Evaluating postcranial macromorphoscopic traits to estimate ancestry among modern South Africans

Submitted by: Nomshado Pearl Bothma

A thesis presented in publication format, submitted to the Department of Anatomy, School of Medicine, Faculty of Health Sciences, University of Pretoria, in fulfilment of the requirements for the degree MSc in Anatomy

30 January 2023

Supervisor: Ms L Liebenberg

Co-supervisor: Prof. E.N. L'Abbé

ACKNOWLEDGEMENTS

Upon receipt of my acceptance letter for the master's degree enrolment, I had lot of doubt and fear. I doubted myself and feared that I would not be able to do it, especially being a full-time mom, running a household and working part-time. But I did it! However, it would have not been possible without the people in both my academic and personal life, and I would like to extend my gratitude to them.

To my supervisor, Ms Leandi Liebenberg, words cannot even describe how grateful I am for you. There is a lot that I have learnt from you such as paying attention to the finest detail. Thank you for all your time, the edits, the reading suggestions, and the advice you gave. You made the journey easy for me. I am here today because you believed in me. Thank you for your endless support and help!

To my co-supervisor, Professor Ericka L'Abbé, thank for all your input, edits, and comments. Thank you for opening a lot of doors for me. Thank you for seeing my potential and letting me gain skills and working experience through the *Bakeng se Afrika* and *Dirisana+* projects. It was and still is a great honour to be part of such great initiatives.

I would also like to thank the University of Pretoria Postgraduate Masters Research Bursary programme for providing funds for my tuition. I would like to also acknowledge the EU Erasmus+ grant [597924-EPP-1-2018] *Bakeng se Afrika* for providing financial support for this project.

Thank you to the curators of the Pretoria Bone and Kirsten Skeletal collections, Ms Gabi Krüger and Dr M Alblas for granting me permission to use the skeletal material in the current study. I would also like to thank the entire Forensic Anthropology Research Centre team!

To my fellow master's student, Ms Miksha Harripershad, thank you for all your support and for lending me an ear whenever I panicked. You were one of the people that encouraged me to believe in myself. And thank you for being my inter-observer!

To the love of my life, my best friend and my husband, Ronnie Bothma, you have been there from the beginning. You know my struggles, my strengths, and my weaknesses and in all that you helped realize my true potential. Thank you for pushing me to always give and do my best. You have been my source of emotional and financial support through-out. Most importantly,

thank you for all your understanding especially when I had to stay up late or when I could not join you for morning walks because I was busy with the write-up.

To my son, Maverick Bothma, you are my main motivation. Thank you for understanding when I could not play games with you or stay in bed with you until you fell asleep. Thank you for being part of the journey. Your random hugs came when I needed them the most.

To my mother, Pastor Thokozile Nhlengethwa, thank you for laying the foundation. Without you and your prayers, I would have not made it this far. Thank you for the good examples you set. At 59 years you went an obtained a master's degree! Now that is admirable. Thank you for being an inspiration. And to my sister, Mrs Nothando Moropodi, thank you for always encouraging me to work hard and give this project my all.

ABSTRACT

When decomposed human remains are recovered, the expertise of a forensic anthropologist is required to assist in the identification of the decedent. Identification involves establishing a biological profile, which includes the estimation of age, sex, ancestry, and stature of the individual. Robust methods are needed to assist in creating an accurate biological profile. While osteometric methods are currently preferred for ancestry estimation for forensic analyses in South Africa, non-metric methods can provide valuable information and need to be further explored. The current study aimed to assess postcranial macromorphoscopic traits as a tool to estimate ancestry among modern South Africans.

A sample of 271 postcranial skeletons belonging to black, white, and coloured South Africans were used to score a series of eleven macromorphoscopic traits. The skeletal material was sourced from the Pretoria Bone Collection (University of Pretoria) and the Kirsten Skeletal Collection (Stellenbosch University). The intra- and inter-observer agreement ranged from fair to almost perfect for all but one trait (accessory transverse foramen of C1*)*. The traits varied in frequency and rarity among the populations, with only seven traits demonstrating significant differences between at least two of the groups. Univariate and multivariate random forest models were created to test the positive predictive performance of the traits to classify ancestry. The univariate models performed poorly, with accuracies that ranged from 33.0% to 53.0%. The overall classification accuracy for the multivariate model incorporating all traits was not much better at 54.6%

The results of the current study indicate that the postcranial macromorphoscopic approach does not outperform current methods employed to estimate ancestry. Furthermore, the low accuracies and Kappa values obtained with the random forest models suggest that the traits are not reliable classifiers, and as such, the method does not currently have practical applicability for medicolegal casework. However, with significant differences observed, more research needs to be conducted to potentially improve the method for use in South Africa.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Acronym	Meaning
AMCF	Anterior and medial calcaneal facets
aNN	Artificial neutral networks
ATF	Accessory transverse foramen
DSAF	Double superior articular facets
KSC	Kirsten Skeletal Collection
MMS	Macromorphoscopics
PB	Posterior bridging
PBC	Pretoria Bone Collection
P-value	Probability value
RFM	Random Forest Modelling
SA	Septal Aperture
SCP	Supracondyloid process
SPB	Spinous process bifurcation
SSF	Suprascapular foramen
STA	Sternal aperture
SUN	Stellenbosch University
SVM	Support vector machines
TT	Third trochanter
UP	University of Pretoria
VN	Vastus notch

CHAPTER 1: INTRODUCTION

Forensic anthropologists in South Africa are often presented with human remains that are inadvertently discovered with little or no context about the decedent (Steyn et al., 2016; Krüger et al., 2018). The remains may be skeletonized or in various stages of decomposition with no form of identification present (Steyn et al., 2016; Krüger et al., 2018). Therefore, alternative methods are needed to identify the unknown individual. Forensic anthropologists assess skeletal remains to provide a biological profile and assist in the process of presumptive identification (L'Abbé and Steyn, 2012). The biological profile consists of sex, age, stature, and ancestry. The parameters of the biological profile are based on skeletal variation that has been quantified and correlated to specific demographic characteristics through meticulous research (L'Abbé and Steyn, 2012).

Ancestry estimation is one of the most important factors in establishing the biological profile as many methods that are employed to assess the other parameters are population-specific and require prior knowledge of ancestry (e.g. Feldesman and Lundy, 1988; Oettlé and Steyn, 2000; Krüger et al., 2015). Ancestry (also denoted as population affinity) refers to the geographical origin and population history of a particular individual and how the combination of both these factors influences skeletal morphology. Essentially, ancestry estimation is the classification of an individual into the population to which the unknown individual had mostly likely belonged. In the forensic anthropological context, different combinations of skeletal traits and osteometric dimensions have been correlated to populations across the globe to estimate ancestry. An ancestry estimate is obtained through the application of statistical analyses and methods, using reference samples of known individuals for comparative purposes (L'Abbé et al., 2013; Liebenberg et al., 2015, 2019).

Methods employed for ancestry estimation are particularly important in countries with diverse populations. South Africa is such a country, with a heterogeneous population that consists of socially identified black, white, coloured, and Indian or Asian individuals. Heterogeneous populations display skeletal variation among the populations; however, there is also substantial group overlap (Stull et al., 2014; Liebenberg et al., 2015, 2019). Because of the diversity, developing various methods of ancestry estimation for South African populations is imperative to effectively quantify group variation.

Metric and non-metric methods have been developed to estimate ancestry using several different skeletal elements (Hefner, 2009; L'Abbé et al., 2013; Liebenberg et al., 2015, 2019). Standard metric methods quantify the size of skeletal elements through the use of measuring tools, like calipers. In comparison, non-metric methods quantify the size and shape of skeletal elements through the visual evaluation of morphological skeletal variants (Ousley and Jantz, 2012; L'Abbé et al., 2013). Non-metric methods have popularly been used to assess and distinguish differences between populations in the past, with a revival of the method in the last decade (Hooton, 1926; Rhine, 1990; Hefner, 2009).

Forensic anthropologists continuously study and explore human variation as a means to progress the non-metric method from a subjective to a more objective approach (Hefner and Ousley, 2005; Hefner, 2009; Klales and Kenyhercz, 2015). For example, Hefner (2009) introduced the standardization of non-metric trait analysis of the cranium to avoid subjectivity when using non-metric traits. The standardization included the introduction of line drawings, better definitions, and robust statistics (Hefner, 2009). As a result, the scoring of the traits on an ordinal scale according to their degree of expression was introduced, commonly known as the macromorphoscopic (MMS) approach.

Currently the cranial MMS traits are the most developed non-metric method for ancestry estimation (Hefner and Ousley, 2014; Klales and Kenyhercz, 2015). In South Africa, cranial MMS traits have been assessed for population variation and repeatability, although further research still needs to be conducted for the method to be useful as a tool for the classification of ancestry (L'Abbé et al., 2011).

Regarding the postcranial skeleton, traits have previously been assessed in the evaluation of biodistance between populations, with very little application in forensic anthropology (Finnegan, 1978; Finnegan and McGuire, 1979; Donlon, 2000). The postcranial MMS approach is fairly new in the forensic context, and currently there are no published studies that have assessed postcranial MMS traits for ancestry estimation in South Africa (Spiros, 2019; Spiros and Hefner, 2020). Therefore, the postcranial MMS traits need to be investigated to ascertain if the method is suitable to estimate ancestry for the South African population.

This purpose of this study was to explore eleven postcranial MMS traits as a tool to estimate ancestry among black, white, and coloured South Africans. The objectives were to test the inter-and intra-observer repeatability of eleven MMS postcranial traits on 271skeletons; to

explore the frequency distribution of each trait between the sexes and among the three socially defined populations; and to define the univariate and multivariate classification accuracy of the eleven MMS postcranial traits to estimate ancestry among three South Africans groups when using random forest modelling (RFM).

CHAPTER 2: LITERATURE REVIEW

To effectively assess the skeletal variation associated with ancestry, numerous approaches are required. This literature review discusses research that has evaluated non-metric traits to assess population relatedness. In addition, the transition from non-metric to MMS trait methods is explored. The literature review also addresses studies that have assessed postcranial MMS traits for the evaluation of ancestry. But first, the histories associated with race and ancestry are revisited to better understand the current view on human variation.

2.1. Race and ancestry

The difference between race and ancestry needs to be clarified as the concepts have different definitions and societal implications. Race is a social construct used to group humans according to arbitrary features including physical traits (such as skin colour, hair texture), culture, temperament and behaviour, with the aim to subdivide humans into a limited number of discrete races, or "types" (Livingstone, 1962; Montagu, 1962; Brace, 1964; Cartmill, 1998; Loue, 2006). Ancestry, on the other hand, is based purely on skeletal variation. Skeletal variation or diversity within populations cannot be represented by the race concept as the social race view has no genetic or biological basis (Livingstone, 1962; Montagu, 1962; Brace, 1964; Cartmill, 1998). The current study does not aim to justify the concept of race but rather, the focus of this study is to expand on the understanding of human variation by assessing skeletal morphology.

François Bernier, a French physician, proposed that humans should be classified into four races based on physical attributes such as facial features and body types (Loomba and Burton, 2007). In the 18th and 19th centuries, humans were classified into categories including taxonomic classifications by Carolus Linnaeus (Goodman et al., 2019). German anatomist, Johann Blumenbach, classified humans into five categories and described them as human varieties and not races (Blumenbach, 1865). Blumenbach's five categories were Caucasian, Mongolian, Ethiopian, American, and Malayan. The categories were described according to colour such as white, yellow, black, red, and brown. The human categorization later became the basis of racial classification according to colour by scientists of later generations. The scientists included Carleton Coon, who also classified humans into five "races" based on physical features, namely: Caucasoid (whites); Negroid (blacks); Mongoloid (Asians); Capoid, and Australoid (Coon, 1962).

Before the introduction of the current understanding of genetics, scientists believed that the categories of humans were also associated with certain behaviours and attitudes (Jackson and Weidman, 2005; Goodman et al., 2019). For example, white individuals were described as the "superior race" that was active and creative, and black individuals were labelled as the "inferior race", careless and lazy (Hudson, 1996; Tishkoff and Kidd, 2004; Takezawa, 2011; Jackson and Weidman, 2005). The racial categorization of humans was a driving tool for the rise of eugenics in the early 20th century (Jackson and Weidman, 2005). Eugenics in terms of "race" was created as justification to sterilize people of socially classified lower races, which included individuals of colour (Jackson and Weidman, 2005).

Typological approaches to investigating population differences function under the assumption that races or "types" are discreet, with no group overlap, and possess a suite of traits unique to that particular group. The typological view of human races led Earnest Hooton, one of the early physical anthropologists in North America, to develop the Harvard list. Hooton's work was focused on racial classifications (Hooton, 1926). The Harvard list used cranial non-metric traits to categorize human groups. With the Harvard list, a trait would be linked to a certain population with no regard for overlap or variation, which contributed to scientific racism (Hooton, 1926). For example, investigators would look for traits that appeared more commonly in a certain racial group and would automatically assume that a skull exhibiting that particular trait belongs to that group with no regard for human variation (Hooton, 1926). Overall, the non-metric trait list qualified as typological.

Forensic anthropologists have shifted from the typological approach to focus on understanding within- and between-population variations (Hefner, 2009; Hefner and Ousley, 2012, 2014; Klales and Kenyhercz, 2015; Spiros and Hefner, 2020). The term "ancestry" has also been introduced instead of "race" when referring to an individual's population (Ousley et al., 2009). "Ancestry" has been the most popular term used in biological anthropology for the last decade (L'Abbé et al., 2011; Klales and Kenyhercz, 2015; Spiros, 2019; Spiros and Hefner, 2020). More recently, a discussion has been initiated within the discipline of biological and forensic anthropology to consider the term "population affinity" in lieu of "ancestry", as it encapsulates a better understanding of how microevolutionary forces influenced human variation (Ross and Pilloud, 2021; Spradley and Jantz, 2021). With this discussion still ongoing, the term "ancestry" is used throughout this dissertation to remain consistent with the bulk of the literature reviewed for this study.

2.2. Ancestry estimation

With the study of human skeletal variation, reference samples have been created for various populations to assist in ancestry estimation (Ousley and Jantz, 2012; L'Abbé et al., 2013; Ousley and Jantz, 2013; Liebenberg et al., 2019). In a forensic anthropological context, reference samples are a collection of skeletal data obtained from individuals with known, age, sex, and ancestry. The reference samples are used to compare an individual to the populations present in the samples to obtain a probable estimate. Currently, reference samples that are available for the South African population include that of black, white and coloured individuals. The skeletal data collected from the three populations specifically for sex and ancestry, were obtained through osteometry (measurement of skeletal elements). Both craniometric and postcraniometric databases have been established for the South African population (L'Abbé et al., 2013; Liebenberg et al., 2019, 2015). Compared to the osteometric method, non-metric methods to estimate ancestry among South Africans are undeveloped. Non-metric methods pertaining to ancestry have not always been used in a forensic context, but rather in a biological sense to study global populations and population histories (Corruccini, 1974; Finnegan, 1978; Buikstra et al., 1990; Donlon, 2000; Reed, 2006; Du Toit, 2014; Asvat, 2012). To date, most analyses pertaining to a non-metric method for ancestry estimation, are mostly based on the assessment of cranial and less of the postcranial macromorphoscopic (MMS) traits (Hefner and Ousley, 2005; Hefner, 2009; Klales and Kenyhercz, 2015; Spiros, 2019; Spiros and Hefner, 2020).

With MMS analyses, traits are scored through physical observation using drawings and descriptions of trait character states. Certain skeletal traits assist forensic anthropologists to estimate ancestry as some traits may, or may not, be more prominent in some populations over others. Thus, standards for cranial MMS traits have been developed for ancestry estimation specifically for a North American sample. However, postcranial MMS trait assessment for ancestry estimation has not received the same level of research as compared to cranial MMS traits. More work and research need to be carried out for non-metric analysis for ancestry estimation purposes specifically for South Africans.

2.3. Non-metric traits and biodistance studies

Non-metric traits are non-pathological skeletal features that may vary in size or expression among populations. Many terms have been used when referring to non-metric traits, such as discrete or dichotomous. Discrete or dichotomous traits are described as either present or

absent. The traits were also considered as epigenetic traits because they were believed to be the result of both genetic and environmental influences interacting with one another (Haas et al., 1994; Hauser and Stefano, 1989). However, the term non-metric is generally used as it encompasses the meaning of the related terms (Saunders and Rainey, 2008).

In the 18th and 19th centuries, cranial and postcranial non-metric traits were observed as anomalies rather than traits that indicate human variation (Cunningham, 1886; Macalister, 1893; Russell, 1900). This perspective later changed to note that the traits can be assessed to investigate population similarities and differences (Finnegan, 1978; Donlon, 2000; Spiros, 2019; Spiros and Hefner, 2020). The first attempt to link cranial non-metric traits to anthropological research occurred in the 1800s with biodistance analysis (Berry and Berry, 1967). A biodistance analysis is the assessment of population relatedness to determine similarities and dissimilarities between populations based on polygenic traits through the application of multivariate statistical methods using data obtained from skeletal remains (Buikstra et al., 1990; Hefner et al., 2016). Biodistance studies demonstrate genetic and environmental variations within and between populations (Larsen and Walker, 2010). For example, dissimilarities within a population would be due to genetic variation, and dissimilarities between populations would be due to the different environmental origins of the different populations (Larsen and Walker, 2010; Stojanowski, 2018).

Many biodistance studies have made comparisons between populations by observing trait frequencies (Berry and Berry, 1967; Hauser and Stefano, 1989). As illustrated by Berry and Berry (1967), a comparison of populations from eight different regions was conducted to assess cranial non-metric trait frequencies found both among and within the populations. In the 1970s, a biodistance analysis between indigenous North Americans and groups of African ancestry was carried out using a Measure of Divergence statistic (Ossenberg, 1976). Modern samples collected from the Terry Collection in North America, have also been examined in biodistance studies (Corruccini, 1974). The studies that conducted biodistance analysis using cranial non-metric traits noted significant distances, and concluded that the traits can be useful in comparing different populations (Berry and Berry, 1967; Corruccini, 1974; Ossenberg, 1976).

Postcranial non-metric traits have also been used to study biological distance (Finnegan, 1978; Donlon, 2000). Finnegan (1978) employed line drawings to score 30 postcranial non-metric traits using black and white North Americans. According to Finnegan (1978), postcranial are better suited for analysis than cranial traits because postcranial traits are located

on skeletal elements that are likely to survive excavation and prolonged burial. The conclusion in the Finnegan (1978) study was that the traits can be effective in comparing populations. The author (Finnegan 1978) further elaborated that postcranial non-metric traits age dependency is not statistically significant enough to affect population difference. Finnegan (1978) also noted some sexual dimorphism among postcranial non-metric traits and concluded that skewed samples may have been the cause, as the sexes were not equally distributed.

The data collected in the Finnegan (1978) study was used in a different study (Finnegan and McGuire, 1979) to test different statistical methods that can be employed in classifying populations using the same 30 postcranial non-metric traits. Finnegan and McGuire (1979) tested the Rubison procedure, Bayes' theorem, linear discriminant functions, tally method, and the weight of evidence procedures. The statistical methods were tested to assess which of the techniques performed best in classifying populations. Overall, the Rubison classification technique performed best in comparison to the other techniques (Finnegan and McGuire, 1979). The accuracies of various methods were about 50% for the lowest, with the best performing method being just over 90% with the Rubison method showing only 9.23% misclassifications when using postcranial traits (Finnegan and McGuire, 1979). Finnegan and McGuire (1979) concluded that statistical methods exist that can classify individuals with great accuracies when using postcranial traits.

In a later study, Donlon (2000), assessed Australian, African, East Asian, European, and Polynesian groups and compared the relationships observed from 19 postcranial traits to those obtained from traditional craniometric when assessing biodistance. The Donlon (2000) study began with 40 postcranial traits that were reduced to 19 after testing intra-observer error, as some traits were removed because of low repeatability. Some of the traits were the suprascapular foramen and the vastus notch (see definitions on Appendix A). Traits that were extremely rare or extremely common among the populations, were removed. With the remaining traits, Donlon (2000) applied the mean measure of divergence and principal component analysis techniques to study the frequency and variation of the traits between the sampled populations. Trait frequencies were further analysed according to presence or absence. The Donlon (2000) study did not specify whether line drawings with definitions were used, which might have been the reason for the poor repeatability of the removed traits. A trait was considered present if observed on either left or right or both sides (Donlon, 2000). Donlon (2000) concluded that the remaining postcranial non-metric traits are not affected by age and side preference (for bilateral traits). Side preference, also known as handedness, refers to the

body that is mostly dominant or preferred over the other. Another conclusion that was drawn in the Donlon (2000) study was that postcranial traits can be used to measure divergence or relatedness of populations that are within the same region (Donlon, 2000).

Biological distances using postcranial traits have also been studied between black and white South Africans (Du Toit, 2014). Du Toit (2014) noted that some of the traits were population-specific, indicating that population differences can be observed when assessing postcranial traits. Asvat (2012) also analysed a sample of black and white South African individuals to investigate the frequency of the spinous process bifurcation, which is one of the postcranial traits that will be used for the current study. A high frequency of bifidity was observed for white South Africans compared to black South Africans; however, an ancestry estimation evaluation was not conducted (Asvat, 2012). Biodistance studies that assessed postcranial traits assisted in determining which traits are applicable in comparing different populations (Finnegan, 1978; Donlon, 2000; Du Toit, 2014; Asvat, 2012). The Du Toit (2014) and Asvat (2012) studies support the necessity for further research in postcranial traits and their adequacy in comparing South African populations. Additionally, ancestry estimation was not examined in either of these studies (Asvat, 2012; Du Toit, 2014), which opens up a gap for analysis like the one proposed for the current study.

2.4. Macromorphoscopics (MMS) and forensic ancestry estimation

Macromorphoscopics (MMS) traits are quasi-continuous variables that are homologous with the adjacent soft tissue traits and are used by forensic anthropologists to estimate ancestry (Hefner and Ousley, 2005, 2012; Hefner, 2009; Klales and Kenyhercz, 2015). Contrary to MMS traits, non-metric traits are discontinuous variables of skeletal features that are used in biodistance studies (Corruccini, 1974; Buikstra et al., 1990; du Toit, 2014; Stojanowski, 2018). In comparison to biodistance studies, forensic anthropology focuses on an individual rather than a population. Prior to discussing how MMS traits have been used in a forensic context, one needs to understand how the transition from non-metric to MMS traits took place in a forensic context.

The application of non-metric traits in biological anthropology was originally introduced by E.A. Hooton (1926). Hooton's goal was to study the differences of and to classify human groups using traits that were both hereditary and not affected by the environment. He referred to this collection of traits as the Harvard list (Hooton, 1926).

The Harvard list utilized forms and standard descriptions such as “narrow”, “sharp”, “small”, “very thick” etc. (Hooton, 1926; Hefner, 2003). The method in which the traits were described on the Harvard List was subjective, for example, “narrow” to one person may not be “narrow” to the next, and therefore repeatability of the trait was in question. The lack of standardized descriptions required that only experienced observers, who had been previously exposed to the skeletal traits, could reliably score them (Hefner, 2003, 2009). Additionally, the frequency of each of the traits among and within the populations could not be accounted for because the descriptions of the traits on the Harvard List were not standardized (Hooton, 1926).

The Harvard trait list was not statistically tested, and an error rate was not specified. An error rate is one of the requirements in a scientific methodology (Grivas and Komar, 2008). Thus, the trait list method demonstrated to be non-reliable and subjective and was based only on observer experience (Hefner and Ousley, 2012; Plemons and Hefner, 2016). In addition, scientific methodologies cannot be based only on observer experience, as the least experienced of observers need to be able to repeat the method. Thus, Hefner (2009) introduced an objective standardized approach with associated line drawings to illustrate trait variation. With this approach, an ordinal scoring system with descriptions was introduced, such that a statistical analysis could be carried out (Hefner and Ousley, 2005; Hefner, 2009). The term macromorphoscopics (MMS) was created to describe the reviewed and standardized trait analysis applicable to a forensic context (Hefner and Ousley, 2005).

Macromorphoscopic traits belonging to African (West, East, and African Americans), European (Europeans and White Americans), Native American, and Asian samples have been studied (Hefner, 2009). Eleven cranial MMS traits were analysed to explore trait variations among the groups. Significant differences in trait frequencies were noted between the sampled populations. According to Hefner (2009), the extreme expression of a trait should not be linked to a certain population, as trait variations were observed within all populations. The observation in the Hefner (2009) study, further demonstrated that the visual assessment of traits cannot be conducted independently, but rather a statistical approach is needed. In addition, significant population differences were observed throughout the frequency distribution of the traits except for one, the malar tubercle. In the same study, classification accuracy was assessed, where a combination of traits was used and an overall correct classification ranging from 84% to 93% of the populations was obtained (Hefner, 2009). Several statistical techniques were employed, namely; the k-Nearest Neighbour, logistic regression, and the naïve Bayesian to obtain classification rates (Hefner, 2009). Hefner (2009) illustrated that a combination of traits worked

better as compared to individual traits, which further supported a multivariate approach as stated by Liebenberg et al. (2015).

Nine of the MMS traits that were described and illustrated by Hefner (2009) were tested on a South African sample (L'Abbé et al., 2011). Hefner (2009) assessed the repeatability of the non-metric method and the frequency distribution of the traits. After assessing frequency distributions, L'Abbé et al. (2011) noted that black and coloured South Africans fell into the African and Asian groups, and the white South Africans fell into the European groups. Traits that could separate the three South African populations were located in the midfacial area when assessing frequency distribution. The traits were the inferior nasal margin, anterior nasal spine, and nasal contour (L'Abbé et al., 2011). Frequency distributions assist in determining which traits are mostly observed in one group in comparison to another, which can further assist in classifying groups.

Phenotypic variation in trait expression has been noted among various groups, specifically black and white North Americans (Klales and Kenyhercz, 2015). When classifying ancestry using the MMS trait as described in the Hefner (2009) study, correct classification of 86.60% was obtained (Klales and Kenyhercz, 2015). Klales and Kenyhercz (2015) also noted significant differences of trait expressions on the midfacial area between the populations as observed by L'Abbé et al. (2011). In the Klales and Kenyhercz (2015) study, the same statistical method, as described by Hefner (2009) was applied. In addition, the ordinal logistic regression method performed best when classifying black and white North Americans.

Hefner et al. (2014), assessed the efficacy of a random forest model for ancestry estimation using both cranial macromorphoscopic and metric data of a North American sample. Correct classification of 89.6% was obtained when a random forest model was employed (Hefner et al., 2014). The findings in the Hefner (2009), Hefner et al (2014), and Klales and Kenyhercz (2015) studies further justify the utilization of MMS traits for ancestry estimation. The studies further illustrate the efficacy of certain statistical models, as high classification accuracies were obtained. Moreover, the random forest model will be employed for the current study.

Postcranial MMS traits have been examined to investigate if they can be used for ancestry analyses (Spiros, 2019; Spiros and Hefner, 2020). The reliability of the scoring technique and the frequency distribution of postcranial MMS traits was tested between black and white North Americans (Spiros, 2019). Spiros (2019) created trait illustrations and definitions (Appendix A) using over 100 skeletons and numerous isolated bones (Spiros, 2019). As a result, eleven

postcranial MMS traits were standardized. The standardization of the postcranial MMS traits was achieved by selecting traits according to how frequent they were in terms of differentiating between populations. All eleven traits demonstrated complete agreement for intra-observer error and nine out of the eleven traits had a high inter-observer agreement. The two remaining traits had substantial agreement which were the vastus notch and the posterior bridging (see Appendix A for trait definitions) (Spiros, 2019). Sex and side differences (excluding the septal aperture) were not statistically significant and therefore sexes were pooled, as well as traits recorded on both left and right sides. Frequency distributions of the traits were calculated for ancestry comparison between the black and white North American populations. Statistically significant frequency distribution differences were noted for some traits such as the third trochanter, spinous bifurcation of the C3 and C4 of the cervical vertebra, septal aperture, and the anterior and middle calcaneal facets (Appendix A). Spiros (2019) also noted that there was no trait specific to one population, as all traits presented in a single skeleton when evaluating black and white North Americans, which further supports variation of traits between and within populations (Hefner, 2009; Hefner and Ousley, 2014; Klales and Kenyhercz, 2015). Spiros (2019) further suggested that the left and right sides of the septal aperture should be scored separately rather than pooled. Spiros (2019) encouraged further investigation of the postcranial MMS traits and their correlation to ancestry estimation.

The classification accuracy of the ordinal scoring technique created by Spiros (2019) was further tested on black and white North Americans using both cranial and postcranial MMS traits (Hefner, 2009; Spiros, 2019; Spiros and Hefner, 2020). The eleven postcranial traits used in the Spiros (2019) study were utilized for the assessment of the postcrania (Spiros, 2019). Bilateral traits were not pooled for the Spiros and Hefner (2020) study. Classification models were created for the postcranial MMS traits using quadratic discriminant function analysis (QDA), support vector machine (SVM), artificial neural networks (aNN), and random forest models (RFM) techniques. Statistical techniques such as the quadratic discriminant function analysis and RFM were employed to measure the effectiveness of the postcranial MMS traits in ancestry estimation. Correct classification accuracy of 77.6% – 81.6% was obtained for models developed for postcranial MMS traits. A combination of cranial and postcranial traits had a correct classification of 89.5 – 92.1% with the spinous process bifurcation being the best discriminator between black and white North Americans.

Due to the typological history associated with the morphological method, Bethard and DiGangi (2020) argue that MMS traits should be suspended in the estimation of ancestry.

However, forensic anthropologists have introduced a statistical approach for both cranial and postcranial MMS methods for ancestry estimation (Hefner and Ousley, 2005; Hefner, 2009; Klales and Kenyhercz, 2015; Plemons and Hefner, 2016; Spiros, 2019; Spiros and Hefner, 2020). The studies that have examined postcranial MMS in a forensic context (Spiros, 2019; Spiros and Hefner, 2020), have also demonstrated that the traits can assist in classifying sampled populations. However, as discussed, the studies were conducted on North American samples. The lack of postcranial MMS analysis for ancestry estimation signifies the need to further investigate if the method can be employed in a forensic context to estimate ancestry for South African populations. Furthermore, a postcranial MMS analysis can assist in quantifying skeletal features that a metric analysis is unable to quantify such as the presence or absence of a feature that can possibly contribute to ancestry estimation.

The addition of reliable postcranial MMS methods for ancestry estimation can assist in learning more about human variation among the populations being assessed and possibly develop reference samples that can be used for future research and forensic case analysis. However, the postcranial MMS method needs to be validated and tested through the application of statistical methods suitable for South Africans.

CHAPTER 3: MATERIALS AND METHODS

3.1. Materials

The sample consists of the postcrania of 271 black, white, and coloured South Africans (Table 1). The sample was obtained from the Pretoria Bone Collection and the Kirsten Skeletal Collection located at the University of Pretoria and the Stellenbosch University, respectively. All individuals in the sample were adults, older than 18 years of age (see Table 2 for age distribution of the sample). Any individuals with excessive post-mortem damage or pathological lesions that prevented the accurate scoring of the traits were excluded. Additionally, individuals that were missing more than four traits were excluded from the sample.

TABLE 1. The sample distribution.

Population	Males	Females	Total
Black	46	41	87
White	49	41	90
Coloured	47	47	94
Total	142	129	271

TABLE 2. Age of sample (average).

Population	Males		Females		Total average	
	Mean	Range	Mean	Range	Mean	Range
Black	48.5	(24-76 years)	44.3	(22-75 years)	46.5	(22-76 years)
White	51.8	(20-85 years)	69.0	(42-88 years)	57.0	(40-88 years)
Coloured	63.7	(40-85 years)	45.2	(18-81 years)	48.5	(18-85 years)

3.1.1. The South African population history

South Africa has a heterogeneous population with more than 59 million individuals, including black (80.8%), white (7.8%) coloured (8.8%), and Indian or Asian (2.6%) populations (Statistics South Africa, 2020). The South African population is also diverse in terms of culture, beliefs, and origins. South Africa is rich in culture and has eleven official languages, namely, Northern and Southern Sotho, Afrikaans, English, Zulu, Xhosa, Ndebele, Swati, Tsonga, Venda, and Tswana. The eleven official languages have a direct link to the different tribes. For

example, the black population is divided into different tribes such as the: Nguni (SiSwati, isiZulu, isiXhosa, isiNdebele); Sotho (Setswana, Sepedi); Tsonga, and Venda (Thompson, 2001; Franklin et al., 2007; L'Abbé et al., 2011). The black South African population initially descended from Bantu-speaking groups that migrated from West Africa to sub-Saharan Africa and settled south of the Limpopo River (by AD 300) (Thompson, 2001; Ribot, 2004).

In the 1600s, the Dutch East India Company established a colony in the Cape of Good Hope which led to the immigration of more European settlers into South Africa. Apart from the Dutch, the European settlers included the British, German, and the French (Davenport and Saunders, 2000; Thompson, 2001). Some of the settlers established farms in the Cape which led to an increase in the immigration of European women. The European settlers then became the parental groups for the English and Afrikaans speaking white South African population (Davenport and Saunders, 2000; Thompson, 2001; Patterson et al., 2010).

The Cape coloured group, also known as the South African coloured population, is a heterogeneous population with the widest variety of global genetic contributions that is both intra- and inter-continental (Patterson et al., 2010). According to Thompson (2001), between the 17th and 18th centuries, populations from India, Malaysia, Madagascar, Mozambique, Indonesia, and Ceylon (modern-day Sri-Lanka) were transported from their homelands as slaves for the European settlers (Thompson, 2001; Patterson et al., 2010). As a result, together with the slave groups, the European settlers, black and Khoesan populations contributed to the South African coloured population. The Khoesan group is the major contributor to coloured South Africans, particularly along the maternal lineage (Quintana-Murci et al., 2010). The great maternal contribution may be a result of the Khoesan population being the first to encounter the European settlers who were initially males.

The introduction of legislature such as the Immorality Act of 1949, Group Areas Act of 1950, and the Natives Amendment Act of 1952, led to the socially defined groups, namely black, white, coloured and Asian/Indian South Africans, which was followed with forced segregation of these groups. Under institutionalized racism, mixed marriages were prohibited, most especially between whites and other groups (Posel, 2001). The segregation of these socially defined groups limited mating among the groups, reduced gene flow and thus contributed to the patterns of observable skeletal variation.

Skeletal variation has been noted among black, white and coloured South Africans (L'Abbé et al., 2013; Stull et al., 2014; Liebenberg et al., 2019). Skeletal traits overlap among the three

populations, with most similarities being noted between black and coloured South Africans when assessing craniometric and post-craniometric data (Liebenberg et al., 2015). The overlap between black and coloured South Africans may have been a result of the segregation laws not being as strict on black and coloured marriages as compared to marriages between whites and other groups (Posel, 2001). Before the segregation laws were introduced, European male settlers were allowed to marry slaves that shared the same religion (Posel, 2001).

3.1.2. The Pretoria Bone Collection

The Pretoria Bone Collection (PBC) was established in 1942 when the Department of Anatomy and the Medical School at the University of Pretoria was established (L'Abbé et al., 2005, 2021). The main purpose of the PBC was to educate dental, medical, and health students. Today, the PBC is used for student teaching and research purposes. The PBC is composed of donated and unclaimed bodies of known sex, age, stature, ancestry, and cause of death. Willed and unclaimed whole body donations to South African Medical Schools are governed by the National Health Act of 2003, which states that anyone can donate their body for tissue transplants, research, and medical training (National Health Act, 2003). For unclaimed bodies, if an individual dies in a public institution and is unclaimed by a relative or spouse, the body can be donated to an academic institution such as a university for medical research purposes. Most of the unclaimed bodies in the Department of Anatomy come from the local City of Tshwane public hospitals, such as the Kalafong and Mamelodi Hospitals.

3.1.3. The Kirsten Skeletal Collection

The Kirsten Skeletal Collection (KSC) is housed in the Department of Biomedical Sciences in the Faculty of Medicine and Health Sciences at the Stellenbosch University, South Africa (Alblas et al., 2018). According to Alblas et al. (2018), the collection has a total of 1161 skeletal elements from cadavers of known sex, age, date of birth, date of death, ancestry, last known residential address, and hospital or funeral home that the body was received from. The KSC was established in the 1970s and was named after Professor J.F. van E. Kirsten, who was a qualified surgeon and the first to collect skeletal material for study at the Stellenbosch University.

Similar to the PBC, the KSC is a collection of willed and unclaimed whole-body donations of persons who died from natural causes (Alblas et al., 2018). The Department of Biomedical Sciences receives bodies from the Western Cape Region, specifically the Northern suburbs of Cape Town and surrounding rural settlements (Alblas et al., 2018). The cadavers are used in dissection halls for the training of undergraduate, postgraduate as well healthcare medical

students. Most of the unclaimed bodies, similar to the PBC, belong to individuals of low socioeconomic status such as migrant labourers. The collection contains 1161 individuals of which approximately 12.0% are white, 16.5% black and 60.0% coloured South Africans (Albas et al., 2018). This collection is unique, as it encapsulates a large portion of the coloured South African sample (skeletal), as well as the genetic and socioeconomic diversity of these groups (Pfeiffer et al., 2016).

3.2. Methods

A total of eleven postcranial MMS traits were visually assessed and scored by applying the methodology described by Spiros (2019), and Spiros and Hefner (2020). The eleven traits are located on the following skeletal elements: the cervical vertebrae, sternum, scapula, humerus, femur, patella, and calcaneus. In the current study, cervical vertebrae C5 and C6 were included for the assessment of the spinous bifurcation process. In the case of bilateral traits, both the left and right sides were scored; however, only the left-side scores were used to create classification models. Table 3 presents the traits and their descriptions.

TABLE 3. Summary of trait descriptions and scoring system.

Trait and location	Abbreviation	Score	Description	Definition
Accessory Transverse Foramen (cervical vertebra)	ATF_ (C1, C3, C4, C6, C7)	0 1 2	Absent Unilateral Bilateral	One to two extra foramina located near the articular facet of the cervical vertebrae (Spiros, 2019).
Posterior Bridging (cervical vertebra)	PB	0 1 2	Absent Unilateral Bilateral	A bony protrusion, extending from the posterior aspect of the superior articular facet of C1 to the arch located posteriorly (Spiros, 2019).
Double Superior Articular Facets (cervical vertebra)	DSAF	0 1 2	Absent Unilateral Bilateral	Forms when a groove forms in the middle of the superior articular facet of C1, resulting in double facets (Spiros, 2019).
Spinous Process Bifurcation (cervical vertebra)	SPB_ (C3, C4, C5, C6)	0 1 2	Non-bifid Partially bifid Completely bifid	A “split” of the most posterior part of the spinous process of the cervical vertebra to form two bony tubercles or projections (Asvat, 2012; Cunningham, 1886).
Suprascapular Foramen (scapula)	SF	0 1	Absent Present	Formed when the suprascapular notch is enclosed, forming a foramen. The suprascapular notch is located on the superior border of the scapula, at the root of the coracoid process (White et al., 2012).

Sternal Aperture (sternum)	STA	0 1	Absent Present	An oval opening or hole located on the sternum that can vary in size (Mann and Hunt, 2019).
Supra-condyloid Process (humerus)	SCP	0 1	Absent Present	A bony projection located above the medial epicondyle of the humerus and is an extension of the supracondylar ridge (Spiros, 2019).
Septal Aperture (humerus)	SA	0 1 2 3	Absent Translucent Small perforation Large perforation	An opening located on the distal part of the humerus, that joins the olecranon fossa to the coronoid fossa resulting in an oval or round shape hole (Mann and Hunt, 2019).
Third Trochanter (femur)	TT	0 1	Absent Present	A tubercle located at the superior end of the gluteal ridge on the posterosuperior aspect of the femur and slightly inferior to the lesser trochanter (Mann and Hunt, 2019).
Vastus Notch (patella)	VN	0 1	Absent Present	A concaved or flattened indentation located on the superolateral or medial angle of the patella (Mann and Hunt, 2019).
Anterior and Middle Calcaneal Facets (calcaneus)	AMCF	0 1 2 3	No anterior facet Single elongated facet Small anterior facet Large anterior facet	Located on the calcaneus; can vary from a single small oval facet to an elongated oval facet, where the middle facet is joined with an anterior or double facet where the two facets are separated (Spiros, 2019).

The inter- and intra-observer agreement was tested with Cohen's kappa, using the Landis and Koch scale to better describe the degree of repeatability (Landis and Koch, 1977). Eleven individuals were selected to test inter- and intra-observer agreement. The Landis and Koch scale provides parameters to classify the kappa statistic values obtained when testing the strength of intra- and inter-observer agreement with acceptable values preferably being $\kappa>0.61$(Byrt, 1996) (Table 4- Landis and Koch scale).

TABLE 4. Landis and Koch scale (Landis and Koch, 1977).

Kappa statistic (κ)	Strength of agreement
<0.00	Poor
0.01 – 0.20	Slight
0.21 – 0.40	Fair
0.41 – 0.60	Moderate
0.61 – 0.80	Substantial
0.81 – 1.00	Almost perfect

Frequency distributions were created for each trait, and Kruskal-Wallis tests were applied to determine if there are significant differences between the sexes and among the groups for each trait. A *post-hoc* Dunn's test was also applied to further investigate group differences and overlap. Random forest modelling (RFM) was employed to create classification models and see if the traits are useful for ancestry estimation. As a classification method, RFM refers to a combination of decision trees that are generated using a non-parametric algorithm integrating random sampling with replacement and majority voting (Breiman, 2001). The sample was divided so that 75% constituted the training set (to create the models), and the remaining 25% was kept as the out-of-bag (OOB) testing set (to validate the models). In the case of missing data, the mode was calculated for each sex and population separately within each trait and the result was used as the score of that missing trait. The mode was used because it represents the centre or the middle of the sex and population data distribution. Both univariate and multivariate models were employed to evaluate the performance of the traits and to determine which model worked best in classifying the groups. The univariate models assist to determine the performance of each trait and multivariate models depict the relationship between the traits. Three different multivariate models were tested: the first included all the traits; for the second model all traits with a variable importance value below one were removed; and finally, the

third model included only the variables that were found to be significantly different when analysed with the Kruskal-Wallis.

A total of 2500 classification trees were used for each model with four variables at each split. The classification accuracy, Kappa values, and variable importance were recorded for each model. Both the classification and Kappa values are measures of model accuracy. The classification accuracy presents the percentage of correctly classified individuals out of all of the individuals, whereas the Kappa value presents the percentage of correctly classified individuals while taking random chance into account. The Kappa value is a particularly useful metric in the case of unbalanced classes (e.g., where traits can be scored as 0 or 1, but a score of 1 is a fairly rare occurrence). With variable importance, the higher the value the more a variable contributes to the classification model.

CHAPTER 4: TRAIT PREVALENCE AND VARIATION

4.1 Introduction

The purpose of this chapter is to evaluate the repeatability with which each of the postcranial MMS traits can be scored. Additionally, this chapter explores the prevalence of each trait among black, white and coloured South Africans. Assessing the within- and among-group variation is an essential step in the validation of the method and provides information required for the creation of classification models.

4.2 Manuscript to be submitted

Exploring the prevalence of postcranial macromorphoscopic traits among modern South Africans.

N.P. BOTHMA, E.N. L'ABBÉ, L. LIEBENBERG

Manuscript to be submitted for publication to the Journal of Forensic Sciences.

Abstract

As research in forensic anthropology is continuously published, new methods need to be evaluated to determine whether they can be added to currently existing standards. Very few studies have assessed the variation of postcranial macromorphoscopic traits (MMS) for ancestry analysis. The lack of postcranial macromorphoscopic trait databases indicate the need to further investigate if the method can be employed repeatably, specifically in a forensic context. The current study aimed to assess the prevalence of eleven postcranial macromorphoscopic traits among black, white and coloured South Africans.

A sample of 271 postcrania of adult black, coloured, and white South Africans housed at the Pretoria Bone and Kirsten Skeletal Collections in the Gauteng and Western Cape Provinces of South Africa were assessed. The intra- and inter-observer agreement ranged from fair to almost perfect except for one trait, the accessory transverse foramen of C1, which had poor agreement between observers. The frequencies of the traits demonstrated substantial group overlap, with only seven traits differing significantly between at least two of the groups. Numerous trait variations were also observed in the sample, particularly pertaining to the accessory transverse foramen and the suprascapular foramen (scapular notch).

With the observed significant differences, more research needs to be conducted with refined definitions to aid in obtaining an optimal intra- and inter-observer agreement.

KEYWORDS: Forensic anthropology; Population affinity; Morphology; Variation; Observer agreement

Highlights

- This study evaluated postcranial macromorphoscopic traits in three South African groups
- Experience with the traits improves the repeatability of the method
- Numerous trait variations were observed that may complicate scoring
- Seven of the eleven traits were significantly different for ancestry

INTRODUCTION

Anthropological practitioners have developed various methods to estimate ancestry, which includes both metric and non-metric approaches using several different skeletal elements (1–4). Standard metric methods quantify the size of skeletal elements through the use of measuring equipment, like calipers. In comparison, non-metric methods quantify the size and shape of skeletal elements through the visual evaluation of morphological skeletal variants (2,5). Non-metric traits are non-pathological skeletal features that may vary in size or expression among populations. Many terms have been used when referring to non-metric traits, such as discrete or dichotomous. Non-metric methods have popularly been used to assess and distinguish differences between populations in the past (biodistance studies), with a revival of the method in the last decade (1,6,7).

Forensic anthropologists continue to develop and test new methods on populations across the globe to ensure that methods are suitable to effectively quantify skeletal variation and classify populations other than the ones used to create the methods (8,9). Many studies have noted differences between modern South Africans and North Americans (2,10), which have prompted ongoing work to modify existing standards before adopting them for skeletal analyses in South Africa. Currently, a number of South African-specific databases exist which contain metric standards for skeletal analyses (2–4). However, non-metric methods to estimate ancestry has not received as much attention in the country. Hefner (1) introduced the standardization of non-metric trait analysis to avoid subjectivity when using non-metric traits of the cranium. The standardization included the introduction of line drawings, better definitions, and robust statistics (1). As a result, the scoring of the traits, commonly known as macromorphoscopic (MMS) traits, was introduced. More recently, Spiros (11) and Spiros and Hefner (12) introduced similar work assessing postcranial MMS traits.

The reliability of the scoring technique and the frequency distribution of postcranial MMS traits was assessed in a sample of black and white North Americans (11). Spiros (11) created trait illustrations and definitions for eleven traits using over 100 skeletons and numerous isolated bones to ensure robust, standardized methodology (11). All eleven traits demonstrated complete agreement for intra-observer error and nine out of the eleven traits had a high inter-observer agreement (13). Sex and side differences (excluding the septal aperture) were not statistically significant and therefore sexes were pooled, as well as traits recorded on both left and right sides. Frequency distributions of the traits were calculated for ancestry comparison

between the black and white North American populations. Statistically significant frequency distribution differences were noted for some traits such as the third trochanter, spinous bifurcation of the C3 and C4 of the cervical vertebra, septal aperture, and the anterior and middle calcaneal facets. Spiros (11) also noted that there was no trait specific to one population, as any combination of these traits may be present in either one of the population groups, which further supports trait variation between and within population groups (1,13,14). Spiros (11) further suggested that the left and right sides of the septal aperture should be scored separately rather than pooled as significant differences between the sides were observed. Spiros (11) encouraged further investigation of the postcranial MMS traits and their correlation to ancestry and the quantification of population variation.

The addition of reliable postcranial MMS methods for ancestry estimation can assist in learning more about human variation among the populations being assessed and possibly develop reference samples that can be used for future research and forensic case analysis. However, the postcranial MMS method needs to be further explored and validated for analyses involving South Africans. This study aims to explore postcranial variation and the prevalence of eleven postcranial MMS traits as a tool to estimate ancestry among black, coloured, and white South Africans.

MATERIALS AND METHODS

The sample consists of the postcrania of 271 black, coloured, and white South Africans (Table 5). More specifically, the cervical vertebrae, sternum, scapula, humerus, femur, patella, and calcaneus were used (Table 6). All individuals in the sample were adults, older than 18 years of age. Any individuals with excessive post-mortem damage or pathological lesions that prevented the accurate scoring of the traits were excluded. Additionally, individuals that were missing more than four traits, were excluded from the sample.

The sample was obtained from two South African collections: the Pretoria Bone Collection (PBC) and the Kirsten Skeletal collection (KSC) located at the University of Pretoria and the Stellenbosch University, respectively. The PBC is composed of donated and unclaimed bodies of known sex, age, stature, ancestry, and cause of death. The KSC is also a collection of willed and unclaimed whole-body donations of persons who died from natural causes. Willed and unclaimed whole body donations to South African medical schools are governed by the National Health Act of 2003, which states that anyone can donate their body for tissue

transplants, research, and medical training (15). Ethical approval was obtained from the Human Research Ethics Committee (HREC) of the Faculty of Health Sciences (Ref 610/2021).

TABLE 5. The sample distribution.

Population	Males	Females	Total
Black	46	41	87
White	49	41	90
Coloured	47	47	94
Total	142	129	271

A total of eleven postcranial MMS traits were visually assessed and scored by applying the methodology described by Spiros (11), and Spiros and Hefner (12). Table 6 presents the traits and their descriptions. In the case of bilateral traits, both the left and right sides were assessed.

TABLE 6. Summary of the trait names and abbreviations and their associated scores taken from Spiros (11).

Trait	Location	Abbreviation	Score	Description
Accessory Transverse Foramen	Cervical vertebra (C1, C3, C4, C5, C6, C7)	ATF	0	Absent
			1	Unilateral
			2	Bilateral
Posterior Bridging	Cervical vertebra	PB	0	Absent
			1	Unilateral
			2	Bilateral
Double Superior Articular Facets	Cervical vertebra	DSAF	0	Absent
			1	Unilateral
			2	Bilateral
Spinous Process Bifurcation	Cervical vertebra (C3, C4, C5, C6)	SPB	0	Non-bifid
			1	Partially bifid
			2	Completely bifid
Suprascapular Foramen	Scapula	SF	0	Absent
			1	Present
Sternal Aperture	Sternum	STA	0	Absent
			1	Present

Supra-condyloid Process	Humerus	SCP (L & R)	0	Absent
			1	Present
Septal Aperture	Humerus	SA (L & R)	0	Absent
			1	Translucent
			2	Small perforation
			3	Large perforation
Third Trochanter	Femur	TT (L & R)	0	Absent
			1	Present
Vastus Notch	Patella	VN (L & R)	0	Absent
			1	Present
Anterior and Middle Calcaneal Facets	Calcaneus	AMCF (L & R)	0	No anterior facet
			1	Single elongated facet
			2	Small anterior facet
			3	Large anterior facet

The inter- and intra-observer agreement was tested with Cohen's kappa, using the Landis and Koch scale to describe the degree of repeatability (Table 4). Eleven individuals were randomly selected to test inter- and intra-observer agreement. The first and second intra-observer agreement rounds were a week apart.

Frequency distributions were created for the traits, and Kruskal-Wallis tests were applied to determine if there are significant differences between the sexes and among the groups for each trait. A *post-hoc* Dunn's test was also applied to further investigate group differences and overlap.

RESULTS

Inter- and intra-observer agreement

The intra-observer agreement ranged between moderate and almost perfect (κ = 0.58 to 1.00) (Table 8 and Figure 1). Overall, six of the eleven traits demonstrated almost perfect agreement (κ = 1.00). The trait with the lowest agreement was the spinous process bifurcation of C4 (κ = 0.58). The inter-observer agreement was substantially lower than the intra-observer agreement, ranging between poor and almost perfect (κ = -0.11 to 1.00). When comparing scores between two different observers, only four of the eleven traits demonstrated almost perfect agreement. Additionally, the accessory transverse foramen of C1 presented with agreement poorer than randomly allocating a score (κ = -0.11). Some traits had a "non-applicable" outcome such as the accessory transverse foramen for C3 and C4, suprascapular foramen, posterior bridging, supracondylar process and the third trochanter. The "non-applicable" Kappa outcome is due to the lack of variation observed among the three population groups, as none of the randomly selected specimens had the traits and thus all received the same score (i.e. a score of 0 to indicate absent) (16).

TABLE 7. Kappa values for the inter- and intra-observer agreement with the associated description following Landis and Koch (17).

Trait	Intra-observer	Description	Inter-observer	Description
ATF_C1	1.00	Almost perfect	-0.11	Poor
ATF_C3	N/A*		N/A*	
ATF_C4	N/A*		N/A*	
ATF_C5	0.84	Almost perfect	0.85	Almost perfect

ATF_C6	1.00	Almost perfect	0.67	Substantial
ATF_C7	0.87	Almost perfect	0.14	Slight
PB	1.00	Almost perfect	N/A*	
DSAF	1.00	Almost perfect	1.00	Almost perfect
SPB_C3	0.86	Almost perfect	0.40	Fair
SPB_C4	0.58	Moderate	0.51	Moderate
SPB_C5	0.72	Substantial	0.63	Substantial
SPB_C6	N/A*		0.64	Substantial
SSF	N/A*		N/A*	
STA	1.00	Almost perfect	1.00	Almost perfect
SCP	N/A*		N/A*	
SA	0.88	Almost perfect	0.68	Substantial
TT	1.00	Almost perfect	N/A*	
VN	0.75	Substantial	0.62	Substantial
AMCF	1.00	Almost perfect	1.00	Almost perfect

*N/A: “not applicable” Kappa value outcome due to lack of variation

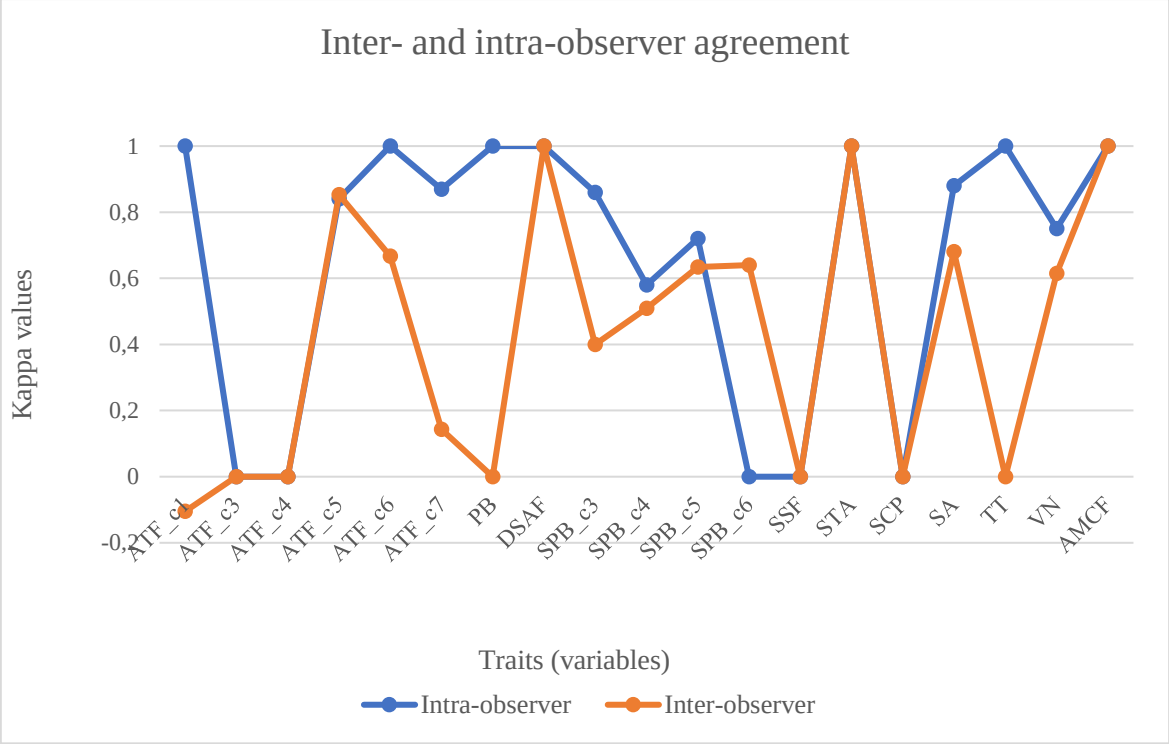

FIGURE 1. Visual representation of Kappa values for inter- and intra-observer agreement.

Frequency distribution

Table 9 presents the frequency distributions for the traits among the three populations and between the sexes. For ancestry, the results of the Kruskal-Wallis test revealed that seven of the traits were significantly different; this includes the accessory transverse foramen (only for C4, C5 and C6), double superior articular foramen, spinous process bifurcation (C3 to C6), suprascapular foramen, septal aperture, vastus notch and the anterior and medial calcaneal facets traits. Notably, for two of the bilateral traits (septal aperture and vastus notch), only the right side was observed to differ significantly among the groups ($p<0.05$). Not a single instance of a supra-condyloid process was recorded in the sample, indicating that it will not be a useful trait to distinguish among the groups.

A Dunn's test was conducted to further explore the variation of the traits among the groups (see Table 11 for the breakdown of population group overlap). None of the traits demonstrated significant differences among all three groups; in other words, at least two of the groups showed overlap for the traits that were noted to differ significantly. More specifically, the black and coloured South Africans demonstrated the most similarities and subsequent group overlap, while the white South Africans typically demonstrated greater differences among at least four traits (see Table 9 for trait frequencies). Overall, the double superior articular facets and bifid

spinous processes of the cervical vertebrae were noted more frequently in white South Africans, while both black and coloured South Africans had single facets and non-bifid spinous processes. Coloured South Africans were more likely to possess a vastus notch on the patella, and a translucent septal aperture on the humerus than the other groups.

The frequencies of the traits were also compared between males and females, with the population groups pooled together (Table 9). Only four traits were noted to be statistically significantly different, namely spinous process bifurcation (C3 to C6), suprascapular foramen, septal aperture and vastus notch. All the above-mentioned traits were also significantly different when comparing among the populations. It should be acknowledged that while only the right septal aperture and vastus notch was significant for ancestry, both the left and right sides for both traits were significant when assessing sex. This indicates substantial differences between the left and right sides. Overall, females were more likely to present with a spinous process bifurcation and a vastus notch on the patella compared to males.

TABLE 8. Trait frequencies among population groups (black, coloured, and white) and between the sexes (males and females). Refer to Table 6 for trait names and abbreviations.

Score	Ancestry						Sex			
	Black		Coloured		White		Females		Males	
	n	%	n	%	n	%	n	%	n	%
ATF (C1)										
0	56	74.67	71	78.02	67	84.8	92	76.67	102	81.60
1	14	18.67	14	15.39	8	10.13	21	17.50	15	12.50
2	5	6.67	6	6.59	4	5.06	7	5.83	8	6.67
ATF (C3)										
0	85	100.00	90	100.00	80	100.00	123	99.19	132	100.00
1	0	0	0	0	0	0	1	0.81	0	0
2	0	0	0	0	0	0	0	0	0	0
ATF (C4)										
0	80	97.56	87	95.60	77	87.50	116	92.80	128	98.46
1	2	2.44	4	4.40	9	10.23	8	6.40	1	0.77
2	0	0	0	0	2	2.27	1	0.80	1	0.77
ATF (C5)										
0	69	83.13	67	72.04	52	60.47	85	67.46	103	75.74
1	12	14.46	21	22.58	25	29.07	32	25.40	26	19.12
2	2	2.41	5	5.38	9	10.47	9	7.14	7	5.15
ATF (C6)										
0	53	63.86	52	59.77	37	44.05	61	51.26	81	60.00
1	18	21.69	25	28.74	24	28.57	35	29.41	32	23.70
2	12	14.46	10	11.49	23	27.38	23	19.33	22	16.30
ATF (C7)										
0	72	90.00	71	92.21	69	81.18	103	89.57	110	85.94
1	7	8.75	6	7.79	12	14.12	10	8.70	15	11.72
2	1	1.25	0	0	4	4.71	2	1.74	3	2.34
PB										
0	62	82.67	74	80.43	68	86.08	103	85.12	101	80.80
1	8	10.67	13	14.13	6	7.60	13	10.74	14	11.20
2	5	6.67	5	5.44	5	6.33	5	4.13	10	8.00

DSAF

0	65	86.67	78	84.78	48	60.76	100	82.64	91	72.80
1	7	9.33	8	8.70	19	24.05	10	8.26	24	19.20
2	3	4.00	6	6.52	12	15.19	11	9.09	10	8.00

SPB (C3)

0	65	80.25	62	73.81	23	29.87	82	67.77	68	56.20
1	11	13.58	9	10.71	18	23.38	18	14.88	20	16.53
2	5	6.17	13	15.48	36	46.75	21	17.36	33	27.28

SPB (C4)

0	72	61.54	60	72.29	16	20.78	82	67.77	50	41.67
1	15	12.82	7	8.43	7	9.09	18	14.88	20	16.67
2	30	25.64	16	19.28	54	70.13	21	17.36	50	41.67

SPB (C5)

0	33	41.25	50	56.82	13	15.66	57	47.11	39	30.00
1	16	20.00	10	11.36	9	10.84	12	9.92	23	17.69
2	31	38.75	28	31.82	61	73.49	52	42.98	68	52.31

SPB (C6)

0	49	61.25	65	71.43	32	38.10	78	69.64	68	50.37
1	8	10.00	12	13.19	8	9.52	13	11.61	15	11.11
2	23	28.75	14	15.39	44	52.38	21	18.75	52	38.52

SSF

0	85	100.00	86	92.47	79	87.78	115	89.84	135	96.43
1	0	0.00	7	7.53	11	12.22	13	10.16	5	3.57

STA

0	65	90.28	72	90.00	63	96.92	95	94.06	105	90.52
1	7	9.72	8	10.00	2	3.08	6	5.94	11	9.48

SCP (L)

0	87	100.00	93	100.0	88	100.00	128	100.00	140	100.00
1	0	0	0	0	0	0	0	0	0	0

SCP (R)

0	87	100.00	93	100.00	90	100.00	128	100.0	142	100.00
1	0	0	0	0	0	0	0	0	0	0

SA (L)

0	27	31.03	8	8.60	26	30.59	18	14.06	43	31.39
1	30	34.48	64	68.82	40	47.06	63	49.22	71	51.83
2	11	12.64	12	12.90	6	7.06	19	14.84	10	7.30
3	19	21.84	9	9.68	13	15.29	28	21.88	13	9.49
					SA (R)					
0	27	31.77	11	11.83	30	34.48	22	17.46	46	33.09
1	29	34.12	62	66.67	45	51.72	63	50.00	73	52.52
2	10	11.77	8	8.60	6	6.90	14	11.11	10	7.19
3	19	22.35	12	12.90	6	6.90	27	21.43	10	7.19
					TT (L)					
0	83	95.40	81	89.01	74	87.06	115	90.56	123	90.44
1	4	4.60	10	10.99	11	12.94	12	9.45	13	9.56
					TT (R)					
0	83	96.51	79	88.76	82	93.18	119	94.44	125	91.24
1	3	3.49	10	11.24	6	6.82	7	5.56	12	8.76
					VN (L)					
0	55	66.27	56	60.87	65	77.38	92	74.19	84	62.22
1	28	33.73	36	39.13	19	22.62	32	25.81	51	37.78
					VN (R)					
0	59	71.08	46	51.69	61	75.31	89	72.36	77	59.23
1	24	28.92	43	48.32	20	24.69	34	27.64	53	40.77
					AMCF (L)					
0	1	1.28	2	2.44	2	2.70	3	2.75	2	1.60
1	57	73.08	60	73.17	38	51.35	75	68.81	80	64.00
2	20	25.64	20	24.39	27	36.49	27	24.77	40	32.00
3	0	0.00	0	0.00	7	9.46	4	3.67	3	2.400
					AMCF (R)					
0	0	0.00	2	2.44	3	3.90	3	2.70	2	1.61
1	59	77.63	64	78.05	42	54.54	83	74.78	82	66.13
2	17	22.37	16	19.51	26	33.77	21	18.92	38	30.64
3	0	0.00	0	0.00	6	7.79	4	3.60	2	1.61

TABLE 9. Kruskal-Wallis test results assessing significant differences in trait frequency among the population groups and the sexes. Refer to Table 3 for trait names and abbreviations.

Trait	Probability value (p-value)	
	Ancestry	Sex
ATF (C1)	0.31	0.39
ATF (C3)	0.34	0.30
ATF (C4)	0.02*	0.67
ATF (C5)	<0.01*	0.14
ATF (C6)	<0.01*	0.19
ATF (C7)	0.07	0.41
PB	0.67	0.33
DSAF	<0.01*	0.10
SPB (C3)	<0.01*	<0.05*
SPB (C4)	<0.01*	<0.01*
SPB (C5)	<0.01*	0.03*
SPB (C6)	<0.01*	0.01*
SSF	<0.01*	0.03*
STA	0.23	0.33
SCP (L)	N/A	N/A
SCP (R)	N/A	N/A
SA (Lt)	0.18	<0.01*
SA (R)	<0.01*	<0.01*
TT (L)	0.15	0.96
TT (R)	0.14	0.32
VN (L)	0.06	0.04*
VN (R)	<0.01*	0.03*
AMCF (L)	<0.01*	0.34
AMCF (R)	<0.01*	0.11

* Indicates significant differences ($p<0.05$).

TABLE 10. Breakdown of group overlap based on the Kruskal-Wallis and Dunn's test results comparing the population groups. Refer to Table 2 for trait names and abbreviations.

No groups overlap	All groups overlap	Black and coloured overlap	Black and white overlap	White and coloured overlap
-	ATF (C1)	ATF (C4)	SA (R)	ATF (C4)
	ATF (C3)	ATF (C5)	VN (R)	ATF (C5)
	ATF (C7)	ATF (C6)		SSF
	PB	DSAF		
	STA	SPB (C3)		
	SA (L)	SPB (C4)		
	TT (L)	SPB (C5)		
	TT (R)	SPB (C6)		
	VN (L)	SSF		
		SA (R)		
		AMCF (R)		
		AMCF (R)		

DISCUSSION & CONCLUSION

This is the first study to assess the usefulness postcranial MMS traits using the methodology proposed by Spiros (11) on modern South African populations. The study aimed to test the repeatability with which the traits can be scored, and to explore the variation of the traits among South Africans.

Overall, the repeatability (particularly intra-observer repeatability) of the postcranial MMS traits was better than reported for the cranial MMS approach when applied to the same population (10). This is largely assumed to be the result of the scoring system itself. More specifically, the recordation scale of the postcranial traits are dichotomous (either present or absent) or related to the bilaterality of the trait. Cranial MMS traits are mostly ordinal and quantify quasi-continuous variation where traits can be classified as either small, intermediate, or large, or with minor shape variations (1). Ordinal traits may be more difficult to score because it potentially introduces more subjectivity to the scoring process; one observer may view a trait as small, whereas another observer may view the same trait expression as

intermediate (17). The amount of overlap between the trait expressions and among the population groups may also further complicate scoring in the case of ordinal variables. Therefore, the postcranial MMS traits seem easier to score reliably than cranial MMS traits.

Despite the intra-observer agreement of the traits being quite high, some of the traits were noted to be much less repeatable when the scores were compared between multiple observers (e.g., accessory transverse foramen for C1 and C7 and the spinous process bifurcation for C3). The reason for the decreased agreement may be due to a lack of experience with the postcranial MMS method or less experience with osteological variations, particularly in the case of traits that are as rare as some of the ones included in this study. Similar research assessing morphoscopic variation has also emphasized the role that experience can play in the scoring process (14,18,19). For the current study, the primary observer underwent a training period to become familiar with the traits. The second observer received a packet with the definitions and the line drawings taken from the Spiros (11) study. Both observers were MSc students from the same cohort, with identical training and similar experience levels with regard to osteological variation and the analysis of skeletal remains. Thus, the major difference between the observers was experience with the traits, and the results indicate that familiarity with the traits affects the repeatability of the method. It is recommended that students, researchers and practicing forensic anthropologists develop the necessary experience with the postcranial MMS methodology before employing it in studies or skeletal analyses.

Spiros (11) demonstrated almost perfect inter-observer agreement for nine out of the eleven traits, with the remaining traits demonstrating substantial agreement. In comparison, the current study yielded inter-observer agreement levels ranging from poor to almost perfect. The difference in inter-observer agreement levels between the two studies can most likely be explained by the fact that Spiros (11) was involved in the development of the method and possess a better understanding of the trait definitions and more subtle expressions, especially in the case of trait variations. For example, the accessory transverse foramen is one of the traits that demonstrated the most variation in terms of inter-observer repeatability, resulting in a poor agreement score. The poor agreement for the accessory transverse foramen is most likely the result of observed variations, such as an incomplete bridge (see Figure 2). The trait definitions do not indicate how to approach such variations, so researchers may resolve scoring the variant morphologies in different ways (18). Incomplete bridges were also observed with the posterior bridging trait.

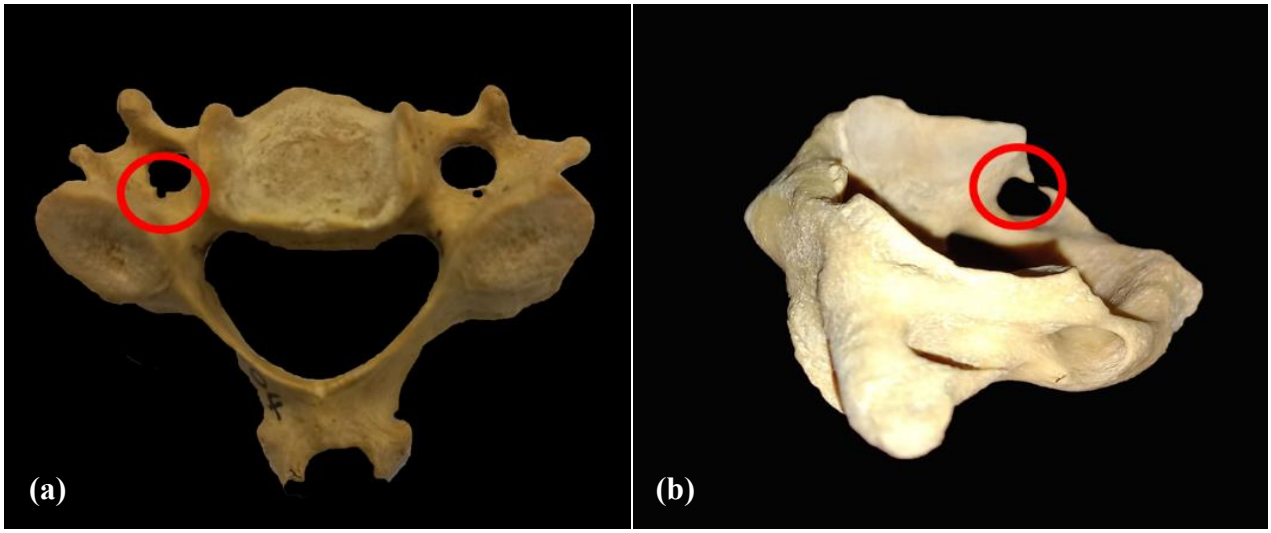

FIGURE 2. Incomplete bridge (encircled) on the (a) accessory transverse foramen (superior view) and (b) posterior bridging (lateral view) traits on the vertebrae (Photo: NP Bothma).

Trait variation needs to be considered as many variations were observed with other traits, specifically with the suprascapular foramen (see Figure 3). For example, a few individuals had both a scapular notch and a suprascapular foramen inferior to the coracoid process. Spiros (11) mentioned that a suprascapular foramen inferior to the coracoid process should be marked as absent, as the trait is rare (20). However, for an inexperienced observer, the trait can be confused with the traditional suprascapular foramen. Furthermore, different degrees of expression for the scapular notch as discussed by Hrdlička (21) were observed with some displaying an incomplete bridge over the suprascapular foramen. The variation observed on the scapula should be considered for future research to investigate whether the variation of the trait can affect the frequency distribution of the trait.

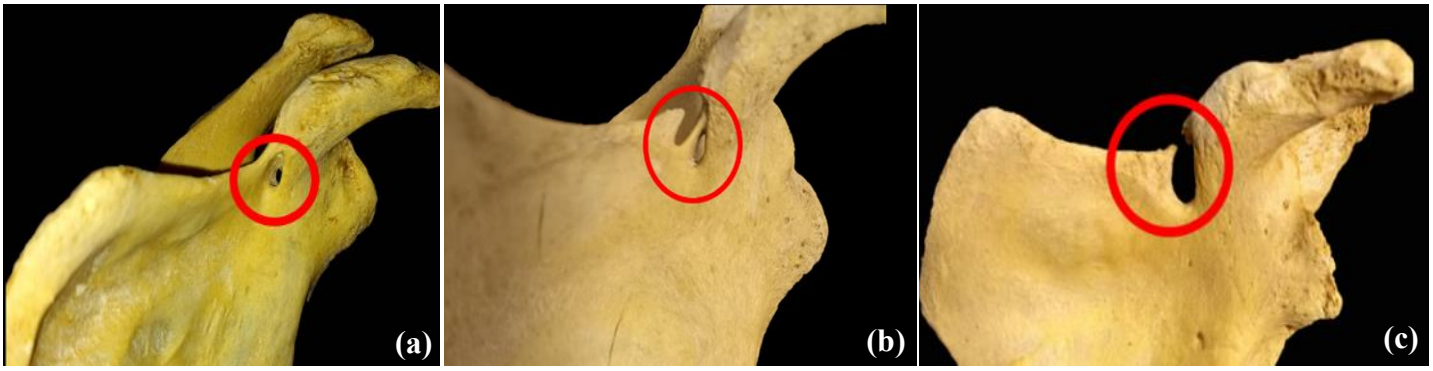

FIGURE 3. Variation observed on the scapula or scapular foramen. (a) Scapular foramen below the coracoid process; (b) suprascapular notch with scapular foramen; (c) incomplete bridge over suprascapular foramen (Photo: NP Bothma).

Spiros (11) also cautioned about enthesophytes being confused as the third trochanter and distinguishing between the two was a challenge with some specimens in the current study (see Figure 4 for images comparing an enthesophyte and a third trochanter). Other indications of pathology, such as myositis ossificans traumatica, were also features to be cognizant of when scoring, specifically with the supracondylar process of the humerus (22,23). The supracondylar process was not observed in any of the specimens, but a bony spur consistent with myositis ossificans traumatica was present on one of the specimens and mimicked the trait (see Figure 5 for the feature consistent with myositis ossificans). One of the deterrents from scoring the anomaly as a present supracondylar process trait was that the feature was located on the lateral side of the humerus and was therefore ruled out. Again, familiarity with the trait definitions and locations is essential to score traits accurately. Furthermore, osteological knowledge and an understanding of normal morphology will contribute to greater accuracy in recognizing the traits and any variations of the traits.

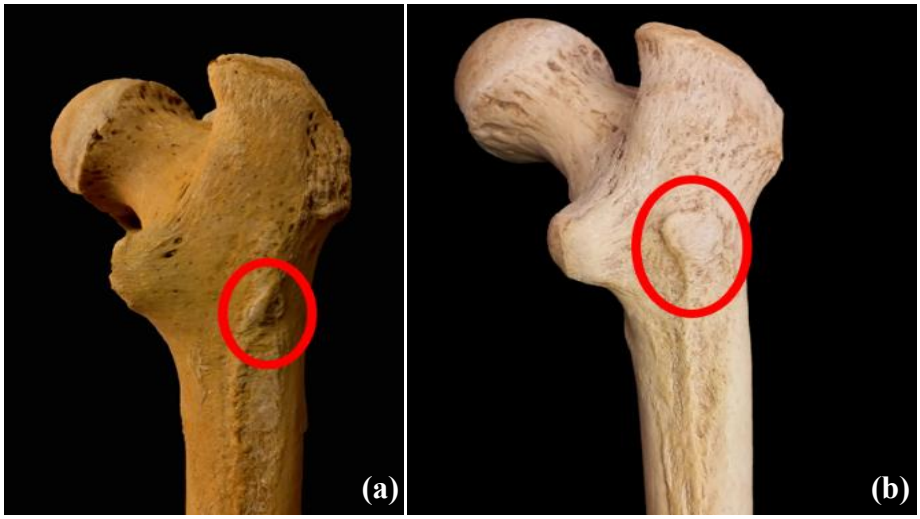

FIGURE 4. Comparison of (a) enthesophyte and (b) third trochanter (posterior view) (Photo: NP Bothma).

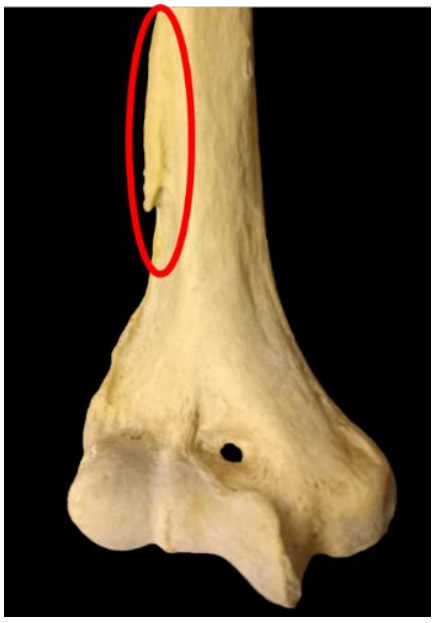

FIGURE 5. A feature consistent with myositis ossificans traumatica on the lateral aspect of the humerus (Photos: NP Bothma)

The trait frequencies observed in the current study demonstrate similar patterns of group overlap among the South African populations as previously noted with macromorphoscopic, craniometric and post-craniometric data (6,7,42) . More specifically, black and coloured South Africans displayed the most overlap, while white South Africans were more dissimilar. This is similar to the patterns of variation observed with osteometric studies in South Africa and has largely been attributed to socio-political circumstances and positive assortative mating. (2,3,24). Historically, mixed marriages between coloured and black South Africans occurred

more frequently in comparison to coloured-white or black-white mixed marriages (25,26). Legislature against mixed marriages such as the Prohibition Act of Mixed Marriages Act No. 55 of 1949, were not enforced by the apartheid government as strictly on mixed marriages between black and coloured South Africans as it was between white South Africans and other populations.

The postcranial MMS method has only been evaluated in a few studies (13,14). The current study reported seven traits that showed significant differences, while Spiros (11) found four traits (spinous process bifurcation for C3 and C4, third trochanter, and the anterior and medial calcaneal facet) with significant differences when assessing black and white North Americans. The spinous process bifurcation was the only trait between the North American and South African samples that demonstrated a common significant difference outcome. Overall, the South African sample yielded more traits with significant differences. Interestingly, Spiros (2019) did not find significant differences between the sexes; however, the current study noted significant differences among four of the traits. Thus, the effects of sex on the traits should be further explored. The difference in frequency distribution and statistical difference between the South African and the North American samples supports the notion that the traits are believed to have a genetic component that can be modified by epigenetic factors resulting from the environment and internal physiology (27,28). The two samples are from different geological environments with different population histories, which can possibly explain the variation of the trait expressions. Furthermore, with the traits showing similar patterns of variation among South Africans compared to osteometric data, which has been shown to reflect genetic relationships and heritability, postcranial MMS traits may be useful in attempting to classify ancestry (29). Spiros and Hefner (12) produced promising results in their assessment of ancestry using combined cranial and postcranial macromorphoscopic models (with accuracies between 89.5% to 92.1%). Further research should be conducted to explore the application of these traits in classification models to estimate ancestry among modern South Africans. The implications of sex and asymmetry on the expression of the traits should also be evaluated.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the EU Erasmus+ grant [597924-EPP-1-2018] *Bakeng se Afrika* for providing financial support for this project.

References

1. Hefner JT. Cranial Nonmetric Variation and Estimating Ancestry*. J Forensic Sci. 2009;54(5):985–95.

2. L'Abbé EN, Kenyhercz M, Stull KE, Keough N, Nawrocki S. Application of Fordisc 3.0 to Explore Differences Among Crania of North American and South African Blacks and Whites. J Forensic Sci. 2013;58(6):1579–83.

3. Liebenberg L, L'Abbé EN, Stull KE. Population differences in the postcrania of modern South Africans and the implications for ancestry estimation. Forensic Sci Int. 2015 Dec 1; 257:522–9.

4. Liebenberg L, Krüger GC, L'Abbé EN, Stull KE. Postcraniometric sex and ancestry estimation in South Africa: a validation study. Int J Legal Med. 2019 Jan 1;133(1):289–96.

5. Ousley SD, Jantz RL. Fordisc 3 and Statistical Methods for Estimating Sex and Ancestry. In: Dirkmaat DC, editor. A Companion to Forensic Anthropology. West Sussex: Wiley-Blackwell; 2012. p. 311–27.

6. Hooton EA. Methods of Racial Analysis. Science. 1926;63(1621):75–81.

7. Rhine S. Non-metric skull racing. In: Skeletal attribution of race: Methods for forensic anthropology. Maxwell Museum of Anthropology; 1990. p. 9–20.

8. Krüger GC, Liebenberg L, Myburgh J, Meyer A, Oettlé AC, Botha D, et al. Chapter 27 - Forensic Anthropology and the Biological Profile in South Africa. In: Latham KE, Bartelink EJ, Finnegan M, editors. New Perspectives in Forensic Human Skeletal Identification. Academic Press; 2018. p. 313–21.

9. Krüger GC, L'Abbé EN, Stull KE, Kenyhercz MW. Sexual dimorphism in cranial morphology among modern South Africans. Int J Legal Med. 2015 Jul 1;129(4):869–75.

10. L'Abbé EN, Van Rooyen C, Nawrocki, SP, Becker PJ. An evaluation of non-metric cranial traits used to estimate ancestry in a South African sample. Forensic Sci Int. 2011; 209:195. e1-195.e7.

11. Spiros MC. Standardization of Postcranial Nonmetric Traits and Their Utility in Ancestry Analysis. Forensic Anthropol. 2019 Jan 16;2(1):29-44-29–44.

12. Spiros MC, Hefner JT. Ancestry Estimation Using Cranial and Postcranial Macromorphoscopic Traits. J Forensic Sci. 2020 May;65(3):921–9.

13. Hefner JT, Ousley SD. Statistical Classification Methods for Estimating Ancestry Using Morphoscopic Traits. J Forensic Sci. 2014;59(4):883–90.

14. Klales AR, Kenyhercz MW. Morphological Assessment of Ancestry using Cranial Macromorphoscopics. J Forensic Sci. 2015;60(1):13–20.

15. National Health Act. National Health Act 61 of 2003 | South African Government Gazette [Internet]. Sect. 62 2003 p. 66. Available from: https://www.gov.za/documents/national-health-act

16. Landis JR, Koch GG. The Measurement of Observer Agreement for Categorical Data. Biometrics. 1977;33(1):159–74.

17. Maier C. Evaluating Mixed-Methods Models for the Estimation of Ancestry from Skeletal Remains. Forensic Anthropol. 2019 Winter;2(1):45–56.

18. Wilczak CA, Mariotti V, Pany-Kucera D, Villotte S, Henderson CY. Training and interobserver reliability in qualitative scoring of skeletal samples. J Archaeol Sci Rep. 2017 Feb 1; 11:69–79.

19. Klales A. Current State of Sex Estimation in Forensic Anthropology. Forensic Anthropol. 2021 Apr 6;4(2):118–27.

20. Voisin JL, Ropars M, Thomazeau H. Anatomical evidence for a uniquely positioned suprascapular foramen. Surg Radiol Anat. 2016 May 1;38(4):489–92.

21. Hrdlička A. The scapula: Visual observations. Am J Phys Anthropol. 1942;29(1):73–94.

22. Thorndike AJ. Myositis Ossificans Traumatica. JBJS. 1940 Apr;22(2):315–23.

23. Beiner JM, Jokl P. Muscle Contusion Injury and Myositis Ossificans Traumatica. Clin Orthop Relat Res 1976-2007. 2002 Oct;403: S110.

24. Stull KE, Kenyhercz MW, Tise ML, L'Abbé EN, Tuamsuk P. Chapter 13 - The Craniometric Implications of a Complex Population History in South Africa. In: Pilloud MA, Hefner JT, editors. Biological Distance Analysis. San Diego: Academic Press; 2016. p. 245–63.

25. Sofer C. Some Aspects of Inter-Racial Marriages in South Africa, 1925–461. Africa. 1949 Jul;19(3):187–203.

26. Hyslop J. White Working-class Women and the Invention of Apartheid: 'Purified' Afrikaner Nationalist Agitation for Legislation Against 'Mixed' Marriages, 1934–9. J Afr Hist. 1995 Mar;36(1):57–81.

27. Saunders SR. Nonmetric skeletal variation. Reconstr Life Skelet. 1989;95–108.

28. Mann RW, Hunt DR. Non-metric traits and anatomical variants that can mimic trauma in the human skeleton. Forensic Sci Int. 2019 Aug 1; 301:202–24.

29. Ousley S, Jantz R, Freid D. Understanding race and human variation: Why forensic anthropologists are good at identifying race. Am J Phys Anthropol. 2009;139(1):68–76.

CHAPTER 5: PREDICTIVE PERFORMANCE OF TRAITS

5.1 Introduction

The aim of the previous chapter was to evaluate the frequency distribution of postcranial MMS traits among black, white and coloured South Africans. The results demonstrated that some traits differed significantly between at least two populations. There was also substantial observed group overlap, particularly between black and coloured South Africans. Finally, the repeatability of the postcranial MMS method was noted to be sufficiently repeatable.

The previous paper was proof of concept that there are significant differences, which led to the assumption that the traits could be useful for ancestry. This paper aimed to test if this was the case, and the predictive performance of a series of RFMs was evaluated to see if the traits should be incorporated in future forensic case analyses and standard operating procedures.

5.2 Manuscript to be submitted

Estimating ancestry using random forest models and postcranial macromorphoscopic traits

N.P. BOTHMA, E.N. L'ABBÉ, L. LIEBENBERG

Manuscript to be submitted for publication to Forensic Science International.

ABSTRACT

Population overlap and the variation within and between populations has been globally observed but is difficult to quantify. To achieve this, methods need to be explored and validated to assist with the creation of an accurate biological profile. The macromorphoscopic approach is fairly popular among many forensic practitioners, and recent publications have assessed population variation on a series of postcranial macromorphoscopic traits. With the postcranial traits, significant differences have been noted among modern South Africans. However, little information is available on the positive predictive performance of these traits, particularly in classifying heterogeneous populations. The current study aimed to assess the postcranial macromorphoscopic traits as a tool to estimate ancestry in South Africa.

A sample of 87 black, 90 white and 94 coloured South Africans from the Pretoria Bone Collection and the Kirsten Skeletal Collection were assessed. Univariate and multivariate random forest models were created to test the positive predictive performance of the traits to classify ancestry. The classification accuracies for the univariate models ranged from 33.3% to 53.0%. The classification accuracies for the multivariate models when using random forest model ranged from 54.6% to 62.1%. With closer inspection, 77.5% of white South Africans in the sample were correctly classified, while black and coloured South Africans in the sample were correctly classified 57.6% and 54.9% of the time, respectively. Many of the traits were fairly rare in the sample and did not contribute much discriminatory information. Based on the variable importance, the traits assessing spinous process bifurcation were the most discriminatory variables.

The results of the current study indicate that the postcranial MMS approach does not outperform current methods employed to estimate ancestry. Furthermore, the low Kappa values obtained with the RFMs suggest that the traits are not reliable classifiers when used on their own, and as such the method does not currently have practical applicability for medicolegal casework in South Africa.

KEYWORDS

Forensic anthropology; Population affinity; Classification accuracy; Variable importance; Machine learning

HIGHLIGHTS

- This study evaluated the performance of postcranial morphological variants to estimate ancestry
- Univariate models performed poorly with accuracies between 33.3% and 53.0%
- Multivariate models presented fair accuracies ranging between 54.6% and 62.1%
- The model that only included the significantly different variables performed the best
- On its own, the postcranial MMS method was not an accurate indicator of ancestry in South Africa

INTRODUCTION

Ancestry is one of the fundamental parameters of the biological profile, and many methods that are employed to estimate the other parameters are population-specific and require prior knowledge of ancestry (1–3). Ancestry refers to the geographical origin and population history of a particular individual and how the combination of both these factors influences skeletal morphology (4–6). Essentially, ancestry estimation is the classification of an individual into the population to which the individual had mostly likely belonged. In the forensic anthropological context, different combinations of skeletal traits and osteometric dimensions have been correlated to populations across the globe to estimate ancestry (5,7–12). As such, it is possible to obtain a fairly accurate ancestry estimate through the application of robust statistical analyses and methods using reference samples of known individuals.

Ancestry estimation methods are particularly important in countries with diverse populations such as South Africa. South Africa has a heterogeneous population that consists of socially identified black, white, coloured, and Indian or Asian individuals, refer to Krüger et al. (6) for more information regarding the population origins and history. Heterogeneous populations display skeletal variation among the populations; however, there is also substantial group overlap (5,10,12). Because of its diversity, developing various methods of ancestry estimation for the South African population is imperative.

The development of various methods enables forensic anthropologists to estimate ancestry by employing metric and non-metric methods using several different skeletal elements (7,10,12,13). With standard metric methods, observers evaluate the size of skeletal elements using measuring tools, such as calipers and osteometric boards. In comparison, non-metric methods quantify the size and shape of skeletal elements through the visual evaluation of

morphological skeletal variants (13,14). In the past, non-metric methods have popularly been used to assess and researchers could determine differences between populations in biodistance studies (7,15–18).

To minimize subjectivity when using cranial non-metric traits for ancestry estimation, Hefner (7) introduced line drawings, improved definitions, and robust statistics to better assess the traits (7). An ordinal scale was also introduced to score the traits according to their degree of expression, commonly known as the macromorphoscopic (MMS) method. Currently, the cranial MMS traits are the most developed non-metric method for ancestry estimation (9,11). However, in South Africa, research on non-metric traits for ancestry estimation has been minimal. Further research needs to be conducted for the method to be useful as a tool for the classification of ancestry, specifically postcranial MMS as they are the least developed in comparison to cranial MMS (19).

To date, postcranial MMS traits have been examined to investigate if they can be used for ancestry analyses only in North America (20,21). Spiros (20) created trait illustrations and definitions using over 100 skeletons and numerous skeletal elements (20). As a result, eleven postcranial MMS traits were standardized. The classification accuracy of the traits was further tested on black and white North Americans using both cranial and postcranial MMS traits in the Spiros and Hefner study (21). Classification models were created using a variety of machine-learning techniques, including random forest models (RFM), support vector machines (SVM) and artificial neural networks (aNN), to measure the performance of the postcranial MMS traits in ancestry estimation. Correct classification accuracies ranging between 77.6% and 81.6% were obtained. Additionally, a combination of cranial and postcranial traits yielded correct classifications between 89.5% and 92.1%, with the spinous process bifurcation being the best discriminator between black and white North Americans.

A few studies have been conducted to assess the use of postcranial MMS traits in a forensic context, and more research needs to be carried out for non-metric analyses for ancestry estimation purposes to become suitable methods for use with South African remains. However, with any scientific method, the validity and predictive performance of a method needs to be tested sufficiently (22). Therefore, the current study is the first to investigate the accuracy with which the postcranial MMS method can predict ancestry in a modern South African sample. To achieve this, random forest modelling (RFM) has been selected as the classification method.

As a classification method, RFM refers to a combination of decision trees that are generated using a non-parametric algorithm integrating random sampling with replacement and majority voting (23). Through a series of nodes or rules, RFM predicts a categorical variable (such as ancestry) from a set of measurements or observations on one or more predictor variables (such as postcranial scores) (24–26). Training data sets are used to create classification models, after which a hold-out (or "out-of-bag") testing set is used to simultaneously evaluate the models using an independent sample. Essentially, a training data set is the known or collected data that is used to fit a classification model, and a testing data set is unseen data that evaluates if the training model works adequately for classification.

In general, single decision trees may be inclined to overfitting of data, meaning that the model captures the data errors of the training data, but does not generalize well and misses important patterns of variation that are necessary for correct classification (26). The advantage of RFM over decision trees is that the algorithm can manage the issue of overfitting. In addition, RFM presents the variable importance, which indicates which variables contribute to the classification of the dataset (26). While more complex machine learning methods, such as aNN or SVM, have been employed in anthropological research and have been shown to outperform RFM in some instances (9,21,27), their inherent complexity poses a problem. Specifically, these methods may be computationally expensive and the results difficult to interpret. The inclusion of robust statistical analyses is essential to anthropological methodology, but many practitioners are reluctant to employ overly complicated computational methods (28). As such, it is important to find a good balance between robust statistics and user-friendly methods with feasible interpretability. While Spiros and Hefner (21) used the aNN algorithm to assess ancestry using both cranial and postcranial MMS traits, RFM was selected for this study given its success in other papers assessing ancestry (21,26). For example, Hefner et al. (25), assessed the use of RFMs for ancestry estimation using both cranial macromorphoscopic and metric data from a North American sample, and reported an accuracy of 89.6%. Similarly, Klales and Kenyhercz (11) achieved 73.3% when classifying black and white North Americans using cranial MMS traits. Finally, Navega et al. (26) also employed the RFM to classify population groups of African and European origins and obtained a correct classification of 93.8%.

Only a few studies have assessed the postcranial MMS method for ancestry estimation. Among South Africans, Bothma and colleagues (refer to Chapter 4 of this volume) identified significant population differences, which suggests that the traits may be useful for ancestry

estimation. This paper aimed to estimate ancestry using RFM and postcranial MMS traits in a modern South African population.

METHOD AND MATERIALS

For the current study, 11 postcranial MMS traits were visually assessed and scored in a sample of 271 black, white and coloured South Africans (see Table 12 for sample distribution). The sample was taken from the Pretoria Bone Collection and the Kirsten Skeletal collection housed at the University of Pretoria and Stellenbosch University, respectively. Ethical clearance was granted from the Human Research Ethics Committee (HREC) of the Faculty of Health Sciences at the University of Pretoria (Ref 610/2021). The postcranial MMS traits were scored by applying the methodology described by Spiros (20) and Spiros and Hefner (21). Table 13 provides a summary of the traits and their descriptions. The 11 traits are located on the following skeletal elements: the cervical vertebrae, sternum, scapula, humerus, femur, patella, and calcaneus. In the current study, cervical vertebrae C5 and C6 were also included for the assessment of the spinous process bifurcation, as significant differences have previously been reported. In the case of bilateral traits, only the left-side scores were used in the classification models.

TABLE 11. The sample distribution.

Population	Black	White	Coloured
Males	46	49	47
Females	41	41	47
Total	87	90	94

TABLE 12. Summary of trait descriptions and scoring system (Trait descriptions and scoring taken from Spiros (18)).

Trait and location	Abbreviation	Score	Description	Definition
Accessory Transverse Foramen (cervical vertebra)	ATF (C1, C3, C4, C6, C7)	0 1 2	Absent Unilateral Bilateral	One to two extra foramina located near the articular facet of the cervical vertebrae
Posterior Bridging (cervical vertebra)	PB	0 1 2	Absent Unilateral Bilateral	Observed as a bony protrusion, extending from the posterior aspect of the superior articular facet of C1 to the arch located posteriorly.
Double Superior Articular Facets (cervical vertebra)	DSAF	0 1 2	Absent Unilateral Bilateral	Forms when a groove forms in the middle of the superior articular facet of C1, resulting in double facets
Spinous Process Bifurcation (cervical vertebra)	SPB (C3, C4, C5, C6)	0 1 2	Non-bifid Partially bifid Completely bifid	A "split" of the most posterior part of the spinous process of the cervical vertebra to form two bony tubercles or projections (29,30).
Suprascapular Foramen (scapula)	SF	0 1	Absent Present	Formed when the suprascapular notch is enclosed, forming a foramen. The suprascapular notch is located on the superior border of the scapular, at the root of the coracoid process (31).

Sternal Aperture (sternum)	STA	0 1	Absent Present	An oval opening or hole located on the sternum that can vary in size (32).
Supra-condyloid Process (humerus)	SCP	0 1	Absent Present	A bony projection located above the medial epicondyle of the humerus and is an extension of the supracondylar ridge.
Septal Aperture (humerus)	SA	0 1 2 3	Absent Translucent Small perforation Large perforation	An opening located on the distal part of the humerus, that joins the olecranon fossa to the coronoid fossa resulting in an oval or round shape hole (32).
Third Trochanter (femur)	TT	0 1	Absent Present	A tubercle located at the superior end of the gluteal ridge on the posterosuperior aspect of the femur and slightly inferior to the lesser trochanter (32).
Vastus Notch (patella)	VN	0 1	Absent Present	A concaved or flattened indentation located on the superolateral or medial angle of the patella (32).
Anterior and Middle Calcaneal Facets (calcaneus)	AMCF	0 1 2 3	No anterior facet Single elongated facet Small anterior facet Large anterior facet	Located on the calcaneus can vary from a single small oval facet to an elongated oval facet, where the middle facet is joined with an anterior or double facet where the two facets are separated.

Random forest modelling (RFM) was employed to create classification models and to see if the traits are useful for ancestry estimation. The sample was divided so that 75% constituted the training set (to create the models), and the remaining 25% was kept as the out-of-bag (OOB) testing set (to validate the models). In the case of missing data, the mode was calculated for each sex and population group separately within each trait and the result was used as the score of that missing trait. The mode was used as an imputation value specifically because a it appears the most in a set of values which in this case, in a population or sex group, most individuals are likely to depict that value. Both univariate and multivariate models were employed to evaluate the performance of the traits when tested both individually and in a group. The univariate models assist to determine the performance of each trait and the multivariate models determine the performance of the traits when tested in combination. Three different multivariate models were tested: the first model included all the traits; for the second model all traits with variable importance below one (as calculated from the first model) were removed; and finally, the third model included only the variables that were previously found to be significantly different (refer to Chapter 4 of this volume). Table 14 presents the traits that showed significant differences between at least two of the three population groups.

A total of 2500 classification trees were used for each model with four variables at each split. The classification accuracy (for the training and testing samples), Kappa values, and variable importance were recorded for each model. Both the classification and Kappa values are measures of model accuracy. The classification accuracy presents the percentage of correctly classified individuals out of all of the individuals; whereas, the Kappa value presents the percentage of correctly classified individuals while taking random chance into account. The Kappa value is a particularly useful metric in the case of unbalanced classes (e.g., where traits can be scored as zero or one, but a score of 1 is a fairly rare occurrence). With variable importance, the higher the value, the more a variable contributes to the classification.

RESULTS

Univariate models

Table 14 presents the classification results for the univariate models. For the training sample, the classification accuracy ranged from 3.9% to 60.0%. When the univariate models were validated on the testing sample, the classification accuracy ranged from 33.3% to 53.0%. Overall, the spinous process bifurcation for C4 had the highest classification accuracy for both training and testing samples. The accessory transverse foramen for C1 had the lowest classification accuracy of 33.3% for the testing sample. The Kappa values for the testing sample ranged from 0.0% to 29.0% (see Table 14 for classification accuracies and Kappa values).

TABLE 13. Positive predictive performance of each trait using univariate RFM to estimate ancestry. Refer to Table 2 for trait abbreviations.

Variable	Training sample accuracy (%)	Testing sample accuracy (%)	Kappa value (%)
ATF_C1	35.8	33.3	0.0
ATF_C3	33.3	34.9	0.0
ATF_C4	36.3	36.4	2.0
ATF_C5*	28.4	42.4	15.0
ATF_C6*	36.3	34.9	1.0
ATF_C7	35.9	42.4	12.0
PB	34.1	34.9	2.0
DSAF*	42.5	34.9	4.0
SPB_C3*	26.5	43.9	17.0
SPB_C4*	60.0	53.0	29.0
SPB_C5*	53.8	50.0	25.0
SPB_C6*	46.3	50.0	24.0
SSF*	3.9	36.4	7.0

STA	13.7	36.4	5.0
SA	35.6	42.4	13.0
TT	6.6	36.4	3.0
VN	39.1	39.4	9.0
AMCF*	39.1	36.4	3.0

*Traits that showed significant differences among the ancestry groups**.**

Multivariate models

There were three multivariate models that were tested for classification. The first model included all the traits; the second model included only the traits that had high variable importance (VarImp values >0) and the third model included only the traits that showed significant differences among the ancestry groups. When the first multivariate model was analysed, the classification accuracy was 63.5% for the training sample and 54.6% for the testing sample with a Kappa value of 32.0% (see Table 18 for the comparison of all three multivariate models). Black South Africans presented with a classification error rate of 42.4%, with 30.3% misclassifying as coloured South Africans (see Table 4 for a classification matrix). White South Africans had the lowest classification error of 22.1%. Most black and coloured South Africans misclassified as one another. White South Africans misclassified equally as both black and coloured with no specific trend. Figure 6 presents the variable importance for the first multivariate model. The trait with the highest variable importance was the spinous process bifurcation of C4. Two traits – supracondylar process and accessory transverse foramen of C3 – had a very low variable importance of 0.0 (i.e., do not contribute any information to the model).

TABLE 14: Confusion matrix showing patterns of overlap and misclassification among the groups for the training model for the first model (all the traits).

		Classifies into:			
		Black	**Coloured**	**White**	**Classification error**
Group:	**Black**	38	20	8	42.4 %
	Coloured	23	39	9	45.1%
	White	8	7	53	22.1%

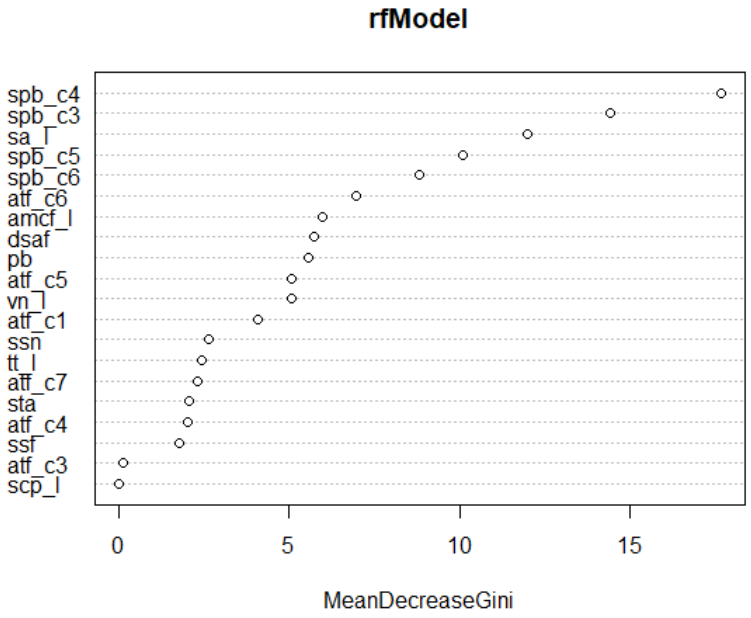

FIGURE 6. Comparison of variable importance for the multivariate model employing all the traits.

For the second multivariate model, traits with a variable importance of zero were removed, these include the accessory transverse foramen of C3 and the supracondylar process. The removal of the variables led to a 0.5% decrease for the training accuracy, while both the testing accuracy and Kappa value increase with 2.0% (see Table 18 for a comparison of variable importance for all three multivariate models). When assessing the confusion matrix (see Table 15), black South Africans had a classification error rate of 42.4%, coloured South Africans had a classification error rate of 46.5% and white South Africans had a classification error rate of 22.1%. Similar patterns of misclassifications were observed for both the first and second

models. The spinous process bifurcation of C4 was still considered the trait with the highest variable importance and the suprascapular foramen had the least variable importance (see Figure 7).

TABLE 15: Confusion matrix showing patterns of overlap and misclassification among the groups for the training model for the traits that demonstrated high variable importance.

		Classifies into:			
		Black	**Coloured**	**White**	**Classification error**
Group:	**Black**	38	20	8	42.4 %
	Coloured	23	38	10	46.5%
	White	7	8	53	22.1%

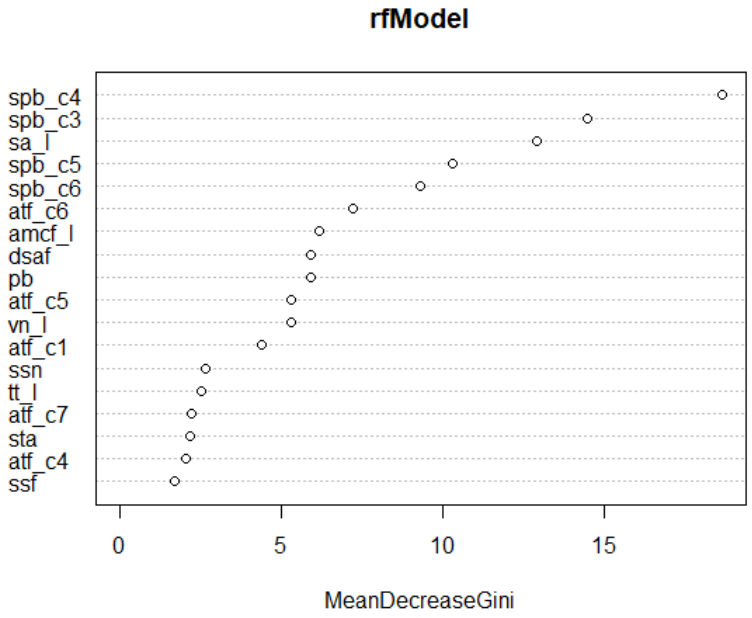

FIGURE 7. Comparison of variable importance for the multivariate model employing traits with a variable importance greater than 0 as indicated by the model including all variables.

For the third multivariate model, only variables that were noted to be significantly different by Bothma et al. (refer to Chapter 4), were selected. This includes the accessory transverse foramen of C3 to C6, spinous process bifurcation of C4 to C6, double superior articular facet,

suprascapular foramen and the anterior and medial articular facets. The training accuracy for the third model was 2.5% lower than the first model and 2.0% lower than the second model. However, there was a marked increase in the testing accuracy and the Kappa value. Once again, similar patterns of misclassification were observed. Similar to the second model, the spinous process bifurcation of C4 had the highest variable importance and the suprascapular foramen had the lowest (see Figure 8).

TABLE 16: Confusion matrix showing patterns of overlap and misclassification among the groups for the training model for the traits that demonstrated significant differences.

		Classifies into:			
		Black	**Coloured**	**White**	**Classification error**
Group:	**Black**	24	25	8	48.5 %
	Coloured	22	39	10	46.0%
	White	9	7	52	23.5%

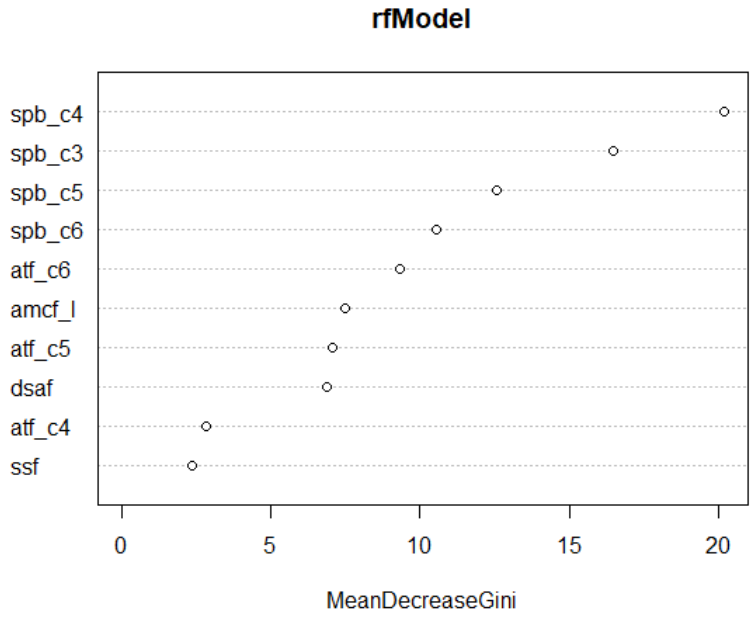

FIGURE 8. Comparison of variable importance for the multivariate model employing traits with significant differences as indicated by the Kruskal-Wallis test.

TABLE 17. A comparison of the training, testing and Kappa values for the three multivariate models.

	Training accuracy (%)	Testing accuracy (%)	Kappa value (%)
All Traits model	63.5	54.6	32.0
Variable importance model	63.0	56.1	34.0
Significant difference model	61.0	62.1	43.4

DISCUSSION AND CONCLUSION

Continuous research is necessary to improve classification accuracies specifically for ancestry in countries such as South Africa which is rich in population diversity. The current study was the first to evaluate the postcranial MMS method to assess ancestry estimation on a modern South African sample. The current study differed from the Spiros (18) study on which it was modelled by the fact that a tripartite sample was used, where with the original study only two groups were compared. Additionally, two cervical vertebrae (C5 and C6) were added for the analysis of the accessory transverse foramen.

The Spiros (18) study only assessed frequency distribution of the traits between black and white North Americans and did not assess predictive performance of the traits. This is likely because limited significant differences were observed in their sample. In a follow-up study, Spiros and Hefner (21) combined cranial and postcranial MMS traits and attempted to estimate ancestry using a variety of statistical methods. When employing RFM, the authors reported an 88.0% correct classification for the testing sample. While it should be acknowledged that their results are not directly comparable to the current study because of the combination of crania and postcrania, these are the only published error rates that included the postcranial MMS traits. The current study used both univariate and multivariate models, where the highest accuracy obtained was for the multivariate model that employed only variables that have been shown to differ among black, white and coloured South Africans. However, the classification accuracy was fairly low with both the testing and training accuracies in the lower 60%. It has been

recommended that classification methods yield accuracies at least 50% better than chance (14). This was not the case with the postcranial MMS traits. This is likely why Spiros and Hefner only presents combined crania and postcrania results rather than presenting postcrania results on their own. Compared to methods currently employed in South African medicolegal casework, the postcranial MMS traits did not perform well. For example, the current craniometric standards yield accuracies of 73.0% and postcraniometric standards yield accuracies of 85.0% when using discriminant analysis on the same population (10,13). Thus, the postcranial MMS traits do not outperform existing methods and should not be added to standard operating procedures simply because it is a novel method. The combination of cranial and postcranial traits does yield higher accuracies (21), thereby suggesting that the postcranial MMS method may have some potential for ancestry estimation, although further research is required.

Similar to the Spiros and Hefner (21) study, the current study observed the spinous process bifurcation trait to be the most discriminatory. However, the Kappa values for the spinous process bifurcation (C3 to C6) in the current study were much lower than the classification accuracies, which may make the trait inadequate for ancestry estimation specifically on its own. Overall, the multivariate models performed better than the univariate models. This is not unexpected and has previously been shown in other studies (10). However, univariate analyses are also necessary to determine how each variable works on its own and if that particular variable can be used in instances where there is limited skeletal material to assess such as missing skeletal elements and fragmented remains. For the current method, the multivariate approach is recommended as the univariate models did not perform well because the classification accuracies were low. A multivariate approach is best in forensic case analysis because there is reduced bias and it encompasses more variation as multiple variables are assessed because skeletal traits are not unique to just one group (14,33–36). The analysis of just one variable or trait in a univariate approach may exclude the variation within a population as the traits cannot be limited to one population group. In addition, traits are also dependent on their frequency distribution within a population. Assessing multiple traits makes up for a trait that may not be present on skeletal remains being assessed in a forensic case analysis. There is an overlap and different groups share similarities and certain traits as previously observed in postcranial metric analyses (10). Therefore, a combination of traits that have different patterns of overlap and dissimilarities is necessary to be able to distinguish between population groups. As such, even when limited skeletal material is available for analysis, it is recommended that

as many traits from as many different methods as possible be used for classification; the use of univariate postcranial MMS traits should be a last resort. The other advantage of the multivariate model particularly for this study is that we could observe which traits had the least variable importance and therefore deduce which of the traits contribute to the classification model. In addition, when the traits that did not contribute to the classification model were removed, the classification accuracy for the testing sample increased slightly. However, whether variables of the least importance are in the model or not, the RFM is still able to select the ones contributing to the classification model. This means that the classification accuracies will not necessarily be affected by variables of lower importance. This supports the recommendation provided by Navega et al. (26) that states that the user can include all variables and allow the model to select the variables rather than modifying models or removing variables manually based on variable importance (26). With that being said, the model that included only the traits that demonstrated significant differences performed better in terms of classification. Therefore, this observation should be considered when creating standards for postcranial MMS traits.

The postcranial MMS approach satisfies the *Daubert* criteria in terms of reporting error rates and classification accuracies (22). However, the accuracies are too low for the method to be used on its own and as it stands may only be used to confirm results obtained with another method. Future research should explore a combination of both cranial and postcranial MMS traits to estimate ancestry in South Africa.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the EU Erasmus+ grant [597924-EPP-1-2018]

Bakeng se Afrika for providing financial support for this project.

REFERENCES

1. Feldesman M, Lundy J. Stature estimates for some African Plio-Pleistocene hominids. J Hum Evol. 1988 Sep 1; 17:583–96.

2. Oettlé AC, Steyn M. Age estimation from sternal ends of ribs by phase analysis in South African Blacks. J Forensic Sci. 2000 Sep;45(5):1071–9.

3. Krüger GC, L'Abbé EN, Stull KE, Kenyhercz MW. Sexual dimorphism in cranial morphology among modern South Africans. Int J Legal Med. 2015 Jul 1;129(4):869–75.

4. Dunn RR, Spiros MC, Kamnikar KR, Plemons AM, Hefner JT. Ancestry estimation in forensic anthropology: A review. WIREs Forensic Sci. 2020;2(4): e1369.

5. Stull KE, Kenyhercz MW, L'Abbé EN. Ancestry estimation in South Africa using craniometrics and geometric morphometrics. Forensic Sci Int. 2014 Dec 1; 245:206. e1-206.e7.

6. Krüger GC, Liebenberg L, Myburgh J, Meyer A, Oettlé AC, Botha D, et al. Chapter 27 - Forensic Anthropology and the Biological Profile in South Africa. In: Latham KE, Bartelink EJ, Finnegan M, editors. New Perspectives in Forensic Human Skeletal Identification. Academic Press; 2018. p. 313–21.

7. Hefner JT. Cranial Nonmetric Variation and Estimating Ancestry*. J Forensic Sci. 2009;54(5):985–95.

8. Hefner J, Ousley S. Morphoscopic Traits and the Statistical Determination of Ancestry. In: Proceedings of the 57th annual meeting of the American Academy of Forensic Sciences. Colorado Springs; 2005. p. 291–2.

9. Hefner JT, Ousley SD. Statistical Classification Methods for Estimating Ancestry Using Morphoscopic Traits. J Forensic Sci. 2014;59(4):883–90.

10. Liebenberg L, L'Abbé EN, Stull KE. Population differences in the postcrania of modern South Africans and the implications for ancestry estimation. Forensic Sci Int. 2015 Dec 1; 257:522–9.

11. Klales AR, Kenyhercz MW. Morphological Assessment of Ancestry using Cranial Macromorphoscopics. J Forensic Sci. 2015;60(1):13–20.

12. Liebenberg L, Krüger GC, L'Abbé EN, Stull KE. Postcraniometric sex and ancestry estimation in South Africa: a validation study. Int J Legal Med. 2019 Jan 1;133(1):289–96.

13. L'Abbé EN, Kenyhercz M, Stull KE, Keough N, Nawrocki S. Application of Fordisc 3.0 to Explore Differences Among Crania of North American and South African Blacks and Whites. J Forensic Sci. 2013;58(6):1579–83.

14. Ousley SD, Jantz RL. Fordisc 3 and Statistical Methods for Estimating Sex and Ancestry. In: Dirkmaat DC, editor. A Companion to Forensic Anthropology. West Sussex: Wiley-Blackwell; 2012. p. 311–27.

15. Hooton EA. Methods of Racial Analysis. Science. 1926;63(1621):75–81.

16. Corruccini RS. An examination of the meaning of cranial discrete traits for human skeletal biological studies. Am J Phys Anthropol. 1974;40(3):425–45.

17. Ossenberg NS. Within and between race distances in population studies based on discrete traits of the human skull. Am J Phys Anthropol. 1976;45(3):701–15.

18. Rhine S. Non-metric skull racing. In: Skeletal attribution of race: Methods for forensic anthropology. Maxwell Museum of Anthropology; 1990. p. 9–20.

19. L'Abbé EN, Van Rooyen C, Nawrocki, SP, Becker PJ. An evaluation of non-metric cranial traits used to estimate ancestry in a South African sample. Forensic Sci Int. 2011; 209:195. e1-195.e7.

20. Spiros MC. Standardization of Postcranial Nonmetric Traits and Their Utility in Ancestry Analysis. Forensic Anthropol. 2019 Jan 16;2(1):29-44-29–44.

21. Spiros MC, Hefner JT. Ancestry Estimation Using Cranial and Postcranial Macromorphoscopic Traits. J Forensic Sci. 2020 May;65(3):921–9.

22. Grivas C, Komar D. Kumho, Daubert, and the Nature of Scientific Inquiry: Implications for Forensic Anthropology. J Forensic Sci. 2008 Aug 1; 53:771–6.

23. Breiman L. Random Forests. Mach Learn. 2001 Oct 1;45(1):5–32.

24. Biau G. Analysis of a Random Forests Model. 2012;33.

25. Hefner JT, Spradley MK, Anderson B. Ancestry Assessment Using Random Forest Modeling. J Forensic Sci. 2014;59(3):583–9.

26. Navega D, Coelho C, Vicente R, Ferreira MT, Wasterlain S, Cunha E. AncesTrees: ancestry estimation with randomized decision trees. Int J Legal Med. 2015 Sep;129(5):1145–53.

27. Nikita E, Nikitas P. On the use of machine learning algorithms in forensic anthropology. Leg Med. 2020 Nov 1; 47:101771.

28. Steadman DW. Who Needs Data? I've Got Experience! Hum Biol. 2018;90(1):77–82.

29. Asvat R. The Configuration of Cervical Spinous Processes in Black and White South African Skeletal Samples. J Forensic Sci. 2012;57(1):176–81.

30. Cunningham DJ. Neural Spines of the Cervical Vertebræ as a Race-Character. J Anat Physiol. 1886 Jul;20(Pt 4):637–40.

31. White TD, Black MT, Folkens PA. Human Osteology. Third Edition. Oxford: Elsevier; 2012.

32. Mann RW, Hunt DR. Non-metric traits and anatomical variants that can mimic trauma in the human skeleton. Forensic Sci Int. 2019 Aug 1; 301:202–24.

33. Maier HR, Dandy GC. Determining Inputs for Neural Network Models of Multivariate Time Series. Comput-Aided Civ Infrastruct Eng. 1997;12(5):353–68.

34. Frelat MA, Mittereocker P. Postnatal ontogeny of tibia and femur form in two human populations: a multivariate morphometric analysis. Am J Hum Biol Off J Hum Biol Counc. 2011 Dec;23(6):796–804.

36. Alladio E, Poggiali B, Cosenza G, Pilli E. Multivariate statistical approach and machine learning for the evaluation of biogeographical ancestry inference in the forensic field. Sci Rep. 2022 May 28;12(1):8974.

CHAPTER 6: DISCUSSION AND CONCLUSION

Ancestry is one of the most difficult parameters of the biological profile to establish mainly because of the association between skeletal morphology that forensic anthropologists need to assess to help identify a decedent and link the individual to their socially defined race. Like any other scientific method, ancestry estimation methods need to satisfy the *Daubert* criteria, which includes documented error rates, method repeatability, and creating population-specific standards (Grivas and Komar, 2008). The postcranial MMS method has only been assessed in a few publications to investigate the prevalence of the traits between populations (Spiros, 2019) and their predictive ability in the context of ancestry estimation (Spiros and Hefner, 2020). This is the first study to assess postcranial MMS traits for ancestry estimation analysis using a South African sample. As such, the results in the current study contribute to our knowledge of ancestry estimation and how we approach the parameter in forensic case analyses. The current study aimed to determine the frequency distribution of the traits based on the Spiros (2019) method and also to investigate whether the postcranial MMS method can be utilized to assess ancestry in the South African population by employing the RFM algorithm.

Evaluation of the postcranial MMS method indicated that observer experience is a factor to consider and that familiarity with the method plays a role in accurately scoring the traits. For example, the overall inter-observer agreement in the current study was lower (substantial agreement following the Landis and Koch descriptions) compared to the Spiros (2019) study which demonstrated substantial to almost perfect agreement. The method was standardized by Spiros (2019) and it is assumed that their greater experience and familiarity with the traits led to the improved repeatability. Another notable observation in terms of the observer agreement was that the spinous process bifurcation had the lowest agreement value (κ=0.58) but the highest variable importance value (17.66), and testing accuracy (53.0%). Thus, the trait will likely always be selected in classification models when it is available, and it needs to be scored accurately. Caution needs to be taken when scoring the trait as one easily can confuse a partially bifid trait with a trait that is completely bifid as both features have a certain degree of bifurcation (see Figure 9).

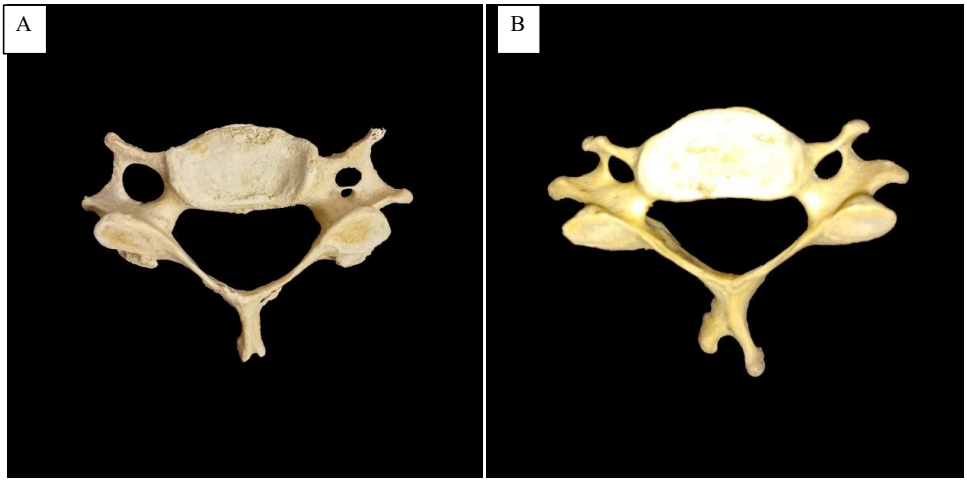

FIGURE 9. Spinous process bifurcation. A. Partially bifid, B. Completely bifid trait. (Photos by NP Bothma)

The significant differences observed in the frequency distribution for seven of the traits indicate the potential for developing new methods to attempt ancestry estimation. Overall, seven traits showed significant differences (accessory transverse foramen of C6 and C7, double superior articular foramen, spinous process bifurcation of C3 to C6, suprascapular foramen, septal aperture, vastus notch and the anterior and medial calcaneal facets); however, only the spinous process bifurcation were correspondingly noted be significant in the North American sample (Spiros, 2019). This shows the population-specificity of the traits and the importance of testing methods between and among populations prior to employing methods in skeletal analyses. Significant differences were also observed between the sexes in the current study, whereas Spiros (2019) did not observe any differences in the traits between the sexes and opted to pool the sexes for further analyses. This observation further demonstrates the uniqueness of populations due to different geographical origins and the need to validate methods for different populations (Saunders, 1989; Mann and Hunt, 2019). With that said, the implication of sex on the postcranial MMS method was beyond the scope of the current study and should be further explored to better understand the variation attributable to sexual dimorphism.

Bilateral traits also need to be explored further as some in the current study exhibited asymmetry where one side of the trait yielded a significant difference such as the septal aperture and the vastus notch traits. In their methodology Spiros (2019) already recommended that the septal aperture be assessed by side because of the significant differences of both the right and left side in their study. However, there was no mention of doing the same for the vastus notch as they did not note any significant differences were observed for the sides of the trait. More

research should be conducted that investigates asymmetry and how it affects bilateral traits when applying the postcranial MMS method. A few studies have looked at asymmetry on the skeleton and how it can influence classification (Call, 2016; Cole et al., 2020). In the current study, for example, the left side of the bilateral traits was utilized for creating classification models. Since the left vastus notch did not yield significant differences, it was excluded from the models. But if the right side of the vastus notch (which was significantly different for ancestry) was included in the classification models, the accuracies may increase, especially the multivariate model consisting of only significant traits. More specifically, the vastus notch, while rare, was observed to be more prevalent among coloured South Africans and may assist in distinguishing between black and coloured South Africans, which are the groups that misclassify more frequently.

The lack of research on the postcranial MMS traits calls for the method to be assessed a bit more specifically for ancestry estimation. The current study tested the classification accuracy of the method using the RFM algorithm to evaluate both univariate and multivariate models. As expected, the multivariate models performed better than the univariate models in terms of classification accuracy (Spradley and Jantz, 2011; Liebenberg et al., 2015; Tabachnick et al., 2019). However, the accuracies for both univariate and multivariate models obtained in the current study were not high enough to establish the postcranial MMS method as a good classifier of ancestry. This indicates that the postcranial MMS method is not useful on its own to classify populations. In addition, this emphasizes the importance of the predictive performance of the traits and how different they perform in each population because relying on p-values alone is not enough. The p-values tell us which variables are significantly different; however, a significant difference does not necessarily mean practical applicability (Solla et al., 2018). For instance, in the current study, some traits yielded significant differences but had low training and testing classification accuracies which mean that they do not have practical applicability for ancestry estimation as significant differences do not always translate to skeletally quantifiable differences between the populations. The methods that are currently in place for ancestry estimation in South Africa outperform the postcranial MMS method, and as such there is not enough justification to include the postcranial traits in the standard operating procedures for skeletal analyses.

In terms of classification, the RFM employed in the current study was observed to be a good method as the algorithm is robust and easily accessible while still being computationally inexpensive. The RFM algorithm gave outputs that are comparable to linear discriminant

analysis (as employed in Fordisc). More specifically, clear training and testing accuracies are presented, which is similar to the classification accuracy and leave-one-out cross validation accuracy which is reported with Fordisc analyses. Furthermore, the number of individuals that were misclassified into different populations could be determined through the confusion matrix outcome. Finally, the RFM algorithm presents the variable importance that helped to determine which traits contributed to the classification of populations which is similar to the stepwise variable selection function in Fordisc (Ousley and Jantz, 2012, 2013). This demonstrates that RFM is easy to interpret and is a good algorithm for ancestry estimation and will likely be embraced by the anthropological community. However, RFM needs to be explored further specifically when combining different datasets. For example, Spiros and Hefner (2020) combined both cranial and postcranial MMS and the classification accuracies when using RFM were high (90.0% for training and 88.0% for testing samples). The increased accuracies when both cranial and postcranial MMS traits were used for classification simultaneously indicate the advantage of using a multifactorial approach (Spiros and Hefner, 2020). Similarly, Hefner et al. (2014) demonstrated how combining cranial MMS traits and craniometrics yielded high classification accuracies (85.5% correct classification). Datasets that capture different types of variation such as size versus shape are expected to better quantify the skeletal variation attributed to population differences. Therefore, a holistic combined approach of different datasets from different parts of the skeleton should be explored.

REFERENCES

Alblas, A., Greyling, L.M., Geldenhuys, E.-M., 2018. Composition of the Kirsten Skeletal Collection at Stellenbosch University. South Afr. J. Sci. 114, 1–6.

Asvat, R., 2012. The Configuration of Cervical Spinous Processes in Black and White South African Skeletal Samples. J. Forensic Sci. 57, 176–181.

Berry, A.C., Berry, R.J., 1967. Epigenetic variation in the human cranium. J. Anat. 101, 361–379.

Bethard, J.D., DiGangi, E.A., 2020. Letter to the Editor—Moving Beyond a Lost Cause: Forensic Anthropology and Ancestry Estimates in the United States. J. Forensic Sci. 65, 1791–1792.

Blumenbach, J.F., 1865. On the natural variety of mankind, third ed. 1795, in: Bendyshe, T. (Ed.), The Anthropological Treatises of Johann Friedrich Blumenbach. Longman, Green, Longman, Roberts, and Green, London, Great Britain, pp. 145–276.

Brace, C.L., 1964. On the Race Concept. Curr. Anthropol. 5, 313–320.

Breiman, L., 2001. Random Forests. Mach. Learn. 45, 5–32.

Buikstra, J.E., Frankenberg, S.R., Konigsberg, L.W., 1990. Skeletal biological distance studies in American Physical Anthropology: Recent trends. Am. J. Phys. Anthropol.

Byrt, T., 1996. How Good Is That Agreement?: Epidemiology 7, 561.

Call, S.J., 2016. Sex estimation in forensic anthropology: A test of the Klales et al. (2012) method with implications of asymmetry (MSc Dissertation). University of Alaska Fairbanks, United States -- Alaska.

Cartmill, M., 1998. The Status of the Race Concept in Physical Anthropology. Am. Anthropol.

Cole, S.J., Hulse, C.N., Stull, K.E., 2020. Chapter 19 - The effects of skeletal asymmetry on accurate sex classification, in: Klales, A.R. (Ed.), Sex Estimation of the Human Skeleton. Academic Press, pp. 307–325.

Coon, C.S., 1962. The Origin of Races, 1st edition. ed. Alfed A. Knopf, Inc, Place of publication not identified.

Corruccini, R.S., 1974. An examination of the meaning of cranial discrete traits for human skeletal biological studies. Am. J. Phys. Anthropol. 40, 425–445.

Cunningham, D.J., 1886. Neural Spines of the Cervical Vertebræ as a Race-Character. J. Anat. Physiol. 20, 637–640.

Donlon, D.A., 2000. The value of infracranial nonmetric variation in studies of modern Homo sapiens: An Australian focus. Am. J. Phys. Anthropol. 113, 349–368.

Du Toit, C., 2014. Postcranial discrete traits and biological distance between two South African groups (BSc Hons dissertation). University of Pretoria.

Feldesman, M., Lundy, J., 1988. Stature estimates for some African Plio-Pleistocene hominids. J. Hum. Evol. 17, 583–596.

Finnegan, M., 1978. Non-metric variation of the infracranial skeleton. J. Anat. 125, 23–37.

Finnegan, M., McGuire, S.A., 1979. Classification systems for discrete variables used in forensic anthropology. Am. J. Phys. Anthropol. 51, 547–553.

Goodman, A.H., Moses, Y.T., Jones, J.L., 2019. Race: Are We So Different? John Wiley & Sons.

Grivas, C., Komar, D., 2008. Kumho, Daubert, and the Nature of Scientific Inquiry: Implications for Forensic Anthropology. J. Forensic Sci. 53, 771–776.

Haas, J., Buikstra, J.E., Ubelaker, D.H., Aftandilian, D., Field Museum of Natural History (Eds.), 1994. Standards for data collection from human skeletal remains: proceedings of a seminar at the Field Museum of Natural History, organized by Jonathan Haas. Arkansas Archeological Survey, Fayetteville, Ark.

Hauser, G., Stefano, G.F.D., 1989. Epigenetic Variants of the Human Skull, 0 edition. ed. Lubrecht & Cramer Ltd, Stuttgart.

Hefner, J., 2003. Assessing Nonmetric Cranial Traits Currently Used in the Forensic Determination of Ancestry.

Hefner, J., Ousley, S., 2005. Morphoscopic Traits and the Statistical Determination of Ancestry, in: Proceedings of the 57th Annual Meeting of the American Academy of Forensic Sciences. Presented at the American Academy of Forensic Sciences, Colorado Springs, pp. 291–292.

Hefner, J.T., 2009. Cranial Nonmetric Variation and Estimating Ancestry*. J. Forensic Sci. 54, 985–995.

Hefner, J.T., Ousley, S.D., 2014. Statistical Classification Methods for Estimating Ancestry Using Morphoscopic Traits. J. Forensic Sci. 59, 883–890.

Hefner, J.T., Ousley, S.D., 2012. Morphoscopic Traits and Assessment of Ancestry, in: Dirkmaat, D.C. (Ed.), A Companion to Forensic Anthropology. Wiley-Blackwell, West Sussex, pp. 287–309.

Hefner, J.T., Pilloud, M.A., Buikstra, J.E., Vogelsberg, C.C.M., 2016. Chapter 1 - A Brief History of Biological Distance Analysis, in: Pilloud, Marin A., Hefner, Joseph T. (Eds.), Biological Distance Analysis. Academic Press, San Diego, pp. 3–22.

Hefner, J.T., Spradley, M.K., Anderson, B., 2014. Ancestry Assessment Using Random Forest Modeling. J. Forensic Sci. 59, 583–589.

Hooton, E.A., 1926. Methods of Racial Analysis. Science 63, 75–81.

Hudson, N., 1996. From "Nation to “Race”: The Origin of Racial Classification in Eighteenth-Century Thought. Eighteenth-Century Stud. 29, 247–264.

Jackson, J.P.J., Weidman, N.M., 2005. The Origins of Scientific Racism. J. Blacks High. Educ. 15.

Klales, A.R., Kenyhercz, M.W., 2015. Morphological Assessment of Ancestry using Cranial Macromorphoscopics. J. Forensic Sci. 60, 13–20.

Krüger, G.C., L’Abbé, E.N., Stull, K.E., Kenyhercz, M.W., 2015. Sexual dimorphism in cranial morphology among modern South Africans. Int. J. Legal Med. 129, 869–875.

Krüger, G.C., Liebenberg, L., Myburgh, J., Meyer, A., Oettlé, A.C., Botha, D., Brits, D.M., Kenyhercz, M.W., Stull, K.E., Sutherland, C., L’Abbé, E.N., 2018. Chapter 27 - Forensic Anthropology and the Biological Profile in South Africa, in: Latham, K.E., Bartelink, E.J., Finnegan, M. (Eds.), New Perspectives in Forensic Human Skeletal Identification. Academic Press, pp. 313–321.

L’Abbé, E.N., Kenyhercz, M., Stull, K.E., Keough, N., Nawrocki, S., 2013. Application of Fordisc 3.0 to Explore Differences Among Crania of North American and South African Blacks and Whites. J. Forensic Sci. 58, 1579–1583.

L’Abbé, E.N., Krüger, G.C., Theye, C.E.G., Hagg, A.C., Sapo, O., 2021. The Pretoria Bone Collection: A 21st Century Skeletal Collection in South Africa. Forensic Sci. 1, 220–227.

L’Abbé, E.N., Loots, M., Meiring, J.H., 2005. The Pretoria Bone Collection: a modern South African skeletal sample. Homo Int. Z. Vgl. Forsch. Am Menschen 56, 197–205.

L’Abbé, E.N., Steyn, M., 2012. Forensic Anthropology in South Africa, in: Dirkmaat, D.C. (Ed.), A Companion to Forensic Anthropology. Wiley-Blackwell, West Sussex, pp. 626–638.

L’Abbé, E.N., Van Rooyen, C., Nawrocki, S.P., Becker, P.J., 2011. An evaluation of non-metric cranial traits used to estimate ancestry in a South African sample. Forensic Sci. Int. 209, 195.e1-195.e7.

Landis, J.R., Koch, G.G., 1977. The Measurement of Observer Agreement for Categorical Data. Biometrics 33, 159–174.

Larsen, C.S., Walker, P.L., 2010. Bioarchaeology: Health, Lifestyle, and Society in Recent Human Evolution, in: A Companion to Biological Anthropology. John Wiley & Sons, Ltd, pp. 379–394.

Liebenberg, L., Krüger, G.C., L'Abbé, E.N., Stull, K.E., 2019. Postcraniometric sex and ancestry estimation in South Africa: a validation study. Int. J. Legal Med. 133, 289–296.

Liebenberg, L., L'Abbé, E.N., Stull, K.E., 2015. Population differences in the postcrania of modern South Africans and the implications for ancestry estimation. Forensic Sci. Int. 257, 522–529.

Livingstone, F.B., 1962. On the Non-Existence of Human Races. Curr. Anthropol. 3, 279–281.

Loomba, A., Burton, J., 2007. François Bernier (1620–88), in: Loomba, A., Burton, J. (Eds.), Race in Early Modern England: A Documentary Companion. Palgrave Macmillan US, New York, pp. 272–274.

Loue, S., 2006. Defining Race, Ethnicity, and Related Constructs, in: Loue, S. (Ed.), Assessing Race, Ethnicity and Gender in Health. Springer US, Boston, MA, pp. 25–53.

Macalister, A., 1893. Notes on the Development and Variations of the Atlas. J. Anat. Physiol. 27, 519–542.

Mann, R.W., Hunt, D.R., 2019. Non-metric traits and anatomical variants that can mimic trauma in the human skeleton. Forensic Sci. Int. 301, 202–224.

Montagu, A., 1962. The Concept of Race. Am. Anthropol. 64, 919–928. https://www.jstor.org/stable/666943

National Health Act, 2003. National Health Act 61 of 2003 | South African Government Gazette.

Oettlé, A.C., Steyn, M., 2000. Age estimation from sternal ends of ribs by phase analysis in South African Blacks. J. Forensic Sci. 45, 1071–1079.

Ossenberg, N.S., 1976. Within and between race distances in population studies based on discrete traits of the human skull. Am. J. Phys. Anthropol. 45, 701–715.

Ousley, S., Jantz, R., 2013. Fordisc 3. Rechtsmedizin 23, 97–99.

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

Ousley, S.D., Jantz, R.L., 2012. Fordisc 3 and Statistical Methods for Estimating Sex and Ancestry, in: Dirkmaat, D.C. (Ed.), A Companion to Forensic Anthropology. Wiley-Blackwell, West Sussex, pp. 311–327.

Pfeiffer, S., Heinrich, J., Beresheim, A., Alblas, M., 2016. Cortical bone histomorphology of known-age skeletons from the Kirsten collection, Stellenbosch university, South Africa. Am. J. Phys. Anthropol. 160, 137–147.

Plemons, A., Hefner, J.T., 2016. Ancestry Estimation Using Macromorphoscopic Traits. Acad. Forensic Pathol. 6, 400–412.

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

Reed, J.C., 2006. The Utility of Cladistic Analysis of Nonmetric Skeletal Traits for Biodistance Analysis (PhD Dissertation). University of Pittsburgh.

Rhine, S., 1990. Non-metric skull racing, in: Skeletal Attribution of Race: Methods for Forensic Anthropology. Maxwell Museum of Anthropology, pp. 9–20.

Ross, A.H., Pilloud, M., 2021. The need to incorporate human variation and evolutionary theory in forensic anthropology: A call for reform. Am. J. Phys. Anthropol. 176, 672–683.

Russell, F., 1900. Studies in Cranial Variation. Am. Nat. 34, 737–745.

Saunders, S.R., 1989. Nonmetric skeletal variation. Reconstr. Life Skelet. 95–108.

Saunders, S.R., Rainey, D.L., 2008. Nonmetric Trait Variation in the Skeleton: Abnormalities, Anomalies, and Atavisms, in: Biological Anthropology of the Human Skeleton. John Wiley & Sons, Ltd, pp. 533–559.

Solla, F., Tran, A., Bertoncelli, D., Musoff, C., Bertoncelli, C.M., 2018. Why a P-Value is Not Enough. Clin. Spine Surg. 31, 385–388.

Spiros, M.C., 2019. Standardization of Postcranial Nonmetric Traits and Their Utility in Ancestry Analysis. Forensic Anthropol. 2, 29-44-29–44.

Spiros, M.C., Hefner, J.T., 2020. Ancestry Estimation Using Cranial and Postcranial Macromorphoscopic Traits. J. Forensic Sci. 65, 921–929.

Spradley, K., Jantz, R., 2021. What Are We Really Estimating in Forensic Anthropological Practice, Population Affinity or Ancestry? Forensic Anthropol. 4, 171–180.

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

Steyn, M., L'Abbé, E.N., Myburgh, J., 2016. Forensic Anthropology as Practiced in South Africa, in: Handbook of Forensic Anthropology and Archaeology. Routledge, pp. 193–206.

Stojanowski, C.M., 2018. Biodistance, in: The International Encyclopedia of Biological Anthropology. American Cancer Society, pp. 1–3.

Stull, K.E., Kenyhercz, M.W., L'Abbé, E.N., 2014. Ancestry estimation in South Africa using craniometrics and geometric morphometrics. Forensic Sci. Int. 245, 206.e1-206.e7.

Tabachnick, B.G., Fidell, L.S., Ullman, J.B., 2019. Using multivariate statistics, Seventh edition. ed. Pearson, NY, NY.

Takezawa, Y., 2011. Problems with the Terms: "Caucasoid", "Mongoloid" and "Negroid" 9.

Tishkoff, S.A., Kidd, K.K., 2004. Implications of biogeography of human populations for "race" and medicine. Nat. Genet. 36, S21–S27.

White, T.D., Black, M.T., Folkens, P.A., 2012. Human Osteology, Third Edition. ed. Elsevier, Oxford.

APPENDIX – TRAIT DEFINITIONS

Accessory Transverse Foramen (ATF)

The ATF (Figure A-1) are one to two extra foramina located near the articular facet of the cervical vertebrae. The ATF will be scored with a scale of zero to two, with zero being absent, one being present on either left or right (unilateral) and two being present on both left and right sides (bilateral) (Spiros, 2019). The atlas (C1) and cervical vertebrae three to seven (C3 – C7) will be assessed.

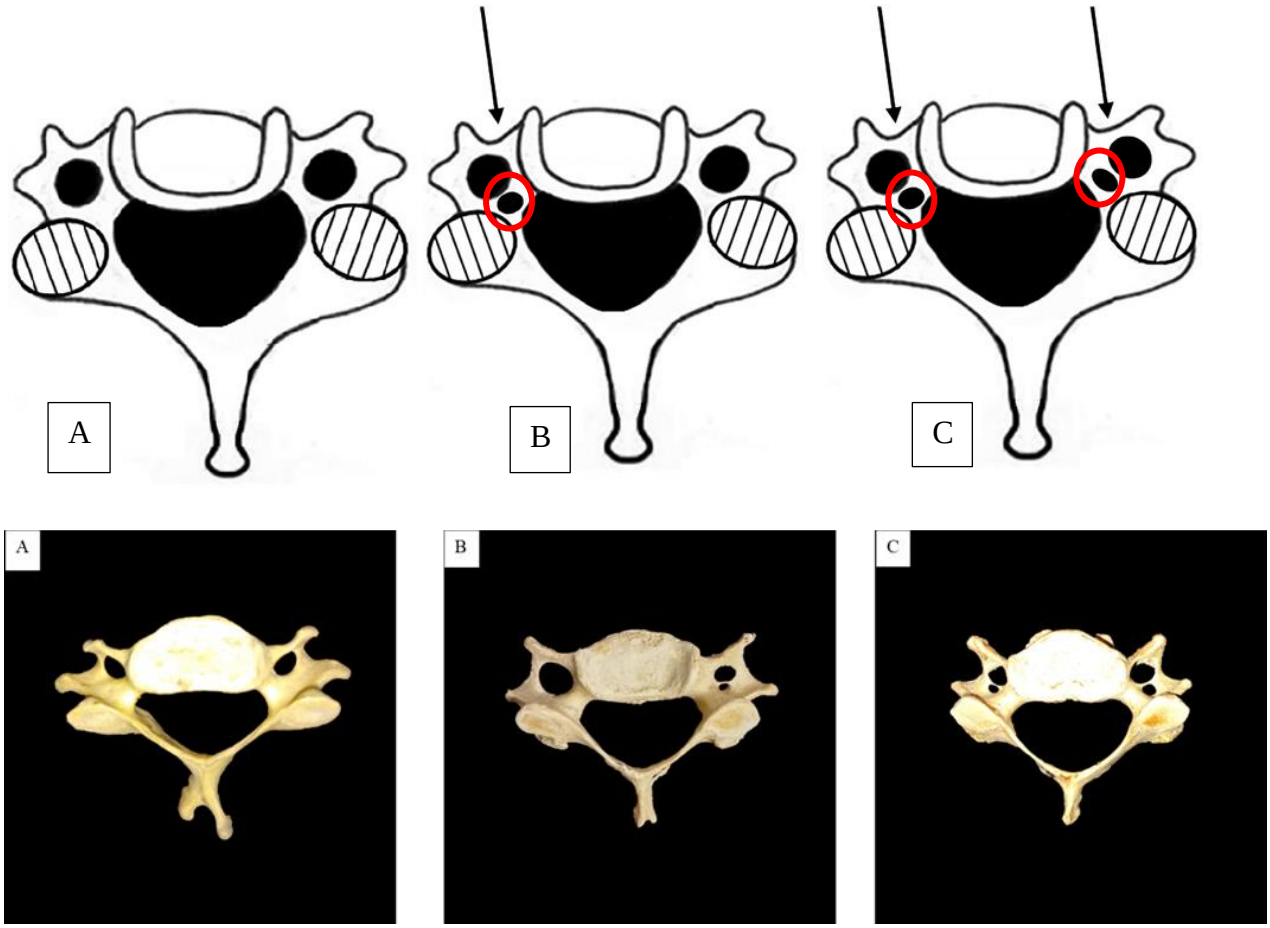

Figure A- 1. Accessory Transverse Foramen (red circle) A. Absent (score = 0), B. Unilateral (score = 1), C. Bilateral (score = 2) *(Spiros, 2019)* (Photos by NP Bothma).

Posterior Bridging (PB)

The PB (Figure A-2) is observed as a bony protrusion, extending from the posterior aspect of the superior articular facet of C1 to the arch located posteriorly. The trait will be scored with a scale of zero to two, with zero defined as absent, one being unilateral and two being bilateral.

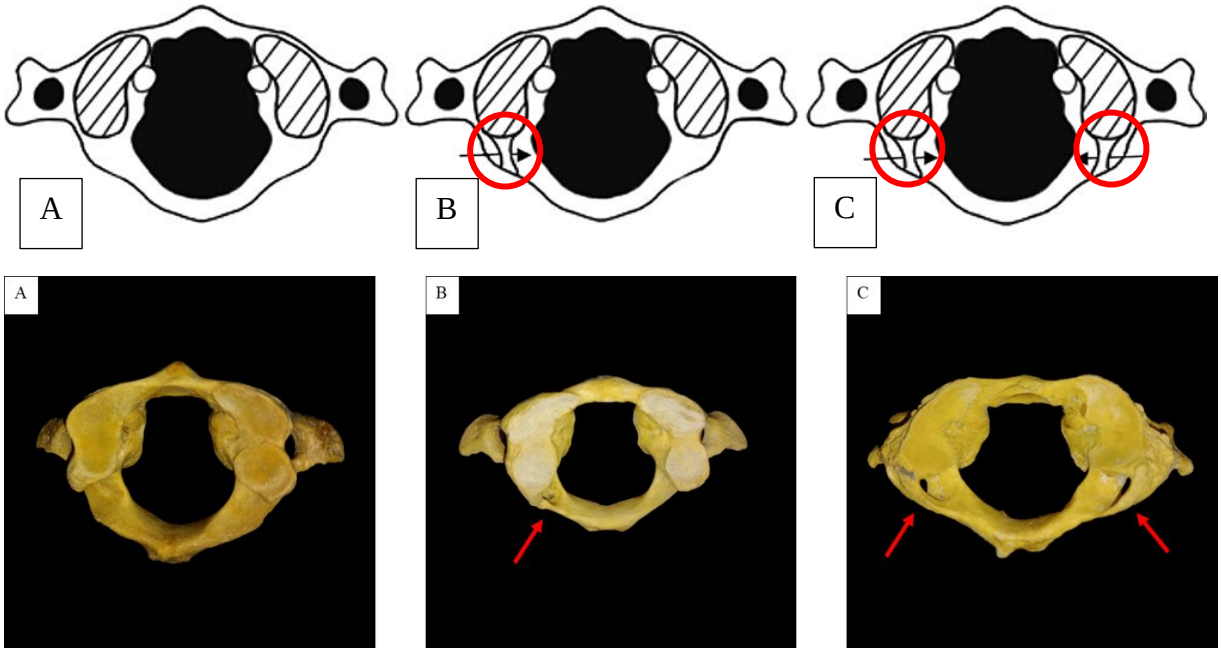

Figure A- 2. Posterior Bridging (red circle). A. Absent (score = 0), B. Unilateral (score = 1) and C. Bilateral (score = 2) *(Spiros, 2019)* (Photos by NP Bothma).

Double Superior Articular Facets (DSAF)

The DSAF (Figure A-3) forms when a groove forms in the middle of the superior articular facet of C1, resulting in double facets (Spiros, 2019). The DSAF can either be absent, unilateral, or bilateral, and will be scored zero, one, and two, respectively.

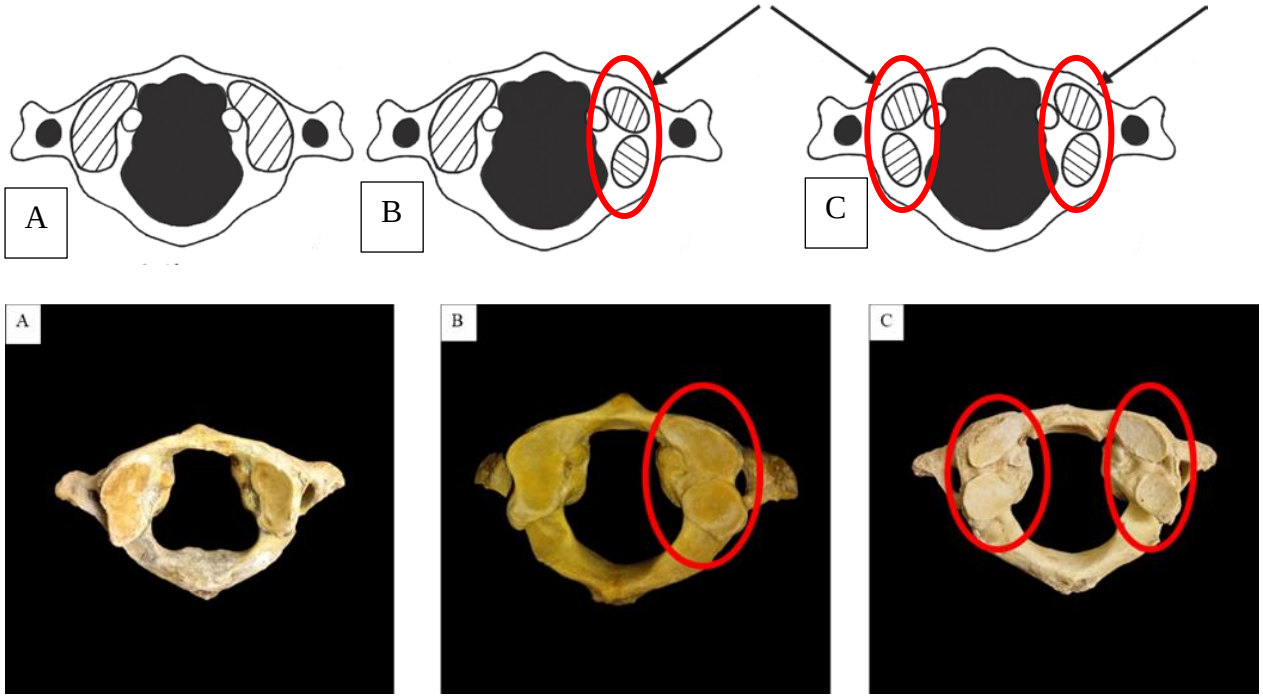

Figure A- 3. Double Superior Articular Facets (red circle). A. Absent (score = 0), B. Unilateral (score = 1) and C. Bilateral (score = 2) *(Spiros, 2019)* (Photos by NP Bothma).

Spinous Process Bifurcation (SPB)

The SPB (Figure A-4) is defined as a "split" of the most posterior part of the spinous process of the cervical vertebra to form two bony tubercles or projections (Cunningham, 1886; Asvat, 2012; Spiros, 2019). Following the Spiros (2019) approach, C3 and C4 will be assessed for this study. The bifidity is scored from zero to two, with zero observed as no bifid, one observed as partially bifid, and two as completely bifid. Partially bifid means that the spinous has no split entirely, instead two bony tubercles can be observed. Completely bifid means the spinous process has a complete split where two bony projections can be observed (Spiros, 2019).

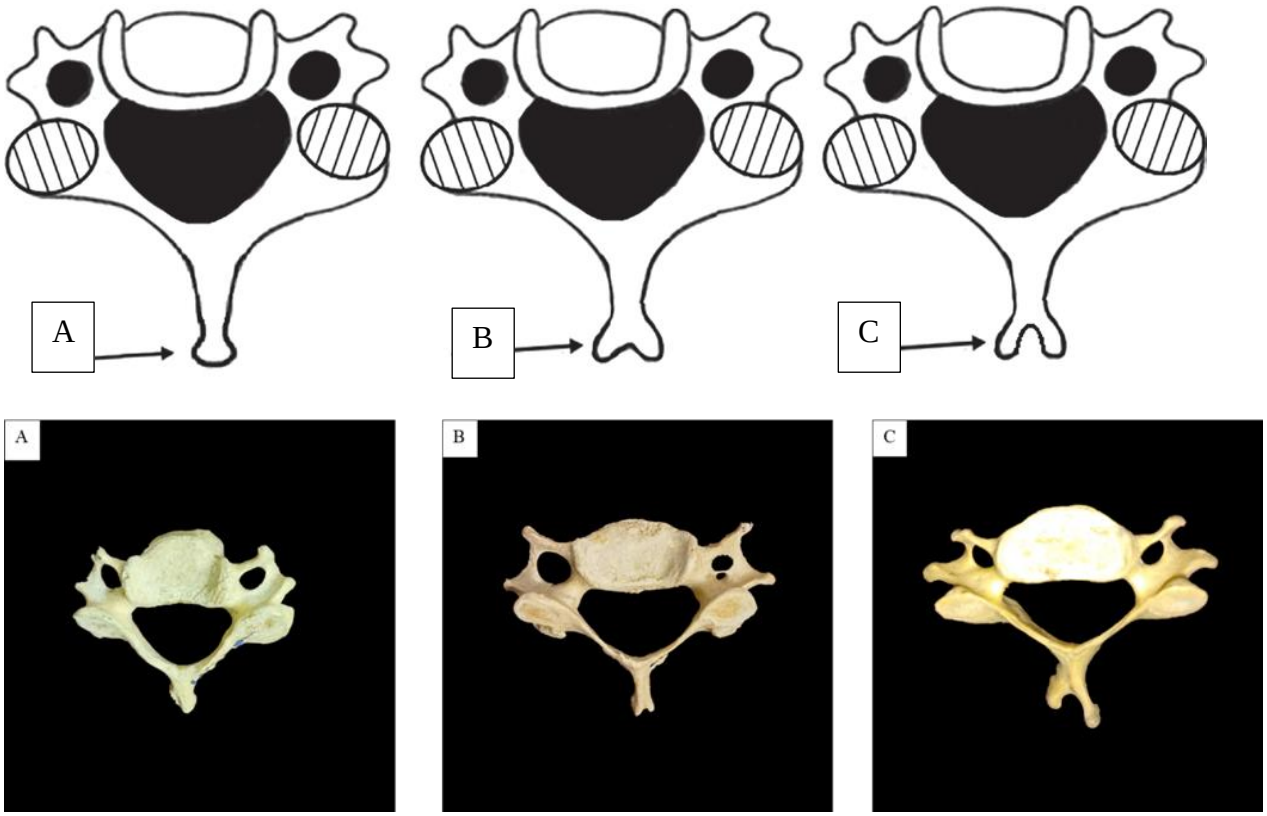

Figure A- 4. Spinous Process Bifurcation (black arrows). A. Non-bifid (score = 0), B. Partially bifid (score = 1) and C. Completely bifid (score = 2) *(Spiros, 2019)* (Photos by NP Bothma).

Suprascapular Foramen (SSF)

The SSF (Figure A-5) is formed when the suprascapular notch is enclosed, forming a foramen. The suprascapular notch is located on the superior border of the scapular, at the root of the coracoid process (White et al., 2012; Spiros, 2019). The SSF will be scored according to absence or presence which is zero to one, respectively. If the sample in this study presents mostly suprascapular notches, the scoring of the trait will be adjusted to fit the South African population.

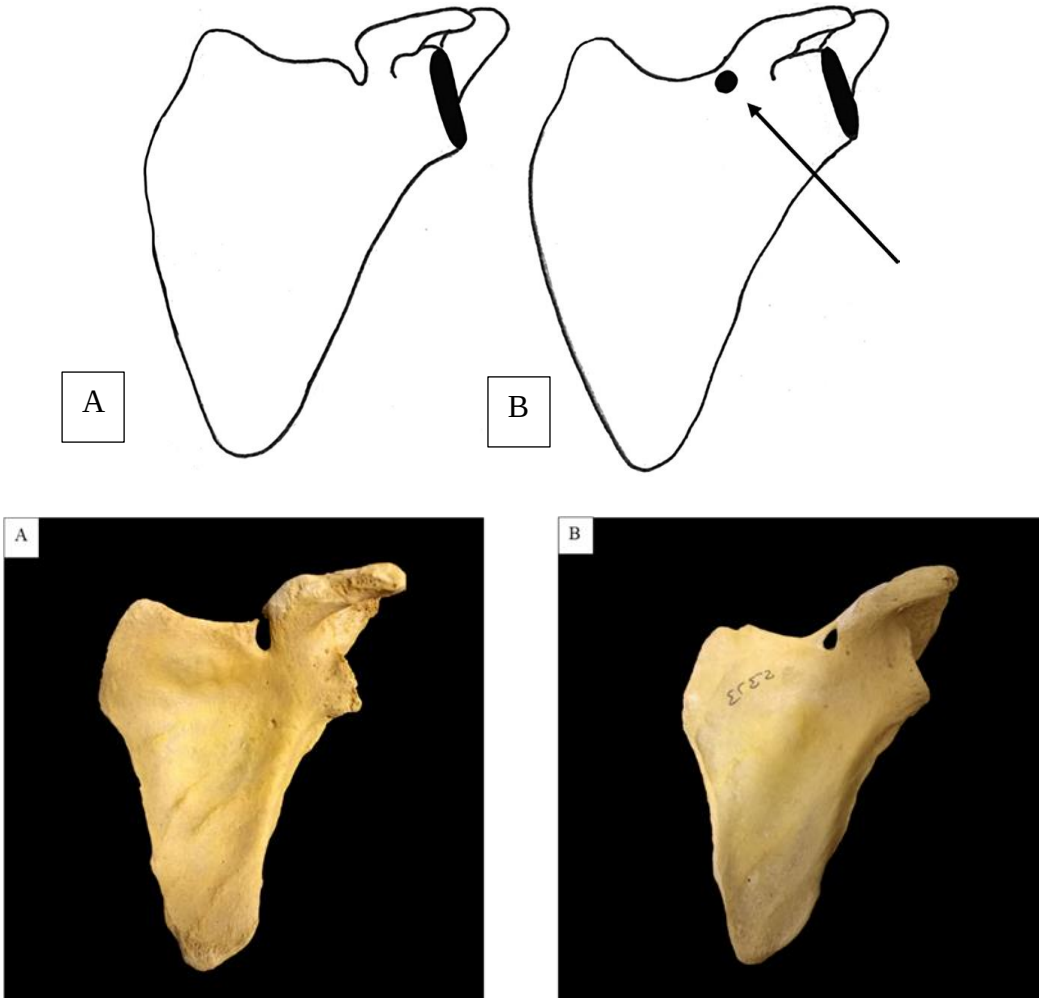

Figure A- 5. Suprascapular foramen (black dot). A. Absent (score = 0) and B. Present (score = 1) *(Spiros, 2019)* (Photos by NP Bothma).

Sternal Aperture

The sternal aperture (Figure A-6) is an oval opening or hole located on the sternum that can vary in size (Mann and Hunt, 2019; Spiros, 2019). Following Spiros (2019) both sternal and xiphoid apertures will be assessed as one. A scale of zero to one will be applied, where zero indicates absence and one represents the presence of the trait.

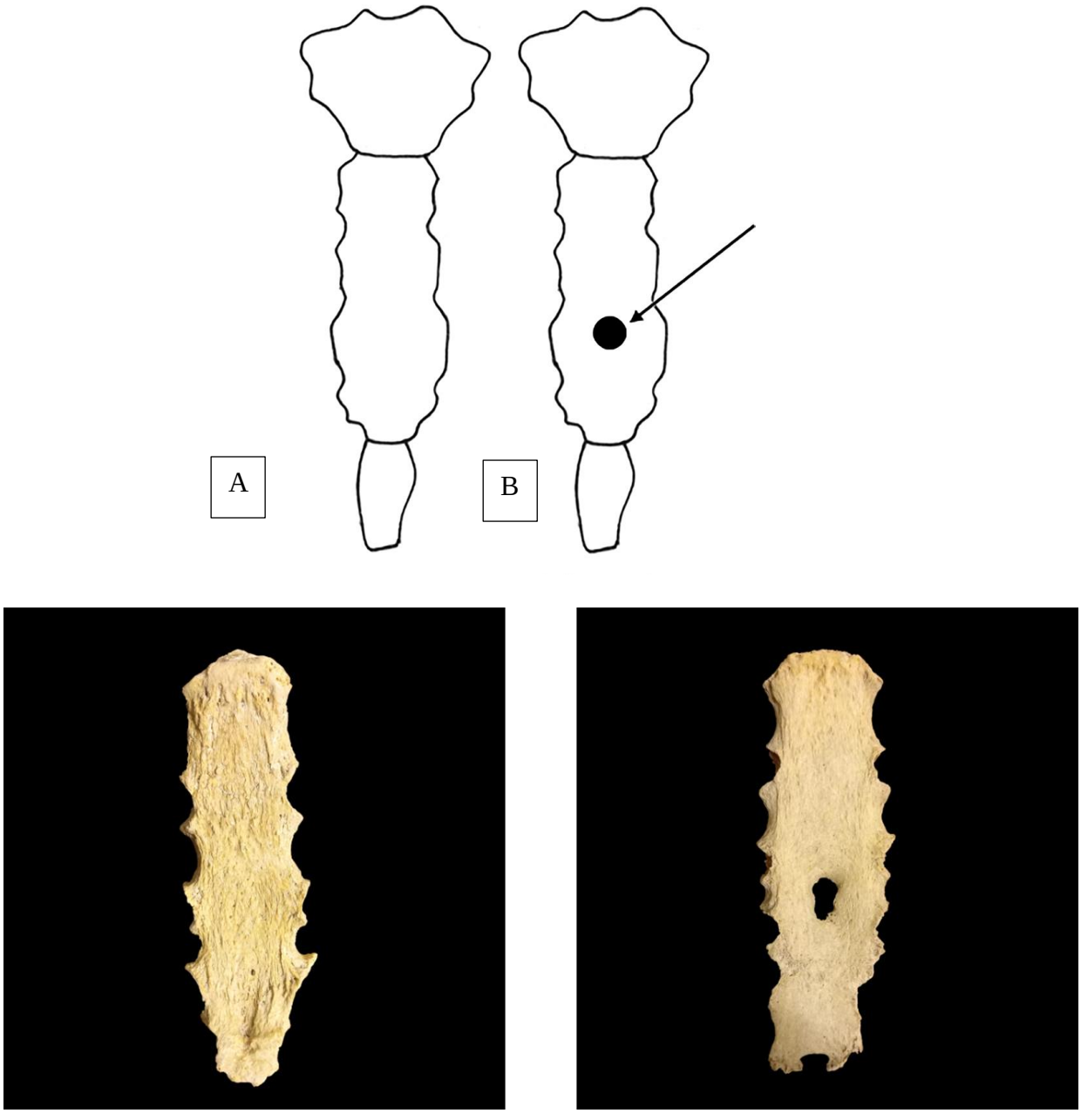

Figure A- 6. Sternal Aperture (black arrow). A. Absent (score = 0) and B. Present (score = 1) *(Spiros, 2019)* (Photos by NP Bothma).

Supra-condyloid process (SCP)

The SCP (Figure A-7) is a bony projection located above the medial epicondyle of the humerus and is an extension of the supracondylar ridge (Spiros, 2019). The trait will be scored according to absence or presence with a scale of zero to one, regardless of length or size.

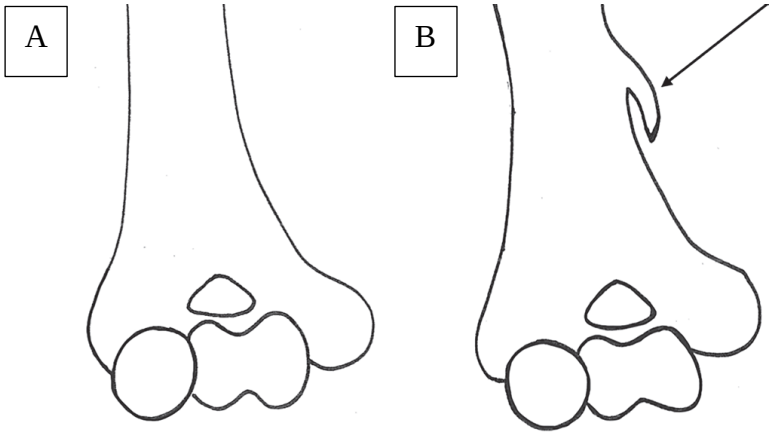

Figure A- 7. Supra-condyloid Process (black arrow). A. Absent (score = 0) and B. Present (score = 1) *(Spiros, 2019)* (Photos by NP Bothma).

Septal Aperture (SA)

The SA (Figure A-8) is an opening located on the distal part of the humerus, that joins the olecranon fossa to the coronoid fossa resulting in an oval or round shape hole (Mann and Hunt, 2019; Spiros, 2019). The trait will be scored with a scale of zero to three, where zero means absent, one means translucent, two indicates a small perforation (pinhole size), and three representing a large perforation (Spiros, 2019).

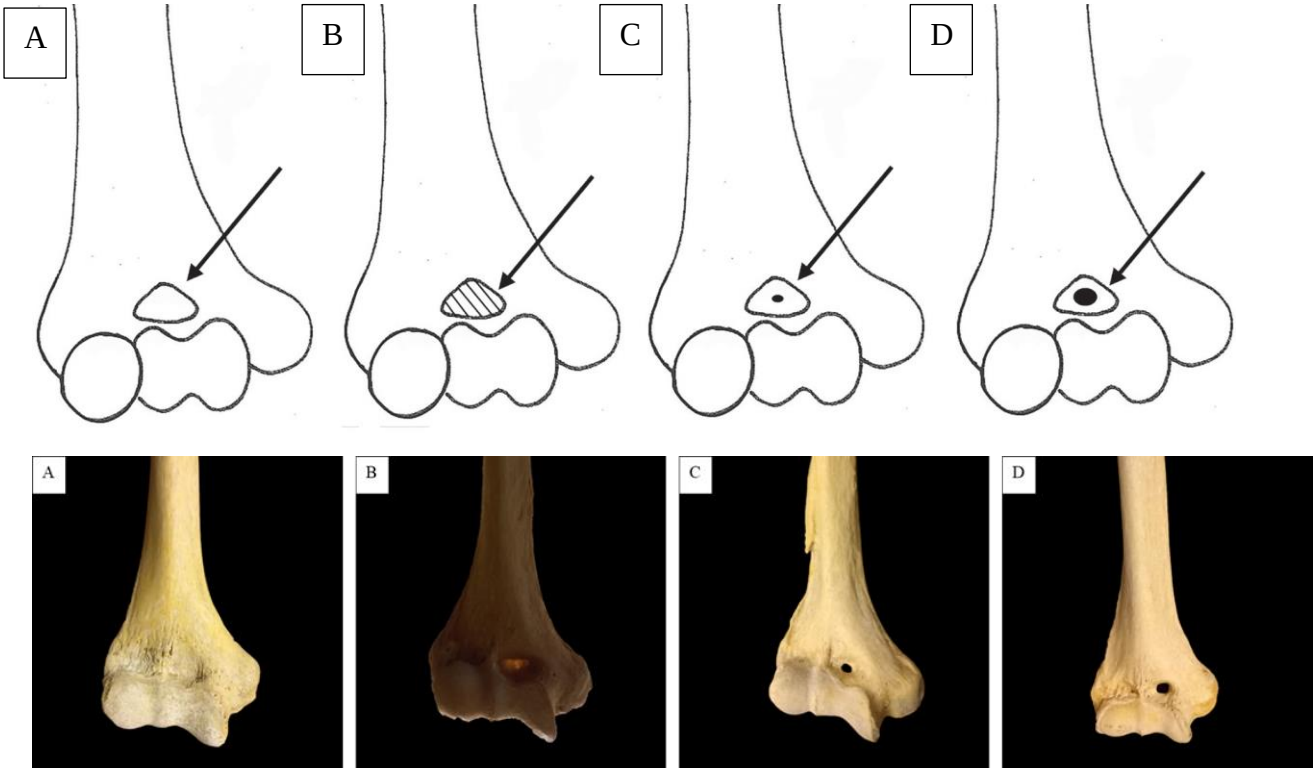

Figure A- 8. Septal Aperture (black arrows). A. Absent (score = 0), B. Translucent (score = 1), C. Small perforation (score = 2) and D. Large perforation (score = 3) (Spiros, 2019) (Photos by NP Bothma).

Third trochanter (TT)

The TT (Figure A-9) is defined as a tubercle located at the superior end of the gluteal ridge on the posterosuperior aspect of the femur and slightly inferior to the lesser trochanter (Mann and Hunt, 2019; Spiros, 2019). The trait will be scored as present or absent, which is zero and one, respectively.

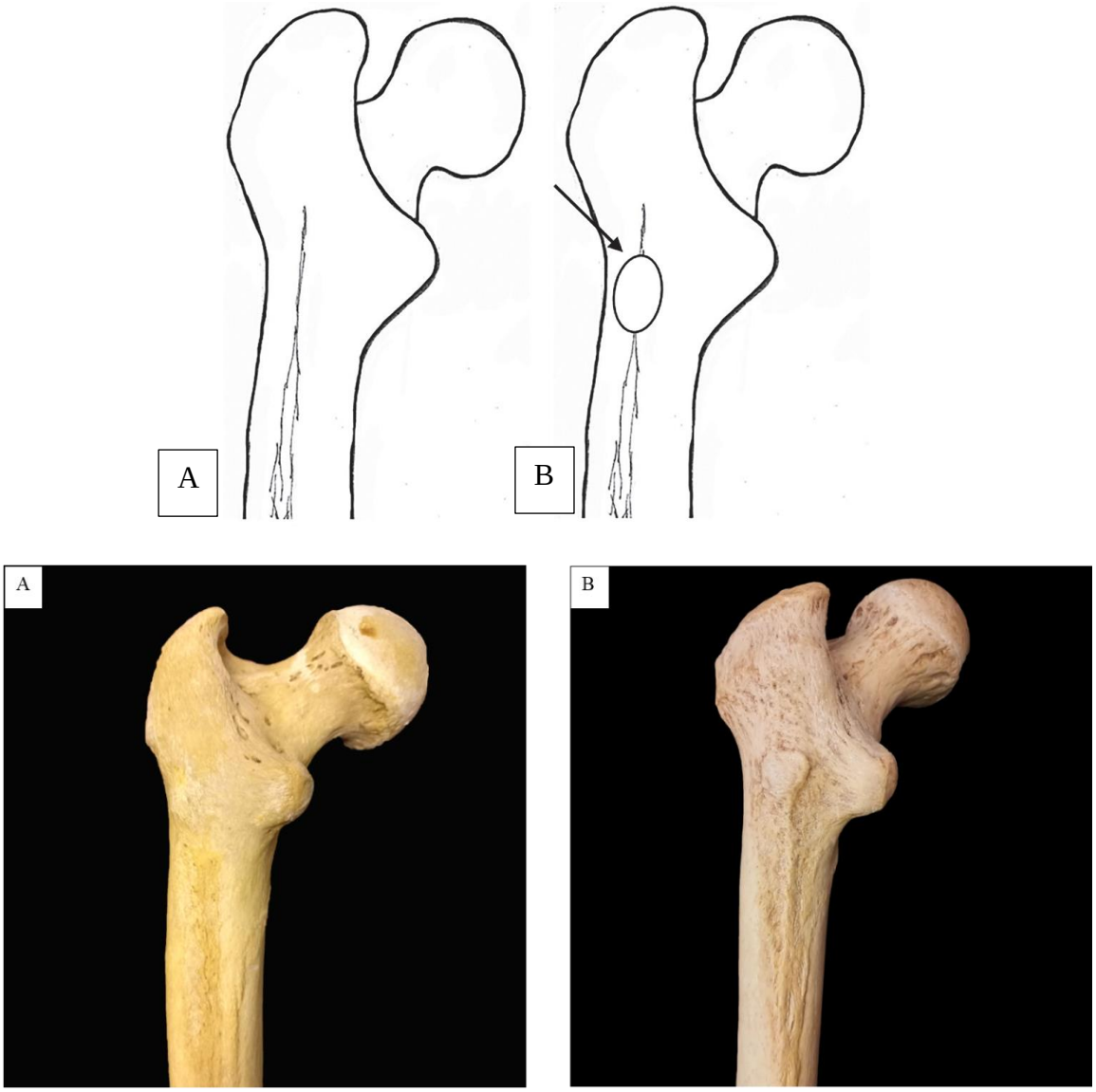

Figure A- 9. Third Trochanter (represented by oval). A. Absent (score = 0) and B. Present (score = 1) *(Spiros, 2019)* (Photos by NP Bothma).

Vastus Notch (VN)

The VN (Figure A-10) is a concaved or flattened indentation located on the superolateral or medial angle of the patella (Mann and Hunt, 2019; Spiros, 2019). The trait will be scored zero if there is no indentation or notch and scored a one if there is, regardless of size. The VN can be confused with a bipartite patella, but in comparison, a bipartite patella has a rough surface that appears like a portion of the patella has been bitten or pulled away and the VN has a smooth surface (Figure A-11).

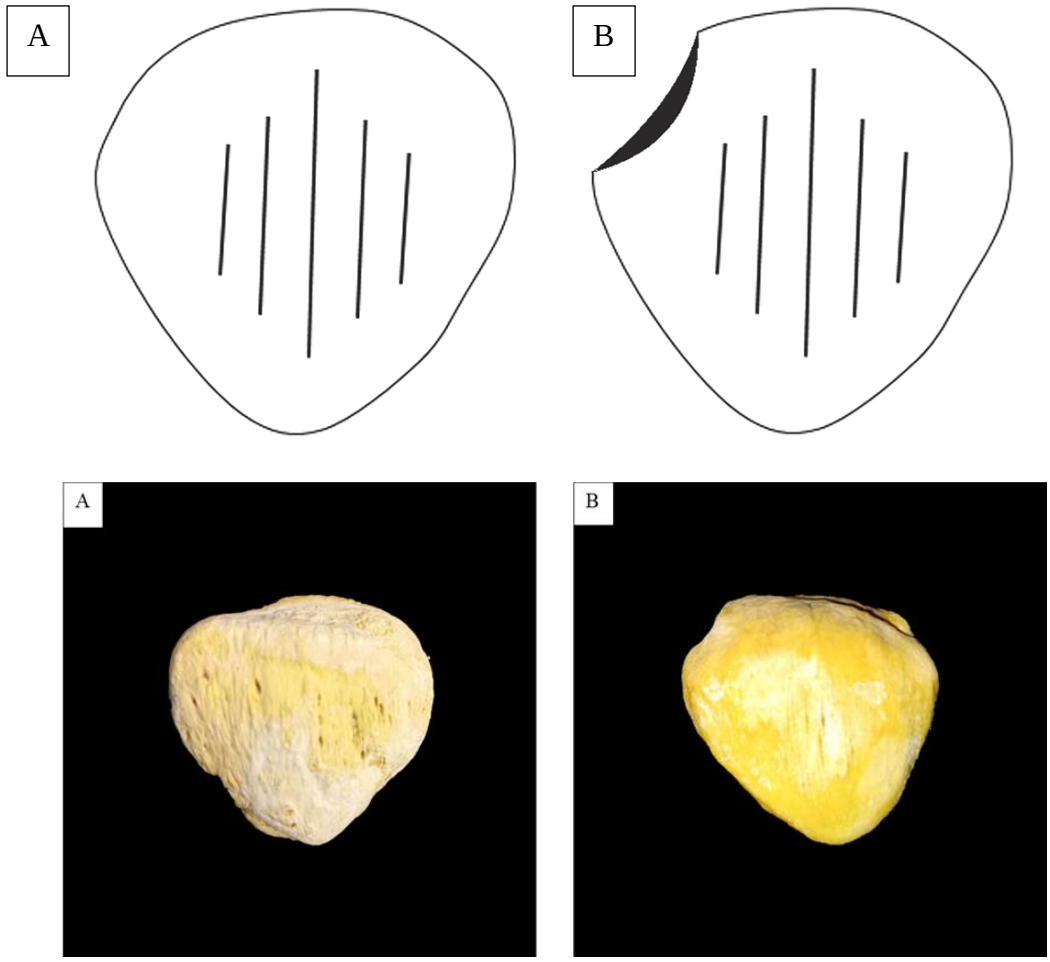

Figure A- 10. Vastus Notch. A. Absent (score = 0) and B. Present (score = 1) *(Spiros, 2019)* (Photos by NP Bothma).

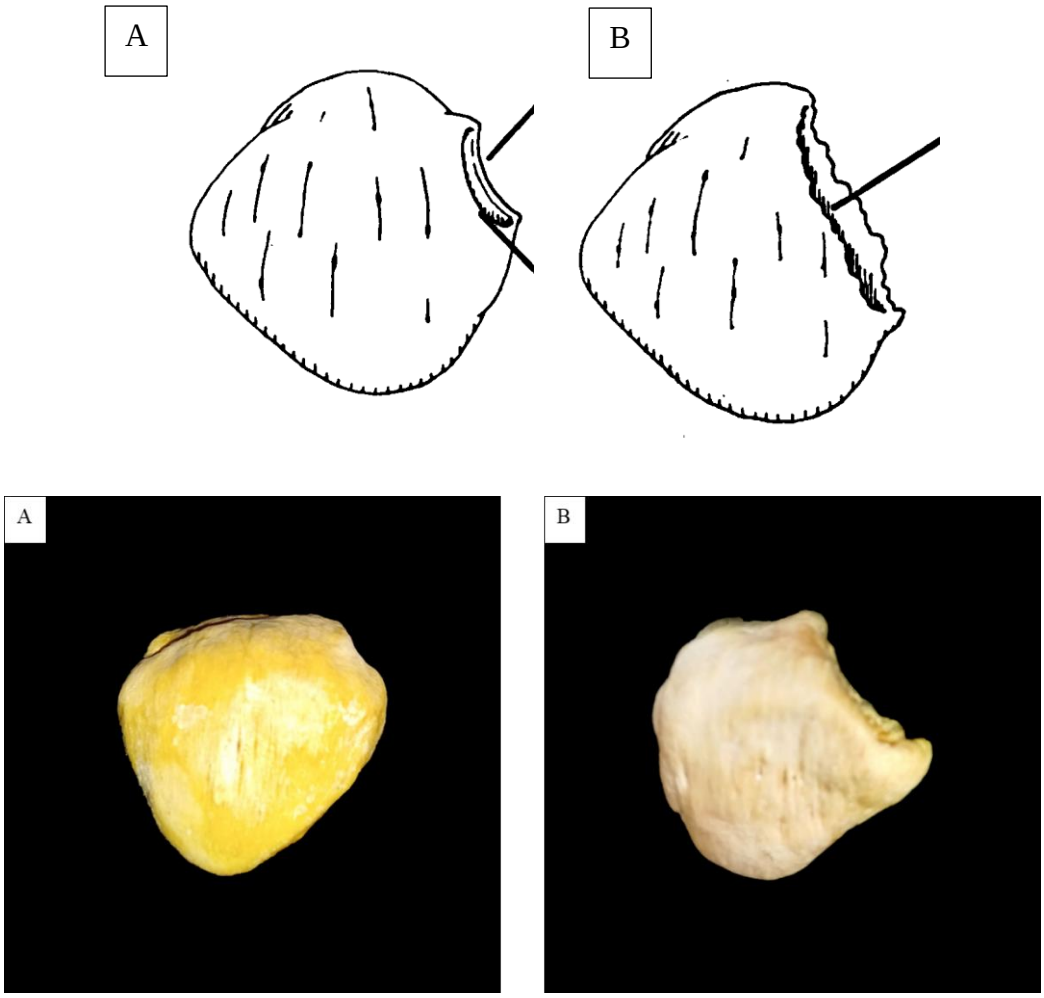

Figure A- 11. A. Patella with a vastus notch and B. Bipartite patella *(Finnegan, 1978)* (Photos by NP Bothma).

Anterior and Medial Calcaneal Facets (AMCF)

The AMCF (Figure A-12) is located on the calcaneus can vary from a single small oval facet to an elongated oval facet, where the middle facet is joined with an anterior or double facet where the two facets are separated (Spiros, 2019). The trait will be scored using a scale of zero to three, where zero represents the absence of an anterior facet, one represents a single elongated facet, two represents a small anterior facet, and three indicating a large anterior facet.

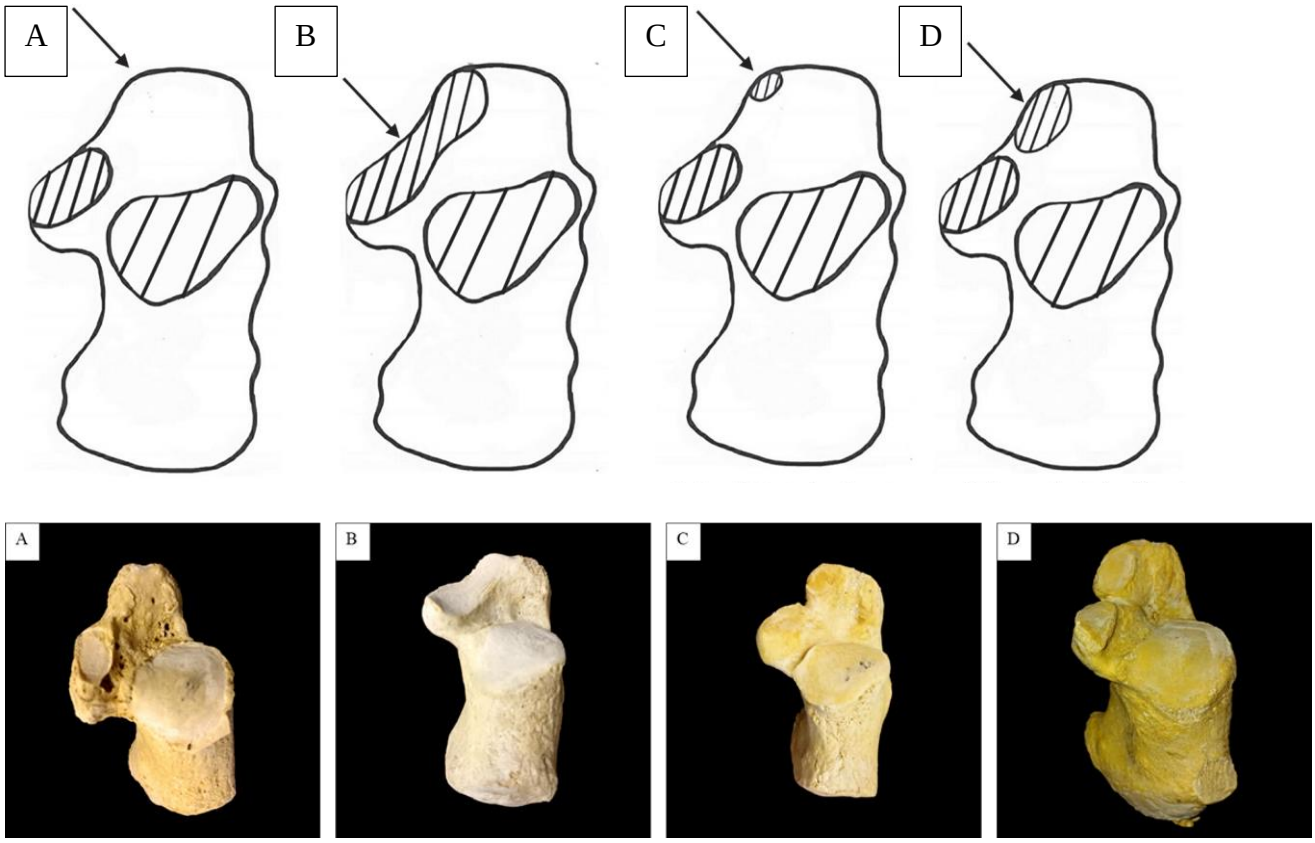

Figure A- 12. Anterior and middle calcaneal facets (black arrow showing Anterior facet and joined facets). A. Absent Anterior Facet (score = 0), B. Single Elongated Facet (score = 1), C. Small Anterior Facet (score = 2) and D. Large Anterior Facet (score = 3) *(Spiros, 2019)* (Photos by NP Bothma).