	196.55	**<0.001**	0.28	0.75
	clavrd	0.73	0.48	152.76	**<0.001**	1.81	0.17
	claapd	1.73	0.18	219.43	**<0.001**	0.79	0.46
Scapula	scapht	46.35	**<0.001**	305.13	**<0.001**	0.76	0.47
	scapbr	33.27	**<0.001**	219.71	**<0.001**	0.04	0.97
Humerus	humxln	63.70	**<0.001**	334.71	**<0.001**	0.72	0.49
	humebr	35.48	**<0.001**	434.72	**<0.001**	2.02	0.13
	humhdd	142.56	**<0.001**	442.46	**<0.001**	1.41	0.24
	hummxd	15.8	**<0.001**	234.2	**<0.001**	0.2	0.82
	hummwd	11.23	**<0.001**	302.96	**<0.001**	1.71	0.18
Radius	radxln	39.36	**<0.001**	319.37	**<0.001**	0.23	0.8
	radapd	4.36	**0.013**	475.45	**<0.001**	0.11	0.900
	radtvd	22.80	**<0.001**	201.97	**<0.001**	1.39	0.25
Ulna	ulnxln	41.83	**<0.001**	285.90	**<0.001**	0.07	0.93
	ulndvd	1.28	0.28	225.60	**<0.001**	2.21	0.11
	ulntvd	16.21	**<0.001**	347.59	**<0.001**	0.21	0.81
	ulnphl	39.23	**<0.001**	264.77	**<0.001**	1.01	0.36
Sacrum	sacaht	11.20	**<0.001**	8.11	**0.0049**	1.01	0.3661
	sacabr	61.72	**<0.001**	24.90	**<0.001**	0.59	0.56
	sacs1b	27.79	**<0.001**	66.88	**<0.001**	0.63	0.53
Pelvis	innoht	170.17	**<0.001**	198.54	**<0.001**	0.11	0.89
	iliabr	211.59	**<0.001**	27.08	**<0.001**	4.29	**0.015**
Femur	femxln	50.43	**<0.001**	199.90	**<0.001**	0.14	0.87
	fembln	40.10	**<0.001**	201.37	**<0.001**	0.16	0.85
	femebr	74.67	**<0.001**	387.28	**<0.001**	0.47	0.62
	femhdd	45.76	**<0.001**	445.08	**<0.001**	1.37	0.26
	femsap	44.85	**<0.001**	146.30	**<0.001**	1.41	0.25
	femstv	20.63	**<0.001**	208.55	**<0.001**	2.81	0.062
	femmap	24.30	**<0.001**	175.49	**<0.001**	0.46	0.63
	femmtv	41.91	**<0.001**	200.89	**<0.001**	0.11	0.89
Tibia	tibxln	30.92	**<0.001**	148.41	**<0.001**	0.19	0.83
	tibpeb	48.96	**<0.001**	439.35	**<0.001**	0.12	0.88
	tibdeb	87.84	**<0.001**	239.23	**<0.001**	0.47	0.62
	tibnfx	20.02	**<0.001**	197.24	**<0.001**	2.46	0.087
	tibnft	23.48	**<0.001**	160.27	**<0.001**	0.31	0.74
Fibula	fibxln	35.09	**<0.001**	157.83	**<0.001**	0.56	0.57
	fibmdm	16.26	**<0.001**	61.42	**<0.001**	0.17	0.85
Calcaneus	calcxl	48.01	**<0.001**	137.59	**<0.001**	4.71	**0.031**
	calcbr	1.59	0.206	122.71	**<0.001**	4.78	**0.009**

APPENDIX III – TUKEY'S HSD

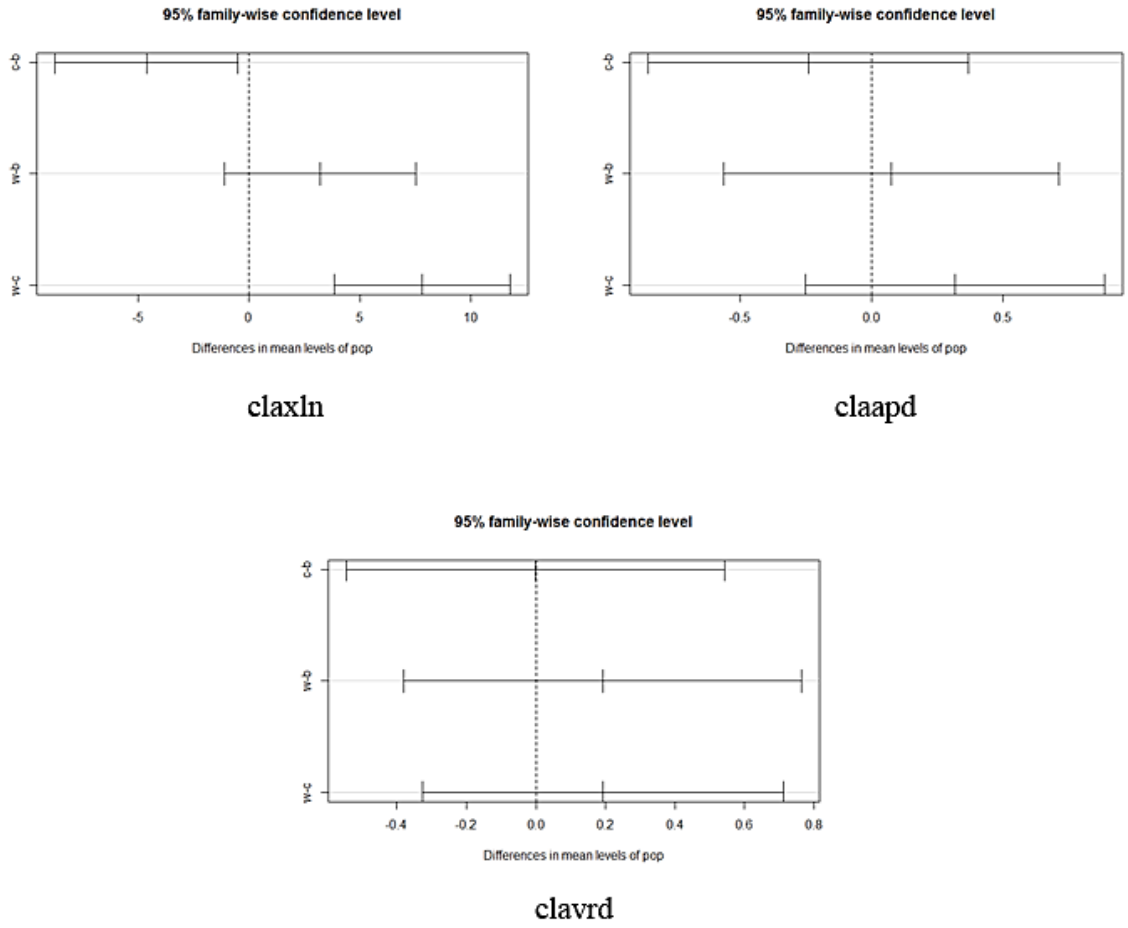
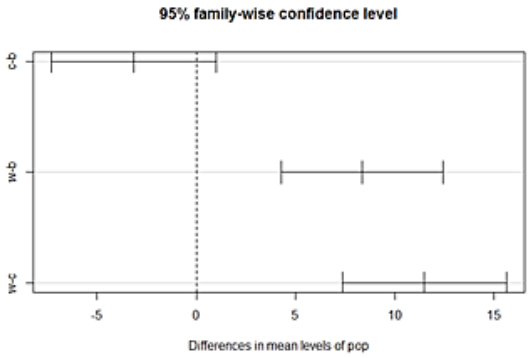
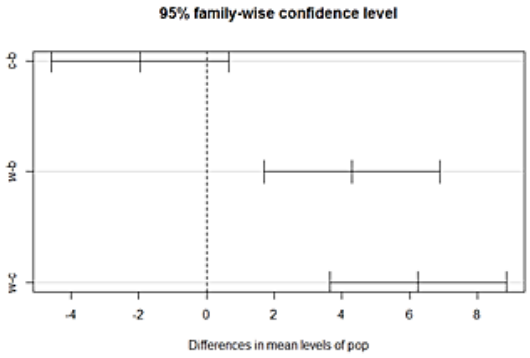


Figure A3.1 - Results for Tukey's HSD illustrating group differences and variable overlap for the measurements of the clavicle.



scapht



scapbr

Figure A3.2 - Results for Tukey's HSD illustrating group differences and variable overlap for the measurements of the scapula.

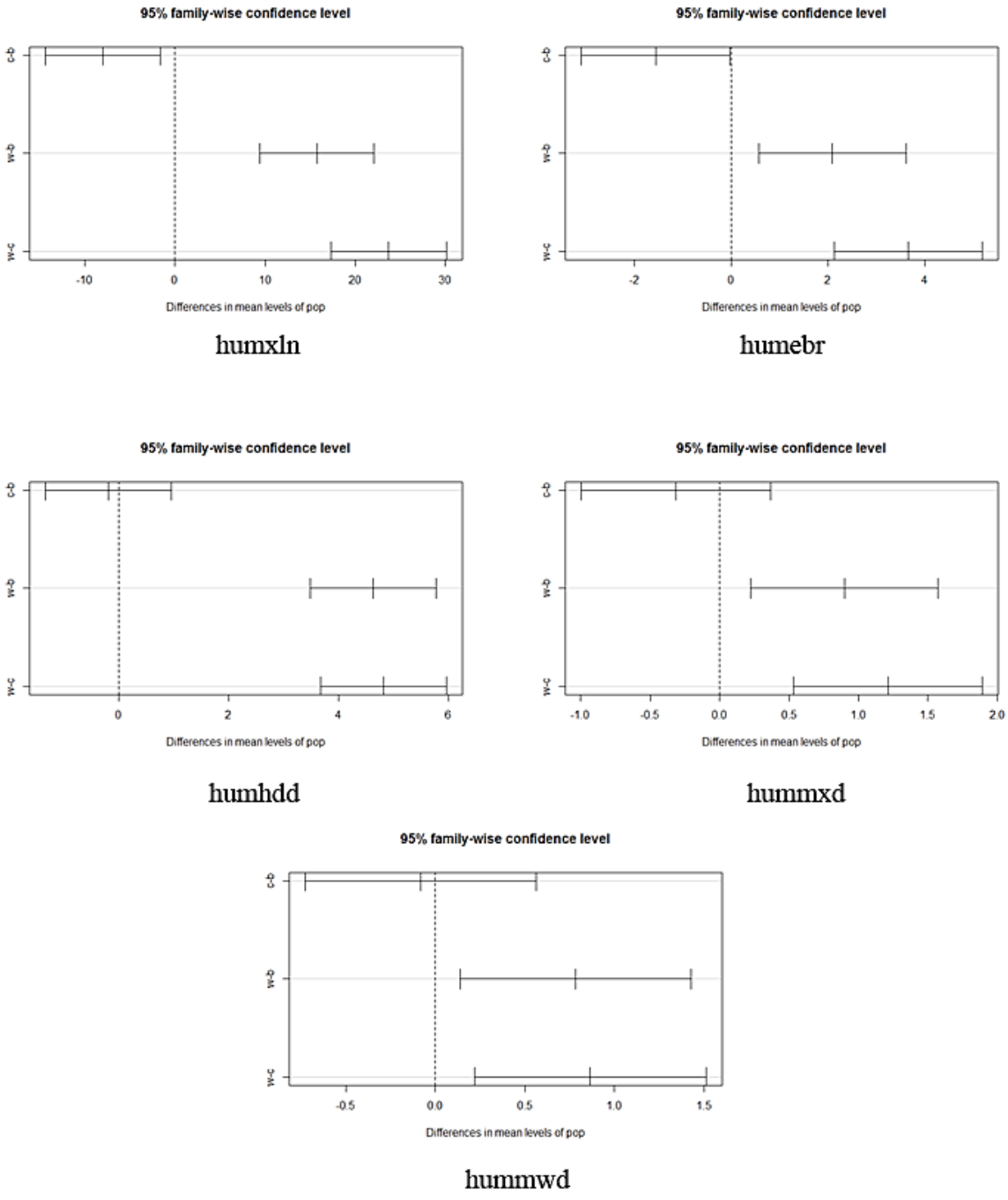
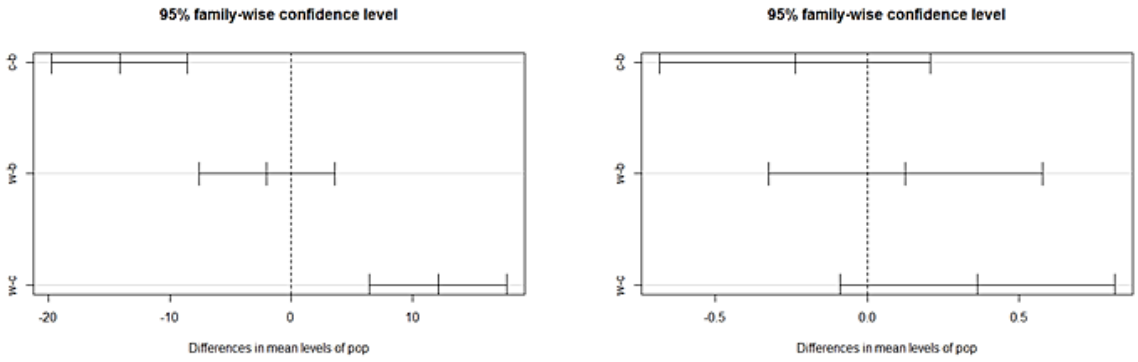
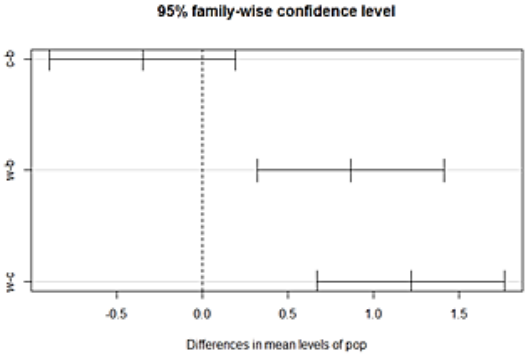


Figure A3.3 - Results for Tukey's HSD illustrating group differences and variable overlap for the measurements of the humerus.



radxln

radapd



radtvd

Figure A3.4 - Results for Tukey's HSD illustrating group differences and variable overlap for the measurements of the radius.

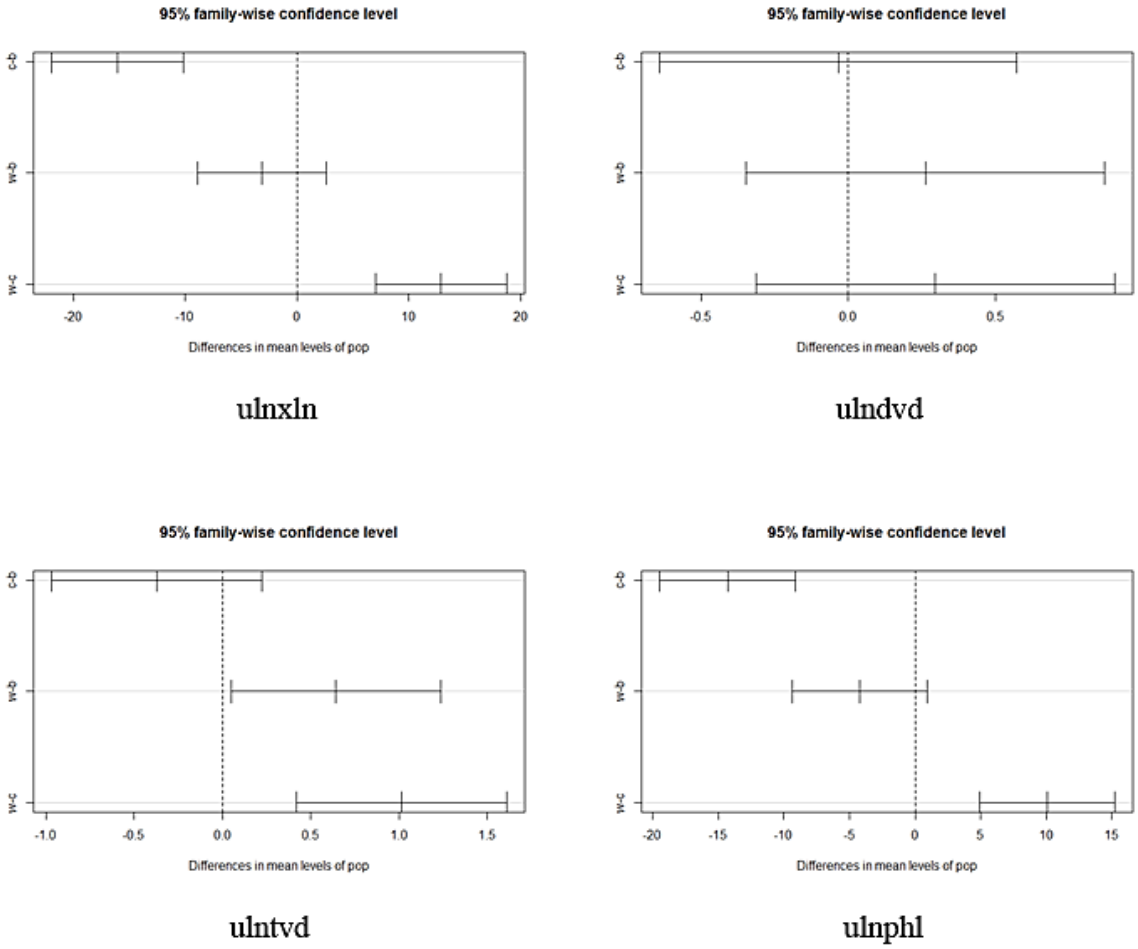


Figure A3.5 - Results for Tukey's HSD illustrating group differences and variable overlap for the measurements of the ulna.

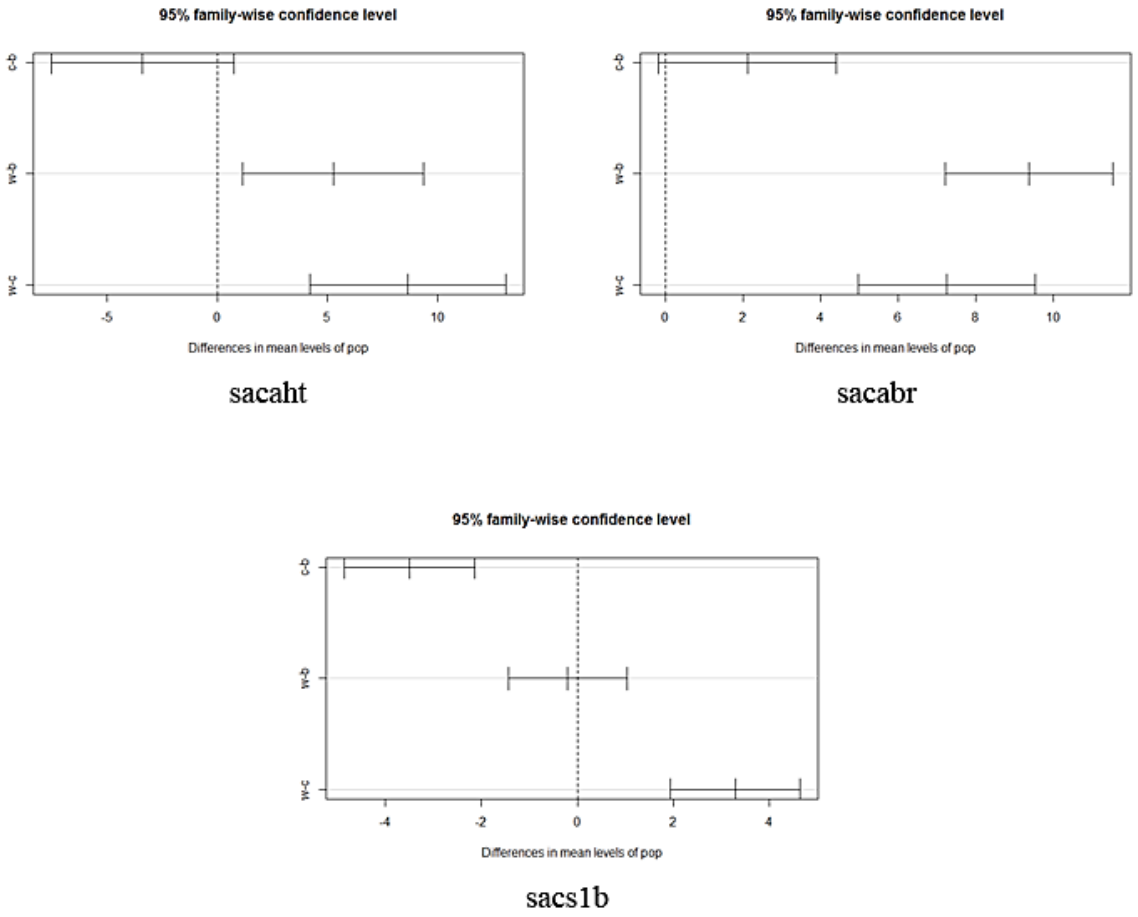
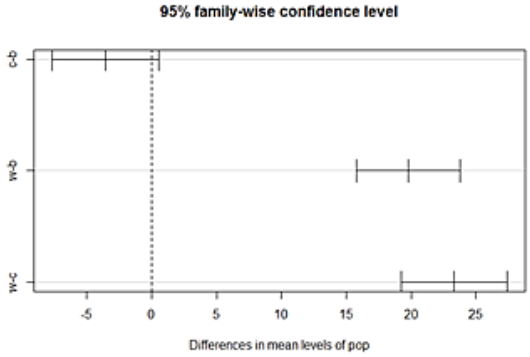
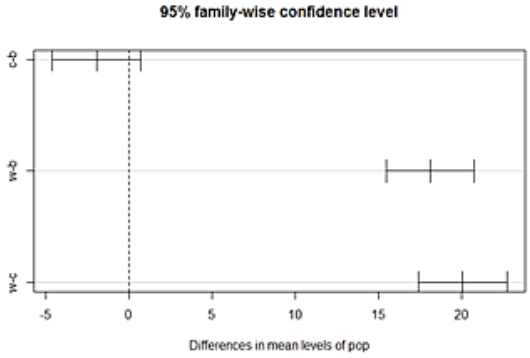


Figure A3.6 - Results for Tukey's HSD illustrating group differences and variable overlap for the measurements of the sacrum.



innoht



iliabr

Figure A3.7 - Results for Tukey's HSD illustrating group differences and variable overlap for the measurements of the pelvis.

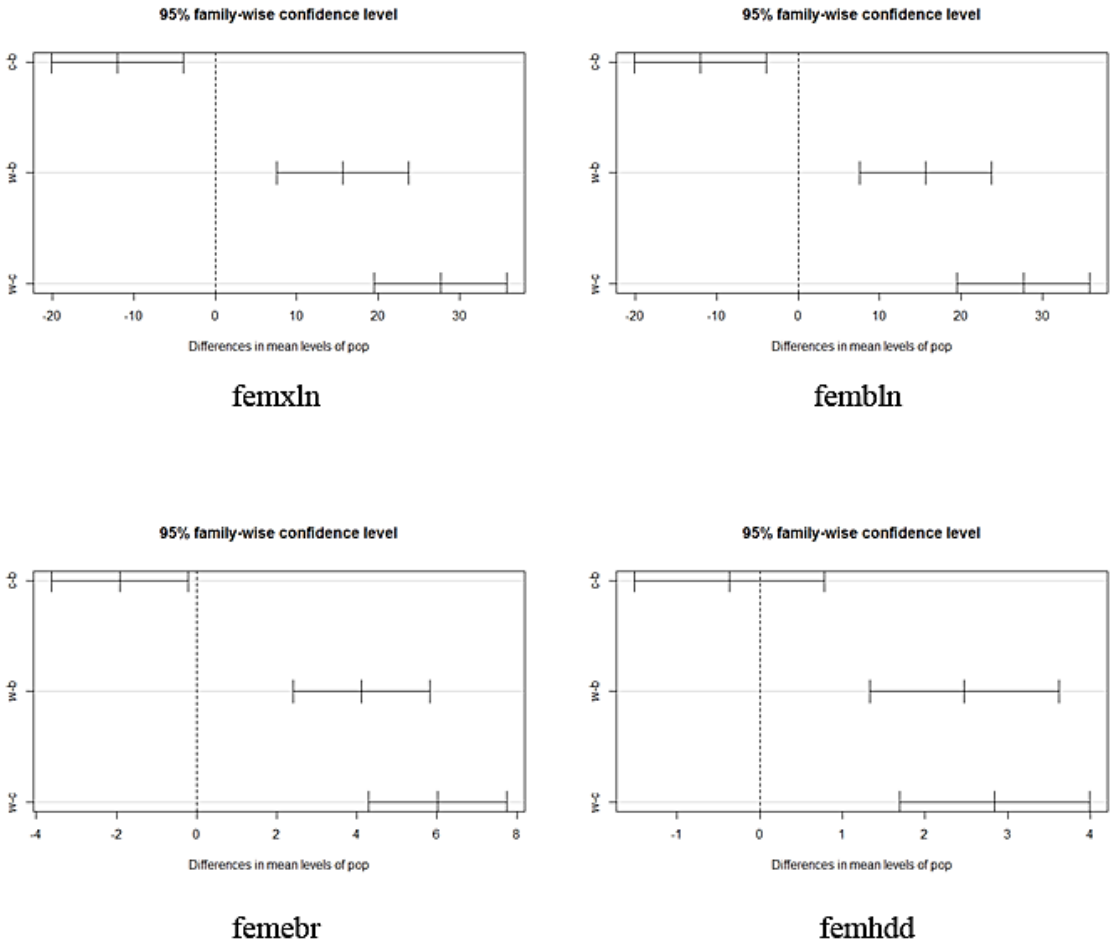


Figure A3.8 - Results for Tukey's HSD illustrating group differences and variable overlap for the measurements of the femur.

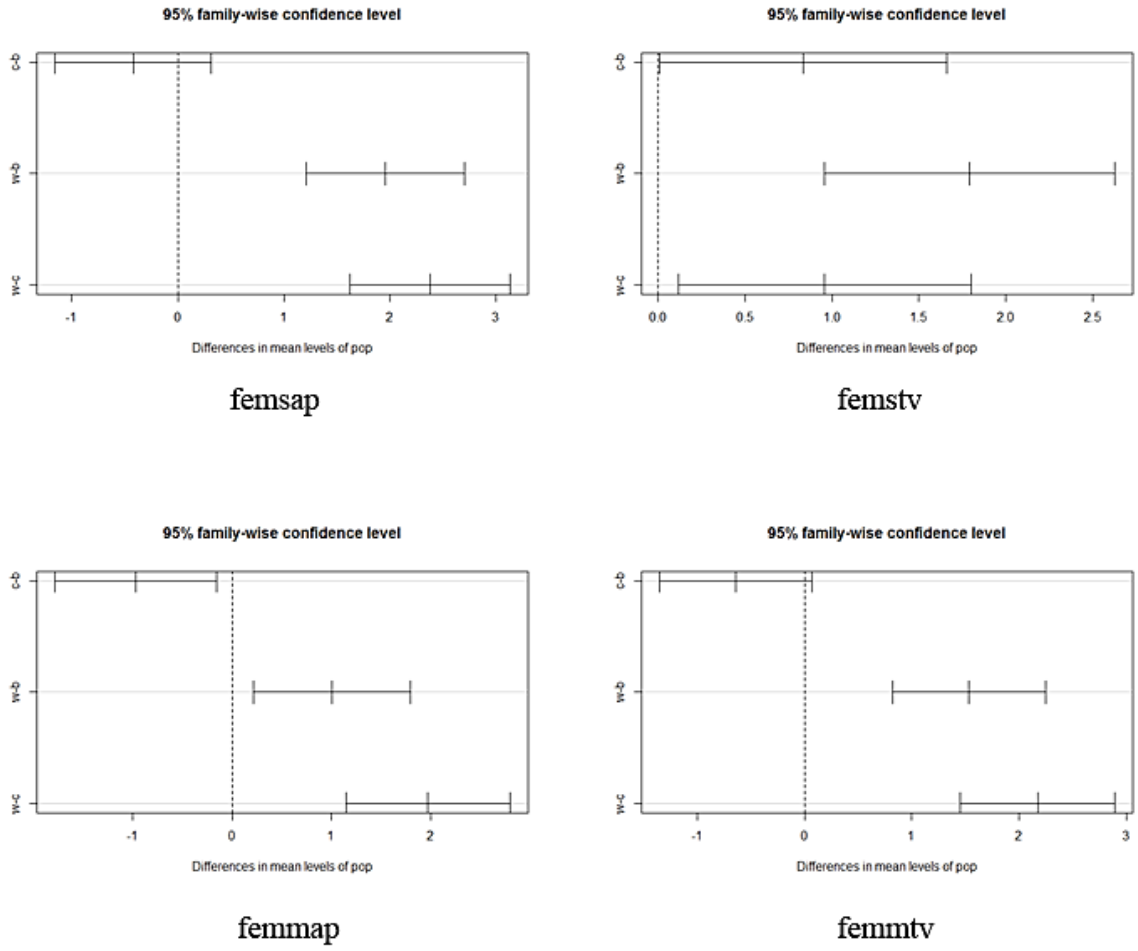


Figure A3.8 (cont.) - Results for Tukey's HSD illustrating group differences and variable overlap for the measurements of the femur.

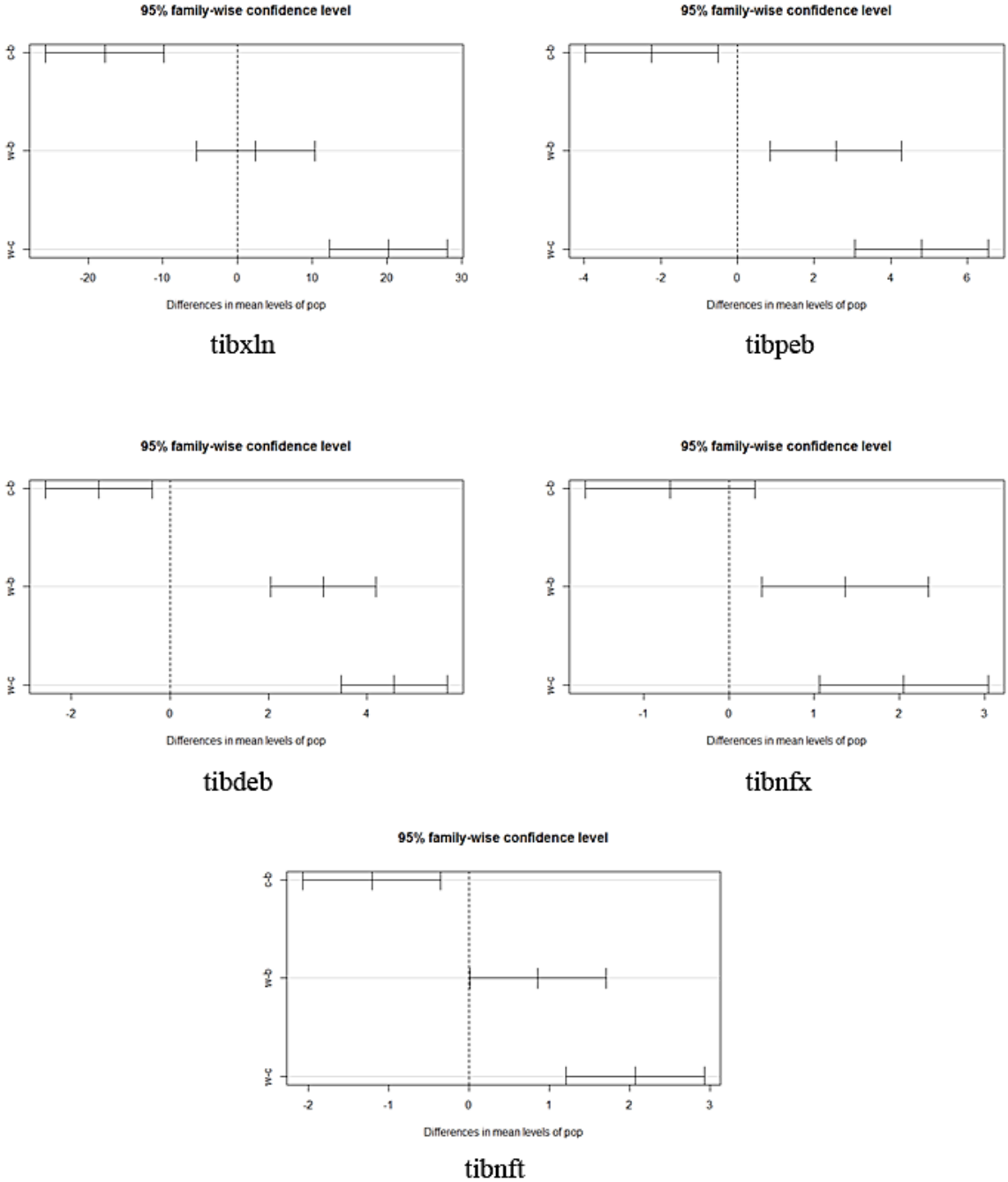
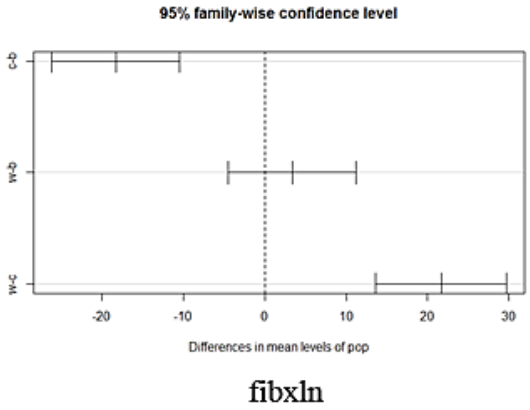
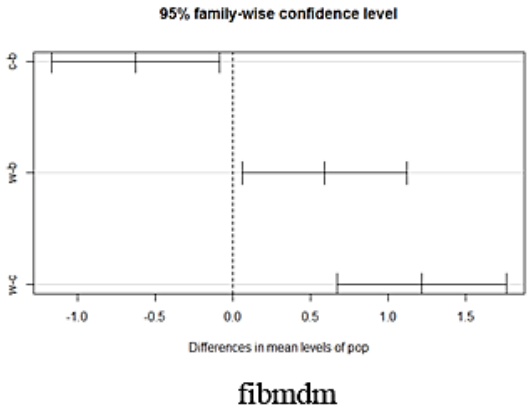


Figure A3.9 - Results for Tukey's HSD illustrating group differences and variable overlap for the measurements of the tibia.



fibxln



fibmdm

Figure A3.10 - Results for Tukey's HSD illustrating group differences and variable overlap for the measurements of the fibula.

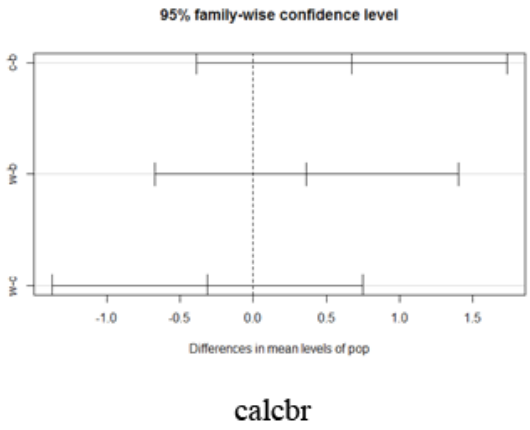


Figure A3.11 - Results for Tukey's HSD illustrating group differences and variable overlap for the measurements of the calcaneus.

APPENDIX IV – BOXPLOTS

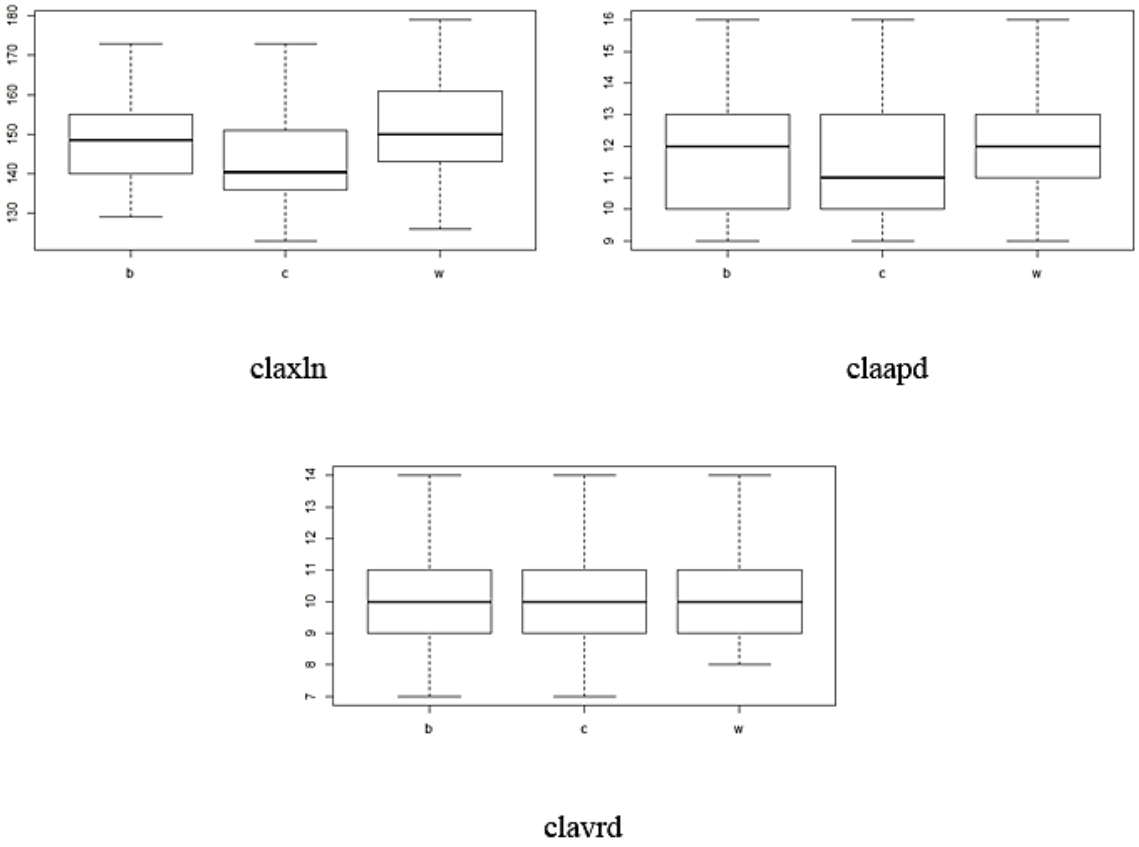
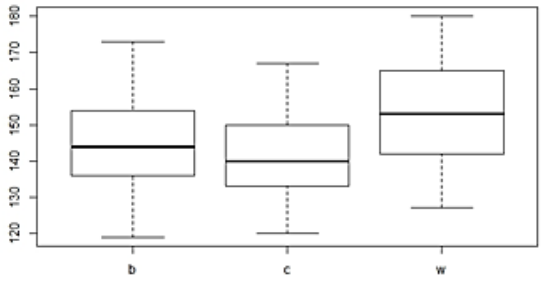
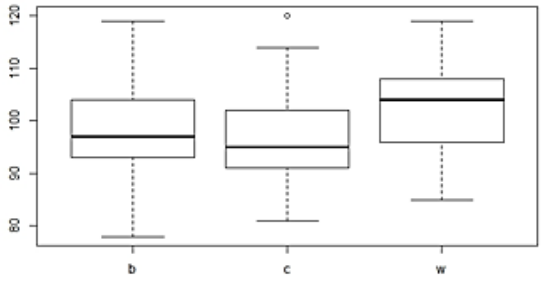


Figure A4.1 - Boxplots illustrating the degree of variation within each group for the measurements of the clavicle.



scapht



scapbr

Figure A4.2 - Boxplots illustrating the degree of variation within each group for the measurements of the scapula.

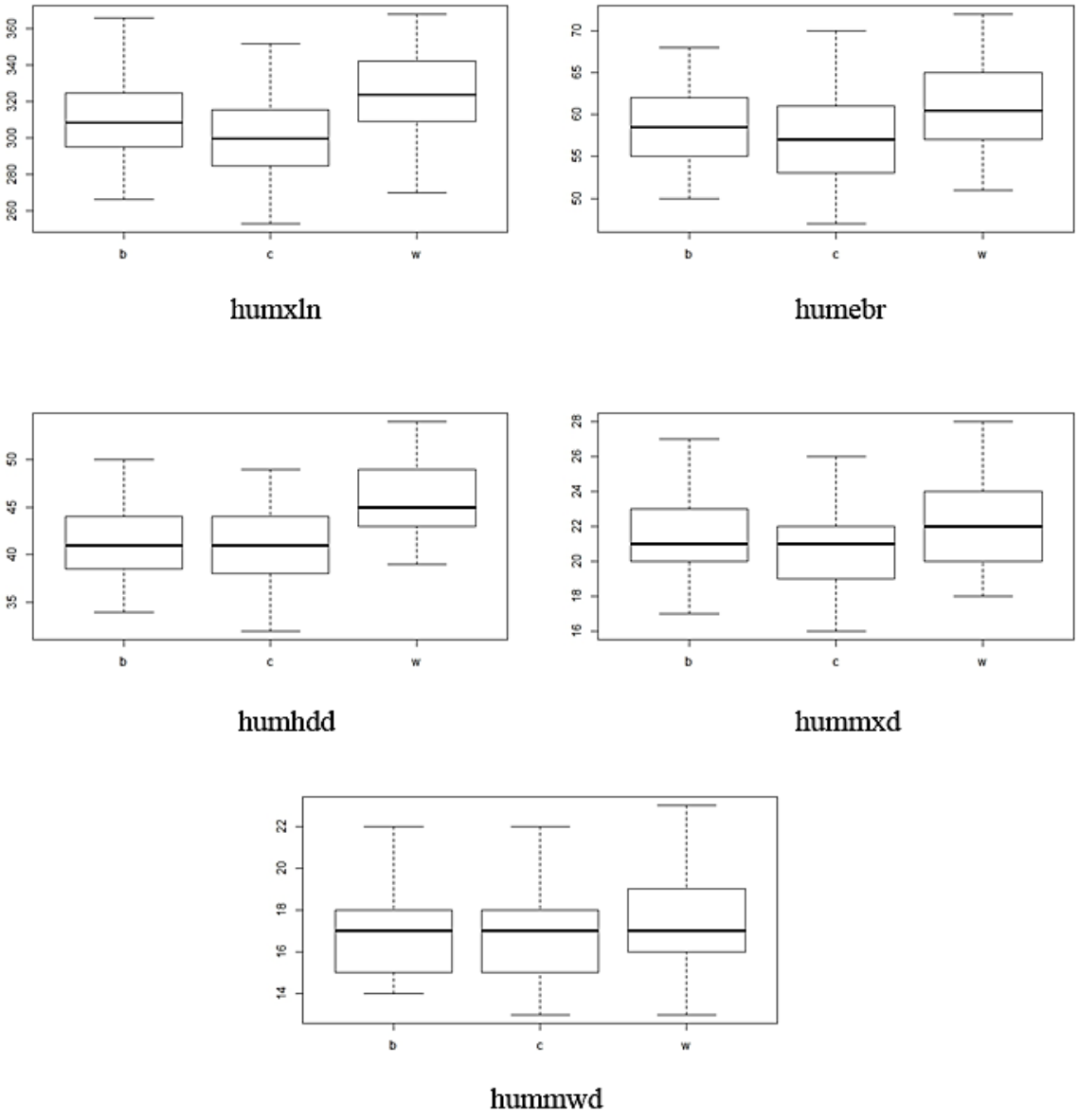


Figure A4.3 - Boxplots illustrating the degree of variation within each group for the measurements of the humerus.

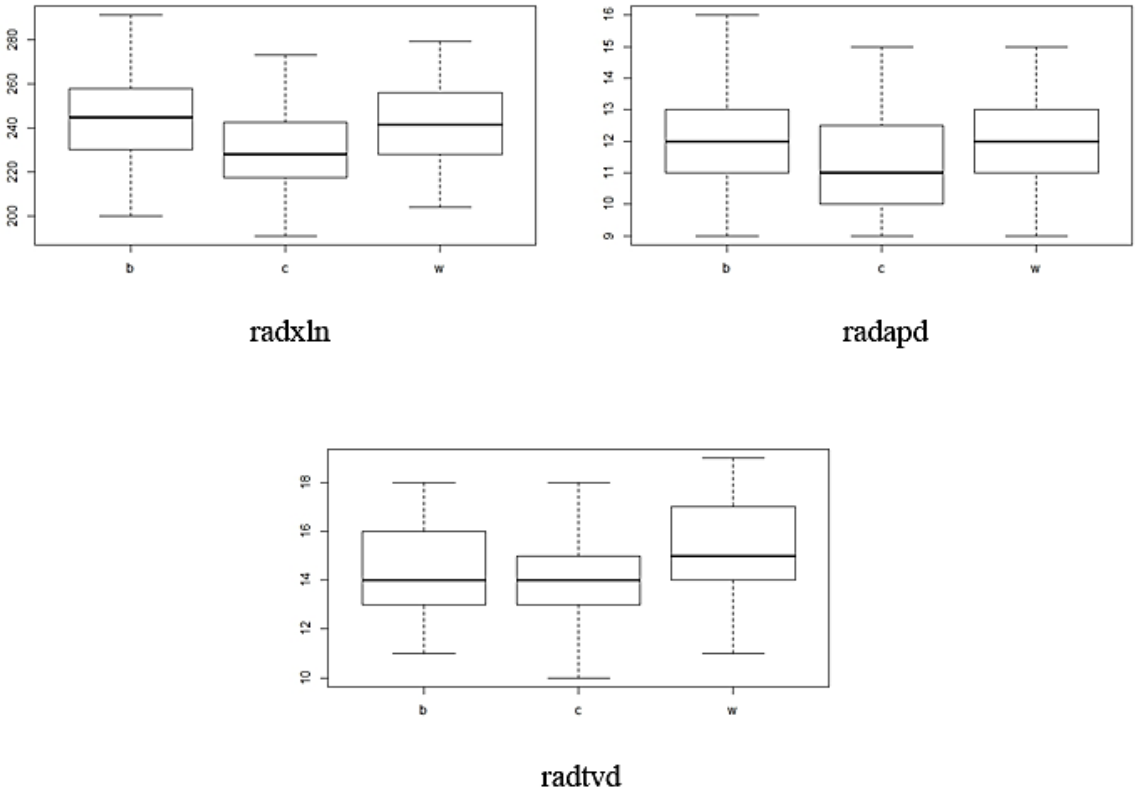


Figure A4.4 - Boxplots illustrating the degree of variation within each group for the measurements of the radius.

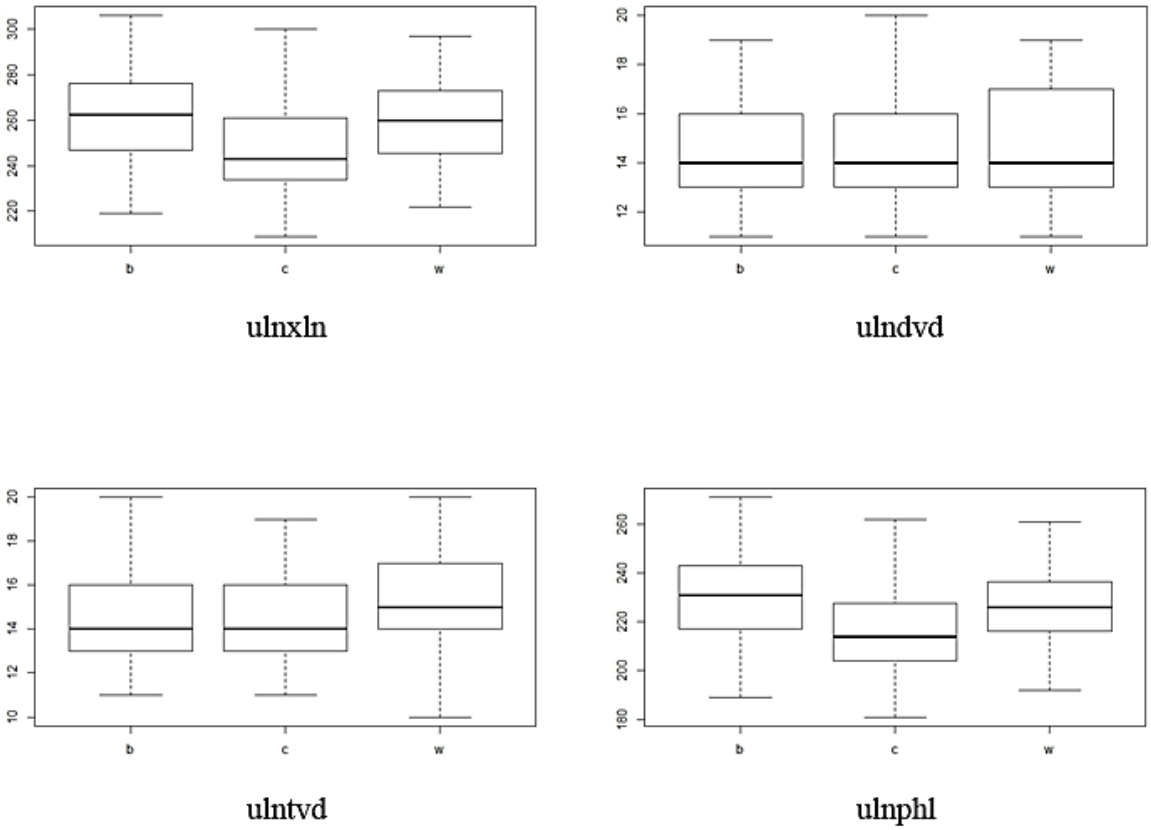


Figure A4.5 - Boxplots illustrating the degree of variation within each group for the measurements of the ulna.

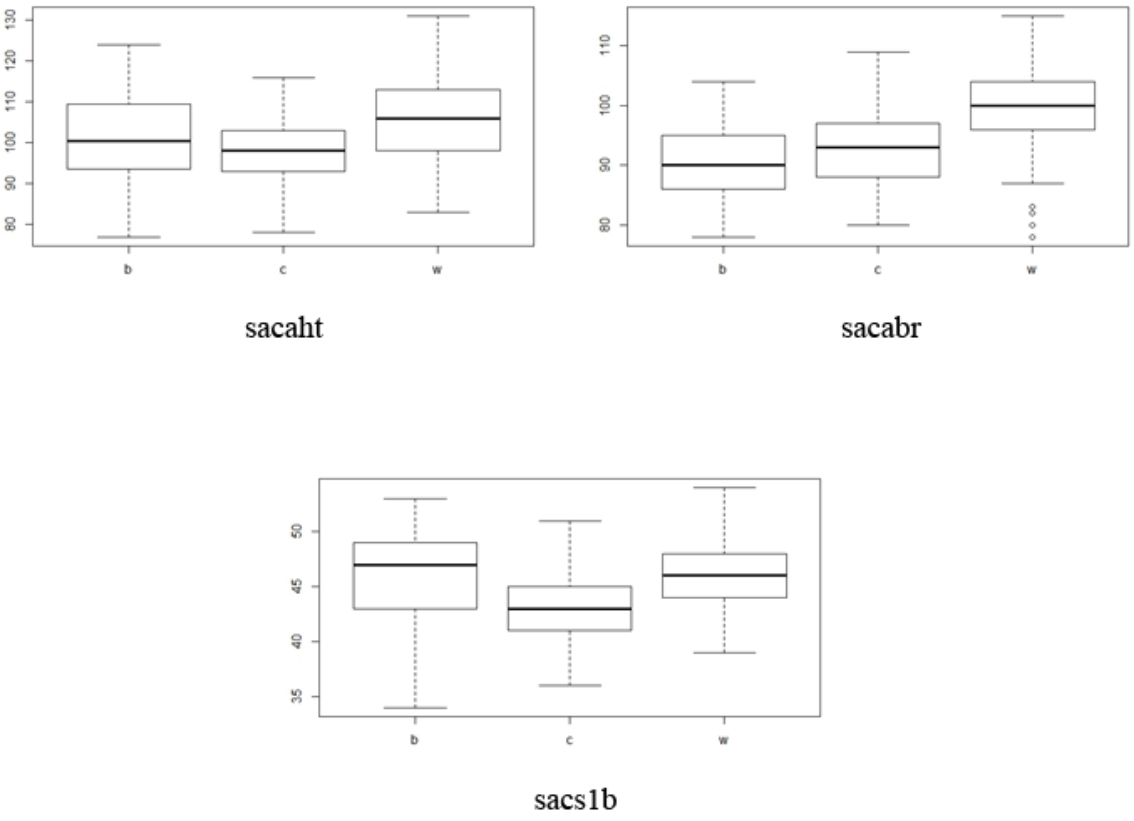


Figure A4.6 - Boxplots illustrating the degree of variation within each group for the measurements of the sacrum.

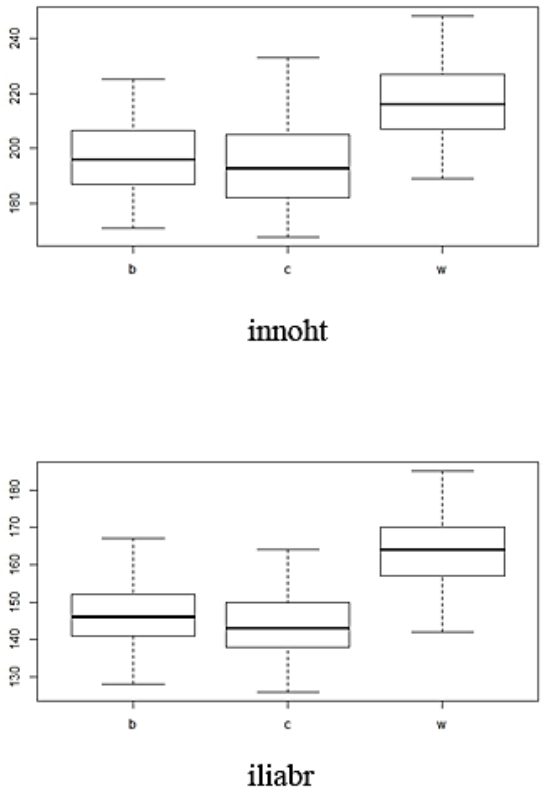


Figure A4.7 - Boxplots illustrating the degree of variation within each group for the measurements of the pelvis.

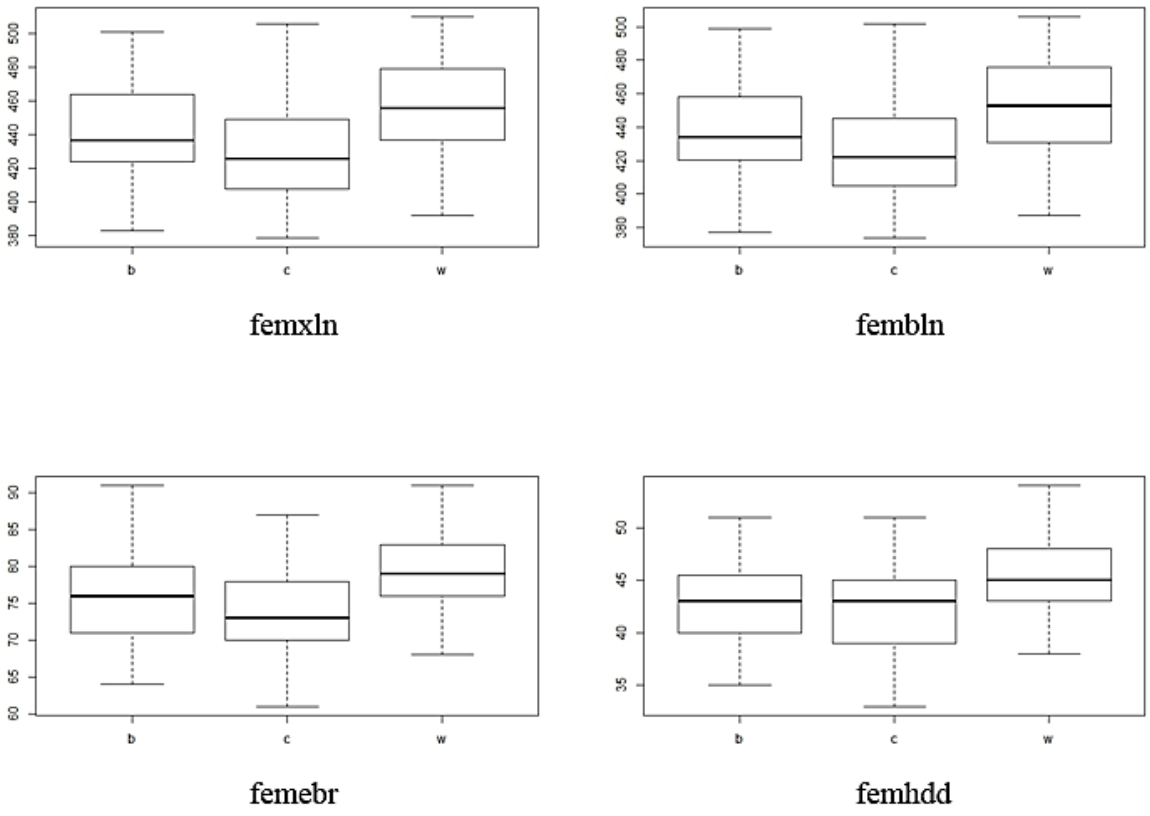


Figure A4.8 - Boxplots illustrating the degree of variation within each group for the measurements of the femur.

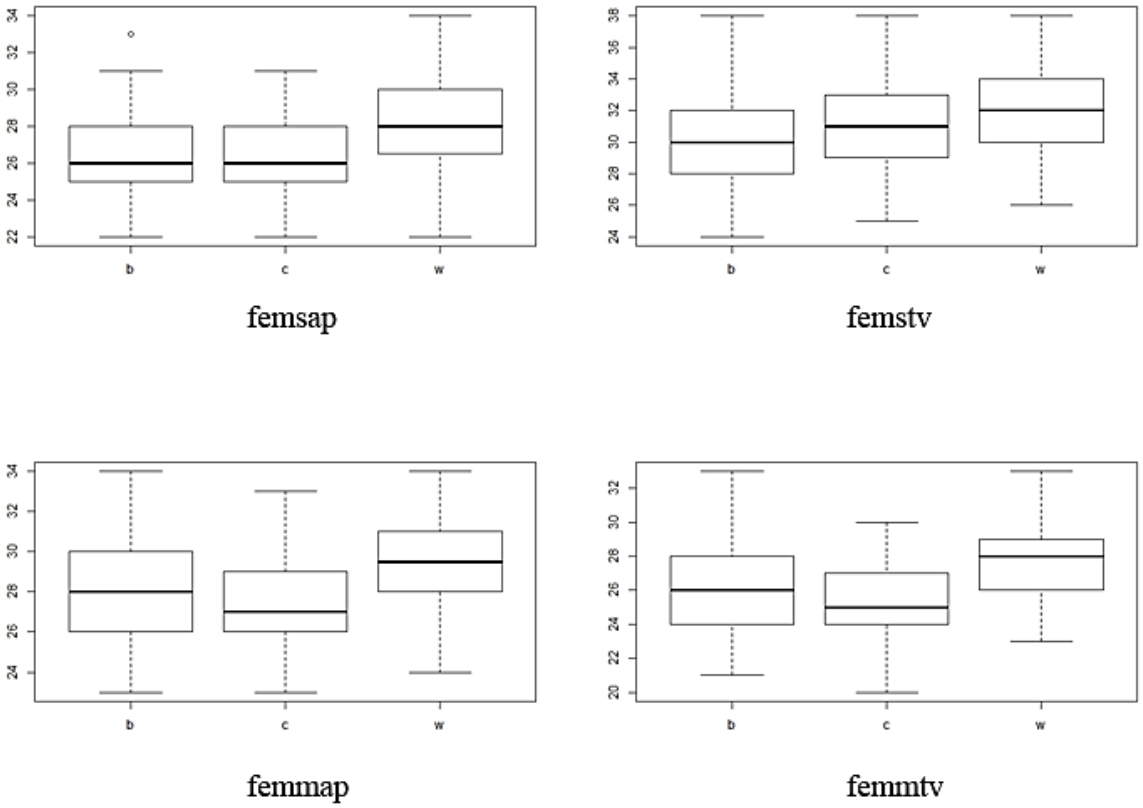


Figure A4.8 (cont.) - Boxplots illustrating the degree of variation within each group for the measurements of the femur.

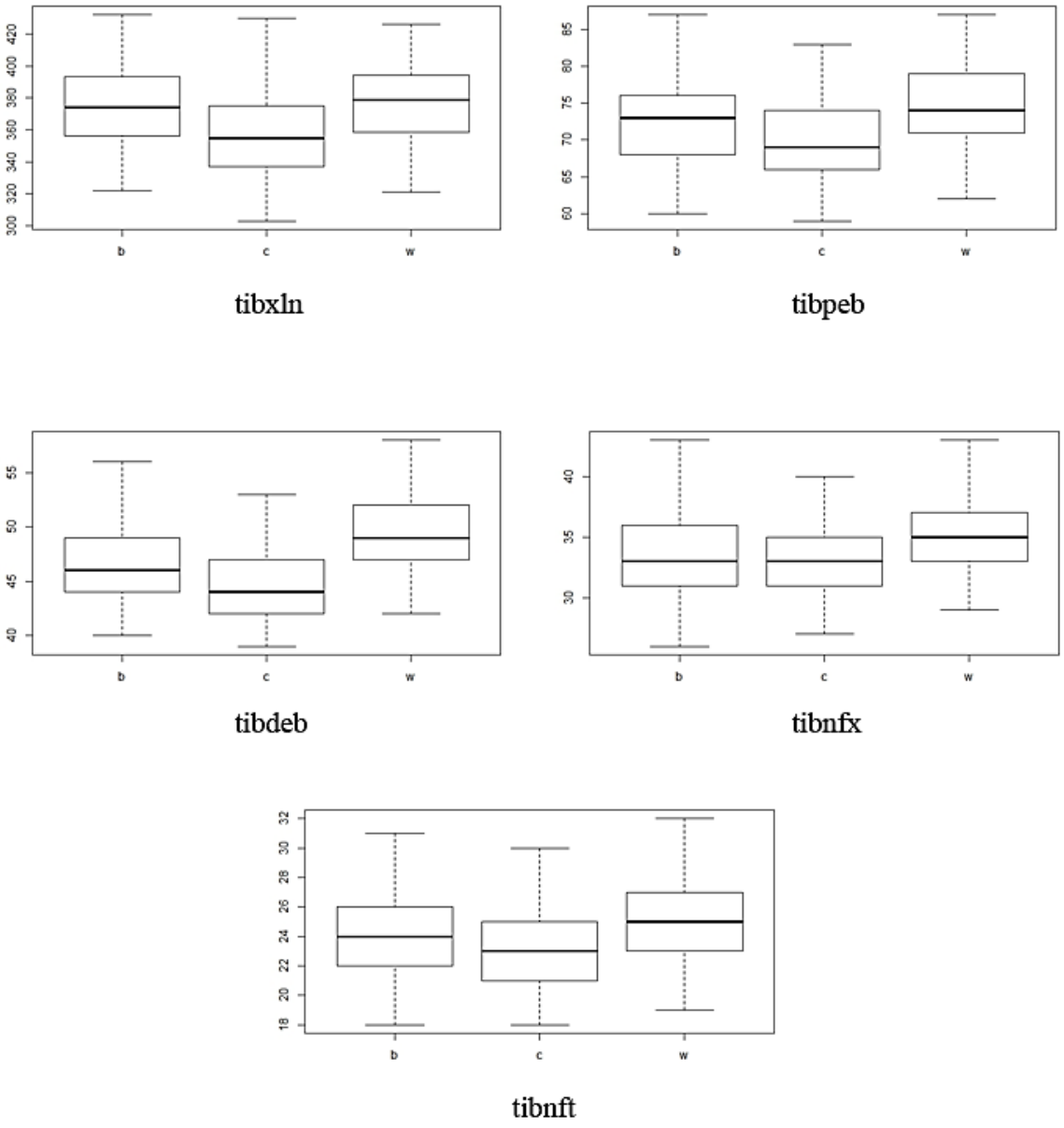
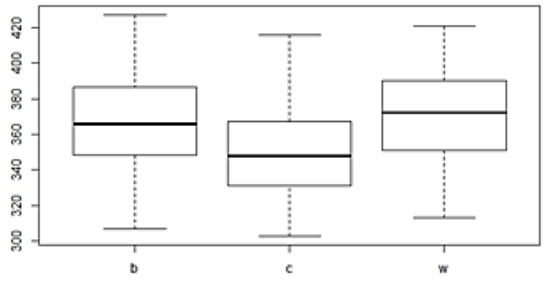
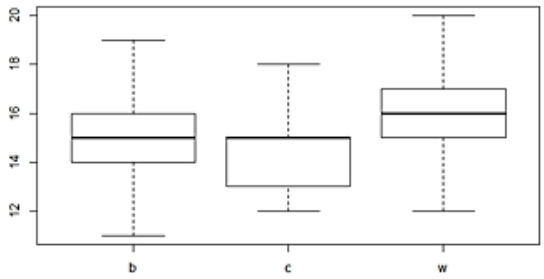


Figure A4.9 - Boxplots illustrating the degree of variation within each group for the measurements of the tibia.



fibxln



fibmdm

Figure A4.10 - Boxplots illustrating the degree of variation within each group for the measurements of the fibula.

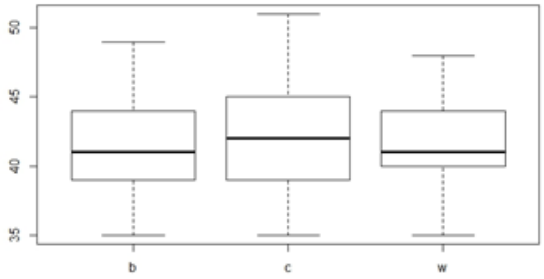


Figure A4.11 - Boxplots illustrating the degree of variation within each group for the measurements of the calcaneus.

APPENDIX V – MANOVA

Table A5 - MANOVA results evaluating the effects of ancestry, sex and the interaction between ancestry and sex for each bone. Bold indicates significant.

	Ancestry		Sex		Ancestry*Sex	
	Pillai-value	**Pr (< F)**	**Pillai-value**	**Pr (< F)**	**Pillai-value**	**Pr (< F)**
Clavicle	0.148	**<0.001**	0.646	**<0.001**	0.027	0.34
Scapula	0.258	**<0.001**	0.570	**<0.001**	0.003	0.91
Humerus	0.543	**<0.001**	0.640	**<0.001**	0.033	0.35
Radius	0.274	**<0.001**	0.652	**<0.001**	0.021	0.32
Ulna	0.314	**<0.001**	0.616	**<0.001**	0.052	**0.025**
Sacrum	0.526	**<0.001**	0.329	**<0.001**	0.027	0.66
Pelvis	0.576	**<0.001**	0.430	**<0.001**	0.039	**0.012**
Femur	0.532	**<0.001**	0.607	**<0.001**	0.094	**0.043**
Tibia	0.460	**<0.001**	0.587	**<0.001**	0.55	0.099
Fibula	0.224	**<0.001**	0.346	**<0.001**	0.002	0.95
Calcaneus	0.216	**<0.001**	0.478	**<0.001**	0.021	0.12

APPENDIX VI – UNIVARIATE SECTIONING POINTS

Table A6 - Sectioning points separating black (B), coloured (C) and white (W) South Africans.

Variable	Sectioning points (mm)
claxln	C < 150 < BW
scapht	CB < 149 < W
scapbr	CB < 100 < W
humxln	C < 306 < B < 317 < W
humebr	CB < 60 < W
humhdd	CB < 44 < W
hummxd	CB < 22 < W
hummwd	CB < 17 < W
radxln	C < 236 < BW
ulnxln	C < 253 < BW
ulnphl	C < 221 < BW
sacaht	CB < 104 < W
sacabr	CB < 96 < W
sacs1b	C < 46 < BW
innoht	CB < 205 < W
iliabr	CB < 155 < W
femxln	C < 435 < B < 449 < W
fembln	C < 432 < B < 445 < W
femebr	C < 75 < B < 78 < W
femhdd	CB < 44 < W
tibxln	C < 367 < BW
fibxln	C < 359 < BW