

https://dspace.org/

Sex and ancestry estimation of South African crania using 3D-ID

Lottering, Tamara Leigh

2020

https://hdl.handle.net/10539/30476

Lottering, Tamara Leigh (2020) Sex and ancestry estimation of South African crania using 3D-ID, University of the Witwatersrand, Johannesburg, <http://hdl.handle.net/10539/30476>

Downloaded from a DSpace repository. https://dspace.org/

SEX AND ANCESTRY ESTIMATION OF SOUTH AFRICAN CRANIA USING 3D-ID

Tamara Leigh Lottering



A dissertation submitted to the Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, in fulfilment of the requirements for the degree of Master of Science in Medicine (Dissertation)

Johannesburg, 2020

CANDIDATES DECLARATION

I, Tamara Leigh Lottering, declare that this Dissertation is my own, unaided work. It is being submitted for the Degree of Master of Science in Medicine (Dissertation) at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at any other University.



(Signature of candidate)

_____25_____day of___March_____2020_____in_Parktown_____

PRESENTATION ARISING FROM THIS STUDY

2019 The 19th Congress of The International Federation of Associations of Anatomists (Excel, London, UK): Oral presentation

Title: Sex and ancestry of South African Coloured crania using 3D-ID

POTENTIAL PAPERS ARISING FROM THIS DISSERTATION

Lottering, T.L., Brits, D. and Small, C. Suggested title: Sex and ancestry estimation of South African Coloured crania using 3D-ID. In preparation

Lottering, T.L., Brits, D. and Small, C. Suggested title: Sex and ancestry estimation of South African Coloured crania using the custom SADB in FORDISC. In preparation

Lottering, T.L., Brits, D. and Small, C. Suggested title: Sex and ancestry estimation of South African Black and White crania using 3D-ID. In preparation

ABSTRACT

Computer programs using osteometrics for sex and ancestry estimation have become more popular. The programs 3D-ID and FORDISC, compare data from an unknown individual with an existing database representing various populations. These programs are traditionally used in American cases, however, there is a need to expand the application of these programs for international use. The aim of this study was thus to assess how accurately 3D-ID estimates sex and ancestry for a South African sample as well as to compare these results from 3D-ID to results from FORDISC. A total of 450 South African Black (SAB), South African Coloured (SAC) and South African White (SAW) crania were digitised using a Microscribe 3-DX digitiser. Sex and ancestry were estimated using 3D-ID and FORDISC. 3D-ID achieved a range of accuracies (44.0% to 93.3%) for sex estimation, with females achieving the highest. It was found that 3D-ID classified SAB individuals as African American or European Southwestern, SAC individuals as European Southwestern, and SAW individuals as European American and European Southwestern. FORDISC achieved sex accuracies ranging from 66.7% to 100.0% and a range of ancestry accuracies with the SAC population achieving the lowest accuracies of 26.7% and 42.7%. It was seen that FORDISC achieved slightly higher accuracies compared to 3D-ID, however this could be because FORDISC has a custom database for South African individuals, whereas 3D-IDs database contains no South Africans. Overall, 3D-ID had moderate classifications showing that it is not an accurate tool for ancestry estimation of South African populations. It would therefore be beneficial to include South African individuals into the 3D-ID database to achieve increased accuracies.

ACKNOWLEDGEMENTS

I would like to thank the following people, whom I would not have been able to complete this Masters without. My supervisors Dr Candice Small and Dr Desiré Brits. Thank you for everything, your input, your advice and most importantly, your patience; you have made these last two years a lot easier for me. Dr Jason Hemingway for all your help with stats, your help was invaluable. Our head of school Professor Maryna Steyn for allowing me access to the Raymond A. Dart Collection of Human Skeletons. As well as Dr Brendon Billings and Mr Mashudu Mulaudzi for your assistance in the Raymond A. Dart Collection of Human Skeletons. Dr Mandi Alblas from the University of Stellenbosch for allowing me access to the Kirsten Collection. Professor Ann Ross for your assistance in showing me how to use 3D-ID and clarifying any questions I had regarding the program. My family and friends for helping pick me up when I was down, and making sure that I always had a plan for the future. The University of the Witwatersrand, Johannesburg, South Africa, for their financial assistance with the Postgraduate Merit Award, the Faculty Research Committee individual research grant and the J.J.J Smieszek Bursary.

TABLE OF CONTENTS

CANDIDATES DECLARATION I

PRESENTATION ARISING FROM THIS STUDY II

POTENTIAL PAPERS ARISING FROM THIS DISSERTATION II

ABSTRACT III

ACKNOWLEDGEMENTS IV

NOMENCLATURE/LIST OF ABBREVIATIONS AND SYMBOLS VIII

LIST OF FIGURES IX

LIST OF TABLES X

CHAPTER 1: INTRODUCTION 1

CHAPTER 2: LITERATURE REVIEW 2

2.1 GEOMETRIC MORPHOMETRICS 3

2.2 TRADITIONAL STUDIES 5

2.2.1 Sex Estimation 5

2.2.2 Ancestry Estimation 10

2.3 MODERN MORPHOMETRICS 21

2.3.2 Geometric Morphometrics for Ancestry Estimation 24

2.3.3 Computer Software Programs 26

2.3.3.1 FORDISC 27

2.3.3.2 3D-ID 32

CHAPTER 3: MATERIALS AND METHODS 39

3.1 MATERIALS 39

3.2 METHODS 40

3.2.1 Data Collection 40

3.2.2 Data Analysis 46

3.2.2.1 Repeatability 46

3.2.2.2 3D-ID *46*
3.2.2.3 FORDISC *48*
CHAPTER 4: RESULTS **51**
4.1 REPEATABILITY **51**
4.2 3D-ID INVESTIGATION **51**
4.2.1 SEX 51
4.2.2 ANCESTRY 53
4.2.2.1 South African Black (SAB) population *53*
4.2.2.2 South African White (SAW) population *56*
4.2.2.3 South African Coloured (SAC) population *59*
4.3 FORDISC INVESTIGATION **62**
4.3.1 SEX 62
4.3.2 ANCESTRY 63
4.3.3 SEX AND ANCESTRY 64
CHAPTER 5: DISCUSSION **66**
5.1 3D-ID INVESTIGATION **67**
5.1.1 SEX 67
5.1.2 ANCESTRY 69
5.1.2.1 South African Black (SAB) population *69*
5.1.2.2 South African White (SAW) population *72*
5.1.2.3 South African Coloured (SAC) population *74*
5.2 FORDISC **75**
5.2.1 SEX 75
5.2.2 ANCESTRY 76
5.2.2.1 South African Black (SAB) population *76*
5.2.2.2 South African White (SAW) population *77*
5.2.2.3 South African Coloured (SAC) population *78*
5.2.3 SEX AND ANCESTRY 79

5.2.3.1 South African Black (SAB) population 79
5.2.3.2 South African White (SAW) population 80
5.2.3.3 South African Coloured (SAC) population 80
5.3 3D-ID VS FORDISC 81
5.3.1 SEX 81
5.3.2 ANCESTRY 82
5.4 LIMITATIONS AND FUTURE DIRECTIONS 83
5.4.1 LIMITATIONS 83
5.4.2 FUTURE DIRECTIONS 84
CHAPTER 6: CONCLUSION 85
REFERENCES 88
APPENDICES 98
APPENDIX 1: ETHICAL CLEARANCE FOR THE RAYMOND A. DART COLLECTION, UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG 98
APPENDIX 2: PLAGIARISM DECLARATION 99
APPENDIX 3: TURN-IT-IN REPORT 100
APPENDIX 4: INTRA- AND INTER-OBSERVER ERROR RESULTS FROM MORPHOJ 102
APPENDIX 5: POSTERIOR PROBABILITIES (PP) AND F TYPICALITY'S (TP) FOR FORDISC ANALYSIS 105

NOMENCLATURE/LIST OF ABBREVIATIONS AND SYMBOLS

GM – Geometric morphometrics

2D – Two-dimensions\dimensional

3D – Three-dimensions\dimensional

SAB – South African Black

SAC – South African Coloured

SAW – South African White

DFA – Discriminant function analysis

LDA – Linear Discriminant analysis

FDB – Forensic Anthropology Databank

HDB – Howells' Database

SADB – South African Database

ILD – Interlandmark Distance

GPA – Generalised Procrustes Analysis

PCA – Principal component analysis

SD – Standard deviation

DCP – Data Collection Procedures

LIST OF FIGURES

Figure 2.1: An example of the output text generated when using FORDISC 29
Figure 2.2: An example of the output graph generated when using FORDISC 30
Figure 2.3: An example of the output generated when using 3D-ID..... 35
Figure 3.1: Norma frontalis illustrating the landmarks input into 3D-ID..... 44
Figure 3.2: Norma lateralis illustrating the landmarks input into 3D-ID 45
Figure 3.3: Norma basilaris illustrating the landmarks input into 3D-ID..... 45
Figure 3.4: Norma occipitalis illustrating the landmarks input into 3D-ID..... 46
Figure 4.1: Visually representation of the most and second most classification for SAB males and females 56
Figure 4.2: Visually representation of the most and second most classification for SAW males and females 58
Figure 4.3: Visually representation of the most and second most classification for SAC males and females 61

LIST OF TABLES

Table 2.1: Breakdown of population reference samples for cranial data in FORDISC 28
Table 2.2: Sample composition of the reference database in 3D-ID (Adapted from Slice & Ross, 2009) 34
Table 3.1: Age distribution of the sample. SD = standard deviation 40
Table 3.2: Description of digitized landmarks used in 3D-ID (amended from Slice & Ross, 2009) 42
Table 3.3: Interlandmark distances derived from 3D-ID landmarks incorporated into FORDISC 49
Table 4.1: Frequency table showing the correct sex classifications for **SAB** males (n = 75) and females (n = 75) using 3D-ID 52
Table 4.2: Frequency table showing the correct sex classifications for **SAW** males and females using 3D-ID (n = 75)..... 52
Table 4.3: Frequency table showing the sex classifications for **SAC** males (n = 75) and females (n = 75) using 3D-ID 53
Table 4.4: Frequency table related to ancestry classification for **SAB** males (n = 75)..... 54
Table 4.5: Frequency table related to ancestry classification for **SAB** Females (n = 75)..... 55
Table 4.6: Frequency table related to ancestry classification for **SAW** males (n = 75) 57
Table 4.7: Frequency table related to ancestry classification for **SAW** females (n = 75) 58
Table 4.8: Frequency table related to ancestry classification for **SAC** males (n = 75) 60
Table 4.9: Frequency table related to ancestry classification for **SAC** females (n = 75) 61
Table 4.10: **Sex** classification for FORDISC for SAB males (n = 75) and females (n = 75), SAW males (n = 75) and females (n = 75) and SAC males (n = 75) and females (n = 75) 62
Table 4.11: **Ancestry** classification for FORDISC for SAB males (n = 75) and females (n = 75), SAW males (n = 75) and females (n = 75) and SAC males (n = 75) and females (n = 75) 63
Table 4.12: **Sex and ancestry** classification for FORDISC for SAB males (n=75) and females (n=75), SAW males (n=75) and females (n=75) and SAC males (n=75) and females (n=75) 64
Table 5.1: Comparison of sex classifications when using 3D-ID and FORDISC 82
Table 5.2: Comparison of "correct" ancestry classifications for SAB and SAW populations when using 3D-ID and FORDISC..... 82

CHAPTER 1: INTRODUCTION

South Africa has an extremely high violent crime rate (Steyn & İşcan, 1997; Dayal *et al*., 2008; Statistics South Africa, 2019a), which has resulted in an abundance of unidentified remains. The skeletal remains are often fragmentary, which could be due to environmental influences or a calculated forensic counter-measure. Forensic anthropologists are called upon to examine the skeletal remains and create a biological profile that could potentially help identify the individual (Krogman & İşcan, 1986; Cattaneo, 2007; Bidmos *et al*., 2010). Sex and ancestry estimation are critical for the biological profile, as other aspects such as age-at-death and stature, cannot be effectively assessed without this information (Kimmerle *et al*., 2008; Abdel Fatah *et al*., 2014; Langley & Tersigni-Tarrant, 2017).

Although traditional methods used by forensic anthropologists have provided some excellent results, they have been shown to have limitations, such as being highly subjective, expensive and often ill-defined making these methods difficult to reproduce. It has been found that these limitations can be overcome when rather using landmark-based analysis (geometric morphometrics), for example, it provides a repeatable method that is very important for forensic anthropologist (Slice, 2007; Mitteroecker & Gunz, 2009; Bidmos *et al.*, 2010).

Computer programs have recently become more popular with forensic anthropologists with programs such as FORDISC and 3D-ID. These programs use craniometrics in the estimation of sex and ancestry; with FORDISC assessing linear measurements and 3D-ID using geometric morphometric methods applied to digitised cranial landmarks. Programs like these have benefits such as large reference databases that is extremely important for forensic anthropologists (L'Abbé *et al*., 2013; Urbanová *et al*., 2014; Manthey *et al*., 2018; Bertsatos *et al*., 2019). Both FORDISC and 3D-ID have been used on the South African Black and White populations with mixed results (L'Abbe et al 2013; King 2015), however no research has been done on the South African Coloured population using these programs. Thus, the current study plans on addressing this lack of research of the South African Coloured population by assessing sex and ancestry of individuals from the South African population when using 3D-ID and then to compare those results to results attained when using the custom South African Database in FORDSIC.

CHAPTER 2: LITERATURE REVIEW

High crime rates in South Africa have led to an increase in the number of unidentified, decomposed or fragmented skeletal remains. South Africa has one of the highest violent crime rates in the world (Steyn & İşcan, 1997; Dayal *et al*., 2008; Statistics South Africa, 2019a), and according to the 2019 statistical release, over the 2018 to 2019 period rates for crimes such as murder and rape have been increasing (Statistics South Africa, 2019a). Remains are often discovered buried in shallow graves in the veldt or in vacant plots; these remains are usually fragmentary due to scavengers, exposure to the elements, or deliberately fragmented as a forensic counter-measure. These conditions have resulted in inefficacy of traditional methods of assessment conducted by forensic pathologists (Small, 2016). This problem is further aggravated as dental information rarely exists and analysing DNA is time-consuming and very expensive (Barrier & L'Abbé, 2008). Forensic anthropologists are called upon mainly to examine the skeletal remains of an individual and compile a biological profile that can potentially identify the individual (Krogman & İşcan, 1986; Cattaneo, 2007; Bidmos *et al*., 2010). A biological profile consists of the estimation of sex, age-at-death, ancestry, stature, trauma, pathologies and other individualising factors; providing practical information that can be used in a medico-legal setting (Krogman & İşcan, 1986; Sauer, 1992; Ousley *et al*., 2009; İşcan & Steyn, 2013; Stull *et al*., 2014; Langley & Tersigni-Tarrant, 2017). Sex and ancestry estimation are crucial components of the biological profile, as other features of the profile such as age and stature, cannot be efficiently assessed without this information (Kimmerle *et al*., 2008; Abdel Fatah *et al*., 2014; Langley & Tersigni-Tarrant, 2017).

Traditionally, forensic anthropologists use two general methods for sex and ancestry estimation, namely non-metric (qualitative) and metric (quantitative) methods (Nafte, 2000; Dayal *et al*., 2008; Bidmos *et al*., 2010). Non-metric methods involve visually observing traits on the bones and describing any shape differences seen. These methods are time- and cost-efficient, are seldom population specific and can be applied to fragmentary remains. Non-metric methods are therefore often used in the field as preliminary analysis. Nevertheless, these methods are highly subjective, require a highly skilled observer and often the methodologies are ill-defined, making them difficult to reproduce (Bidmos *et al*., 2010; Bigoni *et al*., 2010; Anastasiou & Chamberlain, 2013; Abdel Fatah *et al*., 2014). In order to improve on subjectivity, one can use the more objective metric method. The metric method involves collecting a wide range of measurements (using

callipers and an osteometric board as applicable) and then applying multivariate statistics to quantify morphological characteristics for sex and ancestry estimation. There are three commonly used procedures for metric measurements: 1) direct measurements, 2) indices and 3) discriminant functions. Direct measurements are taken from an unknown individual and compared with the known values of different populations/sexes in the case of ancestry or sex estimation. Indices are assessed by dividing one measurement by another and multiplying the result by 100; dividing allows for standardisation of size, while the multiplication gives the practitioner a percentage. The results are then compared with results from multiple groups/sexes. Discriminant function analysis is a statistical technique that uses multiple measurements, which then classifies an unknown individual and gives the probability of their classification into a certain group, for example sex or ancestral group. The metric method can be tested using multivariate statistics, is less subjective, can be applied by observers with less experience and can easily be reproduced given that the methodology is well-defined (Nafte, 2000; Bidmos *et al*., 2010; Moore *et al*., 2016, Murphy & Garvin, 2018). However, metric methods require more time, a lot of information about shape is lost as only linear measurements are taken, and these methods can be more expensive as specialised equipment is necessary (Slice & Ross, 2009). Another limitation is that any functions derived are highly population specific and can be sensitive to any secular trends involving size (De Villiers, 1968; Steyn & İşcan, 1997; Steyn & İşcan, 1998; Steyn & İşcan, 1999; Dayal *et al*., 2008). To combat these disadvantages landmark-based analysis (geometric morphometrics) is being used.

2.1 GEOMETRIC MORPHOMETRICS

Landmark-based analysis has become more common recently, as it provides a potential solution for the limitations of traditional methods, such as providing a repeatable method while still retaining information on shape (Slice, 2007; Mitteroecker & Gunz, 2009; Bidmos *et al.*, 2010). Geometric morphometrics (GM) uses three-dimensional co-ordinates of homologous landmarks, allowing for the variability due to both shape and size to be quantified (Kimmerle *et al.*, 2008; Bigoni *et al*., 2010). One of the advantages of using GM compared to traditional methods is that GM focuses on coordinates of landmarks which doesn't allow for a loss of shape, whereas traditional methods tend to collapse the overall shape of the object into linear measurements and angles (Adams *et al.*, 2004; Franklin *et al*., 2006; Slice, 2007; Green & Curnoe, 2009). GM provides repeatabilities equivalent to those of metric studies and higher, while still preserving

shape information (Gillick, 2012). It has also been found much easier and more efficient to collect data using a digitiser compared to traditional callipers as digitising doesn't require as much time or experience (Latham *et al.*, 2018).

Landmarks can be captured in different ways with a variety of equipment. Two-dimensional landmarks (2D) can be obtained using digital photographs or a digitising table, while three-dimensional (3D) landmarks can be acquired using digitisers such as a Microscribe or collected from virtual reconstructions of scanned specimens and visualised electronically (Mitteroecker & Gunz, 2009). When using GM techniques, it is important which landmarks you use and locating landmarks requires certain criteria. It is suggested by Zelditch and colleagues (2004) that landmarks should be "(1) homologous anatomical loci that (2) do not alter their typological positions relative to other landmarks, (3) provide adequate coverage of the morphology, (4) can be found repeatedly and reliably, and (5) lie within the same plane" (Zelditch *et al.*, 2004:24). They also suggest that landmarks should be chosen specifically to answer the study question, however, they shouldn't be so specific as to stop the detection of novel shape information (Zelditch *et al.*, 2004).

There are three types of landmarks: Type I, Type II and Type III (Bookstein *et al.*, 1991). Type I landmarks are considered homologous landmarks, meaning that all specimens have relatively the same positioned landmark on the bone. These landmarks are located with ease and are considered to be repeatable among different specimens and the researchers assessing them (Bookstein *et al.*, 1991). An example of Type I landmarks are points located at the intersection between three sutures, such as asterion or bregma. Type II landmarks are defined as landmarks not based on locally defined biology, such as points of maximum curvature along a bony edge or tip of a tooth and points of bony projections (such as mastoidale). Type III landmarks are much more difficult to locate as they are usually the more extreme points whose exact location isn't well defined across specimens. An example of a Type III landmark is ectoconchion, which is described as a point with extreme curvature or the furthest point from a certain landmark (Bookstein *et al.*, 1991).

Once the landmark's coordinates are digitised, they are then either superimposed using generalised Procrustes superimposition or partial Procrustes superimposition where they differ in their determination of a reference orientation, with generalised Procrustes analysis being optimally determined while partial Procrustes is arbitrarily chosen (Adams *et al.*, 2004; Mitteroecker &

Gunz, 2009). Procrustes superimposition translates, scales and rotates the landmark configurations to remove all information unrelated to shape. It minimises the differences among the configurations of landmarks without changing their geometry. During the process of translation, all configurations are moved to the same location. During scaling, centroid size (a scaling factor) is calculated, and size is mitigated, such that shape can be independently assessed, after which size can be reintroduced. In the final process, all configurations are rotated to minimise the sum of the squared distances between corresponding landmarks (Rohlf & Slice, 1990). It is possible to perform these three processes as the differences in location, size and orientation do not alter the shape of the configurations.

Traditional morphometrics assess form quantitatively, where form is defined as both shape and size. Shape is described as the geometric information that remained after translation, rotation and scaling have been done (Kendall, 1977; Zelditch *et al*., 2004; Gonzalez *et al.*, 2009; Slice & Ross, 2009). The ability to remove size as a factor is very important in morphological analyses as the size of the specimen can have an effect on any measurements taken and this could have an effect on the results attained. In traditional forensic analyses, shape is not a distinct variable that is separately assessed from size and this could have an influence on any interpretations made about the morphological differences (Gillick, 2012).

One of the many advantages of GM is that the analysis of the results can easily be understood and visualised (Zelditch *et al*., 2004; Franklin *et al*., 2006; Green & Curnoe, 2009; Gillick, 2012). GM techniques allow shape changes that can be seen by the naked eye, as well as those that are more subtle and difficult to identify, to be quantified statistically, graphically expressed and visually represented (Gonzalez *et al*., 2009; Gillick, 2012; Klingenberg, 2013; Murphy & Garvin, 2018).

2.2 TRADITIONAL STUDIES

2.2.1 Sex Estimation

Sex estimation is the assessment of sexually dimorphic characteristics that show differences in shape and size between males and females. These differences are related to different biomechanical functions of locomotion and parturition. These differences are most pronounced in the pelvis, long bone epiphyses and the skull (İşcan & Steyn, 2013). The skull is one of the most studied skeletal elements as not only is variation in cranial morphology associated with sex but the

cranium is also usually more well preserved than other skeletal elements (Krogman & İşcan, 1986; Dawson *et al*., 2011; Anastasiou & Chamberlain, 2013). The variation in the skull can be linked to nutrition, genetics, muscle attachments and cultural variation (İşcan *et al*., 1995).

Size plays an important role in influencing sex differences of the cranium, including robusticity (Rosas & Bastir, 2002; Green & Curnoe, 2009). This forms the foundation of sex estimation from cranial morphology, with larger crania showing more robust features being classified as males while smaller crania showing more gracile/prepubertal features are classified as females (Green & Curnoe, 2009; Christensen & Passalacqua, 2018). However, because sexual dimorphism is not "extreme" there is some overlap in size/robusticity between males and females that can make sex estimation difficult (Green & Curnoe, 2009; White, 2012; Christensen & Passalacqua, 2018). It is rare for forensic anthropologists to receive a complete skeleton for analysis and therefore numerous methods for the estimation of sex have been developed from individual skeletal elements, such as the cranium and pelvis; with studies achieving accuracies as high as 95% for the pelvis and 92% for the skull (İşcan & Steyn, 2013).

Non-metric traits can be used for sex estimation which includes noting the degree that a particular trait is expressed, or assessing the presence or absence of certain characteristics, for example, the presence or absence of the infraorbital suture (Christensen & Passalacqua, 2018). It was recommended by Buikstra and Ubelaker (1994) that the following five characteristics should be used in non-metric estimation of sex when analysing the skull: glabella prominence, supraorbital margin sharpness, projection of the mental eminence, nuchal crest robusticity and the size of the mastoid process. Each of these traits were represented using five-line drawings, these drawings gave an indication of how the traits varied from "clearly" female to "clearly" male features, on a rating scale of 1-5, with 3 being indeterminate. These drawings, with their accompanying descriptions are then used to score the cranial traits to estimate sex. Following this, Rogers (2005) assessed 17 non-metric traits on the skull to determine their efficiency in sex estimation. Of the 17 traits, four of the traits were deemed useful for assessing sex: zygomatic extension, supraorbital ridge (glabella), nasal aperture and malar size/robusticity. Williams and Rogers (2006) evaluated the accuracy and precisions of the 17 traits evaluated by Rogers (2005) using modern European White crania. Of all the traits, they found six that achieved high accuracy and precision: supraorbital ridge, nasal aperture, zygomatic extension, gonial angle of the

mandible, size and architecture of the skull, and mastoid; reporting accuracies ranging from 80% to as high as 92%.

Walker (2008) assessed the five characteristics outlined by Buikstra and Ubelaker (1994) to determine the possibility of correctly estimating sex from 304 White American and English individuals and a test sample of 150 prehistoric Native American individuals. He used ordinal scores to compute discriminant functions for sex estimation, reporting that combinations of characteristics achieved accuracies ranging from 75.8% to 92.7%. Walker's (2008) method has commonly been used by forensic anthropologists as it is quick and easy. Garvin and colleagues (2014) assessed sex in six different populations using Walker's (2008) method. In addition, they also investigated other variables such as age, populations and body size to explore whether or not these variables influence sex estimation. Their sample of 499 individuals consisted of Black and White Americans, medieval Nubians and Native Americans. The results they achieved supported Walker's (2008) findings, with the pooled sample achieving a correct classification of 85% and glabella and mastoid scores having the greatest weight in the function. Garvin and colleagues (2014) found that population has a significant effect on the cranial trait scores whereas age and body size did not. This led to their recommendation that the assessment of sex from an unknown skull should be done using population-specific statistical approaches (such as discriminant function analysis).

Krüger and colleagues (2014) tested the Walker (2008) method to examine sex estimation among the South African population. The study's sample consisted of 245 South African Black and White male and female crania. It was found that South African White males achieved the highest accuracies of 94% to 97%, while South African White females achieved the lowest accuracies of 31% to 62%. Similar to Garvin and colleagues (2014), the authors emphasised that these low accuracies show a need for population specific standards. So the authors modified Walker's (2008) method for a South African population which yielded higher classification rates ranging between 84% and 93%.

Even though non-metric methods have shown reasonably good accuracies they have limitations; such as being highly subjective (Bidmos *et al*., 2010; Bigoni *et al*., 2010; Anastasiou & Chamberlain, 2013; Abdel Fatah *et al*., 2014), thus metric methods are also used. Most of the metric methods used in forensics are considered to be population specific, with a plethora of metric studies available representing many different populations. Giles and Elliot (1963) assessed sex

using discriminant functions by measuring over 400 skulls from the Terry (Washington University School of Medicine) and Hamann-Todd Collections (Western Reserve University School of Medicine). The authors achieved correct classifications of 85% when using 100 measurements from the skull and developed multiple equations using different measurements to accommodate for cases where the crania are incomplete. Recent studies have achieved lower accuracies when using the formulas derived by Giles and Elliot (1963) on more modern populations, however their method and formulas are still used.

İşcan and colleagues (1995) assessed 16 cranial measurements from 84 crania of Japanese individuals. Two functions were developed, the first using five measurements, the second using seven measurements, achieving an average accuracy of 84%. Both these functions included mastoid height, and so they developed a function from mastoid height alone, achieving an average accuracy of 74%. Ogawa and associates (2013) analysed 113 skulls of modern Japanese individuals (73 males and 40 females) to develop population-specific discriminant functions for the estimation of sex. Ten cranial and mandibular measurements were taken for each individual from the records housed at the National Research Institute of Police in Japan and nine discriminant functions were generated. Ogawa and associates (2013) found cranial base length was the most sexually dimorphic measurement while maximum cranial breadth was the least sexually dimorphic. Cross-validated accuracies of 77.8% to 88.1% were achieved when using the total sample of 113 individuals in the formulae, while accuracies of 86.7% to 93.0% were achieved during a validation study of a subsample of 50 individuals. This study found that when combining maximum cranial length, cranial base length, maximum frontal breadth and bizygomatic breadth, the greatest accuracies of 91.3% for males and 90.9% for females were achieved. While, when combining maximum cranial length, cranial base length, upper facial breadth and bizygomatic breadth the least accurate accuracy of 87.0% was achieved for males and 86.4% was achieved for females.

Kranioti and associates (2008) used discriminant function analysis to explore 16 measurements taken from 178 individuals from a Cretan population from two cemeteries; St Konstantinos and Pateles, Heraklion, Crete. The most discriminatory single-variable discriminant function was bizygomatic breadth which achieved an average accuracy rate of 82%. Of the 16 measurements, six demarking measurements were found to have cross-validated accuracies ranging from 70.20 to 81.9%. Kranioti and associates (2008) found that the size of Cretan crania

was more similar to North American crania compared to South African White crania. Spradley and Jantz (2011) analysed 704 individuals (430 males and 274 females) of North American Black and White populations from the Forensic Databank. Similar to the study by Kranioti and colleagues (2008), bizygomatic breadth was the most sexually dimorphic, achieving correct sex classifications ranging from 75% to 78% for both populations. While the least sexually dimorphic trait was orbital height which achieved correct sex classifications ranging from 44% to 47% for both populations. One discriminant function was derived for each population. The function derived for the North American Black population achieved accuracies of 90.57% for males and 90.7% for females, with an overall accuracy of 90.64%, while the function derived for the White population achieved accuracies of 91.53% for males and 88.49% for females with an overall accuracy of 90.07%. The fact that the functions are made up using different measures for the two populations, gives an indication of population variation in the cranium between the North American Black and White populations.

Marinescu and colleagues (2014) assessed 200 crania of modern Romanian individuals from the Rainer collection, Bucharest. Using 11 cranial measurements, they calculated single-variable and multi-variable discriminant functions. Single-variable discriminant functions achieved accuracies ranging from 57.0% (orbital height) to 83.5% (facial breadth), while multi-variable discriminant functions achieved accuracies of 81.0% (cranial vault measurements), 83.5% (facial measurements) and 88.0% (all cranial measurements). They noted that compared to other studies, the Romanian population was more dimorphic in the facial region than the cranial vault.

Regardless of the method used, it has been shown that sex estimation using the cranium is population-specific (Steyn & İşcan, 1998; Walrath *et al*., 2004; Komar & Buikstra, 2008; Green & Curnoe, 2009; Dirkmaat, 2012; White, 2012; Latham *et al*., 2018), as observed in the considerable variability in body size, shape and robusticity that can be seen amongst different populations. Previous research (e.g. Dayal *et al.*, 2008; Walker, 2008; Green & Curnoe, 2009) has shown that males are more robust, and the sites of muscle attachment are larger than in females of the same population, however, population differences do exist. It is very important that anyone estimating sex has an understanding of the potential population history of the individual they are assessing, as well as the composition of the reference sample used when deriving the methods (Latham *et al*., 2018). Applying methods and discriminant functions to individuals that are different from the population used when deriving the functions have shown significant errors

(Walker, 2008; Bigoni *et al*., 2010). These inaccuracies could be a result of variations in sexual dimorphism, as there are some populations where both males and females are robust while other populations have both gracile males and females (Green & Curnoe, 2009). For these reasons, population specific standards need to be developed in order to achieve high accuracies (Steyn & İşcan, 1998; Patriquin *et al*., 2005; Green & Curnoe, 2009; Bidmos *et al*., 2010).

There are a number of studies that assess sex in a South African population using metric analyses (Rightmire, 1971; Steyn & İşcan, 1998; Franklin *et al*., 2005; Dayal *et al*., 2008; L'Abbé *et al*., 2013). Steyn and İşcan (1998) examined sex using 12 cranial and five mandibular measurements from 91 (44 males and 47 females) White South African individuals. Similar to both Kranioti and associates (2008) and Spradley and Jantz (2011), bizygomatic breadth was the most sexually dimoprhic. While the study found that maximum frontal breadth was the least sexually dimoprhic measurement. Four discriminant functions were derived which achieved cross-validated accuracies ranging from 81.1% to 85.7%. Dayal and colleagues (2008) assessed 120 crania and mandibles from South African Black individuals. They used 21 measurements (landmarks outlined in Martin and Knussman 1988) and ran stepwise and direct analysis. Aside from orbital height, all measurements for males were significantly greater than female measurements, indicating the presence of sexual dimorphism in South African Black skulls. Stepwise and direct analysis achieved average accuracies ranging from 78.3% to 85.0%. They ran cross-validation of the functions used in the stepwise and direct analyses and achieved accuracies ranging from 80.0% to 83.9%. The authors do suggest that these equations should only be used on a South African Black population as numerous previous studies have been shown that discriminant function equations are population specific.

2.2.2 Ancestry Estimation

To this day, social race remains part of people's lives wherein they characterise themselves into different groups, either socially or legally, as is the case of the population census (Stull *et al*., 2014). Anthropologists have stressed the importance of distinguishing between race and ancestry, where race is a sociocultural concept used by many to distinguish between people based on geography, culture and religion, and is usually characterized by traits such as skin colour, facial features and language (Relethford, 2009; Gillick, 2012; Stull *et al*., 2014; King, 2015). Ancestry on the other hand, is based on biological and genetic estimates (Sauer, 1992; Hefner, 2003;

Albanese & Saunders, 2006; Relethford, 2009; Christensen *et al*., 2014; King, 2015), and is usually defined geographically with terms such as African, Asian and European (Gillick, 2012; Christensen *et al*., 2014; King, 2015; Liebenberg *et al.*, 2015). Ancestry estimation is possible because people vary in specific ways due to evolutionary processes such as genetic drift, gene flow, natural selection, mutations and recombination (Stull *et al*., 2014; King, 2015). Ancestry is estimated by observing or measuring skeletal variation caused by those processes. However, ancestry not only has biological aspects, but also has social aspects that need to be taken into account (Christensen *et al*., 2014).

Establishing the ancestry of an unknown individual can help the police, as this basic information can help them refine the focus of their search and allow them to focus their attention on specific groups in the community which could help them to match the biological profile with information on the missing person. The ancestral groups are roughly aligned with the ancestry categories provided by the government, which tend to run in concordance with "race" labels used within societies; such as African (or Black) and European (or white) (Ousley *et al*., 2009; Gillick, 2012). This ensures that there is an understanding between investigators and the general public when forensic anthropologist report information to the police, who in turn relay the information to the communities involved (Albanese & Saunders, 2006; Gillick, 2012). Ancestry estimation generally comprises the visual and metric examination of morphological differences; including bone shape, suture shape, the absence or presence of a trait, the protrusion or prominence of a feature and bony feature morphology. These morphological differences are recognisable between different population groups because of factors such as geographical distances, language and cultural differences that limit group interaction and contribute to an increase in variability between groups (Gillick, 2012; Stull *et al*., 2014). Therefore, allowing biological anthropologists to quantify these morphological differences to estimate ancestry.

The cranium, especially the facial skeleton, is the most frequently studied skeletal element for ancestry because it displays the most obvious characteristics for different populations (Randolph-Quinney *et al*., 2009; Gillick, 2012; Murphy & Garvin, 2018). There are some mid-face nonmetric traits that practitioners have focused on due to their contribution to ancestry estimation. The "trait list" method was developed in order to apply these traits in the estimation of ancestry, providing a list of traits corresponding to and representative of certain population groups (Hughes *et al*., 2011; DiGangi & Hefner, 2012). In 1926, Earnest A. Hooton, a professor at Harvard

University, created the "Harvard List" which was a standardised form containing a trait list including both metric and nonmetric observations used by forensic anthropologists. In order to create this list, Hooton collected both metric and nonmetric data from several sources, including the work he did at the Pecos Pueblo archaeological site and from criminals in the United States (Hefner, 2007; Dirkmaat, 2012). The traits were scored by assessing the presence or absence of the trait and by looking at the expression of varying states for each trait. This analysis was based entirely on the experience of the observer when scoring the morphological traits and the quality of the descriptions used (Dirkmaat, 2012), which led to Hooton emphasising the need for standardised non-metric traits because of the subjectivity in their assessments.

Following on from the "Harvard List", Rhine (1990) supplied a list of 45 nonmetric cranial traits for differentiating among "American Caucasoid", "Southwestern Mongoloid" and "American Blacks" (as cited by İşcan & Steyn, 2013). The lists provided by Rhine (1990) is often cited in literature, however, anyone using it should be cautious as not only do they represent a small amount of variability due to the fact that his sample only contained material from the Southwestern United States, but Rhine (1990) assessed a relatively small sample size of 87 skulls which only represents a small portion of the variability in the population (American Whites = 53; American Blacks = 7; Hispanics = 15; American Indians = 12). Also, important to note when discussing Rhine's (1990) trait list, is that the author is not 100% clear on the methodology for using his trait lists and cautions that practitioners should test the lists on their population before applying them. In a study based on the association between morphological traits and different ancestry groups, Hefner (2009) discussed the fact that Rhine (1990) does not expand on his methodology or why he used that method. Hefner (2009) suggests that the study has not been carried out with appropriate scientific and legal considerations, and intra- and inter-observer repeatability had not been considered.

Walker (2008) has shown that detailed line drawings can help to consistently score non-metric traits for sex estimation, and so Hefner (2009) used a similar approach. Hefner (2003, 2007, 2009) collected morphological trait data from five populations, with a total sample of 906 individuals; including 220 American Blacks, 185 American Whites, 364 American Indians; 98 Asians and 39 Hispanics. Using that data, Hefner (2009) supplied illustrations showing the character states for 11 traits as well as written descriptions for each character and character state with frequency tables done by population. Hefner (2009) and Hefner and Ousley (2014) used

statistical analysis of ordinal scores allocated to the morphological cranial traits for ancestry estimation and reported accuracies ranging from 84% to 93%. While this method does use multivariate statistical analyses, the method still encompasses a degree of subjectivity as it uses the visual comparison of traits to standardised line drawings.

Research has been focused on the estimation of sex, stature and age-at-death, with less attention on the methods that estimate ancestry. It has been suggested that anthropologists pay less attention to ancestry estimation because studies assessing ancestry have resulted in a great amount of ambiguity. Ancestry is considered one of the most challenging components to assess due to the controversy regarding how previous research on "race" has been conducted and used both socially and politically (Komar & Buikstra, 2008; Dirkmaat, 2012; DiGangi & Hefner, 2012; Christensen *et al*., 2014). However, it is understood that there is a considerable amount of variation among individuals within populations and that some of the biological variation is distributed among individuals in different populations and between larger population groups. For this reason, ancestry is still one of the components of the biological profile in South African forensic cases (van Rooyen, 2010).

South Africa has a very diverse population representing many different cultures and a vast heritage. The morphological differences seen between South African populations are the result of decades of colonialism, migration and forced separation (L'Abbé *et al*., 2013; Stull *et al*., 2014). The 2019 South African mid-year population estimates statistical release uses the population categories South African African/Black (SAB), South African Coloured (SAC) and South African White (SAW). The SAB population constitutes 80.7%, the SAC population comprises 8.8% and the SAW population consists of 7.9% of the South African population (Statistics South Africa, 2019b). The SAB population represents population of mixed groups, including the Zulu, Ndebele, Xhosa and Venda to name a few. Although it has been shown that there are some cranial differences between these groups, the distinctions between the tribes are swiftly disappearing and SAB individuals are considered to be homogenous, therefore, they are studied as one group (Steyn *et al*., 2004; Franklin *et al*., 2006; Stull *et al.*, 2014). Descendants of European groups such as the British and German are categorised as SAW (L'Abbé *et al*., 2011; King, 2015). The term "Coloured" is a social term used to define individuals of mixed ancestral heritage, including those who descended from slaves brought to South Africa (de Wit *et al.*, 2010; Isaacs-Martin & Petrus, 2012).

Before colonisation, South African land was divided into areas occupied by several native groups. The name Khoesan is the term used to denote the Khoekhoe and San groups, two groups that are considered to have shared physical and linguistical characteristics (Thompson, 2001). The San were hunter-gatherers who made weapons from stone, wood and bone and who did not grow crops nor domesticate animals. The San did not build permanent settlements as they moved around a lot looking for game and plant foods (Nurse *et al*., 1985; Franklin *et al*., 2007a; King, 2015). The Khoekhoe were pastoralists who reared livestock (cattle and sheep). The Khoekhoe were originally based mostly in the northern parts of southern Africa, but then began moving south searching for new pastures. The Khoekhoe spread into the Cape, which resulted in conflict between the Khoekhoe and the San that originally inhabited the area; with one of the major sources of conflict being competition for game. The Khoekhoe were still heavily reliant on game for nutrients, even though they were herders, and on special ceremonial occasions the Khoekhoe slaughtered their cattle. Unfortunately, the Khoekhoe's sheep and cattle were making the problem even worse because they were depleting the resources that the indigenous game, such as wildebeest and zebra, were dependent on. Watching the number of game depleting lead the San to feeling justified in stealing the Khoekhoe's animals that had led to the displacement of the wild game. This ultimately led to continuous conflicts between the two that sometimes escalated into a full-scale war between the groups, which had many effects on both the Khoekhoe and San. The Khoekhoe formed larger and more organised groups to present as a stronger, united front against the San. Meanwhile the San did one of three things: some fled to mountain and desert areas, some formed groups who kept raiding the Khoekhoe for their cattle, while others made peace with the Khoekhoe and joined them as hunters, warriors, herders and servants. Those who chose the third option were gradually accepted as members of the Khoekhoe communities. To this day, the Khoe language is spoken by hunter-gatherers in parts of the Kalahari while the San are found in the dry parts of Angola, Botswana and South West Africa (Thompson, 2001; Franklin *et al*., 2007a; van Rooyen, 2010; King, 2015).

Another group that came into contact with the Khoesan were the Bantu-speaking people, whose community reared cattle, sheep and goats, and produced food. It is clear that the encounter between these two groups was sometimes peaceful while other times violent. Most of the Bantu-speaking groups survived through European colonialism and are still active today, maintaining their status as many distinct groups to create the SAB population. It is thought that Bantu-speaking

groups migrated from West Africa, at least 3000 years ago, and have formed independent tribes through diversification and include tribes such as Xhosa, Swazi, Zulu, Southern Sotho and Tswana (Franklin *et al*., 2005; Franklin *et al*., 2007a; King, 2015).

In 1487, the first Portuguese expedition rounded the Cape of Good Hope, travelling to India. Following this, the Portuguese government sent fleets travelling around the Cape of Good Hope to the Indian Ocean which led to the European trade with Southeast Asia being diverted from the Persian Gulf and the Red Sea to the oceanic route via the Cape. Nearing the end of the sixteenth century, the Dutch, English, French, and Scandinavian ships were also using the oceanic route via the Cape to Asia. Occasionally they stopped on the Cape peninsula for fresh water and to trade iron and copper goods with local Khoekhoe for cattle and sheep. This led to the Dutch East India Company establishing a refreshment/halfway station at the Cape of Good Hope (Table Bay) in 1652, which was to provide fruit, vegetables and meat to Dutch fleets passing through the Cape (Thompson, 2001).

They had not intended for the refreshment station to become anything more, however, it became an independent thriving society. There are three main processes which are considered to have contributed to the development of the colony (Thompson, 2001; van Rooyen, 2010). In the first process, the company gave land to some of its employees giving them the status of "free burghers" (farmers); who later formed a portion of the white South African population that became known as Boers or Afrikaners. These free burghers were tasked with the production of grain and vegetables, and to sell their produce to the company. Over the next few years, not only did the company release more men for the same purpose, but they brought several people from the Netherlands to settle at the Cape, including French Huguenots who fled to the Netherlands. Over time, the free burgher population grew rapidly, with the company records showing that by 1793 there were approximately 13830 free burghers (Thompson, 2001). Although, some free burghers developed a farming system on the growth of cereals and vine, and they raised livestock on their pastures, many free burgers gave up farming due to a lack of money and labour and became traders in a village on Table Bay, which became known as Cape Town, catering for visiting French, English, Scandinavian and Dutch fleets passing through the Cape (Guelke, 1988).

The second process involved the company importing many slaves, to both work in developing the colony, and for agricultural work. The company, including the community of free burghers, became dependent on the labour of slaves (Thompson, 2001). The slaves imported into

the Cape came from very different social, religious, and linguistic backgrounds compared to the slaves imported into the Americas. The first slaves imported to South Africa were originally from present-day Angola by the Portuguese. According to the records of the slave trade, the slaves brought into the Cape were from the east coast of Africa (26.4%), Indian subcontinent (25.9%), Madagascar (25.1%) and Indonesia (22.7%) (Nurse *et al.*, 1985; Thompson, 2001). Between 1652 and the end of the slave trade in the early 1800's, approximately 60000 slaves were imported to the Cape (van Rooyen, 2010).

The third process involved the settlement developing slowly until it expanded onto land owned by the Khoekhoe, who either had to withdraw from their land or stay and become servants for the Dutch. Before the company developed the refreshment station at the Cape, the Khoekhoe who lived in the Cape peninsula had grown accustomed to the occasional European visitor. The Khoekhoe realised the benefit of trading goods from the Europeans, i.e. copper, iron and brass, and so the Khoekhoe became experienced in trading these goods for sheep and cattle. After 1652, the relationship between the Khoekhoe and the Dutch was friendly, where they traded Western goods for sheep and cattle. However, with the expansion of the free burghers onto their land, tensions started to rise leading to warfare in 1673 which continued intermittently until 1677. Unfortunately, the southwestern Cape Khoekhoe communities deteriorated rapidly having lost much of their most valued possessions, their livestock.

In the early 1700's, European settlers from the Dutch Cape Colony had gained control of most of the land between the Cape peninsula and the mountain escarpment, where they took part in livestock farming, with some hunting. These settlers became known as trekboers, who were considered to be semi-migrant farmers. Unfortunately, due to the disintegration of the Khoekhoe communities in the Cape peninsula area and being wrecked by smallpox, the Khoekhoe were unable to stop the advancement of the trekboers onto their land. The trekboers had grown accustomed to using slaves and the indigenous people for labour and so as they kept advancing progressively onto the Khoekhoe's land while, enslaving many of the Khoekhoe people as they had a particular set of skills required by the trekboers.

Outside the Cape, lived the Khoekhoe and San who had been chased off their land, African and Asian slaves who had escaped, men and women of mixed ancestral heritage, and free burghers who had committed crimes. This group of people were also linked to the Cape by trade, farming livestock as well as hunting much like the trekboers. They bartered for guns, ammunition and other

imported products, in exchange for their cattle, sheep, and ivory. Nearing the end of the eighteenth century, some of these people were organising themselves into Xhosa chiefdoms. Unfortunately, not only were the Xhosa chiefs trying to stop the intrusion of the colonists, but they were also warring with one another. The Khoekhoe weren't sure if they wanted to abandon their roles as labourers for the trekboers or if they wanted to join the Xhosa, who had already welcomed numerous Khoekhoe into their tribes. The Khoekhoe were unable to make a decision even after two lots of warfare in the frontier zone in 1779 and 1793, where property was destroyed, and many people were killed. Due to this, the relationship between the trekboers and colonial government was tense because the trekboers didn't think the government was offering adequate support for them to maintain control of their own land.

An especially critical aspect of the colony's social structure was the White colonists' complete dependence on the labour of both slaves and the indigenous people. It was seen in Europe that ethnic chauvinism was already part of the community meaning that at the Cape, the colonists (that were subjected to a government practising in the slave trade and slavery) were already used to a more privileged life (Thompson, 2001). This meant that they distinguished themselves from their slaves by physical, cultural, economic, and legal criteria. Not only did they see themselves as being very different from the indigenous people of the Cape, they were also growing apart from the north-western European communities where the economic and social conditions were greatly different. Even though the White colonists were very diverse themselves (Guelke, 1988), they still thought of themselves as a distinct community despite the fact that some of the individual's behaved quite unconventionally in that they had intimate relations with women that were considered to be part of the "subordinated" classes. For example, Coenraad de Buys lived and had several children with a Khoikhoi woman, Maria van der Horst, a descendent of a slave. After which he also married another woman, Elizabeth, and had children with her as well (www.geni.com/people/Coenraad-de-Buys-c1d4).

Great Britain dominated the sea during the French Revolution, so to prevent the Cape peninsula from being taken over by the French, they occupied the Cape peninsula. The British were able to easily get the Dutch officials to surrender in 1795, although the Dutch regained the Cape in 1803 under the Treaty of Amiens, however the British once again reclaimed control in 1806 (Thompson, 2001). The British were more interested in South Africa because the Dutch had previously occupied the Cape, rather than capturing South Africa for any reason beyond the

peninsula, as the peninsula was on route to Asia, which is where the English East India Company traded profitably. South Africa was of very little significance to the British economy until the late 1860s when it was revealed just how vast South Africa's mineral wealth was. Because the British thought that South Africa had such little significance, only a few individuals from the British Isles settled in southern Africa before 1870 (Thompson, 2001).

During the British's first conquest of the Cape, the Xhosa farming people had expanded beyond the official boundary line at the Fish River. The trekboers, who were used to taking the law into their own hands, started rebellions; while there were groups of Khoekhoe who were trying to maintain some of their freedom. There were also two (radical White) missionaries of the London Missionary Society (an interdenominational Protestant organization) who supported the Khoekhoe and the Xhosa (Thompson, 2001). The British originally saw themselves as temporary custodians of the Cape, and therefore had no intention of interfering, however, after 1806 when they regained the Cape, they made many changes. The first being that they tried to establish control in the eastern frontier zone. In 1811 and 1812, British troops were helped by colonial commandos and their Khoekhoe units to mercilessly force the Xhosa from their land through to the Fish River, whilst burning down villages, burning crops and running off with thousands of their cattle. They reported that they had done enough damage to leave the Xhosa traumatised and in constant fear of the Colony, after which, the government forbade colonists from making further contact with the Xhosa across the Fish River (Nurse *et al.*, 1985; Thompson, 2001). Not only were the British troops forcing the Xhosa south from their own land, but there was also fighting amongst two other indigenous tribes (Zulu and Ndebele) that were causing trouble for the Xhosa. The fighting between the tribes led to destruction and depopulation which allowed the white settlers to move onto the land previously owed by the Xhosa and the other indigenous tribes (Thompson, 2001). In 1820, Parliament voted to use some of their budget to transport settlers from the British Isles to the Cape and have them set up as agricultural farmers. The Colonial Office in London organised that nearly four thousand men, women, and children from England, Wales, Scotland and Ireland be sent to the Cape. They arrived in the Cape that year and were accompanied by another group of about one thousand people who came at their own expense. Many of these settlers were from the lower to middle classes and had absolutely no farming experience but were urban artisans, meaning they were workers in a skilled trade, usually a trade that involves making things by hand. This group of settlers, known as the 1820 settlers, did not fare as well as the government had hoped.

Within a few years of settling in the Cape, at least half of them had abandoned their land as the soil was not suited to agricultural farming. Many rather became merchants and artisans either at the military post at Grahamstown, in the settlement in Port Elizabeth or other colonial villages. While there were others who started trading with the African farmers beyond the frontier which only became legal in 1828. Those that chose to remain on the land eventually thrived because they were able to acquire more land and started providing wool for the market. What these settlers did not know was that the government had placed them on land belonging to the Africans, specifically Xhosa communities who had lost their land in the clearances in 1811-12 and 1817-19 (Thompson, 2001). The 1820 current and future settlers from Britain further complicated the already complex colonial society with their different language, traditions and religious affiliations, making them very distinct from the earlier settlers whom they did not assimilate with. Well into the twentieth century there was very little social mixing and intermarriages. The British settlers started referring to the previous settlers as Boers, which means farmers and has derogatory overtones. While during the Dutch colonisation, Dutch-speakers had used the word Afrikaner, which is the term that gained popularity and became the label that was recognised universally in the twentieth century. The British settlers, like the Afrikaners, were living in an exposed frontier zone that the African people considered rightfully theirs. This led to intermittent warfare (1834, 1846, and 1850) where the British settlers defended and expanded their territory (Thompson, 2001).

The British settlers, like the Afrikaners, also had an interest in obtaining and owning indigenous labour (Thompson, 2001). In 1807, the British government decided to ban all slave trade to British colonies, which included the colony at the Cape, meaning that slaves could no longer be brought into the Cape, however, this did not mean an end to slavery. In 1833, the British parliament decided to emancipate the slaves in the British Empire. After a transitional period following the passing of the law to emancipate slaves, they became legally free in 1838. These former slaves, including the Khoesan servants, became known as “Coloured” people. This group included descendants from relationships between the indigenous African and European people, and a small group of Muslim people who came from the Indonesian archipelago and were later known as the “Cape Malays”. The term ‘Coloured’ has since expanded to include all individuals of mixed ancestry and autonomous ethnic groups including those with a genetic combination of native Africans, European colonists, and individuals from Asia who were imported as labourers by the British (Nurse *et al*., 1985; de Wit *et al.*, 2010; Quintana-murci *et al*., 2010; Isaacs-Martin

& Petrus, 2012; King, 2015). Therefore, the SAC population has a large amount of genetic diversity that could affect ancestry estimation.

Racial tension reached climax during Apartheid when a series of laws, such as the Population Registration Act (1950), the Prohibition of Mixed Marriages Act (1949) and the Immorality Act (1950), were passed that imposed institutionalised segregation restricting the actions of the SAB and SAC populations. In addition, any interaction between SAB, SAC and SAW individuals that was not business related was forbidden. These laws were put into place from 1948 by the National Political Party and only eradicated in 1994 (L'Abbé *et al*., 2013; King, 2015). The years of segregation led to a reduction in intermixing between SAB, SAC and SAW individuals, leading to a decrease in the variation within groups and an increase in the variation between the groups (Stull *et al*., 2014; King, 2015). This segregation contributed towards the distinct differences seen between the three groups in South Africa, the SAB, SAC and SAW populations.

It has been shown that many of the methods used by forensic anthropologists are population specific. Therefore, it is imperative to develop population-specific standards for the South African populations in order to achieve higher accuracies when identifying individuals from skeletal material (Bidmos *et al*., 2010). İşcan and Steyn (1999) derived discriminant function formulae in order to estimate ancestry from 13 cranial and four mandibular dimensions for SAB and SAW individuals. The use of cranial discriminant function equations achieved higher average accuracies (98%) compared to the mandibular equations (74% for males and 87% for females). İşcan and Steyn (1999) also compared their data with standards developed for North American individuals to test the applicability of North American equations when using data from South African populations. Accuracies achieved when using formulae developed for North American individuals were considerably lower, demonstrating that osteometric standards are population specific. They concluded that standards developed for South African samples will result in higher accuracies when assessing South African individuals.

L'Abbé and colleagues (2011) assessed whether non-metric traits could be used effectively to differentiate between SAB, SAC and SAW individuals. They scored 13 non-metric cranial traits, which were chosen for two main reasons. First, nine of these traits were examined by Hefner (2009) and found to have a positive correlation with ancestry of North American groups and secondly, because of their inclusion in forensic reports in South Africa. They found that for three

of the traits; nasal bone contour, anterior nasal spine, and the inferior nasal margin; SAB and SAC individuals were distributed into the African and Asian groups while SAW individuals were distributed into the European groups. The distribution for the other traits was shown to be highly variable in a sample of South African individuals. They stated that the use of this traditional non-metric method was not reliable when assessing ancestry in South African forensic cases.

2.3 MODERN MORPHOMETRICS

2.3.1 Geometric Morphometrics for Sex Estimation

While there have been some studies using GM methods that have assessed sexual dimorphism using postcranial elements such as the scapula (Scholtz *et al*., 2010), humerus (Kranioti *et al*., 2009; Vance & Steyn, 2013) and femur (Cavaignac *et al*., 2016); most GM studies investigate the cranium (Pretorius *et al*., 2006; Franklin *et al*., 2007a; Green & Curnoe, 2009; Bigoni *et al*., 2010), mandible (Pretorius *et al*., 2006; Franklin *et al*., 2007b; Franklin *et al*., 2008) and/or pelvis (Steyn *et al*., 2004; Pretorius *et al*., 2006). Kimmerle and colleagues (2008) used GM to assess size and shape dimorphism from the craniofacial region of American Black and White individuals. They achieved better results when comparing their results with other studies. They achieved accuracies as high as 93%, while traditional metrics achieved accuracies ranging between 83% and 88%.

Green and Curnoe (2009) employed GM in the analysis of sex of Southeast Asian crania using 35 cranial landmarks. Using shape variables only they achieved relatively low accuracies ranging from 55.6% to 77.1%, whereas discriminant function analysis (DFA) using both shape and size achieved higher accuracies ranging from 78.5% to 86.8%. Gonzalez and colleagues (2011) analysed 125 adult skulls from the Museu Anthropologico de Coimbra, University of Coimbra, Coimbra, Portugal. This was evaluated by taking photographs of the crania and assessing the digital images, where 12 landmarks and 25 semilandmarks were digitised by means of tpsDIG software. Similar to Green and Curnoe (2009), they found that when using shape variables alone, the correct classification was relatively low with accuracies ranging from 60.75% to 64.15%. Whereas when they assessed both shape and size variables, they achieved higher correct classifications ranging between 76.9 and 78.48%. They concluded that this gave a good indication that the traits they analysed showed noticeable sex differences relating to size and robusticity. Bigoni *et al.* (2010) investigated numerous craniofacial landmarks and semi-landmarks (points

that are placed along outlines or surfaces that provide information about curvature (Zelditch *et al*., 2004)) in the estimation of sex of Central European individuals. They found that there were no significant differences in the cranium, but they found accuracies of up to 100% when using the midsagittal curve, orbits, nasal aperture and palate. Kimmerle and colleagues (2008) and Bigoni and associates (2010) concluded that GM shows precise patterns of sexual dimorphism that are not necessarily discernible by the traditional methods.

Chovalopoulou and colleagues (2016) and Bertsatos and associates (2018) assessed 176 crania of Greek individuals that form part of the Athens Collection housed in the Department of Animal and Human Physiology, Faculty of Biology, National and Kapodistrian University of Athens. Chovalopoulou and colleagues (2016) digitised 31 ecto-cranial landmarks and 30 semi-landmarks, and assessed shape, size and form variables by running logistic regression and discriminant function analyses. Shape differences achieved better classifications of 79% in the vault compared to the 68.8% achieved for the midsagittal curve of the neurocranium. Size achieved better accuracies (82%) for the cranial vault, while the midsagittal curve of the vault achieved a poorer accuracy (68.1%). Unsurprisingly, the classification results improved when combining both shape and size, to 89.2% for the cranial vault and 79.4% for the midsagittal curve of the vault. Bertsatos and associates (2018) digitised 80 landmarks (12 midline, 34 bilateral) and ran linear discriminant analysis. Looking at the univariate results they found 60 distances which had correct classifications higher than 92% and 61 angles with correct classifications higher than 75%. These univariate results are reasonably high and can be valuable in cases where forensic anthropologists are examining fragmentary remains. They then focused on the most dimorphic variables for multivariate analysis. These variables included 13 distances that achieved accuracies higher than 85% and seven angles that achieved accuracies higher than 75%. Using these variables for multivariate analysis they were able to achieve accuracies ranging from 85.2% to 95.2%.

Oettlé and colleagues (2005) assessed whether or not there were significant sex differences between SAB male and female mandibular rami. Initially three landmarks were marked on the mandibular ramus and then photographs were taken. A further eight landmarks were placed on the image of the mandibular ramus which was then analysed using the computer program tpsDIG. They found that females were more variable in shape compared to males who clustered around the center point of the graph. However, there was considerable overlap between the sexes and they concluded that although there was some variance, the extent of these differences was not

significant in distinguishing between male and female. Pretorius and associates (2006) analysed the greater sciatic notch, the ramus flexure of the mandible and orbit shape using thin plate splines, canonical variants and relative warp plots to estimate sex from SAB crania. They found that the shape of the greater sciatic notch was the most dimorphic, achieving accuracies of 93.1% for males and 87.09% for females. It was seen that orbit shape performed better than ramus flexure, with the orbit shape achieving accuracies of 73.3% and 80.0% for males and females respectively, while the ramus flexure achieved accuracies of 69.6% and 67.8% for males and females respectively. This revealed that the pelvis and orbit are more sexually dimorphic than the ramus flexure of the mandible in a SAB population.

Franklin and associates (2005) assessed sexual dimoprhism of 332 (182 males and 150 females) crania of indigenous South African individuals using GM. This study analysed cranial dimorphism by collecting 3D landmark coordinates using a digitiser, calculating eight measurements from those landmarks and deriving discriminant functions. It was found that bizygomatic breadth was the most sexually dimorphic measurement, while cranial breadth was the least sexually dimorphic measurement. From these measurements, eight discriminant functions were derived, and cross-validated accuracies ranging from 77% to 80% were achieved. A function containing maximum cranial length, basi-bregmatic height, bizygomatic breadth, mastoid length and maxillo-alveolar breadth achieved the highest accuracies of 76% for males and 84% for females with an overall accuracy of 80%. While a function containing maximum cranial length and bizygomatic breadth achived the lowest accuracies of 77% for males and 80% for females with an overall accuracy of 78%.

Small and associates (2018) derived novel ILDs of the cranium from 3D GM methods. 45 landmarks were digitised from 227 South African White crania from the Raymond A. Dart Collection, University of Witwatersrand South Africa. From these 45 landmarks, 990 ILDs were derived for six regions of the cranium (basicranium, basipalate, zygomatic, orbits, nasomaxilla and global cranium) and then the ILDs were subjected to direct and stepwise discriminant function analyses. The discriminant equations derived for each region achieved accuracies ranging between 71.8% and 88.2%. Small and colleagues (2018) found that a large portion of the derived ILDs are novel, which they believe demonstrates how effective GM methods are and emphasize the need to re-evaluate previous forensic methods as they may not assess biological differences as effectively as GM methods.

2.3.2 Geometric Morphometrics for Ancestry Estimation

Lately, studies have been focusing on the understanding of complex groups biological variation (i.e., having two or more parental groups) to estimate ancestry for these groups with better accuracies (Ross *et al*., 2004; Spradley *et al*., 2008; Spradley, 2013; Spradley & Jantz, 2016). This is possible because technical advances have allowed researchers to collect many different types of skeletal data, including nonstandard interlandmark distances (ILDs) which can be used to capture the variation better between these complex groups (Spradley & Jantz, 2016; Spradley & Stull, 2018). ILDs are the linear distances between landmarks, with non-standard ILDs being distances that are not traditionally taken when measuring skeletal elements.

Xing and colleagues (2012) used a novel standardization technique to assess the variation in the orbit shape of contemporary Asian (n=40), African (n=39) and European (n=40) populations. The Asian samples were collected from the Institute of Vertebrate Paleontology and Paleoanthropology, while the African and European samples were collected from the Raymond A. Dart Collection of Human Skeletons. Each specimen was photographed after obtaining a standardised orbital plane. The authors found that there was regional variation in the shape of the orbit. The orbital contour for the Asian population was found to generally be taller and rounder and the inferior contour was found to be symmetrical, while the contour for the European population tended to be square and more inclined and the contour for the African population was found to be shorter. They also found that the orbit shape overlapped for some of the individuals from the three regions. It was noted that similarities were smaller between the Asian and European samples compared to the similarities between African and Asian samples, and African and European samples. The study found that the most variable area of the orbit was focused around the internal and lateral aspects of the upper margin, and the internal aspects of the lower margins. Using the superior contour of the orbit they achieved cross-validated accuracies of 41% to 60%, with African individuals achieving the lowest accuracy. When using the inferior contour of the orbit they achieved cross-validated accuracies of 72.5% to 80% with European individuals achieving the lowest accuracy.

Spradley and Jantz (2016) tested the discriminant ability of both standard and nonstandard cranial measurements and data derived using GM analysis. They assessed American Black, White and Hispanic males and females. The authors digitised 31 cranial landmarks using a Microscribe

digitiser and from these landmarks generated 465 nonstandard ILDs. The cranial landmarks also generated a subset of 20 cranial measurements from the Data Collection Procedures (DCP) for Forensic Skeletal Material (Moore-Jansen *et al*., 1994) which was referred to as DCP ILDs in the manuscript. The GM-derived data, nonstandard ILDs and the DCP ILDs were analysed using DFA to see which type of data would be able to estimate ancestry with higher accuracy. The DFAs included shape variables only and additional analyses which included both shape and size. The nonstandard ILDs with size achieved the highest classification rates for males (88.2%), while the nonstandard ILDs using shape variables generated the highest classification rates for females (90.5%). It was shown that using nonstandard ILDs instead of GM-derived data or common measurements resulted in better ancestry estimates in the American Black, White and Hispanic populations.

Murphy and Garvin (2017) used a morphometric outline analysis of cranial shape for the estimation of ancestry. They assessed 198 American Black and White crania by taking 3D surface scans. From these scans, 2D images showing the superior, posterior and left lateral aspects of the crania were obtained. They chose these views because they are the most likely views to encompass all the traits which are traditionally used for ancestry estimation. These views were quantified using elliptical Fourier analysis. The elliptical Fourier analysis coefficients were subject to principal component analysis (PCA) which allows for shape variation to be quantified and finally DFA was run. When looking at ancestry, American Blacks were shown to have a more dolichocephalic vault and greater maxillary prognathism compared to American Whites. They achieved overall ancestry classification rates ranging between 70.6% and 92.4%, with the lateral view achieving the highest ancestry estimation rates for both Black and White classifications, with 92.4 and 92.5% respectively. Not only have they shown the potential this method has in possibly improving ancestry estimation, but they have also shown that shape analysis using lateral, superior, and posterior cranial outlines supported the ancestral traits traditionally used.

Franklin and colleagues (2007a) reported on geometric data examining cranial variation in 12 self-identified subpopulations of SAB. In total, they examined 298 male Bantu-speaking individuals, and a small series of Khoesan. PCA was used to explore the relationships between the populations. It was found that although Southern African Bantu populations are closely related, the individual tribes exhibit group-specific characteristics and can be separated from one another with varying degrees of success (46% to 87%). The Khoesan demonstrate characteristics that

differentiate them from Bantu populations with very high degrees of accuracy (Khoekhoe 94% and San 100%). Stull and colleagues (2014) examined 377 crania of SAB (n=158), SAC (n=107) and SAW (n=112) individuals using craniometrics and GM. The authors chose 44 cranial landmarks representing the cranium (with focus on the mid-face) because they previously achieved high accuracies when distinguishing between groups in South Africa. The craniometric data and landmark data were subjected to linear discriminant analysis (LDA). The LDA for the craniometric data achieved 84% accuracy while the LDA on the landmark data achieved the highest accuracy of 89%, emphasising that the relative location of cranial landmarks was better able to differentiate between South African populations than standard linear distances, as the shape data were better at demonstrating the relationships between populations that reside in the same geographic area in South Africa. They also showed that although there are high levels of admixture, SAC are a homogenous group, and by using craniometrics and GM they were able to correctly identify group differences with high accuracy (89%).

2.3.3 Computer Software Programs

Computer software programs using craniometrics in the estimation of sex and ancestry have become more popular with forensic anthropologists, as one of the benefits is that some of the computer programs have large reference databases; and some of these computer programs have been shown to provide precise statistical data, which is of great importance for forensic anthropologists (Elliott & Collard, 2009; Guyomarc'h & Bruzek, 2011; L'Abbé *et al*., 2013; Urbanová *et al*., 2014; King, 2015; Dudzik & Jantz, 2016; Manthey *et al*., 2018; Bertsatos *et al*., 2019). Examples of these programs are FORDISC, 3D-ID, CRANID and COLIPR. The program FORDISC assesses linear measurements taken from crania and postcrania, and the program 3D-ID employs GM applied to digitised cranial landmarks. Both programs assess sex and ancestry by comparing data from an unknown individual with that of an existing database of known individuals representative of many population groups. Both of these programs were developed in the United States and are traditionally used in American forensic cases, however, there is a pressing need to expand the application of these programs to other countries for international forensic use. FORDISC is more widely used in forensic anthropology due to its use of traditional linear measurements and large comparative database, however, FORDISC is not freely available. 3D-ID is a freely available program with a smaller database of reference samples, but the use of GM

methods provides the ability to assess the shape of an individual and not just linear measurements (King, 2015). 3D-ID has the potential to accurately estimate ancestry better than FORDISC because of its application of GM methods.

2.3.3.1 FORDISC

The program FORDISC was developed in 1990 by Richard Jantz and Steven Ousley and is based on the Forensic Anthropology Databank (FDB) (Jantz & Ousley, 2005). FORDISC can be used for sex estimation and stature calculations from both cranial and post-cranial elements, focusing mainly on ancestry estimation using craniometrics. The FDB has a larger number of reference populations of nearly 16000 documented individuals, with the number of reference populations for cranial data exceeding those for postcranial elements. Cranial data in the FDB comes from Black and White American, Native American, Hispanic, Japanese, Chinese, Guatemalan and Vietnamese populations (see table 2.1). While postcranial data comes from Black and White Americans and Hispanic populations. Of these groups, only White and Black Americans, Native Americans Japanese and Hispanic data represent both males and females; the other population groups only have male data. FORDISC also contains additional cranial data from Howells database (HDB), which has worldwide archaeological samples; as well as datasets from the Terry and Hamann-Todd collection of Black and White Americans from the 19th and 20th century (Komar & Buikstra, 2008; Langley & Tersigni-Tarrant, 2017). The South African database (SADB) was established by digitising 107 standard craniometric landmarks which were converted into 37 standard linear measures. The data were collected from crania from the Pretoria Bone Collection at the University of Pretoria and the Raymond A. Dart Collection of Human Skeletons at the University of Witwatersrand (L'Abbé *et al*., 2013). The database contains 141 Black, 177 Coloured and 109 White South Africans (see table 2.1) (Ms. G. Krüger, personal communication, 1 July 2019).

The program inputs metric data (up to 34 cranial measurements and 39 postcranial measurements) and uses LDA to classify the unknown individual into a reference population group represented by a database of known reference samples (Jantz & Ousley, 2005; İşcan & Steyn, 2013; Spradley & Jantz, 2016). Depending on the number of measurements available for a particular specimen, FORDISC will customise each formula so that if certain skeletal elements are incomplete a function can still be calculated (İşcan & Steyn, 2013). The output specifies the group

the specimen is classified into, Mahalanobis squared distances, cross-validated classification accuracies as well as posterior, typicality and ranked probabilities; allowing the observer to come to an informed conclusion (Figure 2.1 and 2.2) (Jantz & Ousley, 2005; İşcan & Steyn, 2013; Langley & Tersigni-Tarrant, 2017).

Table 2.1: Breakdown of population reference samples for cranial data in FORDISC

Population	Males	Females	Total
Forensic Anthropology Databank			
American Blacks	224	137	361
American Indians	59	32	91
American Whites	737	454	1191
Chinese	80	-	80
Guatemalan	83	-	83
Hispanics	281	74	355
Japanese	84	58	142
Vietnamese	51	-	51
South African Database			
South African Blacks	90	51	141
South African Coloureds	115	62	177
South African Whites	62	47	109

The smallest Mahalanobis squared distance gives an indication of the group membership. Posterior probability is a measure of the group membership with the assumption that the unknown specimen is represented in the reference sample and selected for in the analysis. It is the probability that the unknown specimen belongs to a particular group based on the distance from the unknown specimen to all groups in the reference sample. Posterior probability ranges from 0 to 1, the higher the posterior probability, the more likely it is that the specimen belongs to the group it has been allocated to. The typicality probability gives a measure of whether the unknown specimen could belong to any of the reference populations used in the analysis. If the value is above 0.05 it can be accepted that the specimen belongs to the group it was assigned, however, if the value is less than 0.05, it should be rejected as it gives an indication that the specimen is not associated with the

allocated group. The ranked typicality gives an indication of where the specimen would lie relative to all the other specimens in the reference sample (Jantz & Ousley, 2005; İşcan & Steyn, 2013).

```
FORDISC 3.1.316 Analysis of Current Case
Using SA_CRAN_20160725.adt

DFA results using 7 Forward Wilks selected (min: 1 max: 7, out of 12) measurements:
 ZYB   DKB   BBH   FOL   OCC   EKB   UFBR
 ----------------------------------------------------------------
Measurement Checks and Group Means
                               Group Means
                          BF     BM     CF     CM     WF     WM
 Current Case    Chk      51     79     39     57     24     46
 ----------------------------------------------------------------
  ZYB      115     -    120.5  128.9  119.9  127.6  120.0  128.7
  DKB       18     -     22.4   24.6   22.1   22.7   19.2   20.3
  BBH      126          125.4  133.4  127.0  131.9  130.7  137.1
  FOL       34     -     35.6   37.7   36.5   38.4   37.6   39.4
  OCC       91     -     94.3   96.3   94.8   95.9   98.3   99.8
  EKB       90    --     95.7  100.1   96.3   98.5   93.8   98.0
 UFBR       95    --    102.0  107.0  102.4  105.4   99.8  104.4
 ----------------------------------------------------------------
 +/- measurement deviates higher/lower than all group means; ++/-- deviates 1 to 2 STDEVs
 +++/--- deviates two to three STDEVs; ++++/---- deviates at least 3 STDEVs
 ----------------------------------------------------------------
 Outliers in the reference groups: 2

 Natural Log of VCVM Determinant =  16.4375
 ------------------------------------------------------------------------------
Classification Table

-------------------------------------------------------------------------------
 From     Total            Into Group (counts)
 Group    Number     BF     BM     CF     CM     WF     WM    Correct
 ------------------------------------------------------------------------------
     BF     51       22      4     10      8      6      1     43.1 %
     BM     79        6     50      2     13      4      4     63.3 %
     CF     39       14      1     13      5      6      0     33.3 %
     CM     57        5     18      7     14      3     10     24.6 %
     WF     24        3      2      1      0     15      3     62.5 %
     WM     46        0      3      0      4      6     33     71.7 %
 ------------------------------------------------------------------------------
 Total Correct:  147 out of 296 (49.7 %) *** CROSS-VALIDATED ***

 ------------------------------------------------------------------------------------------
Multigroup Classification of Current Case

 ------------------------------------------------------------------------------------------
   Group      Classified      Distance          Probabilities
                 into           from     Posterior   Typ F   Typ Chi   Typ R
 ------------------------------------------------------------------------------------------

     WF         **WF**           3.2        0.415    0.884    0.866    0.920 (2/25)
     CF                          4.3        0.237    0.764    0.741    0.700 (12/40)
     BF                          4.5        0.220    0.744    0.724    0.731 (14/52)
     CM                          6.8        0.069    0.481    0.451    0.603 (23/58)
     WM                          7.9        0.040    0.377    0.343    0.213 (37/47)
     BM                          9.3        0.019    0.255    0.229    0.225 (62/80)
 ------------------------------------------------------------------------------------------
  Current Case is closest to WFs
 ------------------------------------------------------------------------------------------
```

Figure 2.1: An example of the output text generated when using FORDISC

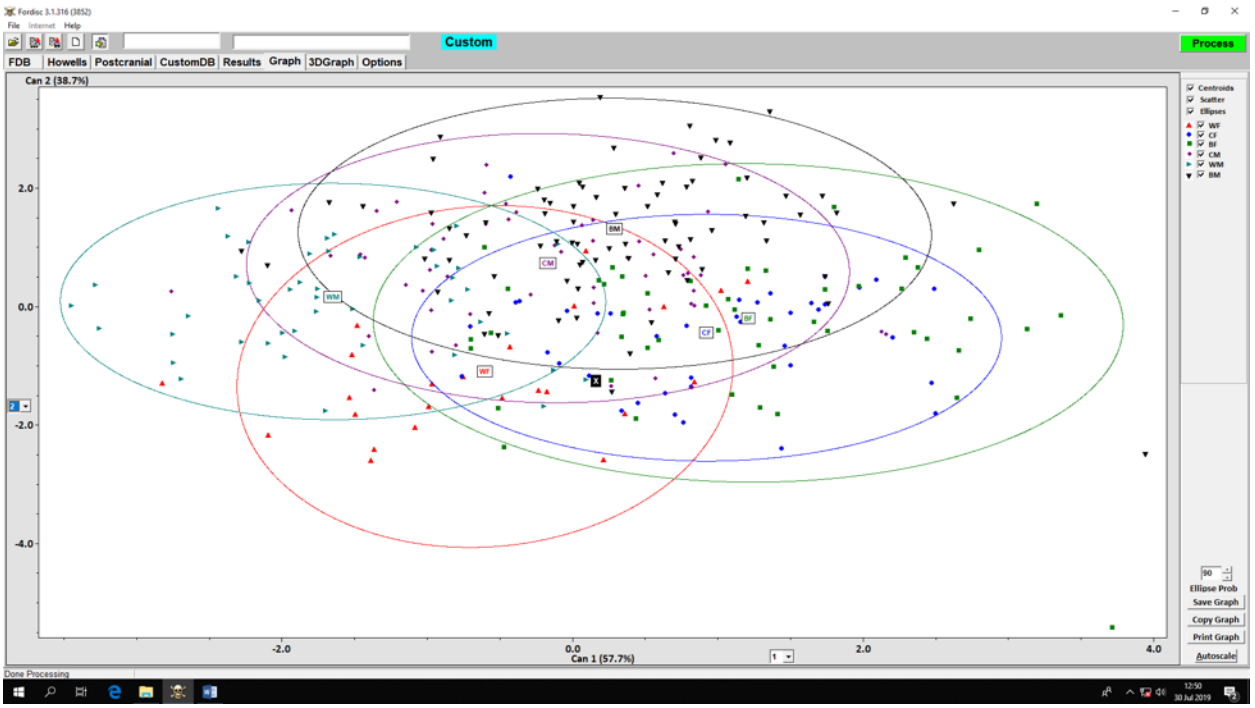


Figure 2.2: An example of the output graph generated when using FORDISC

FORDISC is widely used by forensic anthropologists for the estimation of sex and ancestry, however, the limitations of the programme need to be understood. Anyone using FORDISC needs to understand that the program can misclassify an individual into the incorrect reference population. This is because when using DFA, all groups in the database will overlap and the programme is forced to allocate the specimen into a group. DFA gives a sectioning point and, depending on where the discriminant function score of the unknown specimen lies relative to the sectioning point, an assignment will be given. In the groups that overlap there is a chance that the specimen would lie on the incorrect side of the overlap causing the specimen to be assigned incorrectly. This is also a problem when the specimen does not belong to any of the reference groups in the database as the programme will still allocate the individual into a population group, and so they introduced a measure of typicality to give the user an indication that the unknown individual may not belong to any of the groups represented in FORDISC. Therefore, it is important that the observer must interpret the data they receive and conclude whether the probabilities either support or oppose the classification of the specimen into one of the reference groups. This tends to be more of a problem for ancestry estimation compared to sex estimation (Komar & Buikstra, 2008; Langley & Tersigni-Tarrant, 2017).

Elliott and Collard (2009) assessed a test dataset of 56 variables from ten male and ten female crania from the following populations: Berg (Europe), Hokkaido Japanese (Asia), Santa Cruz (Americas), Tasmanians (Australia and Pacific) and Zulus (Africa). The values for the test dataset come from Howells' (1996) global craniometric dataset, which is one of the reference databases that FORDISC uses. Ten different analyses were run, with some including all 56 variables while others had a reduced number of variables; some of the analyses specified sex while others did not; and in some cases, all populations were included while the source population was excluded from some of the analyses. The results of these analyses showed that using FORDISC can be problematic for the estimation of ancestry. The program was able to correctly classify more than 70% of the specimens in some of the analyses, however, the analyses in these cases used all 56 variables that were available, and the source population was included in the reference sample in these analyses. Whereas, all the other analyses achieved correct classifications of less than 40%. Elliott and Collard (2009) suggest that the use of FORDISC is limited, in that it is useful when the unknown specimen (individual) is complete and the individual belongs to one of the populations that is represented in FORDISC's reference dataset.

Guyomarc'h and Bruzek (2011) used two sub-samples from French and Thai skeletal collections to test the reliability of FORDISC 3.0 to estimate sex. The data for the French skeletal collection were collected by digitising the specimens and then running the landmarks through Morpheus to generate ILDs. While the data from the Thai skeletal collection was collected using digital callipers. They used all of the available groups with the assumption that the unknown individual belongs to one of the groups in the FDB. Select groups within the FDB were used where the French sample was compared with the White groups and the Thai sample compared to the Asiatic-related groups (Chinese, Vietnamese, Japanese and American Indians). The first part of analysis assumed that the individual belonged to one of the groups available in the FDB. The authors determined that the French samples classified more reliably than the Thai sample, where seven out of the 12 measurements were significantly different between French males and females compared to the three out of 12 measurements in the Thai sample. The second part of analysis assessed the sex of the samples while comparing the samples to specific groups from the FDB. Sex classification accuracies ranging between 52.2% and 77.8% were achieved when using all available groups in the FDB, whereas accuracies of 89.3% for the French sample and 60.7% for the Thai sample were achieved when doing group-specific analysis. Similar to Elliott and Collard

(2009), it was concluded that FORDISC is probably more applicable for forensic cases in the United States compared to Europe, as cases in the United States are more comparable with the individuals in the FDB.

FORDISC 3.0 has recently been used on a South African sample with mixed results. L'Abbé and colleagues (2013) have created a custom South African database (SADB) and used the custom database as well as the FDB and HDB to assess 86 SAB and 101 SAW crania. Importantly, the authors only looked at SAB and SAW crania, they did not assess how SAC crania would classify using the SADB in FORDISC. When using the FDB, SAB males where classified correctly more often than SAW males (85.5% and 75.4% respectively), while SAW females classified more correctly than SAB females (72.7% and 48.4% respectively). Using the SADB, they were able to better classify SAB females compared to SAW females (74.2% and 72.7% respectively), and better classify SAW males compared to SAB males (71.9% and 67.3% respectively). Although these accuracies are acceptable, some are below the recommended 75% for use in forensic contexts (L'Abbé *et al*., 2013), and it is thought that the inclusion of more South African individuals in the SADB will help increase the accuracies achieved.

2.3.3.2 3D-ID

3D-ID was developed by Slice and Ross to assess sex and ancestry using landmark data with GM methods (Slice & Ross, 2009). 3D-ID's use of GM allows for an alternate method to 2D methods that are unable to quantify curves and between-point differences. One of the advantages of 3D-ID is that the program can eliminate the confounding effects associated with size differences without affecting the shape differences (Slice & Ross, 2009). The use of 3D-ID also provides the potential for generating higher accuracies for ancestry estimation compared to FORDISC, as the use of GM methods allows for the assessment of which landmark is affecting change, how the shape changes, and the angle of the measurement line by examining the coordinates of the landmarks themselves instead of the line between the landmarks. Not only does the use of GM methods potentially improve the efficacy of 3D-ID, but the current version of 3D-ID is also capable of assessing both shape and form variables, meaning that 3D-ID can assess shape without size affecting the results, but can also look at how the effect of size could influence the results.

The program takes the 3D landmark coordinate data that the user provides and constructs classification functions based on the available reference database of over 2000 individuals.

However, the reference database that 3D-ID employs contains a small sample of only 510 individuals of African descent with 27 African, 272 African American, 55 African Brazilian, 36 East African, 30 Nigerian and 93 West African individuals (Table 2.2). Unfortunately, it is not always specified where these individuals came from (Slice & Ross, 2009). This could increase the potential of misclassifications in the SAB population as the database represents a smaller amount of variation within the population.

3D-ID uses GM methods (including Procrustes superimposition and Principal component analysis) to process the reference database and appropriate reference samples are then extracted. Individuals are classified into population groups found in the 3D-ID database by comparing the GM data of the unknown individual with that of the reference samples. The reference samples are pooled into a mean, after which all the data are scaled and size is removed. Mahalanobis squared distances are calculated between the unknown sample and the mean of each of the reference population groups. The unknown individual is then classified into the population group whose mean is the smallest distance from the unknown individual. Then both posterior probability and typicality are calculated. Posterior probability is calculated for each group to assess how close the unknown sample is to the mean of each population group. The greater the probability, the greater the likelihood that the individual is a member of that population. The typicality (F) calculation is done to demonstrate the likelihood of a true assignment into the given population (see Figure 2.3 for output provided by 3D-ID).

Table 2.2: Sample composition of the reference database in 3D-ID (Adapted from Slice & Ross, 2009)

Sample	N
African	27
African American	272
African Brazilian	55
Brazilian	125
Circumcaribbean	26
Colombia	71
East African	36
East Asia	28
European American	378
European Central	412
European Eastern	2
European Southeastern	266
European Southwestern	446
Japanese Brazilian	27
Mesoamerican	89
Nigeria	30
South American	82
Syrian	43
West African	93
Total	2372

Assessing group membership...
--
Using 80 shape dimensions.

Summary...

==========	D2	Posterior	Typicality
African - female (4):	123.0533	0.0000	0.6572
African - male (6):	96.6010	0.0004	0.9070
African_American - female (1):	154.7724	0.0000	0.9529
African_American - male (2):	128.2759	0.0000	0.8757
East_Asian - female (2):	116.6275	0.0000	0.9505
East_Asian - male (6):	131.8045	0.0000	0.3649
European_American - female (14):	84.7682	0.0514	0.9406
European_American - male (34):	101.5957	0.0001	0.6522
European_Southwestern - female (85):	79.6581	0.5940	0.9435 <===
European_Southwestern - male (95):	81.2687	0.3169	0.9298
Mesoamerican - female (1):	146.6996	0.0000	0.9743
Mesoamerican - male (7):	108.5248	0.0000	0.7252
South_American - female (22):	87.3869	0.0231	0.8985
South_American - male (29):	88.8684	0.0141	0.8707
Syrian - female (5):	117.0099	0.0000	0.6759
Syrian - male (19):	104.5210	0.0000	0.6400

==========
*** PROCESSING COMPLETE ***

Figure 2.3: An example of the output generated when using 3D-ID

Bertsatos and colleagues (2019) used 3D-ID to analyse 158 crania from the Athens Collection, a documented Greek population sample. Shape and form variables were tested using 2 distinct groups which were based on the number of landmarks used. The first group consisted of up to a 25-landmark configuration using the full reference sample that 3D-ID contains. Although 3D-ID can employ 34 landmarks, the authors of the program do recommend that certain landmarks be removed from analysis due to high inter-observer error, hence Bertsatos and colleagues (2019) using a 25-landmark configuration. The second group consisted of up to a 19-landmark configuration using only the European reference samples in 3D-ID. Sex classifications using 3D-ID ranged between 74.1% and 85.4%, with form variables doing better than the shape variables, and the 19-landmark configuration performing better than the 25-landmark configuration. When assessing sex and ancestry collectively, all European population groups were considered one group. Once again, the 19-landmark configuration performed better than the 25-landmark configuration. The 25-landmark-confugruation achieved a correct classification of 61.4% using

shape variables and 63.9% using form variables, while the accuracies achieved for the 19-landmark configuration were 70.3% using shape variables and 70.9% using form variables. When ancestry was assessed separately from sex, the results were similar to those seen when assessing both sex and ancestry collectively, the only significant difference being the distribution of misclassified specimens which followed a somewhat different pattern. The authors concluded that 3D-ID displayed moderate reliability in the estimation of sex and ancestry of a Greek sample. It was clearly shown that the Greek sample does differ from the reference samples found in 3D-ID and that the results using 3D-ID should be interpreted carefully when using the program on an unknown individual.

Urbanová and colleagues (2014) compared the software programs FORDSIC and 3D-ID to assess sex and ancestry. The sample consisted of 174 crania of Japanese Brazilian, Afro-Brazilian, European Brazilian and admixed (mostly Afro-European admixture) individuals. They digitised 27 standard landmarks and calculated 15 ILDs that correspond to traditional cranial measurements. They assessed sex estimation and achieved accuracies ranging from 60% to 70%. When assessing ancestry using FORDISC, they looked at how the HDB compared to the FDB. Using HDB and a configuration of 14 ILDs, 14 specimens were excluded as they did not meet the typicality probability criteria. The program classified 44.5% of the individuals into an applicable ancestral group, with individuals of Asian ancestry achieving the highest accuracy (65% classified correctly). Using the FDB and a configuration of 12 ILDs, 28 specimens were excluded. FORDISC correctly classified about 50% of the individuals, with crania of Asian ancestry achieving the highest accuracies (68% classified correctly). The number of ILDs was increased to 14 and the Japanese reference groups were removed and the analysis provided slightly different results. Only 15 specimens were excluded with 50% of the individuals correctly classified. Individuals of Asian and European ancestry were classified with the highest accuracy (54.2% and 55.1%, respectively). 3D-ID was assessed in two ways, first with a configuration of 16 landmarks then using a configuration of 22 landmarks. Using 16 landmarks, 43 specimens were excluded from the analysis. 3D-ID correctly classified 50% of the crania, with individuals from European ancestry achieving the highest accuracy (87% classified correctly). Using 22 landmarks, the program correctly assigned 61.3% of the individuals, with crania from European and Asian ancestry's achieving the highest accuracy (88%). A larger proportion of individuals were excluded from 3D-ID compared to FORDISC suggesting that Brazilian individuals differ from the individuals in the

3D-ID reference groups to a larger extent. This could show that firstly, the Brazilian ancestral group differs significantly making the 3D-ID reference sample less appropriate, or secondly that the 3D coordinates in conjunction with GM is more sensitive to subtle changes among differing groups (Urbanová *et al.*, 2014). They stated that the poor classification results for sex estimation is likely because none of the reference databases showed similar patterns of sexual dimorphism as seen in the Brazilian sample used. Therefore, they concluded that when using computer-based programs for identification, one should interpret the results with caution when an unknown individual is compared with reference populations.

In King's (2015) master's dissertation, the accuracy of ancestry estimation using the programs FORDISC and 3D-ID were compared for SAB and SAW individuals, but not for SAC individuals. In the current version of 3D-ID, one can assess both shape and form variables, however when King's (2015) research was completed, a previous version of 3D-ID was used which only allowed for the analysis of shape variables. When assessing FORDISC, King (2015) used the FDB to assess sex and ancestry for the South African population and achieved accuracies of 95.7% for the SAB population and 85.4% for the SAW population; which are higher than the results obtained by L'Abbé and colleagues (2013) seen in the previous section on FORDISC. While using 3D-ID for the classification of SAB and SAW individuals, King (2015) used proxy populations where SAB individuals were considered correctly classified if they were assigned as either African or African American while SAW individuals were considered to be classified correctly if they were assigned as either European or European American. In the newer version of 3D-ID the European classification is now made up of European Central, European Eastern, European Southeastern and European Southwestern.

King (2015) achieved an overall 3D-ID classification accuracy of 63.1% after which the individuals were divided into four subsections; SAB males, SAB females, SAW males, SAW females; with accuracy rates ranging from 46.5% to 69.7%. When the influences of sex were removed from the analysis, an overall accuracy of 86% was achieved. King (2015) then analysed SAB and SAW groups separately, achieving accuracies of 81.7% for SAB individuals while SAW individuals achieved an accuracy of 89.9%. Urbanová and colleagues (2014) and King (2015) assessed their data using a previous version of 3D-ID which had a much smaller reference database, and could only assess shape variables not form variables. It was hypothesised that the low sample size seen in the comparative databases could account for the low accuracy rates

achieved using the previous version of 3D-ID, concluding that the inclusion of more samples in the comparative databases would result in higher accuracies (Urbanová *et al*., 2014; King, 2015).

Some have made the assumption that methods derived from American populations can be applied to South African populations because of the similarities between the SAB population and African Americans and the SAW population and White Americans (King, 2015). However, cranial morphology has been shown to be distinctly different between American and South African populations (İşcan & Steyn, 1999). The differences seen between these populations can be a consequence of environmental and genetic influences, but it has been theorised that the differences can be traced back to the parent populations of the groups (L'Abbé *et al*., 2013). An example of this is that many of the African American individuals had originated from the West African populations that were taken during the African Diaspora as slaves to America and are both genetically and geographically distinct from the South African population (King, 2015). So, knowing that the SAB population is very distinct from African American individuals, it can infer that the SAC population would also be distinctly different from all populations. And so, it is important to assess programs that are based on different populations from the individual being analysed as it is known that population differences do exist, especially when assessing the SAC population because programs like 3D-ID do not have any representation of these individuals in its database.

The aim of this study was thus to assess how accurately 3D-ID estimates sex and ancestry for a South African sample compared to FORDISC analyses utilising a custom SADB. To achieve this aim, the following objectives were completed:

1) To determine the ability of the software program 3D-ID to accurately estimate sex from SAB, SAW and SAC crania using both shape and form variables;
2) To examine the ability of the software program 3D-ID to accurately estimate ancestry from SAB and SAW crania using both shape and form variables;
3) To assess the group classification assigned to SAB, SAW and SAC crania from the 3D-ID program based on the current reference material;
4) To establish the group classification (sex and ancestry) assigned to SAB, SAW and SAC crania from FORDISC using landmark data.

CHAPTER 3: MATERIALS AND METHODS

3.1 MATERIALS

This research was conducted under the ethical waiver reserved for the School of Anatomical Sciences, University of the Witwatersrand (W-CJ-140604-1). A total of 450 SAB, SAC and SAW crania (75 males and 75 females per group) were selected from the Raymond A. Dart Collection of Human Skeletons, University of the Witwatersrand, and the Kirsten collection, University of Stellenbosch. The Raymond A. Dart Collection of Human Skeletons represents individuals from the 19th and 20th century (Dayal *et al.,* 2009) and the Kirsten Collection represent individuals from the early to mid-20th century (Alblas *et al.,* 2018). The individuals in these collections were donated or unclaimed bodies provided to the universities for research under the Human Tissues Act, Act 65 of 1983 (South African National Department of Health, 1983), and the more recent National Health Act, Act 61 of 2003 (South African National Department of Health, 2004). Many of the unclaimed bodies that the collections receive are of lower socio-economic status (Steyn & Patriquin 2009; Alblas *et al*., 2018). Each individual in the collections is given an accession number, that is catalogued in a database with the individual's associated information that includes for example sex, ancestry, population affinity, age- and date-of-death.

SAB individuals can be divided into two main groups: 1) the Nguni, including the Zulu, Xhosa, Ndebele and Swazi; and 2) the Sotho-Tswana, which comprises of the Southern, Northern and Western Sotho (Tswana), the Tsonga and Venda. Even though it has been shown that these groups have some cranial differences, the distinctions between the groups are considered to be disappearing and therefore, individuals of the SAB population are considered to be homogenous and are studied as one group (Steyn *et al*., 2004; Franklin *et al*., 2006; Stull *et al*., 2014). As mentioned in the literature review, the SAC population is categorised as individuals of mixed ancestral heritage including individuals from groups such as the Cape Coloured peoples, African peoples, European colonists and Asian peoples, where some were imported by the British labourers and others willingly travelled to South Africa (de Wit *et al.,* 2010; Isaacs-Martin & Petrus, 2012). As previously seen, the SAW population is made up of individuals of European descent, including Afrikaners and the English-speaking British descendants, Dutch, German and French Huguenots, resulting from the Dutch and British colonization (Guelke, 1988; Steyn & İşcan, 1998; Steyn & İşcan, 1999; L'Abbé *et al*., 2011). Although there has been intermixing between the cultural

groups, they have also remained relatively distinct due to racial discrimination and segregation (L'Abbé *et al.*, 2011; King, 2015).

This study assessed modern individuals from the 20th century, equally distributed by sex and population group. The individual ages ranged between 19 and 90 years with a mean of 55.22 ± 16.49 years (Table 3.1), ensuring that adult crania displaying a variety of morphologies were included. Incomplete crania (i.e. missing either the calotte or calvarium) and crania with pathology/trauma resulting in disfiguration were excluded from the study. However, crania exhibiting varying degrees of tooth loss (edentulous, partially edentulous and fully dentate) were included in the study as a large number of the SAW crania in South African collections are edentulous. Crania with missing landmarks were not excluded from this study because 3D-ID and FORDISC can run analyses without a complete list of landmarks. Many of the individuals in the collections come from cadaver programmes, and as such there were many cases where the calotte had been removed from the crania during the dissection process.

Table 3.1: Age distribution of the sample. SD = standard deviation

Age range (years)	**SAB**		**SAC**		**SAW**		**Total**
	Males	**Females**	**Males**	**Females**	**Males**	**Females**	
19-34	13	10	8	17	2	1	51
35-54	35	35	35	41	9	7	162
55-74	19	27	31	14	40	41	172
75-90	7	2	1	1	24	26	61
Mean age	50.32	48.38	50.25	44.77	68.19	68.99	55.22
SD	15.69	14.44	11.59	12.23	13.64	12.45	16.49

3.2 METHODS

3.2.1 Data Collection

In line with King (2015) and Murphy and Garvin (2018), crania with a separate calotte were realigned and secured using medical tape which was immediately removed once data collection were completed. Similar to these papers, no spacer was used, as they have successfully used this method without any mention of added error. Thirty-four (34) homologous landmarks

were marked with a pencil on each skull, where possible, as outlined by the 3D-ID help file (see Table 3.2 and Figures 3.1-3.4). Bi-lateral landmarks are indicated in the landmark column of table 3.2 as some landmarks do not explicitly show left or right in the abbreviations (such as obhi and obhisr). In some cases, certain landmarks (left and right upper and lower orbital border, left and right nasomaxillary suture pinch, and left and right zygion) rely on measurements and as such, these were determined instrumentally and marked with the pencil (see descriptions in Table 3.2). The pencil marks were removed once data collection was completed. The cranium was then secured in a fixed position using modelling clay to ensure that all landmarks were digitised without repositioning the skull. The landmarks were then digitised using a Microscribe 3-DX digitiser, belonging to the University of the Witwatersrand, with a reported accuracy of 0.13mm (Revware, Raleigh, NC). The catalogue number of each individual and their three-dimensional Cartesian coordinates were recorded directly into the software program 3D-ID (2018 version) using Revware's Microscribe Utility Software. Any missing landmarks were noted and left blank when data were entered into the 3D-ID program. Once the landmarks had been recorded, the data were run through 3D-ID using shape only to ensure no mistakes had been made while digitising. As suggested by Professor A. Ross (Personal communication, 18 July 2018), no skull should have a Mahalanobis squared distance higher than 450, as another higher means there was an error when digitising. In the case where the results for Mahalanobis squared distances were too high (more than 450), the landmarks were digitised again and run through the software again to see if the Mahalanobis squared distances had lowered to a more acceptable number (between 0 and 450). Once the results were considered acceptable, the data were saved as a 3D-ID file.

Table 3.2: Description of digitized landmarks used in 3D-ID (amended from Slice & Ross, 2009)

Number	Landmark	Abbreviation	Definition
1&2	Left and right asterion	astl/astr	The intersection of parietal, temporal, and occipital bones. If there are wormian bones or the sutures are indistinct, project suture lines until they intersect.
3	Basion	bas	The point in the midline of the anterior foramen magnum margin where it is intersected by the mid-sagittal plane, opposite the opisthion.
4	Bregma	brg	The point in the midline where the sagittal and coronal sutures intersect. The suture lines are projected in cases where the intersection is interrupted, such as with fontanelle bones.
5&6	Left and right dacryon	dacl/dacr	The point on the medial border where the frontal, lacrimal, and maxillary bones meet, also noted as the intersecting point of the lacrimo-maxillary suture and frontal bone. A small foramen is often present.
7&8	Left and right ectomalare	ecml/ecmr	The most lateral point on the lateral surface of the alveolar crest, found on the maxilla along the second molar.
9&10	Left and right ectoconchion	ectl/ectr	The intersecting point of the most anterior surface of lateral border and an imaginary horizontal line that bisects the orbit.
11&12	Left and right frontomalare anterior	fmal/fmar	The most anterior projecting point on the frontomalare suture.
13&14	Left and right frontomalare temporale	fmtl/fmtr	The most lateral point on the frontomalare suture.

Table 3.2: Continued…

Number	Landmark	Abbreviation	Definition
15	Glabella	glb	Most projecting midline point on the frontal bone above the frontonasal suture.
16	Lambda	lam	The point where the sagittal and lambdoidal sutures meet. Project the suture lines to their intersection point if there are wormian bones present.
17&18	Left and right mastoidale	mastl/mastr	The most inferior point on the mastoid process.
19	Nasion	nas	Midline intersecting point of the frontonasal suture and mid- sagittal plane.
20&21	Left and right lower orbital border	obhi/obhir	The midpoint on the inferior orbital margin.
22&23	Left and right upper orbital border	obhs/obhsr	The midpoint on the superior orbital margin.
24	Opisthion	ops	The point on the posterior foramen magnum margin where the mid-sagittal plane intersects. It is found opposite to basion.
25	Prosthion-Howells estimated	pr/proHEST	The most anterior, midline point on the alveolar process of the maxilla between the central incisors.
26	Subspinale	ssp	The deepest point below the anterior nasal spine.
27&28	Left and right nasomaxillary suture pinch	wnbl/wnbr	The narrowest section of the midline of the nose to the left/right naso-maxillary suture. The minimum distance between wnbl-wnbr forms the simotic chord.

Table 3.2: Continued…

Number	Landmark	Abbreviation	Definition
29&34	Left and right zygion	zygl/zygr	Lateral most point on the zygomatic arch. This is determined by measuring bizygomatic breadth.
30&31	Left and right zygomaxillare	zygoml/zygomr	Point of intersection between zygomaxillary suture and the most medial attachment site of the masseter muscle.
32&33	Left and right zygoorbitale	zygool/zygoor	Intersection of zygomaxillary suture and orbit.

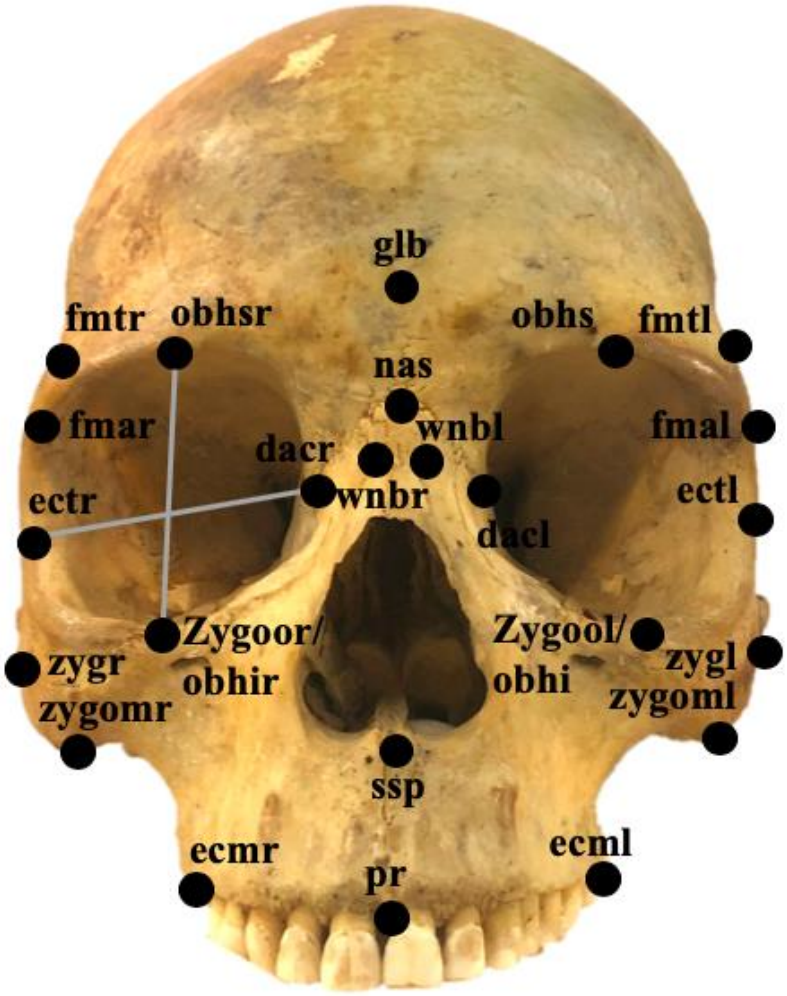


Figure 3.1: Norma frontalis illustrating the landmarks input into 3D-ID

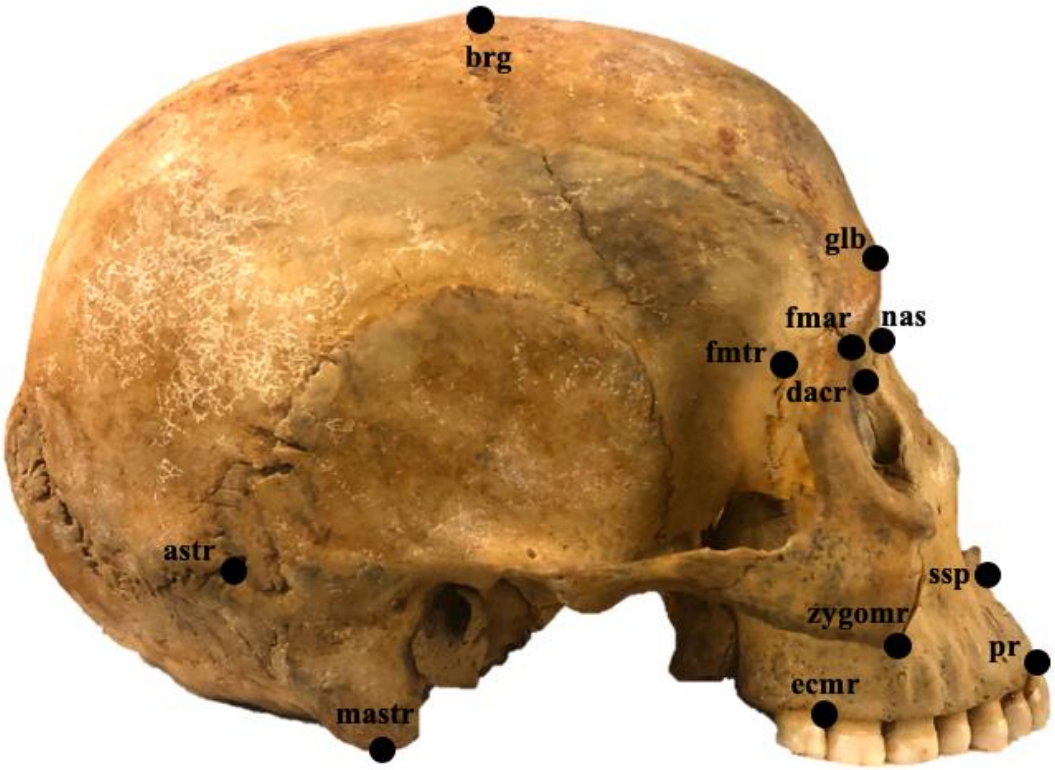


Figure 3.2: Norma lateralis illustrating the landmarks input into 3D-ID

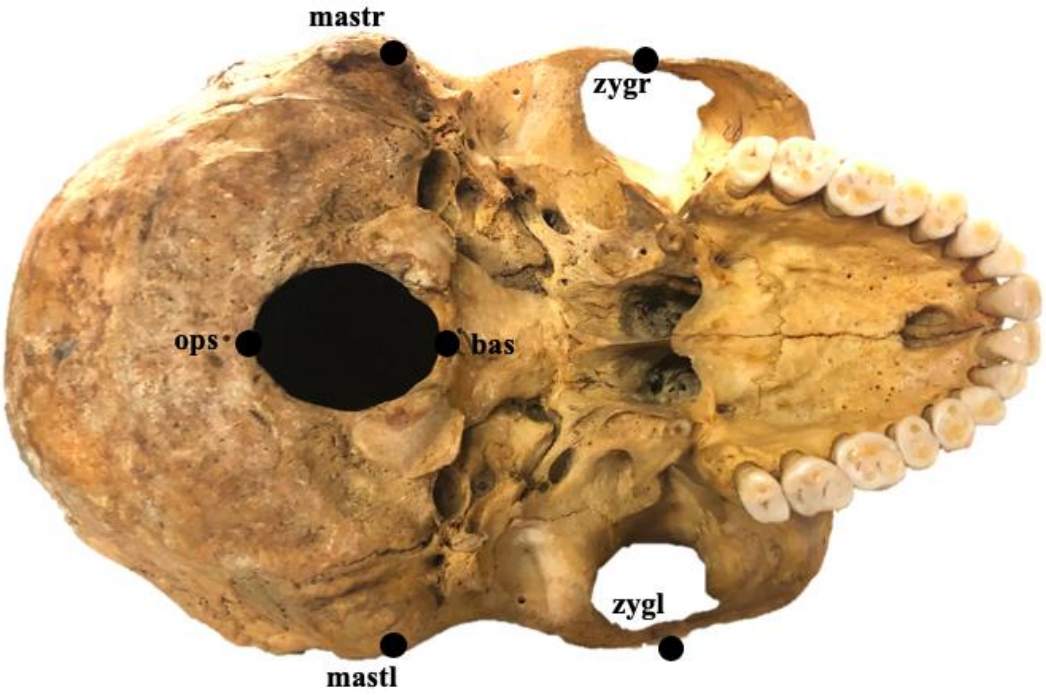


Figure 3.3: Norma basilaris illustrating the landmarks input into 3D-ID

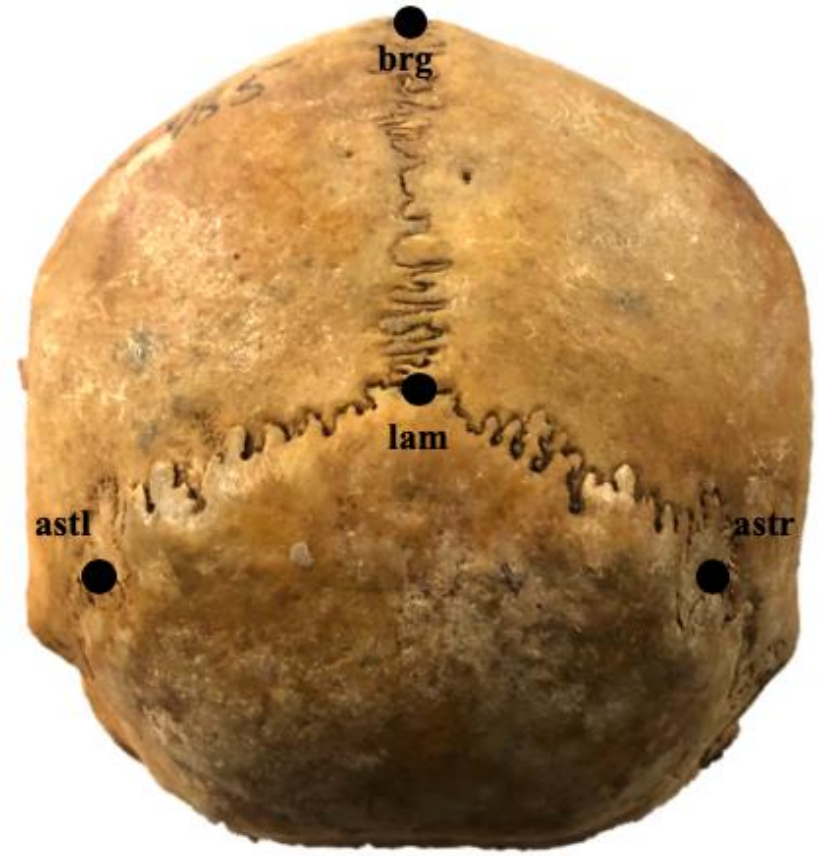


Figure 3.4: Norma occipitalis illustrating the landmarks input into 3D-ID

3.2.2 Data Analysis

3.2.2.1 Repeatability

Intra-observer error was assessed by digitising ten crania twice on different days. Inter-observer error was assessed by using ten crania digitised by myself and comparing the results with data collected by a colleague on those same ten crania. Both intra- and inter-observer error were analysed using Procrustes ANOVA in the computer program MorphoJ (Klingenberg, 2011).

3.2.2.2 3D-ID

To assess the accuracy of the program 3D-ID in a South African population, the landmark data representative of each cranium was analysed to estimate sex and ancestry. The program utilises generalised Procrustes Analysis (GPA) to superimpose the landmark data. After superimposing, the landmarks were subjected to DFA where after sex and ancestry classifications

were assigned. The unknown specimen was compared to the reference database, containing over 2000 individuals, by assigning the unknown to the group from which it has the smallest Mahalanobis squared distance (standardised distances accounting for the covariance structure of the data). Posterior probabilities were used to examine how likely the individual was aligned with a certain reference group with the assumption that the unknown individual is represented and selected for in the reference database. Finally, typicality was calculated to evaluate how likely it is that the unknown individual belongs to any population (Slice & Ross, 2009; İşcan & Steyn, 2013).

All individuals were analysed separately and compared to both males and females from all population groups in the 3D-ID database:

- African
- African American
- African Brazilian
- Brazilian
- Circumcaribbean
- Colombian
- East African
- East Asian
- European American
- European Central
- European Eastern
- European Southeastern
- European Southwestern
- Japanese Brazilian
- Mesoamerican
- Nigerian
- South American
- Syrian
- and West African

When inputting the data, two methods were used. For the first method all landmarks

collected from an individual were run through 3D-ID. For the second method five landmarks were removed from the analysis as recommended by Professor A. Ross (Personal communication, 18 July 2018; www.3d-id.org). These five landmarks include the right lower orbital border, right upper orbital border, right ectomalare, left ectomalare and prosthion. Ectomalare and prosthion were removed from the analysis because the landmarks introduce a lot of inter-observer error as data for these landmarks was gathered by many different individuals for the reference sample. The right upper and lower orbital landmarks were removed as they significantly reduce the sample size. This is because not all researchers who have submitted data to the 3D-ID program, collect both left and right orbital landmarks which also introduces error (Professor A. Ross, personal communication, 18 July 2018; www.3d-id.org). The first method will be referred to as the "Full set" while the second method will be referred to as the "Ross subset" throughout this dissertation. For each specimen, data were run using shape variables, in which size is mitigated, and then run using form variables, which includes a size component.

Once the data were run through 3D-ID, the following were recorded: the population classification, Mahalanobis squared distance, posterior probability and typicality. These results were compared with the recorded sex and ancestry of the specimen after which the accuracy and precision of the assignment of sex as well as ancestry for each individual was estimated using a frequency table. In keeping with King (2015) and to make results comparable, we considered results to be "accurately estimated" when they are 75% or higher.

When noting population classification, a correct classification of African was assigned for SAB individuals and European Southwestern or European Central for SAW individuals. However, for SAC individuals no correct classification was assigned because, not only is this population extremely diverse, they are not represented in the 3D-ID database. Instead for the SAC population it was noted into which population groups 3D-ID assigned the individuals to.

3.2.2.3 FORDISC

To assess the program FORDISC in a South African population, ILDs were derived from the cranial GM data to estimate sex and ancestry. FORDISC assesses linear measurements taken from the cranium and uses LDA to classify the unknown individual into a reference population group. Depending on the number of measurements available for a particular specimen, FORDISC will customise each formula so that if certain skeletal elements are incomplete a function can still

be calculated (İşcan & Steyn, 2013). The output specifies the group classification for the individual, Mahalanobis squared distances, cross-validated classification accuracies as well as posterior, typicality and ranked probabilities; allowing the observer to come to an informed conclusion.

Using the statistical program R (R Core Team, 2018), ILDs were derived by using Pythagorean geometry; the distances between the coordinates were calculated. The ILDs derived from the GM data were then analysed in FORDISC 3.1. From the 34 digitised landmarks we were able to derive 13 ILDs adapted from Moore-Jansen and colleagues (1994) (see Table 3.3). Maximum alveolar length (MAL) and maximum alveolar breadth (MAB) were excluded from analysis because the database has a very small sample of individuals with this information.

Table 3.3: Interlandmark distances derived from 3D-ID landmarks incorporated into FORDISC

Interlandmark distances	Abbreviation	Definition
Basion-bregma height	BBH	The distance from basion to bregma
Basion-nasion length	BNL	The distance from basion to nasion
Basion-prosthion length	BPL	The distance from basion to prosthion
Bi-orbital breadth	EKB	The distance from left to right ectoconchion
Bizygomatic breadth	ZYB	The maximum breadth across the zygomatic arches
Foramen magnum length	FOL	The mid-sagittal distance from the anterior most point of the foramen magnum margin to opisthion
Frontal chord	FRC	The distance from bregma to nasion
Inter-orbital breadth	DKB	The distance between left and right dacryon
Occipital chord	OCC	The distance from lambda to opisthion
Orbital breadth	OBB	The distance from dacryon to ectoconchion
Orbital height	OBH	The distance between the upper and lower orbital border
Parietal chord	PAC	The distance from bregma to lambda
Upper facial breadth	UFBR	The distance between the left and right frontomalare temporale

The data were run through FORDISC 3.1 using the custom South African Database (SADB) to assess sex, ancestry, and combined sex and ancestry for each cranium. The cranial measurements (table 3.3) were imported into FORDISC 3.1 where DFA was run for each individual in the sample. DFA is a useful technique because it can help quantify differences both within and between populations (L'Abbé *et al*., 2013). Once the data were run through FORDISC 3.1, the following output was recorded for each specimen: population classification, posterior probability (PP), typicality probability (TP) and the cross-validated classification accuracy. During the analysis PP and TP were looked at for each individual to allow the researcher to assign a classification to the individual. The results achieved using the SADB was compared with the documented sex and ancestry. The results when using the SADB were also then compared to the results achieved in 3D-ID.

CHAPTER 4: RESULTS

4.1 REPEATABILITY

Using Procrustes ANOVA in MorphoJ (Klingenberg, 2011) it was found that the measurement error for both intra- and inter-observer error was found to be negligible (see Appendix 4).

4.2 3D-ID INVESTIGATION

The ability of 3D-ID to assess sex and ancestry from crania of SAB, SAW and SAC individuals was examined. When capturing the cranial data in 3D-ID, two main methods were used. First, all the available landmarks were included which will be referred to as the "Full set" throughout this dissertation. Then the following five landmarks were removed; right lower orbital border, right upper orbital border, right ectomalare, left ectomalare and prosthion. This subset will be referred to as the "Ross subset" throughout. For each specimen, data were run using shape variables, in which size is mitigated, and then run using form variables, which includes a size component.

4.2.1 Sex

The results of the sex classification for the SAB population are summarised in Table 4.1. The analysis of the Full set for SAB males using form variables achieved higher accuracies for sex classification compared to shape variables, while SAB females achieved the same accuracy of 80.0% for both shape and form variables. The accuracies for the analysis using the Full set achieved low accuracies for SAB males, 44.0% for shape variables and 50.7% for form variables. The analysis of both shape and form variables, using the Ross subset achieved higher accuracies for SAB males, 56.0% and 69.3% respectively, while the analysis for SAB females using only form variables achieved the highest accuracy of 89.3%. Overall, it was seen that SAB females achieved higher sex classification accuracies when using 3D-ID compared to SAB males.

The classification results for sex for the SAW population are summarised in Table 4.2. Form variables achieved higher accuracies for SAW males (65.3% for the Full set and 85.3% for the Ross subset) compared to shape variables (52.0% for the Full set and 76.0% for the Ross subset). However, analysis using shape variables achieved higher accuracies for SAW females

(93.3% for the Full set and 89.3% for the Ross subset) compared to form variables (81.3% for the Full set and 88.0% for the Ross subset). Form and shape variable analysis for SAW males achieved higher accuracies (76.0% and 85.3% respectively) when using the Ross subset compared to the analysis using the Full set which achieved accuracies of 52.0% and 65.3% respectively.

Table 4.1: Frequency table showing the correct sex classifications for **SAB** males (n = 75) and females (n = 75) using 3D-ID

Sex	3D-ID analysis	Shape variables		Form variables	
		Frequency	**Percentage**	**Frequency**	**Percentage**
SAB Males	Full set	33	44.0%	38	50.7%
	Ross subset	42	56.0%	52	69.3%
SAB Females	Full set	60	80.0%	60	80.0%
	Ross subset	62	82.7%	67	89.3%

Blue colour coding - highest correct sex classification frequency

Form variables for SAW females achieved a higher accuracy of 88.0% when using the Ross subset compared to the accuracy of 81.3% when using the Full set. However, shape variable analysis for SAW females using the Full set achieved a higher accuracy of 93.3% compared to the analysis using the Ross subset with an accuracy of 89.3%.

Table 4.2: Frequency table showing the correct sex classifications for **SAW** males and females using 3D-ID (n = 75)

Sex	3D-ID analysis	Shape variables		Form variables	
		Frequency	**Percentage**	**Frequency**	**Percentage**
SAW Males	Full set	39	52.0%	49	65.3%
	Ross subset	57	76.0%	64	85.3%
SAW Females	Full set	70	93.3%	61	81.3%
	Ross subset	67	89.3%	66	88.0%

Blue colour coding - highest correct sex classification frequency

The results of the classification for sex for the SAC population are summarised in Table 4.3. For both shape and form variables the accuracies from the Ross subset were higher for SAC males (70.7% and 78.7% respectively) while the accuracies for the analysis using the Full set were lower (50.7% and 52.0% respectively). However, the accuracies when using the Full set were higher for SAC females, 70.7% for shape variables and 73.3% for form variable; while the accuracies achieved when assessing the analysis for the Ross subset were lower for form variables (58.7% for shape variables and 66.7%). For both SAC males and females, the analysis using form variables achieved higher accuracies for sex classifications with accuracies ranging between 52.0% and 78.7%.

Table 4.3: Frequency table showing the sex classifications for **SAC** males (n = 75) and females (n = 75) using 3D-ID

Sex	**3D-ID analysis**	**Shape variables**		**Form variables**	
		Frequency	**Percentage**	**Frequency**	**Percentage**
SAC Males	Full set	38	50.7%	39	52.0%
	Ross subset	53	70.7%	59	78.7%
SAC Females	Full set	53	70.7%	55	73.3%
	Ross subset	44	58.7%	50	66.7%

Blue colour coding - highest correct sex classification frequency

4.2.2 Ancestry

4.2.2.1 South African Black (SAB) population

Ancestry classification for SAB males are summarised in Table 4.4. When assessing both shape and form variables using the Full set, SAB males were classified mostly as European Southwestern with frequencies of 30.7% and 32.0% respectively, followed by African (26.7% for shape variables and 18.7% for form variables). Using only form analysis for the Full set, SAB males were also classified as South American (18.7%). When assessing both shape and form variables of the Ross subset, SAB males were classified mostly as African American with frequencies of 70.7% and 68.0% respectively, followed by African with frequencies of 9.3% for shape variables and 8.0% for form variables.

Table 4.4: Frequency table related to ancestry classification for **SAB** males (n = 75)

3D-ID reference population	Number of Individuals			
	Full set		Ross subset	
	Shape variables	Form variables	Shape variables	Form variables
African	20 (26.7%)	14 (18.7%)	7 (9.3%)	6 (8.0%)
African American	0	0	53 (70.7%)	51 (68.0%)
Circumcaribbean	0	0	2 (2.7%)	2 (2.7%)
Colombian	0	0	1 (1.3%)	2 (2.7%)
East Asian	5 (6.7%)	7 (9.3%)	0	0
European American	17 (22.7%)	12 (16.0%)	1 (1.3%)	1 (1.3%)
European Southeastern	0	0	0	2 (2.7%)
European Southwestern	23 (30.7%)	24 (32.0%)	6 (8.0%)	3 (4.0%)
Mesoamerican	4 (5.3%)	4 (5.3%)	3 (4.0%)	3 (4.0%)
South American	5 (6.7%)	14 (18.7%)	2 (2.7%)	5 (6.7%)
Syrian	1 (1.3%)	0	0	0

Blue colour coding - highest classification frequency

Ancestry classification for SAB females are summarised in Table 4.5. When assessing both shape and form variables using the Full set SAB females were classified mostly as European Southwestern with frequencies of 45.3% and 40.0% respectively, followed by African with frequencies of 24.0% for shape and 17.3% for form. While when assessing both shape and form variables when using the Ross subset SAB, females were classifying mostly as African American with frequencies of 73.3%, followed by European Southwestern with frequencies of 9.3% for shape and 6.7% for form. Using only form variables whilst using the Ross subset, SAB females also classified as South American with a frequency of 6.7%.

Not only did this study look at which groups SAB males and females would classify as when using 3D-ID, it also investigated how often SAB males and females would correctly classify as African (it was seen that of all the classifications from African countries [i.e. African, East

African, West African and Nigerian] South Africans were only classifying as African, and so only African was used for the correct classification). When assessing the Full set, SAB males were correctly classified as African with accuracies of 26.7% and 18.7% (Table 4.4), while SAB females were correctly classified as African with accuracies of 24.0% and 17.3%, depending on shape and form variables respectively. To visually represent the results for most and second most classification of SAB males and females the following graph was designed (figure 4.1).

Table 4.5: Frequency table related to ancestry classification for **SAB** Females (n = 75)

3D-ID reference population	**Number of Individuals**			
	Full set		**Ross subset**	
	Shape variables	**Form variables**	**Shape variables**	**Form variables**
African	18 (24.0%)	13 (17.3%)	5 (6.7%)	3 (4.0%)
African American	0	1 (1.3%)	55 (73.3%)	55 (73.3%)
Circumcaribbean	0	0	1 (1.3%)	1 (1.3%)
East Asian	2 (2.7%)	7 (9.3%)	0	1 (1.3%)
European American	13 (17.3%)	12 (16.0%)	0	0
European Southeastern	0	0	2 (2.7%)	2 (2.7%)
European Southwestern	34 (45.3%)	30 (40.0%)	7 (9.3%)	5 (6.7%)
Mesoamerican	0	1 (1.3%)	3 (4.0%)	3 (4.0%)
South American	5 (6.7%)	8 (10.7%)	2 (2.7%)	5 (6.7%)
Syrian	3 (4.0%)	3 (4.0%)	0	0

Blue colour coding - highest correct sex classification frequency

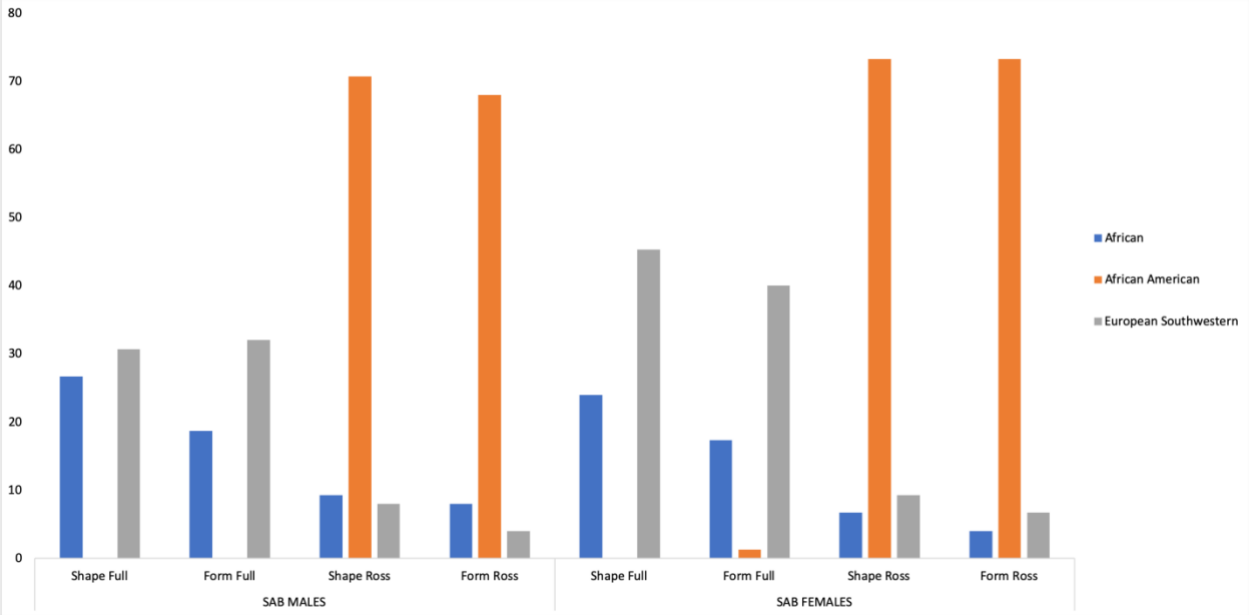


Figure 4.1: Visually representation of the most and second most classification for SAB males and females

When using the Ross subset, SAB males were correctly classifying as African with accuracies of 9.3% and 8.0% (Table 4.4), while SAB females were correctly classified as African with accuracies of 6.7% and 4.0%, once again depending on shape and form variables respectively. Overall, when SAB males and females classified as African, the Full set achieved better classifications compared to the Ross subset, and classifications were better when using shape variables compared to form variables.

4.2.2.2 South African White (SAW) population

Ancestry classification for SAW males are summarised in Table 4.6. Overall, when considering both shape and form and using the Full set and the Ross subset, SAW males classified mostly as European American with frequencies ranging from 34.7% to 54.7%. The second most classification was European Southwestern with frequencies ranging from 22.7% to 37.3%.

Table 4.6: Frequency table related to ancestry classification for **SAW** males (n = 75)

3D-ID reference population	Number of Individuals			
	Full set		Ross subset	
	Shape variables	Form variables	Shape variables	Form variables
African	1 (1.3%)	2 (2.7%)	0	0
African American	0	0	4 (5.3%)	7 (9.3%)
Circumcaribbean	0	0	6 (8.0%)	9 (12.0%)
East Asian	0	1 (1.3%)	1 (1.3%)	1 (1.3%)
European American	41 (54.7%)	36 (48.0%)	30 (40.0%)	26 (34.7%)
European Southeastern	0	0	4 (5.3%)	4 (5.3%)
European Southwestern	22 (29.3%)	28 (37.3%)	19 (25.3%)	17 (22.7%)
Mesoamerican	2 (2.7%)	1 (1.3%)	0	1 (1.3%)
South American	8 (10.7%)	7 (9.3%)	11 (14.7%)	10 (13.3%)
Syrian	1 (1.3%)	0	0	0

Blue colour coding - highest correct sex classification frequency

Ancestry classification for SAW females are summarised in Table 4.7. Analysis of SAW females using shape variables for the Full set were classifying mostly as European American with a frequency of 49.3% followed by European Southwestern with a frequency of 44.0%. While the analysis of SAW females using form variables for the Full set were classified mostly as European Southwestern with a frequency of 52.0% followed by European American with a frequency of 37.3%. For both shape and form variables, when using the Ross subset, this study found that SAW females classified mostly as European Southwestern with frequencies of 57.3% and 58.7% respectively. The analysis of shape variables using the Ross subset for SAW females were second most classifying as South American with a frequency of 16.0% and form variable analysis using the Ross subset were second most classifying as European American and South American with frequencies of 14.7%.

For ease of understanding, the results achieved for SAW males and females when using 3D-ID were visually demonstrated (figure 4.2).

Table 4.7: Frequency table related to ancestry classification for **SAW** females (n = 75)

3D-ID reference population	Number of Individuals			
	Full set		Ross subset	
	Shape variables	Form variables	Shape variables	Form variables
African	0	0	1 (1.3%)	1 (1.3%)
African American	0	0	5 (6.7%)	4 (5.3%)
Circumcaribbean	0	0	2 (2.7%)	1 (1.3%)
European American	37 (49.3%)	28 (37.3%)	10 (13.3%)	11 (14.7%)
European Southeastern	0	0	2 (2.7%)	3 (4.0%)
European Southwestern	33 (44.0%)	39 (52.0%)	43 (57.3%)	44 (58.7%)
South American	5 (6.7%)	7 (9.3%)	12 (16.0%)	11 (14.7%)
Syrian	0	1 (1.3%)	0	0

Blue colour coding - highest correct sex classification frequency

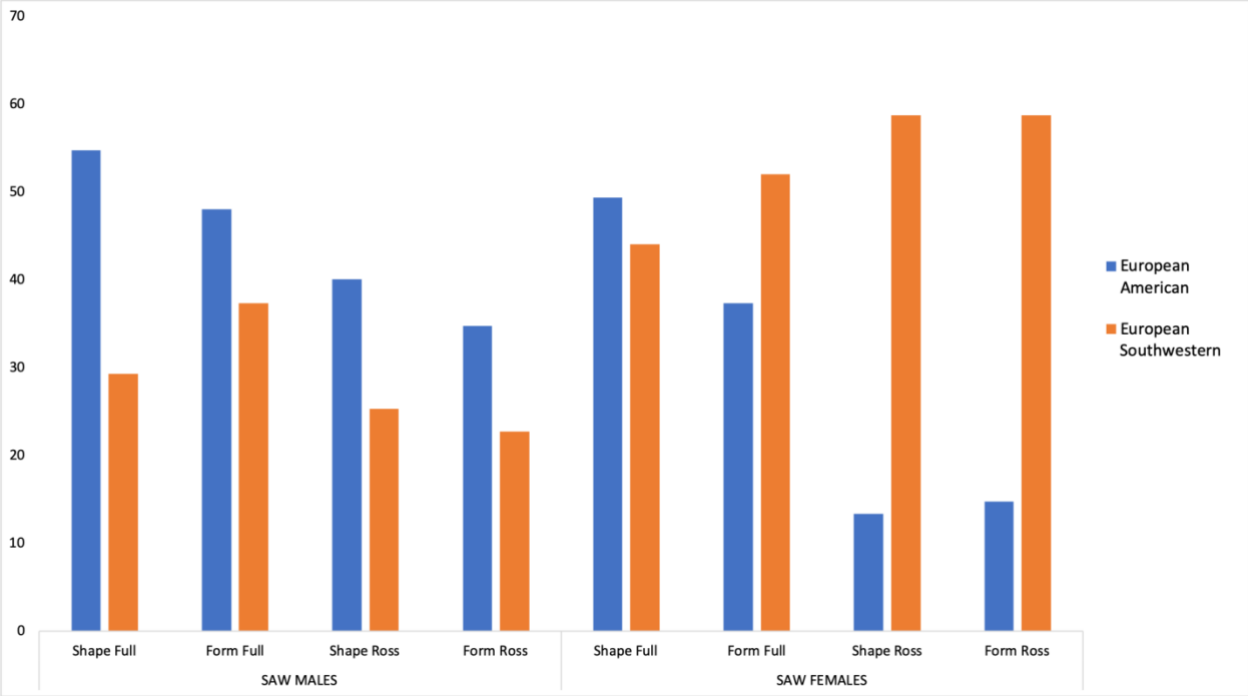


Figure 4.2: Visually representation of the most and second most classification for SAW males and females

Not only did this study look at which groups SAW males and females would classify as in 3D-ID, it also looked to see how often they would correctly classify as European Southwestern (it was noted that no South African individuals classified as European Central). When assessing the Full set, SAW males were correctly classified as European Southwestern with accuracies of 29.3% and 37.3% depending on shape and form variables respectively, while SAW females were correctly classified as European Southwestern with accuracies of 44.0% and 52.0% depending on shape and form variables respectively. When using the Ross subset, SAW males were correctly classifying as European Southwestern with accuracies of 25.3% for shape and 22.7% for form variables, while SAW females were correctly classified as European Southwestern with accuracies of 57.3% for shape and 58.7% for form variables. Overall, this study saw that when SAW males are classified as European Southwestern, they achieved better classifications when using the Full set, while when SAW females are classified as European Southwestern, they achieved better classifications when using the Ross subset. The study also found that for SAW males form performed better for the Full set while shape performed better when using the Ross subset; while SAW females achieved better classifications overall when using form variables.

4.2.2.3 South African Coloured (SAC) population

Unlike with SAB and SAW individuals, no correct classification was assigned for SAC males and females because the SAC population have an extremely complex genetic admixture and are not represented in the 3D-ID reference database. Instead this study was interested in finding which groups the SAC population would classify as in 3D-ID. Ancestry classification for SAC males are summarised in Table 4.8. The analysis of shape and form variables for SAC males using the Full set and the Ross subset classified individuals mostly as European Southwestern with frequencies ranging from 34.7% to 50.7%. SAC males classified as African second most using shape variables for both the Full set (14.7%) and the Ross subset (25.3%) and when using form variables with the Ross subset (28.0%). Using form variable for the Full set, SAC males were second most often classified as South American (18.7%).

Table 4.8: Frequency table related to ancestry classification for **SAC** males (n = 75)

Population group	Number of Individuals			
	Full set		Ross subset	
	Shape variables	Form variables	Shape variables	Form variables
African	11 (14.7%)	5 (6.7%)	19 (25.3%)	21 (28.0%)
African American	1 (1.3%)	4 (5.3%)	10 (13.3%)	10 (13.3%)
Circumcaribbean	0	0	10 (13.3%)	10 (13.3%)
East Asian	8 (10.7%)	5 (6.7%)	4 (5.3%)	2 (2.7%)
European American	10 (13.3%)	10 (13.3%)	1 (1.3%)	0
European Southeastern	0	0	0	0
European Southwestern	38 (50.7%)	29 (38.7%)	27 (36.0%)	26 (34.7%)
Mesoamerican	2 (2.7%)	0	1 (1.3%)	2 (2.7%)
South American	5 (6.7%)	14 (18.7%)	3 (4.0%)	4 (5.3%)
Syrian	0	8 (10.7%)	0	0

Blue colour coding - highest correct sex classification frequency

The ancestry classification for SAC females are summarised in Table 4.9. Similar to SAC males, SAC females most frequently classified as European Southwestern for the analysis when using the Full set (57.3% for shape variables and 50.7% for form variables) and when using the Ross subset (37.3% for shape variables and 40.0% for form variables). This was followed by an African classification (16.0% - 32.0%) except for form analysis using the Full set which classified SAC females as South American (20.0%). The results for SAC males and females using 3D-ID have been visually represented in figure 4.3 for ease of understanding.

Table 4.9: Frequency table related to ancestry classification for **SAC** females (n = 75)

Population group	Number of Individuals			
	Full set		Ross subset	
	Shape variables	Form variables	Shape variables	Form variables
African	12 (16.0%)	5 (6.7%)	24 (32.0%)	21 (28.0%)
African American	1 (1.3%)	1 (1.3%)	8 (10.7%)	10 (13.3%)
Circumcaribbean	0	0	4 (5.3%)	4 (5.3%)
East Asian	2 (2.7%)	1 (1.3%)	0	0
European American	7 (9.3%)	7 (9.3%)	2 (2.7%)	1 (1.3%)
European Southeastern	0	0	2 (2.7%)	2 (2.7%)
European Southwestern	43 (57.3%)	38 (50.7%)	28 (37.3%)	30 (40.0%)
Mesoamerican	1 (1.3%)	0	3 (4.0%)	3 (4.0%)
South American	7 (9.3%)	15 (20.0%)	4 (5.3%)	4 (5.3%)
Syrian	2 (2.7%)	8 (10.7%)	0	0

Blue colour coding - highest correct sex classification frequency

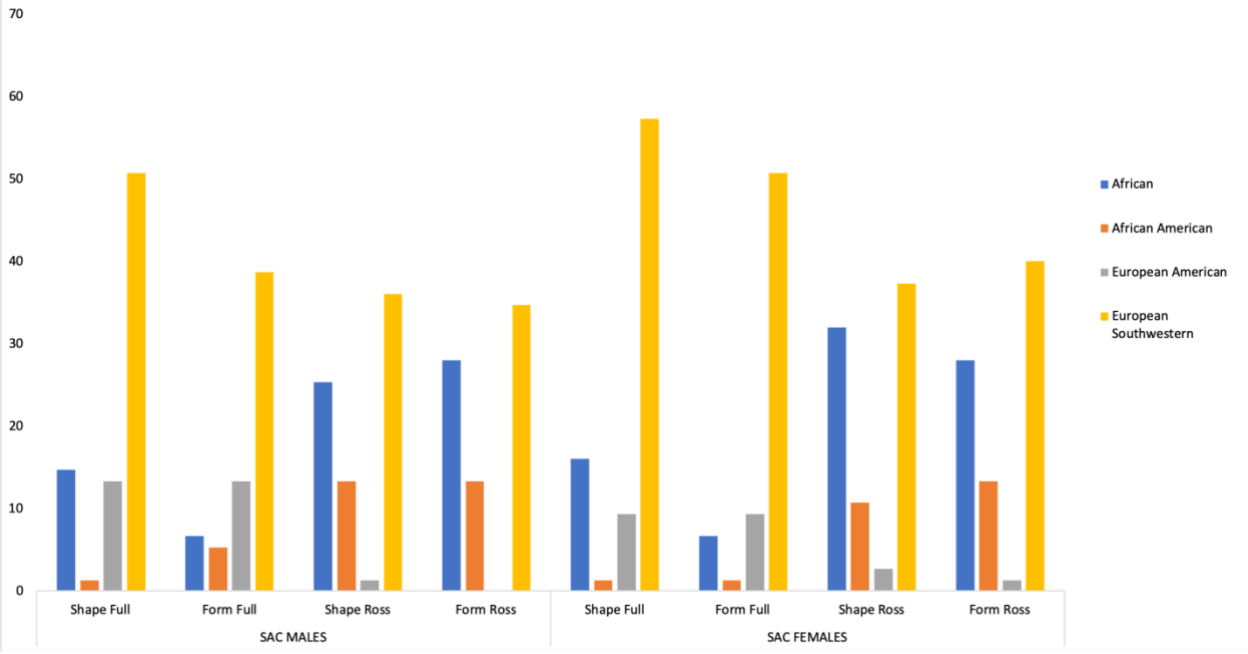


Figure 4.3: Visually representation of the most and second most classification for SAC males and females

4.3 FORDISC INVESTIGATION

From the GM data collected using the Microscribe, ILDs were derived and used to assess the accuracy of FORDISC when assessing South African crania. This study was able to derive 13 ILDs which were then run through FORDISC using the custom South African Database (SADB). When assessing each individual in FORDISC, three analyses were done; sex only, ancestry only, as well as, sex and ancestry.

4.3.1 Sex

The results of sex classification for SAB, SAW and SAC males and females are summarised in Table 4.10. SAB males achieved a correct classification accuracy of 82.7% while SAB females achieved a correct classification accuracy of 85.3%. SAW males achieved a correct classification accuracy of 78.7% while SAW females achieved a correct classification accuracy of 100.0%. SAC males achieved a correct classification accuracy of 66.7% while SAC females achieved a correct classification accuracy of 89.3%.

Table 4.10: **Sex** classification for FORDISC for SAB males (n = 75) and females (n = 75), SAW males (n = 75) and females (n = 75) and SAC males (n = 75) and females (n = 75)

	Sex	Frequency	Percentage
SAB males	Males	62	82.7%
	Females	13	17.3%
SAB females	Males	11	14.7%
	Females	64	85.3%
SAW males	Males	59	78.7%
	Females	16	21.3%
SAW females	Males	0	0.0%
	Females	75	100.0%
SAC males	Males	50	66.7%
	Females	25	33.3%
SAC females	Males	8	10.7%
	Females	67	89.3%

4.3.2 Ancestry

The ancestry classification results for SAB, SAW and SAC males and females are summarised in Table 4.11. SAB males achieved a correct classification of 64.0% while SAB females achieved a lower correct classification of 62.7%. SAW males achieved a correct classification of 69.3% while SAW females achieved a correct classification of 62.7%. SAC males achieved a correct classification of 26.7% while SAC females achieved a higher correct classification of 42.7%. The SAC population were misclassified as SAB individuals more often than they were classified as SAC individuals.

Table 4.11: **Ancestry** classification for FORDISC for SAB males (n = 75) and females (n = 75), SAW males (n = 75) and females (n = 75) and SAC males (n = 75) and females (n = 75)

	Ancestry	**Frequency**	**Percentage**
SAB males	SAB	48	64.0%
	SAC	21	28.0%
	SAW	6	8.0%
SAB females	SAB	47	62.7%
	SAC	22	29.3%
	SAW	6	8.0%
SAW males	SAB	12	16.0%
	SAC	11	14.7%
	SAW	52	69.3%
SAW females	SAB	7	9.3%
	SAC	21	28.0%
	SAW	47	62.7%
SAC males	SAB	40	53.3%*
	SAC	20	26.7%
	SAW	15	20.0%
SAC females	SAB	34	45.3%*
	SAC	32	42.7%
	SAW	9	12.0%

4.3.3 Sex and Ancestry

The sex and ancestry classification results for SAB, SAW and SAC males and females are summarised in Tables 4.12. SAB males achieved a correct sex and ancestry classification of 60.0% while SAB females achieved a correct sex and ancestry classification of 49.3%. SAW males achieved a correct sex and ancestry classification of 49.3% while SAW females achieved a correct sex and ancestry classification of 66.7%. SAC males achieved a correct sex and ancestry classification of 28.0% while SAC females achieved a correct sex and ancestry classification of 38.7%. Interestingly SAC females were also classified as SAB females in 38.7% of the cases.

Table 4.12: **Sex and ancestry** classification for FORDISC for SAB males (n=75) and females (n=75), SAW males (n=75) and females (n=75) and SAC males (n=75) and females (n=75)

	Sex and Ancestry	Frequency	Percentage
SAB males	SAB males	45	60.0%
	SAB females	3	4.0%
	SAC males	13	17.3%
	SAC females	8	10.7%
	SAW males	4	5.3%
	SAW females	2	2.7%
SAB females	SAB males	10	13.3%
	SAB females	37	49.3%
	SAC males	3	4.0%
	SAC females	15	20.0%
	SAW males	1	1.3%
	SAW females	9	12.0%
SAW males	SAB males	11	14.7%
	SAB females	0	0.0%
	SAC males	5	6.7%
	SAC females	7	9.3%
	SAW males	37	49.3%
	SAW females	15	20.0%

Table 4.12: Continued…

	Sex and ancestry	Frequency	Percentage
SAW females	SAB males	0	0.0%
	SAB females	10	13.3%
	SAC males	4	5.3%
	SAC females	11	14.7%
	SAW males	0	0.0%
	SAW females	50	66.7%
SAC males	SAB males	17	22.7%
	SAB females	9	12.0%
	SAC males	21	28.0%
	SAC females	10	13.3%
	SAW males	7	9.3%
	SAW females	11	14.7%
SAC females	SAB males	3	4.0%
	SAB females	29	38.7%*
	SAC males	2	2.7%
	SAC females	29	38.7%
	SAW males	2	2.7%
	SAW females	10	13.3%

* - correct classification

Blue colour coding - the outliers where SAC pop classified with higher % as SAB

CHAPTER 5: DISCUSSION

Computer software programs have become increasingly popular in the field of forensic anthropology, with studies highlighting the value of using programs such as 3D-ID and FORDISC for the estimation of sex and ancestry. However, these programs do not perform very well when the individual being assessed is not included in the reference database (Elliott & Collard, 2009; Guyomarc'h & Bruzek, 2011; Bertsatos *et al.*, 2019). Therefore, the current study assessed the applicability of 3D-ID for the estimation of sex and ancestry from three modern South African populations which are not included in the 3D-ID reference database, SAB, SAW and most importantly, SAC individuals. The study also tested the accuracy of FORDISC for sex and ancestry estimation on the same three South African populations with the custom SADB, which has been validated for the SAW and SAB populations but not for the SAC population (L'Abbé *et al.*, 2013). Given that the FORDISC program has a custom South African database whereas the 3D-ID program does not, the study also compared the results for sex and ancestry between the two programs. Results were considered to be accurate when they were 75% or higher.

This study does not aim to judge how effectively 3D-ID or FORDISC assess sex and ancestry, as this has been shown to be useful in cases where the unknown individual is represented in the reference databases. Rather, this study aims to test how these programs work when assessing an individual who is potentially not represented in the reference database (as is the case for SAC individuals not being represented in 3D-ID). The reason for this is because the study aims to evaluate how well 3D-ID and FORDISC perform when classifying unknown individuals, and to determine if there is a possibility of classifying individuals in forensic cases when the individual is potentially not included in the reference database. This is a very high possibility, as in forensic practice there is no previous knowledge of whether the individual is represented in the reference database of these programs. Using 3D-ID, FORDISC or any other program for the estimation of ancestry in such a context, where there is no prior knowledge of the individual, is not confirmation of ancestry, but rather should be used as an indication towards potential identification. It is, however, interesting to see the population group results of 3D-ID as this can provide information on the biological association between the populations.

5.1 3D-ID INVESTIGATION

5.1.1 Sex

3D-ID is more accurate for SAB females than it is for SAB males for all analyses, as it was found that SAB females achieved higher accuracies (80.0% to 89.3%) compared to SAB males (44.0% to 69.3%). 3D-ID was also found to be more accurate for all variables when assessing SAW females (81.3% to 93.3%) compared to SAW males (52.0% to 85.3%). There were some analyses for SAW males that achieved high accuracies of 76.0% and 85.3% respective for shape and form analyses when using the Ross subset. As seen, SAB and SAW males achieved lower accuracies compared to SAB and SAW females which follows a standard pattern for the estimation of sex for most human skeletal elements where females are more easily recognisable than those for males (Urbanová *et al*., 2013; Urbanová *et al*., 2014). Not only could this be related to females having a smaller range of variation compared to males, but it could also be related to the general health of the population as it is believed that females tend to be less affected by environmental stress (Humphries & Ross, 2011; Urbanová *et al*., 2014). These results were also similar to results found by Urbanová and colleagues (2014). The authors assessed a multi-ancestral Brazilian sample using 3D-ID and found there was a greater frequency of misclassification in the male sample compared to the female sample.

3D-ID performed moderately when assessing sex for the SAC population, as only SAC males achieved an accuracy over 75.0%. L'Abbé and associates (2013) assert that South African populations are not as sexually dimorphic as American populations, which makes up a large portion of 3D-ID's reference database, and this difference in sexual dimorphism could relate to the SAC population achieving lower accuracies for sex estimation.

3D-ID's classification accuracy for SAB and SAW males increased when using form variables instead of shape variables, which is similar to results by Bertsatos and colleagues (2019) who found that form variables achieve higher sex accuracies for a modern Greek sample. This could be related to the fact that size plays a very important role in variation for SAB and SAW males. While females differed, SAW females achieved higher accuracies when using shape variables, giving an indication that shape has a major contribution to variation in SAW females. SAB females on the other hand showed a slightly different pattern, when using the Ross subset form variables achieved a higher accuracy compared to shape variables. However, when using the Full set, they achieved the same accuracy when form and shape variables are used. The analysis

for the Ross subset corroborates that size is a major contributing factor to variation. However, because the analysis for the Full set achieved the same accuracy for both shape and form, it may indicate that size and shape provide an equal contribution to the variation in SAB females. Overall, the results for form variables performed better than the results for shape variables when assessing both SAC males and females which agrees with the proposal that most of the variation seen between males and females is related to size (Lynnerup, 2013; Bertsatos *et al*., 2019). The highest accuracy in the analysis for shape variables was 70.7%, which is considered relatively low in a forensic context, while the highest accuracy in the analysis for form variables was 78.7% which is higher than the universally accepted standard of 75.0% for use in forensics (King, 2015).

The accuracies for the analysis using the Full set was lower than the analysis when using the Ross subset for both SAB, SAW and SAC males and SAB females. This is in line with the guidelines provided for 3D-ID that the Ross subset results in increasing the reference sample size which would give a larger sample to compare the unknown individual to, increasing the accuracy for correct classification (www.3d-id.org). However, shape variables for SAW females when using the Full set and both analyses for SAC females when using the Full set achieved a higher accuracy compared to the analysis when using the Ross subset. By using the Ross subset, the number of individuals in the sample for comparison increased. It is, therefore, possible that when assessing SAW females, the increased reference sample introduces too much variation for shape, resulting in decreased accuracies. Similarly, the difference seen for SAC females could be related to the increased sample size when using the Ross subset as it is possible that the features for SAC females are not as well represented as the SAC males and so when increasing the sample size SAC females achieved lower accuracies.

When testing Walker's (2008) method (the visual assessment of the mental eminence, glabellar area, orbital margin, mastoid process and nuchal area, as outlined by Buikstra and Ubelaker (1994)) on a South African sample, Krüger and colleagues (2014) achieved higher accuracies for both SAB males (84% - 92%) and SAW males (94% - 97%) compared to this study's results of 44.0% to 69.3% for SAB males and 52.0% to 85.3% for SAW males. Whereas, the current study achieved higher accuracies for both SAB females (80.0% – 89.3%) and SAW females (81.3% - 93.3%) compared to Krüger and colleague's (2014) SAB females (55% - 90%) and SAW females (31% - 62%). While using metric methods for a South African sample, Steyn and İşcan (1998) achieved accuracies ranging between 81.1% and 85.7% for SAW individuals

whereas Dayal and associates (2008) achieved accuracies ranging from 80.0% to 83.9% for SAB individuals. These results are similar to the results for SAW and SAB females. In the current study, however, SAW and SAB males had much lower accuracies (44.0% and 52.0%) compared to the lowest accuracies for Steyn and İşcan (1998) as well as, Dayal and collaborators (2008) (81.1% for SAW individuals and 80.0% for SAB individuals). It can be seen that similar results are achieved for a non-metric method, metric methods and the current computer program-based method for SAB and SAW females, however, SAB and SAW males in the current study are not achieving great accuracies. Similar results can be seen with these studies using GM methods. Franklin and colleagues (2005) achieved accuracies ranging from 77% to 80% when assessing sex for indigenous South African individuals while Small and associates (2018) achieved accuracies ranging from 71.8% to 88.2% for SAW individuals. Once again, the lowest accuracies seen in these studies is much higher than the lowest accuracy in the current study. It can be seen that, in some cases, using 3D-ID (a GM based program) is achieving similar and even higher accurate results compared to traditional methods used in forensics, however, 3D-ID is lacking in its accuracy when assessing males. It is possible that South African males are a lot more variable compared to the population groups in 3D-ID's reference database.

SAC male's classification ranged from 50.7% to 78.7%, while SAC females range from 58.7% to 73.3%. This is contrary to what the current study found with the SAB and SAW populations where females achieved higher accuracies. This study also found that SAC females achieved lower accuracies than SAB and SAW females. This result is different to what was found by Urbanová and associates (2014), who found that individuals of admixed ancestry were more accurately sexed than individuals assigned an origin. Urbanová and colleagues (2014) suggest that it is possible that individuals of admixed ancestry follow a universal pattern of sexual dimorphism whereas individuals who were assigned a specific ancestry/origin need ancestry specific sex estimation standards, however, the results from the current study could indicate that this pattern is not the same for South African populations.

5.1.2 Ancestry

5.1.2.1 South African Black (SAB) population

When looking at the ancestry results for the SAB population, it could be argued that the SAB population is classifying as African, African American, European Southwestern and South

American because of phenotypic biological similarity between them. SAB individuals classifying as European is expected as not only is there a very large reference sample of European individuals, but also because historically Europeans used the Cape as a refreshment station and some of the Europeans settled in South Africa where admixture occurred between them and the indigenous people (Thompson, 2001). This means that SAB individuals will share some common genetics with individuals of European ancestry. While SAB individuals classifying as African American is an expected result as it is known that slaves from southern Africa were taken to America, where admixture would have occurred between the South African individuals and the American population (Thompson, 2001; van Rooyen, 2010). This could mean that SAB individuals would possibly share some genetic material with the individuals in America. It is interesting to note that some of the SAB individuals were classifying as South American. This could be due to the fact that not only was there European colonisation in South Africa, but in the 16th century European colonisation occurred in South America, where many African slaves were taken to South America which led to admixture between European, Native American and African individuals (Homburger *et al*., 2015). This admixture could have led to shared genes between these groups. So, it is possible that because of European colonisation, slave trade and admixture in South America, some South African individuals classified as South American.

When looking at the correct classifications for SAB males and females (when SAB individuals classified as African) it was seen that shape variables performs a little better than form variables. Shape variables achieved accuracies of 26.7% and 9.3% for SAB males, and 24.0% and 6.7% for SAB females when using the Full set and the Ross subset, respectively. Form variables only achieved accuracies of 18.7% and 8.0% for SAB males, and 17.3% and 4.0% for SAB females when using the Full set and the Ross subset, respectively. This shows that when including size into the analysis lower results are achieved. Similar to L'Abbé and colleagues (2013) results; this could indicate that sexual dimorphism plays an essential role in the ancestry estimation for the South African population.

3D-ID was more accurate when the analysis used the Ross subset, whereas the analysis using the Full set was not accurate. It is interesting to note that when using the Ross subset (where one of the landmarks removed was prosthion) the current study achieved higher accuracies for the SAB population. The prosthion landmark can give an indication of prognathism, which in the SAB population is quite a relevant landmark for ancestry estimation (İşcan & Steyn, 1999; van Rooyen,

2010; L'Abbé *et al*., 2011) as the SAB population has a much larger degree of prognathism compared to the other South African populations. This could also correspond with the suggestion by the 3D-ID guidelines that a reduced set of landmarks would improve accuracies (www.3d-id.org), in this case when assessing SAB individuals. In total, 40.0% to 45.0% of SAB females were classifying mostly as European Southwestern for both shape and form using the Full set, thus just short of half the SAB females. However, when using the Ross subset it was seen that 68.0% to 73.0% of SAB males and females (more than two-thirds of the population) were classifying as African American. Due to some of these differences in population being linked to the Ross subset, where some landmarks are removed, this begs the question of whether these differences could be specific to the individual; if it could be because of the size of the reference populations in 3D-ID; or if it could be related to the removal of certain landmarks, such as prosthion.

Unfortunately this study shows that 3D-ID is not an accurate tool when assessing for correct ancestry classification of SAB individuals (classifying as African only) as it only achieved accuracies ranging from 4.0% to 26.7%. Whereas other studies; such as İşcan and Steyn (1999), Franklin and colleagues (2007a) as well as Stull and associates (2014); all achieved higher accuracies ranging from 46% to as high as 95.3%. However, when comparing what the SAB individuals most classified as in 3D-ID (African American and European Southwestern) to these other studies, the current study achieved accuracies ranging from 30.7% to 73.3%, which is similar to the results achieved when using traditional methods.

King (2015) was only able to use shape variables at the time as the version of 3D-ID that was used did not have the option to examine form as well. King (2015) removed only the right orbit landmarks because at the time of the study the 3D-ID website only recommended the removal of those landmarks (right lower and right upper orbital borders). Thus the current study compared its results for the Ross subset with the shape variables with King's (2015) results. This author also included African American into the correct classifications. Thus, for comparison, the current study investigated the combined frequency of the SAB population classified as African and African American. King (2015) found that SAB female results showed an accuracy of 46.5% while SAB males resulted in an accuracy of 63.0%. Whereas, the combination of the frequencies for African and African American, the current study achieved much higher results, with SAB female and male results both exhibiting a frequency of 80.0%. The difference seen between these results could be because the current study used a newer version of 3D-ID containing a much larger reference

database; or it could be because King (2015) only removed two of the five landmarks that were omitted in the current study, affecting the reference database.

Previously, studies have applied methods developed for American samples to South African individuals because of the resemblances seen between SAB individuals and African Americans, as well as SAW individuals and White Americans. However, it has been shown that there are distinct differences between South African individuals and North Americans such as the fact that SAB and SAW individuals' cranial dimensions are less sexually dimorphic compared to the North American population. This difference is possibly due to genetic and environmental differences (İşcan & Steyn, 1999; L'Abbé *et al.*, 2013; King, 2015). Therefore, in the current study SAB individuals who are classified as African are considered correct. By doing this, the correct classification accuracies for both SAB males and females were below 10%. This very low accuracy is expected as there are unfortunately no South African individuals and only a small sample of African individuals (n = 27) in the 3D-ID reference database.

5.1.2.2 South African White (SAW) population

Overall, 3D-ID was not very accurate when assessing the SAW individuals, however, it was noted that 3D-ID was more accurate for the assessment of SAW females compared to males. SAW males, for both shape and form, using the Full set and the Ross subset were classifying mostly as European American (34.7% - 54.7%), with a second highest classification of European Southwestern (22.7% - 37.3%). SAW females, for shape variables using the Full set, were classifying mostly as European American (49.3%). While form variables using the Full set were classifying mostly as European Southwestern (52.0%). For both shape and form variables the Ross subset were classifying mostly as European Southwestern (57.3% and 58.7% respectively). It was seen that SAW males were only second most classified as what would be considered their "correct" classification as European Southwestern. While SAW females were mostly classifying as European Southwestern, their "correct" classification, except for the analysis using the Full set with shape variables. It is possible that by including the five landmarks that are removed in the Ross subset, the results for SAW females are negatively affected. It is important to note that a large number of the SAW individuals in this sample are edentulous or partially edentulous, which led to a reduced number of landmarks in that region. This could have affected the results when using 3D-ID for the SAW population as when these landmarks are available to digitise they are used in the

analysis for the Full set while they are excluded when using the Ross subset. As they are supposed to be included in the analysis when using the Full set, it's possible that those results are skewed more towards the results when excluding those landmarks.

SAW females for shape variables using the Full set classified second most as European Southwestern (44.0%). While form variables using the Full set were classifying mostly as European American (37.3%). Shape variables using the Ross subset were second most classifying as South American (16.0%) and form variables using the Ross subset were second most classifying as European American (14.7%) and South American (14.7%). As said before, it could be argued that this is partly a result of biological similarity between the SAW population and European Southwestern, European American and South American populations. SAW individuals classifying as European Southwestern is expected as it is known that SAW individuals are made up of the descendants of European groups that settled in South Africa during the European colonisation (Guelke, 1988; Steyn & İşcan, 1998; Thompson, 2001; L'Abbé *et al*., 2011; King, 2015). Similar to the results for the SAB population, it was seen that some of the SAW individuals are classifying as South American.

Compared to traditional methods, such as craniometrics, the results from the current study were rather low. İşcan and Steyn (1999) achieved accuracies ranging between 71% and 96.1%, while Stull (2014) achieved an accuracy of 80% for SAW individuals when using craniometrics. It is possible that the results in the current study are this low because the program used does not have South African individuals in the reference database. King (2015) was only able to assess shape variables and only removed the right orbit landmarks. The current study compared the results for the Ross subset and shape variables with King's (2015) results. King (2015) combined the classification of European and European American to give a correct classification for White South Africans. In the updated version of 3D-ID the European classification is now made up of European Central, European Eastern, European Southeastern and European Southwestern. So for comparison purposes with King's results, the frequencies for European American, European Southeastern and European Southwestern were combined as "correct" classifications for SAW males and females. King (2015) found that SAW females achieved an accuracy of 69.0% and SAW males achieved an accuracy of 69.7%. Whereas the current study achieved slightly higher results than King (2015), with females achieving an accuracy of 73.3% and SAW males achieving an accuracy of 70.7%.

5.1.2.3 South African Coloured (SAC) population

3D-ID's results were not good for either SAC males or females as the highest ancestry classification percentage achieved was 57.3%. These low results are likely due to the fact that the SAC population is not represented in the 3D-ID database and the population is also very genetically complex. For both SAC males and females, shape and form variables using the Full set and the Ross subset, individuals classified mostly as European Southwestern with a frequencies ranging between 34.7% and 57.3%. For both SAC males and females, when using shape variables, for both the Full set and the Ross subset, individuals are second most classified as African (14.7% for the Full set and 25.3% for the Ross subset). While when using form variables, the individuals are second most classified as South American (18.7%) when using the Full set and African (28.0%) when using the Ross subset. It is very interesting to see that the SAC population also classify as the same two correct classifications for SAB and SAW individuals, African and European Southwestern, which gives an indication of the genetic contribution of the populations. Similar to Stull (2014), it was seen that SAC individuals misclassified more as SAW individuals (European Southwestern classification) than SAB individuals (African classification).

Once again, the results could be on account of biological similarity between the SAC population and European Southwestern, African and South American populations. SAC individuals classifying as European Southwestern is not surprising as previously mentioned some Europeans settled in South Africa, where there was admixture between them and the indigenous people (Nurse *et al*., 1985; Thompson, 2001; de Wit *et al*., 2010; van Rooyen, 2010). This means that SAC individuals will share some common genetics with individuals of European ancestry (de Wit *et al*., 2010). While SAC individuals classifying as African is not an unexpected result as it is known that there were relations between colonists, indigenous Africans and African slaves resulting in the SAC population sharing genetic material with the African individuals. Similar to the results for the SAB population, it was seen that some of the SAC individuals are classifying as South American. The differences in populations are between the analysis for shape and form variables. This could give an indication that there is a difference in size between these different populations.

5.2 FORDISC

It is important to note that when the custom SADB was developed, individuals from the Raymond A. Dart Collection of Human Skeletons were digitised, which could mean that the same individuals who are included in the database could have been included in the current study, which could affect the classifications that were achieved when using FORDSIC.

5.2.1 Sex

The results for both SAB and SAW males and females achieved accuracies higher than 77% and in a forensic context these accuracies would be considered accurate as they are higher than the universally accepted standard of 75% (King, 2015). The results achieved for SAC females would also be considered accurate in a forensic context at 89.3%, however, SAC males achieved a much lower accuracy (66.7%) which would be considered a moderate accuracy.

In this study it was seen that the SAB males (82.7%) performed better than SAC males (66.7%) and SAW males (78.7%) while SAB females (85.3%) performed worse than SAC females (89.3%) and SAW females (100.0%). It was also found in this study that SAB, SAW and SAC females achieved a higher sex accuracy (85.3%, 100.0% and 89.3% respectively) compared to SAB, SAW and SAC males (82.7%, 78.7% and 66.7% respectively). This is similar to Urbanová and colleagues (2014) who found that females achieved higher accuracies, and L'Abbé and associates (2013) who found that when using the custom SADB in FORDISC, South African females achieved higher accuracies compared to South African males.

The results in this study for SAB females (85.3%) are similar and the results for SAW females (100.0%) are much higher compared to Krüger and colleagues (2014) who achieved accuracies ranging from 55% to 90% for SAB females and 31% to 62% for SAW females. Whereas, this study's results for SAB males (82.7%) and SAW males (78.7%) are lower than the accuracies achieved by Krüger and collaborators (2014) who achieved 84% to 92% for SAB males and 94% to 97% for SAW males. Similar to the current study, when using craniometrics, İşcan and Steyn (1999) found that SAW females (83.0% to 96.1%) achieved higher accuracies compared to SAW males (78.8% to 93.5%), however, İşcan and Steyn (1999) found that SAB males (83.7% to 95.3%) had higher accuracies compared to SAB females (80.0% to 91.1%), which is different to what was found in this study (82.7% for SAB males and 85.3% for SAB females). This study saw that similar results for non-metric method, metric methods and FORDISC were achieved for

SAB and SAW females, however, it is seen that SAB and SAW males in the current study tend to lower accuracies compared to the females.

When comparing the results from this study to other studies using GM on SAB or SAW populations, it is seen that FORDISC achieved higher accuracies (82.7% and 85.3% for SAB individuals and 78.7% and 100.0% for SAW individuals) compared to both Franklin and colleagues (2005), who assessed indigenous South African individuals and achieved accuracies ranging from 77% to 80%, and Small and associates (2018), who assessed SAW individuals and achieved accuracies ranging from 71.8% to 88.2%. It was also very interesting to see that SAW females achieved a correct classification accuracy of 100%. As mentioned above, during the development of the custom SADB, individuals from the Raymond A. Dart Collection of Human Skeletons were assessed, which could mean that the same individuals who are included in the database could have been included in this study, which could have led to high classifications that were achieved when using FORDSIC.

5.2.2 Ancestry

5.2.2.1 South African Black (SAB) population

The results for the estimation of ancestry for both SAB males and females achieved accuracies between 62% and 64%. King (2015) achieved an accuracy of 95.7% for the SAB population, which is considerably higher than the accuracies achieved in the current study. This is quite a surprising result as King (2015) used the FDB for analyses while the current study used the custom SADB, where one would think the custom database would achieve higher accuracies compared to the FDB. The reason for this is unknown and needs to be assessed further, however, similar to the discussion above for ancestry in 3D-ID, SAB individuals share genetic similarity with African Americans, which could have affected King's (2015) results. The current study found that SAB males achieved a higher ancestry accuracy compared to SAB females (64.0% and 62.7% respectively). SAB males and females misclassified mostly into the SAC population, which is likely due to the contribution of genes from SAB individuals into the SAC population, and in some cases, such as the assessment of stature, little to no difference is seen between SAB and SAC individuals (Steyn and Smith, 2007).

Similar to the current study, Franklin and colleagues (2007a) achieved accuracies ranging from 46% to 87% for indigenous South Africans. While studies using craniometrics, like İşcan

and Steyn (1999) and studies using both craniometrics and landmark data, such as Stull and colleagues (2014), achieved much higher accuracies for individuals from the SAB population (77.8% to 95.3%). This shows that when assessing the SAB population, FORDISC has moderate accuracies, but there are other traditional forensic methods which achieve higher accuracies.

Another thing noted when analysing the data in FORDSIC (see Appendix 5) was that 20.0% of SAB males achieved posterior probabilities that were within 10% of the posterior probabilities for the correct ancestry classification, while 26.7% of SAB females had posterior probabilities within 10% of the posterior probabilities for the correct ancestry classification. If this were to happen in forensic cases, the forensic anthropologist would err on the side of caution and not necessarily exclude the other classification as one would not be able to confidentially say that the unknown individual belongs to the assigned group. In this instance, other means of ancestry estimation would need to be looked at to come to a more conclusive answer.

5.2.2.2 South African White (SAW) population

Both SAW males and females achieved moderate accuracies of 69.3% and 62.7%, respectively, with SAW males achieving slightly higher accuracies compared to SAW females. King (2015) achieved an accuracy of 85.4% for the SAW population when using the FDB, which is higher than the accuracies achieved in the current study for the SAW population when using the custom SADB which the study's authors assumed would have achieved the higher accuracies. A large number of the SAW individuals in this study's sample are edentulous or partially edentulous which led to a reduced number of landmarks in that region. The reduced number of landmarks has also led to fewer values for ILDs and has thus reduced the number of distances that could be included in the FORDISC analysis. By reducing the number of measurements that could be included into the FORDISC analysis, there is a reduction of the sample size in the SADB that could have been used as comparison against the current study's sample; this reduced sample could mean that the amount of variation seen across the population has been reduced, which could have led to the misclassifications for SAW individuals. SAW males were mostly misclassifying as SAB individuals and SAW females were mostly misclassifying as SAC individuals. Once again, this is different to what L'Abbé and colleagues (2013) found where misclassification occurred more between the sexes rather than the ancestry classification.

An interesting observation was that SAW females are second to classify as SAC (28%) whereas SAC females second most classify as SAB (45.3%) and classify third as SAW (12%). This gives an indication that SAW females are more similar to SAC females, while strangely, SAC females are more similar to SAB females. Further research is needed to better understand this.

It is seen that FORDISC is not as accurate for ancestry estimation of SAW individuals as compared to some traditional methods that have previously been used. İşcan and Steyn (1999) achieved accuracies ranging from 71.7% to 93.5% for SAW males and 83.0% to 96.1% for SAW females, while Stull and associates (2014) achieved accuracies ranging from 80% to 93% depending on whether or not the analysis used craniometric data or landmark data.

When assessing the data in FORDSIC (see Appendix 5), it was found that 13.3% of SAW males and 10.7% of SAW females had posterior probabilities within 10% of the correct ancestry classification. As mentioned before for SAB individuals, the assessor would be cautious and not necessarily exclude the other classification which would lead observers to possibly do more analyses to estimate ancestry as these results can't necessarily be used with confidence for ancestry assessment.

5.2.2.3 South African Coloured (SAC) population

Neither SAC males nor females achieved an accuracy of over 50% for ancestry, with SAC females being correctly classified with 42.7% and SAC males only being classified correctly with 26.7%. It was noted that SAC males and females misclassified mainly into the SAB population, with SAC males being misclassified as SAB males with a frequency over 50%. This, and the fact that the accuracy rates for SAC population were so low, could be a reflection of the history of South Africa and the admixture between SAB and SAW individuals.

When comparing the FORDISC results with a study that looked at both craniometric and landmark data, it was seen that Stull and colleagues (2014) achieved higher accuracies ranging between 66% and 88%, depending on the data used, when assessing individuals from the SAC population. These results were quite surprising as these individuals were most likely included in the custom SADB during development, and so there was an assumption that the results would be skewed towards higher classifications, not lower results.

This study found that 24.0% of SAC males and 22.7% of SAC females had posterior probabilities within 10% of the correct classification for ancestry analysis (see Appendix 5). As

mentioned before, if this were to happen in forensic cases, one would err on the side of caution and would not necessarily exclude the other classification as one could not confidentially say that the unknown individual belongs to the assigned group.

5.2.3 Sex and ancestry

5.2.3.1 South African Black (SAB) population

For the joint assessment of sex and ancestry in FORDISC, SAB males only achieved an accuracy of 60% while SAB females achieved an accuracy of 49.3%. These results would not be considered accurate. Considering that the sex results for the SAB population were more accurate compared to the ancestry results, it's likely that the ancestry estimation is affecting these accuracies, which is as predicted because of the genetic variation of the SAB population. Similar to the ancestry results, SAB males achieved a higher accuracy compared to SAB females.

SAB males and females were mostly misclassified as SAC males and females. This is different to L'Abbé and colleagues (2013) who found that misclassification in the SADB occurred more between the sexes rather than the ancestry classification. When assessing SAB and SAW individuals using the custom SADB in FORDISC, L'Abbé and colleagues (2013) found that SAB males achieved an accuracy of 67.3% and SAB females achieved an accuracy of 74.2%. Not only did the current study find that females achieved a lower accuracy, L'Abbé and associates (2013) had a much higher accuracy from SAB females. Interestingly, King (2015) who assessed the SAB population using the FDB in FORDISC achieved higher accuracies (75% for SAB males and 80.2% for SAB females) than both the current study and L'Abbé and colleagues (2013), both of which used the custom SADB instead. The reason for this is unknown and most definitely needs to be assessed further.

It was found that 21.3% of SAB males and 38.7% of SAB females had posterior probabilities within 10% of the correct classification for sex and ancestry analysis (see Appendix 5). As mentioned above (4.1.2.2), this would cause some uncertainty in forensic cases where caution would be taken and the assessor would not necessarily disregard the other classification and might consider doing additional analysis before the unknown individual can be assigned a population group.

5.2.3.2 South African White (SAW) population

SAW females achieved a higher accuracy of 66.7%, while SAW males achieved an accuracy of less than 50%, which in a forensic context wouldn't be considered accurate. SAW males were mostly misclassifying as SAW females, which is similar to what L'Abbé and colleagues (2013) found that misclassification in the SADB occurred more between the sexes rather than ancestry classification. While SAW females were mostly misclassifying as SAC females, which differs from what L'Abbé and colleagues (2013) found.

When assessing SAW individuals using the custom SADB, L'Abbé and associates (2013) found that SAW males achieved an accuracy of 71.9% and SAW females achieved an accuracy of 72.7%; which are higher compared to those attained in the current study. When King (2015) assessed SAW individuals using the FDB in FORDSIC, SAW males achieved an accuracy of 66.7% and SAW females achieved an accuracy of 81.0%. All three of these studies saw that SAW males achieved poorer accuracies compared to SAW females.

Previously when looking at ancestry assessment it was seen that SAW females were classifying more as SAC females (28%), while SAC females were classifying much more as SAB females (45.3%) compared to SAW females (12%). Interestingly enough, when assessing sex and ancestry it was found that SAW females have an almost equal classification as SAC (14.7%) and SAB (13.3%); while SAC females are still more heavily classifying as SAB females (38.7%) compared to SAW females (13.3%). Is it possible that this could give an indication that the inclusion of sex in the analysis is playing a vital role in the distinction between SAW, SAB and SAC females.

When doing analysis for FORDISC it was seen that 30.7% of SAW males and 28.0% of SAW females had posterior probabilities within 10% of the correct classification for sex and ancestry analysis (see Appendix 5). As mentioned, if this were to happen in forensic cases, one would need to be extremely cautious and not exclude the other classification.

5.2.3.3 South African Coloured (SAC) population

Similar to the ancestry results, neither SAC males nor females achieved an accuracy of over 40% for the joint estimation of sex and ancestry. Because this study demonstrates low ancestry results, it can be said that these low accuracies once again are a reflection of South African history. Not only did SAC males achieve a lower accuracy compared to SAC females, but it was

seen that the SAC population achieved very low accuracies compared to both SAB and SAW individuals. This study saw that SAC males were misclassifying mostly as SAB males, while SAC females were misclassifying as SAB female with the same frequency as they were classifying as SAC females (38.7%). This is not an unexpected result as it has been shown that 32% to 45% of the genes in the SAC population come from Khoesan individuals and 20% to 36% of the genes in the SAC population come from Bantu-speaking African individuals (de Wit *et al*., 2010).

It was noted that 37.3% of SAC males and 25.3% of SAC females had posterior probabilities within 10% of the correct classification for sex and ancestry analysis (see Appendix 5). As previously mentioned, this would cause some uncertainty in forensic cases where it would be necessary to be cautious and not disregard the other classification and one might consider doing additional analysis before the unknown individual can be assigned a population group.

5.3 3D-ID VS FORDISC

5.3.1 Sex

This study found that when assessing sex for SAB males and SAC females FORDISC achieved higher accuracies (82.7% for SAB males and 89.3% for SAC females) compared to 3D-ID (44.0% to 69.3% for SAB males and 58.7% to 73.3% for SAC females) (Table 5.1). While for SAW females FORDISC (100.0%) achieved higher accuracies, if ever so slightly, compared to 3D-ID (81.3% to 93.3%). Sex estimations for SAB females and SAW and SAC males was a bit different. The study found that for SAB females and SAW males the analyses using the Ross subset with form variables in 3D-ID performed better (89.3% for SAB females and 85.3% for SAW males) compared to FORDISC (85.3% for SAB females and 78.7% for SAW males). It's possible that these analyses in 3D-ID achieved higher accuracies because it was the analysis using form variables, which could give an indication that size plays a vital role in these populations which is not necessarily assessed in FORDISC as FORDISC uses linear measurements. Additionally, this study found that for SAC males the analyses using the Ross subset with both shape and form variables in 3D-ID performed better with accuracies of 70.7% and 78.7% compared to FORDISC's accuracy of 66.7%. It is possible that the sample collected for the custom SADB does not capture all the possible variation for SAC males, whereas 3D-ID's database may contain more variation that is similar to the variation seen in SAC males which is leading to high classifications in 3D-ID.

Table 5.1: Comparison of sex classifications when using 3D-ID and FORDISC

SEX	3D-ID	FORDISC
	SAB population	
Males	44.0% – 69.3%	82.7%
Females	80.0% – 89.3%	85.3%
	SAW population	
Males	52.0% – 85.3%	78.7%
Females	81.3% – 93.3%	100.0%
	SAC population	
Males	50.7% – 78.7%	66.7%
Females	66.7% – 73.3%	89.3%

5.3.2 Ancestry

3D-ID achieved accuracies between 4.0% to 58.7% for correct classification for SAB and SAW individuals that is lower than FORDISC (Table 5.2), which achieved accuracies ranging between 62.7% and 69.3% for SAB and SAW individuals. This was expected as FORDISC has a custom database for South African individuals whereas 3D-ID does not.

Table 5.2: Comparison of "correct" ancestry classifications for SAB and SAW populations when using 3D-ID and FORDISC

SEX	3D-ID	FORDISC
	SAB population	
Males	8.0% – 26.7%	64.0%
Females	4.0% – 24.0%	62.7%
	SAW population	
Males	22.7% – 37.3%	69.3%
Females	44.0% – 58.7%	62.7%

It was noted that SAC males and females where being classified mostly as European Southwestern in 3D-ID, while when running SAC males and females in FORDISC they were classifying mostly as SAB males and females. It would have been expected that 3D-ID would rather have classified the SAC individuals as African or African American if the similarities were shared, as seen in FORDISC. This could give an indication that maybe the African and African American sample in 3D-ID are not comparable with the SAB sample in the custom SADB in FORDISC. This would then support studies, such as L'Abbe and colleagues (2013) as well as King (2015), that have found that American individuals and South African individuals have distinctly different cranial characteristics.

Oddly, it was seen that FORDISC achieved low accuracies of 26.7% and 42.7% for SAC individuals which was rather unexpected as FORDISC has SAC individuals in its custom SADB, so an assumption was made that higher accuracies would have been achieved. When assessing the SAC population in 3D-ID, no correct classification was given as there are no populations in 3D-ID's reference database that represent the SAC population. However it was noted that when using 3D-ID SAC individuals classified mostly as European Southwestern with frequencies of 34.7% to 57.3%, which is moderately higher than the accuracies achieved when using FORDISC. Overall, for the SAC population, neither FORDISC nor 3D-ID provided adequate accuracies for ancestry estimation. This is most likely due to the fact that the SAC population has such a diverse genetic history, and even though FORDISC has SAC individuals in the custom SADB, it would still be a challenge to assess such a complex group.

5.4 LIMITATIONS AND FUTURE DIRECTIONS

5.4.1 Limitations

A large number of the SAW individuals in this sample are edentulous or partially edentulous which led to a reduced number of landmarks in that region. This could have affected 3D-ID's results for the SAW population as when these landmarks are available to digitise they are used in the analysis for the Full set while they are excluded when using the Ross subset. As they are supposed to be included in the analysis using the Full set, it's possible that those results are skewed more towards the results for the Ross subset. The reduced number of landmarks has also led to fewer values for ILDs and has thus reduced the number of distances that could be included in the FORDISC analysis, however, just because there is a reduced number of measurements for

the edentulous individuals, does not necessarily mean that including edentulous individuals in the current sample led to lower accuracies. It is possible that the mandibular and maxillary areas may have a strong ability to differentiate between population groups, but because these landmarks are not available it is not possible to know for sure.

Another limitation is that there are a number of individuals of unknown sex and some of unknown ancestry that are included in the 3D-ID reference database. It is possible that this would have some effect on the estimation of sex and ancestry by the program. Unfortunately, the program doesn't mention when these individuals are used in the analysis, so it cannot be conclusively stated when they are used and if they are contributing to the study's misclassifications. It's believed that removing these individuals from the database and including more individuals with known demographics would possibly result in a better comparative database in 3D-ID.

Another limiting factor was that the individuals in the custom SADB could be the same individuals that were used in the current study's sample as data were collected using the same collections. This could have led to the current studies results being skewed towards correct classifications in FORDISC, but unfortunately because of the limited number of individuals for SAW population, all data collected had to be included in order to have a sufficient sample size for comparison purposes.

5.4.2 Future directions

The authors of this study have submitted the South African data collected during this study to the developers of 3D-ID to be included as part of the reference database, after which a validation study could be conducted to assess how including individuals from the South African population would change the results currently being seen in 3D-ID. Furthermore, the database could be expanded even further to include more African populations, not just South African populations. The inclusion of postcrania could also be considered. Another future consideration would be to examine why the FDB in FORDISC achieves better classifications compared to the custom SADB when assessing South African populations.

CHAPTER 6: CONCLUSION

The results from this study have led to the following conclusions:

1. For sex estimation using shape variables:
    a. 3D-ID was able to accurately estimate sex for SAB females (80.0% and 82.7%) and SAW females (89.3% and 93.3%), but not for SAB males (44.0% and 56.0%).
    b. 3D-ID achieved a good sex estimate for SAW males when using the Ross subset (76.0%) whereas, when using the Full set, the accuracy decreased (52.0%).
    c. 3D-ID accurately estimate sex for SAC females when assessing the Full set (70.7%) and for SAC males when using the Ross subset (70.7%).
    d. 3D-ID was not able to accurately estimate sex for SAC females for the Ross subset (58.7%) and for SAC males using the Full set (50.7%).
2. For sex estimation using form variables:
    a. 3D-ID was able to accurately estimate sex for SAB females (80.0% and 89.3%) and SAW females (81.3% and 88.0%), whereas it was not able to accurately estimate sex for SAB males (50.7% and 69.3%).
    b. 3D-ID was able to achieve a good sex estimate for SAW males when using the Ross subset (85.3%), whereas when using the Full set the accuracy decreased (65.3%).
    c. 3D-ID was able to accurately estimate sex for SAC females when assessing the Full set (73.3%) and for SAC males when using the Ross subset (78.7%), whereas it was not able to accurately estimate sex for SAC females for the Ross subset (66.7%) and for SAC males using the Full set (52.0%).
3. For ancestry estimation using shape variables:
    a. Correct ancestry classification for SAB and SAW individuals was poor with accuracies ranging from 6.7% to 26.7% for SAB individuals and 25.3% to 57.3% for SAW individuals.
4. For ancestry estimation using form variables:
    a. SAB and SAW individuals classified poorly with accuracies ranging from 4.0% to 18.7% for SAB individuals, and from 22.7% to 58.7% for SAW individuals.
5. This study assessed the group classification assigned to SAB, SAC and SAW individuals using the current 3D-ID database and found that SAB males and females classified as European Southwestern and African American with the highest frequency depending on

shape and form variables. SAC males and females and SAW females classified with the highest frequency as European Southwestern. SAW males classified as European American.

6. FORDISC correctly classified SAB males and females with 62.7% and 64.0%, correctly classified SAW males and females with 62.7% and 69.3%, and correctly classified SAC males and females with 26.7% and 42.7%. These classifications are all low in comparison to other studies using FORDISC for the South African population, especially the results for the SAC population.
7. The study found that when assessing sex for SAB males, SAC females and SAW females, FORDISC achieved higher accuracies compared to 3D-ID. It was found that for SAB females and SAW males the analyses using the Ross subset with form variables in 3D-ID performed better than FORDISC. It was also found that in SAC males the analyses in 3D-ID using the Ross subset with both shape and form variables performed better than FORDISC.
8. It was found that 3D-ID achieved lower accuracies for SAB and SAW individuals compared to FORDISC. This could be due to the fact that FORDISC has a custom database for South African individuals, whereas 3D-ID does not.
9. This study also determined that for the SAC population overall, neither FORDISC nor 3D-ID provided adequate accuracies for ancestry estimation. This is likely due to the fact that the SAC population is extremely diverse, and even though FORDISC has SAC individuals in the custom SADB, it is still a challenge to assess such a complex group.

This study demonstrated that FORDISC achieved slightly higher accuracies compared to 3D-ID. However, FORDISC has a custom database for South African individuals, whereas 3D-IDs reference database contains no South African data. Overall, the performance of the software 3D-ID had moderate classifications indicating that it is not an accurate tool for the sex and ancestry estimation of South African populations. These low accuracies are likely because 3D-ID is forced to assign an individual to one of the reference population groups in the software, and so the program may not be performing very well because the South African population is not included in the reference samples. It is assumed that 3D-ID would perform much better when classifying

individuals that are included in the reference populations, and so further research is needed to test this assumption.

REFERENCES

Abdel Fatah, E.E., Shirley, N.R., Jantz, R.L. and Mahfouz, M.R. (2014). Improving sex estimation from crania using a novel three-dimensional quantitative method. *Journal of Forensic Sciences* **59(3)**: 590–600.

Adams, D.C., Rohlf, F.J. and Slice, D.E. (2004). Geometric morphometrics: ten years of progress following the 'revolution'. *Italian Journal of Zoology* **71(1)**: 5–16.

Alblas, A., Greyling, L.M. and Geldenhuys, E.M. (2018). Composition of the Kirsten skeletal collection at Stellenbosch University. *South African Journal of Science* **114(1-2)**:1–6.

Albanese, J. and Saunders, S.R. (2006). Is it possible to escape racial typology in forensic identification? In *Forensic Anthropology and Medicine* (pp. 281-316). Humana Press.

Anastasiou, E. and Chamberlain, A.T. (2013). The sexual dimorphism of the sacro-iliac joint: An investigation using geometric morphometric techniques. *Journal of Forensic Sciences* **58**: 126–134.

Barrier, I.L.O and L'Abbé, E.N. (2008). Sex determination from the radius and ulna in a modern South African sample. *Forensic Science International* **179(1)**: 85.e1–85.e7.

Bertsatos, A., Papageorgopoulou, C., Valakos, E. and Chovalopoulou, M.E. (2018). Investigating the sex-related geometric variation of the human cranium. *International Journal of Legal Medicine* **132(5)**: 1505–1514.

Bertsatos, A., Christaki, A. and Chovalopoulou, M.E. (2019). Testing the reliability of 3D-ID software in sex and ancestry estimation with a modern Greek sample. *Forensic Science International* **297**: 132–137.

Bidmos, M.A., Gibbon, V.E. and Štrkalj, G. (2010). Recent advances in sex identification of human skeletal remains in South Africa. *South African Journal of Science* **106(11–12)**: 1–6.

Bigoni, L., Veleminska, J. and Bruzek, J. (2010). Three-dimensional geometric morphometric analysis of cranio-facial sexual dimorphism in a Central European sample of known sex. *HOMO- Journal of Comparative Human Biology* **61(1)**: 16–32.

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

Buikstra, J.E. and Ubelaker, D.H. (1994). Standards for Data Collection from Human Skeletal Remains: Proceedings of a Seminar at the Field Museum of Natural History. Fayetteville: Arkansas Archaeological Series.

Bytheway, J.A. and Ross, A.H. (2010). A geometric morphometric approach to sex determination of the human adult os coxa. *Journal of Forensic Sciences* **55(4)**: 859–864.

Cattaneo, C. (2007). Forensic anthropology: developments of a classical discipline in the new millennium. *Forensic Science International* **165(2–3)**: 185–193.

Cavaignac, E., Savall, F., Faruch, M., Reina, N., Chiron, P. and Telmon, N. (2016). Geometric morphometric analysis reveals sexual dimorphism in the distal femur. *Forensic Science International* **259**: 246.e1–246.e5.

Chovalopoulou, M.E., Valakos, E.D. and Manolis, S.K. (2016). Sex determination by three-dimensional geometric morphometrics of the vault and midsagittal curve of the neurocranium in a modern Greek population sample. *HOMO - Journal of Comparative Human Biology* **67(3)**: 173–187.

Christensen, A.M., Passalacqua, N.V. and Bartelink, E.J. (2014). Forensic anthropology: current methods and practice. Elsevier.

Christensen, A.M. and Passalacqua, N.V. (2018). A Laboratory Manual for Forensic Anthropology. Academic Press.

Dawson, C., Ross, D. and Mallett, X. (2011). Sex determination. In Black, S.M. and Ferguson, E. (eds.) *Forensic Anthopology 2000 to 2010*. CRC Press.

Dayal, M.R., Spocter, M.A. and Bidmos, M.A. (2008). An assessment of sex using the skull of black South Africans by discriminant function analysis. *HOMO- Journal of Comparative Human Biology* **59(3)**: 209–221.

Dayal, M.R., Kegley, A.D.T., Štrkalj, G., Bidmos, M.A. and Kuykendall, K.L. (2009). The history and composition of the Raymond A. Dart Collection of Human Skeletons at the University of the Witwatersrand, Johannesburg, South Africa. *American Journal of Physical Anthropology* **140(2)**: 324–335.

De Villiers, H. (1968). Sexual dimorphism of the skull of the South African bantu-speaking negro. *South African Journal of Science* **64(2)**: 118–124.

de Wit, E., Delport, W., Rugamika, C.E., Meintjes, A., Möller, M., van Helden, P.D., Seoighe, C. and Hoal, E.G. (2010). Genome-wide analysis of the structure of the South African Coloured Population in the Western Cape. *Human Genetics* **128(2)**: 145–153.

DiGangi, E.A. and Hefner, J.T. (2012). Ancestry estimation. In DiGangi, E.A. and Moore, M.K. (eds.) *Research Methods in Human Skeletal Biology.* Academic Press.

Dirkmaat, D. (2012) A companion to forensic anthropology. Wiley-Blackwell.

Dudzik, B. and Jantz, R.L. (2016). Misclassifications of Hispanics using FORDISC 3.1: comparing cranial morphology in Asian and Hispanic populations. *Journal of Forensic Sciences* **61(5)**: 1311–1318.

Elliott, M. and Collard, M. (2009). FORDISC and the determination of ancestry from cranial measurements. *Biology Letters* **5(6)**: 849–852.

Franklin, D., Freedman, L. and Milne, N. (2005). Three-dimensional technology for linear morphological studies: A re-examination of cranial variation in four southern African indigenous populations. *HOMO- Journal of Comparative Human Biology* **56(1)**: 17–34.

Franklin, D., Freedman, L., Milne, N. and Oxnard, C.E. (2006). A geometric morphometric study of sexual dimorphism in the crania of indigenous southern Africans. *South African Journal of Science* **102(5-6)**: 229–238.

Franklin, D., Freedman, L., Milne, N. and Oxnard, C.E. (2007a). Geometric morphometric study of population variation in indigenous southern African crania. *American Journal of Human Biology* **19(1)**: 20–33.

Franklin, D., Oxnard, C.E., O'Higgins, P. and Dadour, I. (2007b). Sexual dimorphism in the subadult mandible: quantification using geometric morphometrics. *Journal of Forensic Sciences* **52(1)**: 6–10.

Franklin, D., O'Higgins, P., Oxnard, C.E., Dadour, I. (2008). Discriminant function sexing of the mandible of Indigenous South Africans. *Forensic Science International*. **179(1)**: 84.e1–84.e5.

Garvin, H.M., Sholts, S.B. and Mosca, L.A. (2014). Sexual dimorphism in human cranial trait scores: effects of population, age, and body size. *American Journal of Physical Anthropology* **154(2)**: 259–269.

Giles, E. & Elliot, O. (1963). Sex determination by discriminant function analysis of crania. *American Journal of Physical Anthropology* **21**: 53–68.

Gillick, H. (2012). Ancestry determination using geometric morphometrics.[Master of Science Thesis]. University of Dundee, Scotland.

Gonzalez, P.N., Bernal, V. and Perez, S.I. (2009). Geometric morphometric approach to sex estimation of human pelvis. *Forensic Science International* **189(1–3)**: 68–74.

Gonzalez, P.N., Bernal, V. and Perez, S.I. (2011). Analysis of sexual dimorphism of craniofacial

traits using geometric morphometric techniques. *International Journal of Osteoarchaeology* **21(1)**: 82–91.

Green, H. and Curnoe, D. (2009). Sexual dimorphism in Southeast Asian crania: A geometric morphometric approach. *HOMO- Journal of Comparative Human Biology* **60(6)**: 517–534.

Guelke, L. (1998). The anatomy of a Colonial Settler Population: Cape Colony 1657-1750. *The International Journal of African Historical Studies* **21(3)**: 453–473.

Guyomarc'h, P. and Bruzek, J. (2011). Accuracy and reliability in sex determination from skulls: a comparison of FORDISC® 3.0 and the discriminant function analysis. *Forensic Science International* **208(1-3)**: 180-e1–180.e6.

Hefner, J.T. (2003). Assessing nonmetric cranial traits currently used in forensic determination of ancestry. [Doctoral Dissertation]. University of Florida.

Hefner, J.T. (2007). The statistical determination of ancestry using cranial nonmetric traits. [Doctoral Dissertation]. University of Florida.

Hefner, J.T. (2009). Cranial nonmetric variation and estimating ancestry. *Journal of Forensic Sciences* **54(5)**: 985–995.

Hefner, J.T. and Ousley, S.D. (2014). Statistical classification methods for estimating ancestry using morphoscopic traits. *Journal of Forensic Sciences* **59(4)**: 883–890.

Homburger, J.R., Moreno-Estrada, A., Gignoux, C.R., Nelson, D., Sanchez, E., Ortiz-Tello, P., Pons-Estel, B.A., Acevedo-Vasquez, E., Miranda, P., Langefeld, C.D. and Gravel, S. (2015). Genomic insights into the ancestry and demographic history of South America. *PLoS Genetics* **11(12)**: e1005602.

Hughes, C.E., Juarez, C.A., Hughes, T.L., Galloway, A., Fowler, G. and Chacon, S. (2011). A simulation for exploring the effects of the "trait list" method's subjectivity on consistency and accuracy of ancestry estimations. *Journal of Forensic Sciences* **56(5)**: 1094–1106.

Humphries, A.L. and Ross, A.H. (2011). Craniofacial sexual dimorphism in two Portuguese skeletal samples. *Anthropologie (1962-)* **49(1)**: 13–20.

Isaacs-Martin, W. and Petrus, T. (2012). The multiple meanings of coloured identity in South Africa. *Africa Insight* **42(1)**: 87–102.

İşcan, M.Y., Yoshino, M. and Kato, S. (1995). Sexual dimorphism in modern Japanese crania. *American Journal of Human Biology* **7(4)**: 459–464.

İşcan, M.Y. and Steyn, M. (1999). Craniometric determination of population affinity in South

Africans. *International Journal of Legal Medicine* **112(2)**: 91–97.

İşcan, M.Y. and Steyn, M. (2013). *The human skeleton in forensic medicine*. Charles C Thomas Publisher.

Jantz, R.L. and Ousley, S.D. (2005). FORDISC 3.1: Personal computer forensic discriminant functions. Knoxville, TN: University of Tennessee.

Kendall, D. (1977). The Diffusion of Shape. *Advances in Applied Probability* **9(3)**: 428–430.

Kimmerle, E.H., Ross, A. and Slice, D. (2008). Sexual dimorphism in America: Geometric morphometric analysis of the craniofacial region. *Journal of Forensic Sciences* **53(1)**: 54–57.

King, R.E. (2015). Estimating ancestry in South Africa: a comparison of geometric morphometrics and traditional craniometrics [Master of Science thesis] Boston University, Massachusetts.

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

Klingenberg, C.P. (2013). Visualizations in geometric morphometrics: How to read and how to make graphs showing shape changes. *Hystrix, the Italian Journal of Mammalogy* **24(1)**: 15–24.

Komar, D.A & Buikstra, J.E (2008). *Forensic Anthropology: Contemporary Theory and Practice.* Oxford University Press.

Kranioti, E.F., İşcan, M.Y. and Michalodimitrakis, M. (2008). Craniometric analysis of the modern Cretan population. *Forensic Science International* **180(2–3)**: 1–5.

Kranioti, E.F., Bastir, M., Sánchez-Meseguer, A. and Rosas, A. (2009). A geometric-morphometric study of the Cretan humerus for sex identification. *Forensic Science International* **189(1–3)**: 111.e1–111.e8.

Krogman, W.M. and İşcan, M.Y. (1986). *The human skeleton in forensic medicine*. Springfield: Charles C Thomas.

Krüger, G.C., L'Abbé, E.N., Stull, K.E. and Kenyhercz, M.W. (2015). Sexual dimorphism in cranial morphology among modern South Africans. *International Journal of Legal Medicine* **129(4)**: 869–875.

L'Abbé, E.N., Van Rooyen, C., Nawrocki, S.P. and Becker, P.J. (2011). An evaluation of non-metric cranial traits used to estimate ancestry in a South African sample. *Forensic Science*

International **209(1–3)**: 195.e1–195.e7.

L'Abbé, E.N., Kenyhercz, M., Stull, K.E., Keough, N. and Nawrocki, S. (2013). Application of FORDISC 3.0 to explore differences among crania of north american and south african blacks and whites. *Journal of Forensic Sciences* **58(6)**: 1579–1583.

Langley, N.R. and Tersigni-Tarrant, M.A. (2017). *Forensic anthropology: A comprehensive introduction*. CRC Press.

Latham, K.E., Bartelink, E.J. and Finnegan, M. eds., 2018. New perspectives in forensic human skeletal identification. Academic Press.

Liebenberg, L., Stull, K.E., L'Abbé, E.N. and Botha, D. (2015). Evaluating the Accuracy of Cranial Indices in Ancestry Estimation Among South African Groups. *Journal of Forensic Sciences* **60(5)**: 1277–1282.

Lynnerup, N. (2013). Forensic anthropology and human identification. *Scandinavian Journal of Forensic Science* **19(1)**: 16–38.

Manthey, L., Jantz, R.L., Vitale, A. and Cattaneo, C. (2018). Population specific data improves FORDISC's performance in Italians. *Forensic Science International* **292**: 263.e1–263.e7.

Marinescu, M., Panaitescu, V., Rosu, M., Maru. N. and Punga, A. (2014). Sexual dimorphism of crania in a Romianian population: Discriminant function analysis approach for sex estimation. *Romanian Journal of Legal Medicine* **22(1)**: 21–26.

Mitteroecker, P. & Gunz, P. (2009). Advances in Geometric morphometrics. *Evolutionary Biology* **36(2)**: 235–247.

Moore-Jansen, P.M., Ousley, S.D., Jantz, R.L. (1994). Data collection procedures for forensic skeletal material. Report of Investigations No. 48. Department of Anthropology, University of Tennessee, Knoxville.

Moore, M.K., DiGangi, E.A., Ruíz, F.P.N., Davila, O.J.H. and Medina, C.S. (2016). Metric sex estimation from the postcranial skeleton for the Colombian population. *Forensic Science International* **262**: 286.e1–286.e8.

Murphy, R.E. and Garvin, H.M. (2018). A morphometric outline analysis of ancestry and sex differences in cranial shape. *Journal of Forensic Sciences* **63(4)**: 1001–1009.

Nafte, M. (2000). *Flesh and bone: An introduction to forensic anthropology*. Carolina Academic Press.

Nurse, G.T., Weiner, J.S. and Jenkins, T. (1985). *The peoples of southern Africa and their*

affinities. Oxford University Press.

Oettlé, A.C., Pretorius, E. and Steyn, M. (2005). Geometric morphometric analysis of mandibular ramus flexure. *American Journal of Physical Anthropology* **128(3)**: 623–629.

Ogawa, Y., Imaizumi, K., Miyasaka, S. and Yoshino, M. (2013). Discriminant functions for sex estimation of modern Japanese skulls. *Journal of Forensic and Legal Medicine* **20(4)**: 234–238.

Ousley, S., Jantz, R. and Freid, D. (2009). Understanding race and human variation: why forensic anthropologists are good at identifying race. *American Journal of Physical Anthropology* **139(1)**: 68–76.

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

Pretorius, E., Steyn, M. and Scholtz, Y. (2006). Investigation into the usability of geometric morphometric analysis in assessment of sexual dimorphism. *American Journal of Physical Anthropology* **129(1)**: 64–70.

Quintana-Murci, L., Harmant, C., Quach, H., Balanovsky, O., Zaporozhchenko, V., Bormans, C., van Helden, P.D., Hoal, E.G. and Behar, D.M. (2010). Strong maternal Khoisan contribution to the South African coloured population: a case of gender-biased admixture. *The American Journal of Human Genetics* **86(4)**: 611–620.

Randolph-Quinney, P.S., Mallett, X. and Black, S.M. (2009). Forensic Anthropology. *Wiley Encyclopedia of Forensic Science* 1–27.

Relethford, J.H. (2009). Race and global patterns of phenotypic variation. *American Journal of Physical Anthropology* **139(1)**: 16–22.

Rhine, S. (1990). Non-metric skull racing. *Skeletal attribution of race: Methods for forensic anthropology* **4**: 9–20.

Rightmire, G.P. (1971). Discriminant function sexing of Bushman and South African Negro crania. *The South African Archaeological Bulletin* **26(103/104)**: 132–138.

Rohlf, F.J. and Slice, D.E. (1990). Extensions of the procrustes method for the optimal superimposition of landmarks. *Systematic Zoology* **39(1)**: 40–59.

Rogers, T.L. (2005). Determining the sex of human remains through cranial morphology. *Journal of Forensic Scienc*e **50(3)**: 1–8.

Rosas, A. and Bastir, M. (2002). Thin-plate spline analysis of allometry and sexual dimorphism

in the human craniofacial complex. *American Journal of Physical Anthropology* **117(3)**: 236–245.

Ross, A.H., Slice, D.E., Ubelaker, D.H. and Falsetti, A.B. (2004). Population affinities of 19th century Cuban crania: implications for identification criteria in south Florida Cuban Americans. *Journal of Forensic Science* **49**: 1–6.

Sauer, J. (1992). Forensic anthropology and the concept of race: If races don't exist, why are forensic anthropologists so good at identifying them? *Social Science & Medicine* **34**: 107–111.

Scholtz, Y., Steyn, M. and Pretorius, E. (2010). A geometric morphometric study into the sexual dimorphism of the human scapula. *Homo* **61(4)**: 253–270.

Slice, D.E. (2007). Geometric morphometrics. *Annual Review of Anthropology* **36**: 261–281.

Slice, D.E. and Ross, A. (2009). 3D-ID: geometric morphometric classification of crania for forensic scientists. Available from: http://www.3d-id.org/downloads.

Small, C. (2016). Sexual dimorphism in white South African crania [Doctor of Philosophy thesis] University of the Wiwatersrand, Johannesburg.

Small, C., Schepartz, L., Hemingway, J. and Brits, D. (2018). Three-dimensionally derived interlandmark distances for sex estimation in intact and fragmentary crania. *Forensic Science International* **287**: 127–135.

South African National Department of Health (1983). Human Tissues Act. *Government Gazette* **216**:1–34.

South African National Department of Health. National policy for Healtyh Act, 2003 (Act No. 61 of 2003)(2004). *Government Gazette* **469**: 66–79.

Spradley, M., Jantz, R., Robinson, A. and Peccerelli, F. (2008). Demographic change and forensic identification: problems in metric identification of Hispanic skeletons. *Journal of Forensic Science* **53**: 21–28.

Spradley, M.K. and Jantz, R.L. (2011). Sex estimation in forensic anthropology: skull versus postcranial elements. *Journal of Forensic Sciences* **56(2)**: 289-296.

Spradley, M.K. (2013). Project IDENTIFICATION: developing accurate identification criteria for Hispanics. *Washington, DC: US Department of Justice.*

Spradley, M.K. and Jantz, R.L. (2016). Ancestry estimation in forensic anthropology: geometric morphometric versus standard and nonstandard interlandmark distances. *Journal of*

Forensic Sciences **61(4)**: 892–897.

Spradley, M.K. and Stull, K.E. (2018). Advancements in sex and ancestry estimation. In *New Perspectives in Forensic Human Skeletal Identification* pp. 13–21. Academic Press.

Statistics South Africa, (2019a). Governance, Public Safety and Justice Survey: Statistical release. Statistics South Africa, Pretoria.

Statistics South Africa, (2019b). Mid-year population estimates: Statistical release. Statistics South Africa, Pretoria.

Steyn, M. and İşcan, M.Y. (1997). Sex determination from the femur and tibia in South African whites. *Forensic Science International* **90(1–2)**: 111–119.

Steyn, M. and İşcan, M.Y. (1998). Sexual dimorphism in the crania and mandibles of South African whites. *Forensic Science International* **98(1–2)**: 9–16.

Steyn, M. and İşcan, M.Y. (1999). Osteometric variation in the humerus: sexual dimorphism in South Africans. *Forensic Science International* **106**: 77–85.

Steyn, M., Pretorius, E., Hutten, L. (2004). Geometric morphometric analysis of the greater sciatic notch in South Africans. *HOMO- Journal of Comparative Human Biology* **54(3)**:197–206.

Steyn, M. and Smith, J.R. (2007). Interpretation of ante-mortem stature estimates in South Africans. *Forenisc Science International* **171**: 97–102.

Steyn, M. and Patriquin, M.L. (2009). Osteometric sex determination from the pelvis – does population specificity matter?. *Forensic Science International* **191(1-3)**: 113.e1–113.e5.

Stull, K.E., Kenyhercz, M.W. and L'Abbé, E.N. (2014). Ancestry estimation in South Africa using craniometrics and geometric morphometrics. *Forensic Science International* **245**: 206.e1–206.e7.

Thompson, L.M. (2001). *A history of South Africa*. Yale University Press.

Urbanová, P., Hejna, P., Zátopková, L. and Šafr, M. (2013). What is the appropriate approach in sex determination of hyoid bones?. *Journal of Forensic and Legal Medicine* **20(8)**: 996–1003.

Urbanová, P., Ross, A.H., Jurda, M. and Nogueira, M.I. (2014). Testing the reliability of software tools in sex and ancestry estimation in a multi-ancestral Brazilian sample. *Legal Medicine* **16(5)**: 264–273.

van Rooyen, C. (2010). Evaluating standard non-metric cranial traits used to determine ancestry

on a South African sample [Doctoral dissertation] University of Pretoria, South Africa.

Vance, V.L. and Steyn, M. (2013). Geometric morphometric assessment of sexually dimorphic characteristics of the distal humerus. *Homo* **64(5)**: 329–340.

Walker, P.L. (2008). Sexing skulls using discriminant function analysis of visually assessed traits. *American Journal of Physical Anthropology* **136(1)**: 39–50.

Walrath, D.E., Turner, P. and Bruzek, J. (2004). Reliability test of the visual assessment of cranial traits for sex determination. *American Journal of Physical Anthropology* **125(2)**: 132–137.

White, T.D., Black, M.T and Folkens, P.A. (2012). *Human Osteology* (3rd ed.). San Diego: Academic Press.

Williams, B.A. and Rogers, T.L. (2006). Evaluating the accuracy and precision of cranial morphological traits for sex determination. *Journal of Forensic Sciences* **51(4)**: 729–735.

Zelditch, M.L., Swiderski, D.L., Sheets, H.D. and Fink, W.L. (2004). *Geometric Morphometrics For Biologists: A Primer*. New York: Elsevier Academic Press.

APPENDICES

Appendix 1: Ethical clearance for the Raymond A. Dart Collection, University of the Witwatersrand, Johannesburg.

UNIVERSITY OF THE WITWATERSRAND JOHANNESBURG | Anatomical Sciences

SCHOOL OF ANATOMICAL SCIENCES ETHICS WAIVER CLEARANCE LETTER

Faculty of Health Sciences
School of Anatomical Sciences
University of the
Witwatersrand Johannesburg

Re: In terms of Chapter 8, sections 62-64 of the National Health Act No 61 of 2003 donated bodies and their tissues may be used for, among other purposes, health and research. Use of such Material is subject only to permission from the responsible person in the School of Anatomical Sciences – the Head or person designed by the Head.

Human Research Ethics Committee (Medical) Clearance Certificate:

W-CJ-140604-1

This letter serves to confirm that the Head of School, based in the School of Anatomical Sciences, Faculty of Health Sciences, has reviewed the research proposal entitled: Sex and ancestry estimation for South African crania using 3D-ID and has granted clearance to access the blanket ethics waiver to conduct the abovementioned research study.

28/10/2019
Dated

Professor Maryna Steyn
Head of School
School of Anatomical Sciences
Health Sciences Faculty

Faculty of Health Sciences
Medical School | 7 York Road Parktown 2193, South Africa | T 27 (0) 11 717-2713 | www.wits.ac.za/anatomicalsciences

Appendix 2: Plagiarism Declaration



PLAGIARISM DECLARATION TO BE SIGNED BY ALL HIGHER DEGREE STUDENTS

SENATE PLAGIARISM POLICY: APPENDIX ONE

I Tamara Leigh Lottering (Student number: 815691) am a student registered for the degree of Master of Science in Medicine (Dissertaton) in the academic year 2020.

I hereby declare the following:

- I am aware that plagiarism (the use of someone else's work without their permission and/or without acknowledging the original source) is wrong.
- I confirm that the work submitted for assessment for the above degree is my own unaided work except where I have explicitly indicated otherwise.
- I have followed the required conventions in referencing the thoughts and ideas of others.
- I understand that the University of the Witwatersrand may take disciplinary action against me if there is a belief that this is not my own unaided work or that I have failed to acknowledge the source of the ideas or words in my writing.
- I have included as an appendix a report from "Turnitin" (or other approved plagiarism detection) software indicating the level of plagiarism in my research document.

Signature: _____ Date: 25/03/2020

Appendix 3: Turn-it-in report

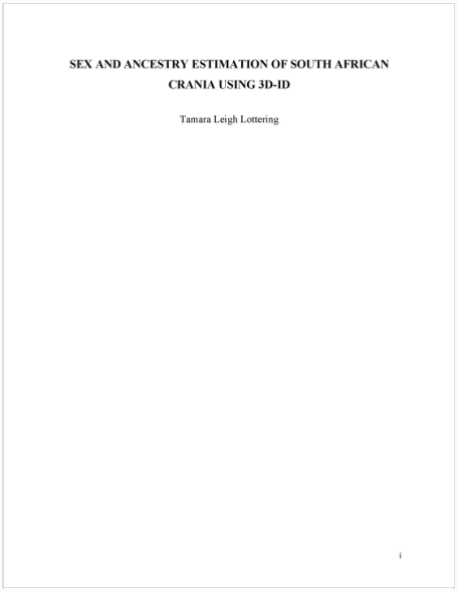
turnitin

Digital Receipt

This receipt acknowledges that Turnitin received your paper. Below you will find the receipt information regarding your submission.

The first page of your submissions is displayed below.

Submission author:	Tamara Lottering
Assignment title:	/ANAT8001A _ 2019
Submission title:	815691:Tamara_Lottering_-_Turn-i…
File name:	80ce-8591512fc2ed_Tamara_Lotte…
File size:	199.04K
Page count:	81
Word count:	27,689
Character count:	149,910
Submission date:	03-Dec-2019 01:45PM (UTC+0200)
Submission ID:	1225938144



SEX AND ANCESTRY ESTIMATION OF SOUTH AFRICAN CRANIA USING 3D-ID

Tamara Leigh Lottering

Copyright 2019 Turnitin. All rights reserved.

815691:Tamara_Lottering_-_Turn-it-in_draft_2019_.docx

ORIGINALITY REPORT

13%	6%	6%	7%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	open.bu.edu Internet Source	1%
2	jaimeseniorproject.blogspot.com Internet Source	1%
3	www.sahistory.org.za Internet Source	1%
4	api.research-repository.uwa.edu.au Internet Source	1%
5	Submitted to University of Cape Town Student Paper	1%
6	Andreas Bertsatos, Aikaterini Christaki, Maria-Eleni Chovalopoulou. "Testing the reliability of 3D-ID software in sex and ancestry estimation with a modern Greek sample", Forensic Science International, 2019 Publication	1%
7	Petra Urbanová, Ann H. Ross, Mikoláš Jurda, Maria-Ines Nogueira. "Testing the reliability of software tools in sex and ancestry estimation in	1%

Appendix 4: Intra- and Inter-observer error results from MorphoJ

Intra-observer error analysis was done in MorphoJ, by first running a new Procrustes fit and then running Procrustes ANOVA as follows:

New Procrustes fit:
34 landmarks in 3 dimensions.
The dataset contains 20 observations, of which 20 are included for analyses.

Average shape:

Lmk.	Axis 1 (x)	Axis 2 (y)	Axis 3 (z)
1	0.24606741	0.01820029	-0.13177839
2	0.24384920	0.01039399	0.13912300
3	0.12489356	0.07147539	0.00610721
4	0.11184914	-0.26333027	-0.01204153
5	-0.07180145	-0.03347624	-0.03359298
6	-0.07365772	-0.03543446	0.02936254
7	-0.02949360	0.10349927	-0.07625210
8	-0.03445805	0.10137573	0.08492874
9	-0.04691568	-0.00908136	-0.12715074
10	-0.05161257	-0.01698590	0.12525987
11	-0.04726853	-0.03686391	-0.12953456
12	-0.05089861	-0.04559834	0.12611676
13	-0.03828416	-0.04531753	-0.13762350
14	-0.04342008	-0.05610675	0.13342626
15	-0.09703540	-0.07991745	-0.00495459
16	0.35924042	-0.10828317	-0.00251212
17	0.15326666	0.09375133	-0.12156128
18	0.14900866	0.08789453	0.13205349
19	-0.09232379	-0.05651461	-0.00367779
20	-0.07328478	0.01810788	-0.08721547
21	-0.07859382	0.01177371	0.08124447
22	-0.07228022	-0.06715108	-0.09371139
23	-0.07605295	-0.07462098	0.08277044
24	0.21736926	0.08814247	0.00692439
25	-0.12478734	0.11159173	0.00322673
26	-0.11190821	0.07056425	0.00247987
27	-0.09108384	-0.03105263	-0.01274822
28	-0.09105971	-0.03222868	0.00938145
29	0.03022516	0.03094959	-0.15849301
30	-0.05190090	0.06525994	-0.11601822
31	-0.05679988	0.05823947	0.12129064
32	-0.07671454	0.01649150	-0.07379898
33	-0.07992656	0.01197855	0.07484926
34	0.02579290	0.02227375	0.16411974

Procrustes sums of squares: 0.048966225337184616
Tangent sums of squares: 0.04883456791351239

Procrustes ANOVA: Procrustes ANOVA ...
Classifiers used for the Procrustes ANOVA:
Individuals: individual

Centroid size:

Effect	SS	MS	df	F	P (param.)
Individual	2884.753879	320.528209	9	1623.57	<.0001
Residual	1.974221	0.197422	10		

Shape, Procrustes ANOVA:

Effect	SS	MS	df	F	P (param.)	Pillai tr.	P (param.)
Individual	0.04358888	0.0000509812	855	9.23	<.0001		
Residual	0.00524568	0.0000055218	950				

Inter-observer error analysis was done in MorphoJ, by first running a new Procrustes fit and then running Procrustes ANOVA as follows:

New Procrustes fit:
34 landmarks in 3 dimensions.
The dataset contains 20 observations, of which 20 are included for analyses.

Average shape:

Lmk.	Axis 1 (x)	Axis 2 (y)	Axis 3 (z)
1	0.24727969	0.01908177	-0.13129994
2	0.23980160	0.01071158	0.14126772
3	0.12491893	0.07144074	0.00630580
4	0.11267766	-0.26333924	-0.01338570
5	-0.07120311	-0.03304312	-0.03432197
6	-0.07345116	-0.03575803	0.02975411
7	-0.03378855	0.10465588	-0.07589758
8	-0.03483801	0.10162518	0.08433089
9	-0.04652359	-0.00993775	-0.12746531
10	-0.05064672	-0.02109009	0.12481334
11	-0.04720639	-0.03650616	-0.13016544
12	-0.05087703	-0.04566483	0.12581528
13	-0.03672037	-0.04541628	-0.13807094
14	-0.04407798	-0.05565759	0.13337892
15	-0.09676730	-0.07795348	-0.00574649
16	0.36002528	-0.10714168	-0.00254313
17	0.15296507	0.09446968	-0.12098957
18	0.14812793	0.08791757	0.13273098
19	-0.09230428	-0.05730576	-0.00417208
20	-0.07531754	0.01694872	-0.08010230
21	-0.07942832	0.01158305	0.07816810
22	-0.07208753	-0.06725904	-0.09545322
23	-0.07595031	-0.07501862	0.08375402
24	0.21690188	0.08858238	0.00770178
25	-0.12581369	0.10939219	0.00317809
26	-0.11340884	0.07474275	0.00215543

27	-0.09116313	-0.03157473	-0.01379983
28	-0.09089428	-0.03289656	0.00910710
29	0.03362625	0.03071467	-0.15931580
30	-0.04951145	0.06618694	-0.11606439
31	-0.05433239	0.05869704	0.12130736
32	-0.07676611	0.01634340	-0.07442626
33	-0.08021779	0.01168732	0.07525415
34	0.02697157	0.02078214	0.16419689

Procrustes sums of squares: 0.057044621356386195
Tangent sums of squares: 0.05686550364434854

Procrustes ANOVA: Procrustes ANOVA ...
Classifiers used for the Procrustes ANOVA:
Individuals: individual

Centroid size:

Effect	SS	MS	df	F	P (param.)
Individual	2957.897042	328.655227	9	323.77	<.0001
Residual	10.150813	1.015081	10		

Shape, Procrustes ANOVA:

Effect	SS	MS	df	F	P (param.)	Pillai tr.	P (param.)
Individual	0.04412678	0.0000516103	855	3.85	<.0001		
Residual	0.01273872	0.0000134092	950				

Appendix 5: Posterior Probabilities (PP) and F Typicality's (TP) for FORDISC analysis

FORDISC	RECORDED		ANCESTRY				SEX AND ANCESTRY			
Assigned no.	Record-ed sex	Recorded ancestry	Estimated ancestry	Posterior probability (PP)	Typicality F	Within 10% of correct PP	Estimated sex + ancestry	Posterior probability (PP)	Typicality F	Within 10% of correct PP
COLOURED FEMALES										
STELLIES1	Female	Coloured	Coloured	0.544	0.966		Coloured Female	0.579	0.868	
STELLIES2	Female	Coloured	Black	0.775	0.622		Black Female	0.602	0.726	
STELLIES3	Female	Coloured	Black	0.662	0.861		Black Female	0.541	0.868	
STELLIES4	Female	Coloured	Black	0.505	0.819	Coloured PP 0,462	Coloured Female	0.614	0.728	
STELLIES5	Female	Coloured	Coloured	0.71	0.092		Coloured Female	0.483	0.164	
STELLIES6	Female	Coloured	Black	0.764	0.709		Coloured Female	0.364	0.617	Black Female PP 0,341
STELLIES7	Female	Coloured	Black	0.803	0.652		Black Male	0.491	0.337	
STELLIES8	Female	Coloured	White	0.523	0.957		White Female	0.535	0.926	
STELLIES9	Female	Coloured	Black	0.836	0.265		Black Female	0.525	0.46	
STELLIES10	Female	Coloured	Black	0.66	0.699		Coloured Female	0.556	0.589	
STELLIES11	Female	Coloured	Black	0.761	0.79		Black Female	0.381	0.804	Black Male PP 0,321
STELLIES12	Female	Coloured	Black	0.583	0.531		Black Female	0.279	0.577	Coloured Male PP 0,230
STELLIES13	Female	Coloured	Black	0.429	0.497		Black Female	0.313	0.835	Coloured Female PP 0,264
STELLIES14	Female	Coloured	Black	0.524	0.037		Black Female	0.653	0.107	
STELLIES15	Female	Coloured	Coloured	0.578	0.849		Coloured Female	0.566	0.683	
STELLIES16	Female	Coloured	Black	0.546	0.601		Black Female	0.471	0.479	
STELLIES17	Female	Coloured	Coloured	0.529	0.672		Coloured Female	0.452	0.684	
STELLIES18	Female	Coloured	Black	0.513	0.305		Black Female	0.439	0.471	Coloured Female PP 0,422
STELLIES19	Female	Coloured	Coloured	0.519	0.674		Coloured Female	0.434	0.718	
STELLIES20	Female	Coloured	Black	0.912	0.452		Black Female	0.669	0.294	
STELLIES21	Female	Coloured	Black	0.549	0.864		Black Female	0.446	0.888	Coloured Female PP 0,381
STELLIES22	Female	Coloured	Coloured	0.657	0.84		Coloured Female	0.693	0.569	

FORDISC	RECORDED		ANCESTRY				SEX AND ANCESTRY			
Assigned no.	Record-ed sex	Recorded ancestry	Estimated ancestry	Posterior probability (PP)	Typicality F	Within 10% of correct PP	Estimated sex + ancestry	Posterior probability (PP)	Typicality F	Within 10% of correct PP
STELLIES23	Female	Coloured	Coloured	0.527	0.347		Coloured Female	0.406	0.583	Black Female PP 0.307
STELLIES24	Female	Coloured	Coloured	0.609	0.965		Coloured Female	0.522	0.686	
STELLIES25	Female	Coloured	Coloured	0.426	0.095	Black PP 0.358	White Female	0.447	0.535	
STELLIES26	Female	Coloured	Coloured	0.514	0.342	Black PP 0,479	Coloured Female	0.52	0.468	
STELLIES27	Female	Coloured	Coloured	0.54	0.897		Coloured Female	0.403	0.951	
STELLIES28	Female	Coloured	Coloured	0.404	0.75	White PP 0,363	White Female	0.498	0.923	
STELLIES29	Female	Coloured	Black	0.826	0.307		Black Female	0.379	0.685	Black Male PP 0,323
STELLIES30	Female	Coloured	White	0.56	0.517		White Female	0.43	0.589	
STELLIES31	Female	Coloured	Black	0.908	0.182		Black Female	0.532	0.065	
STELLIES32	Female	Coloured	Coloured	0.504	0.992		Coloured Female	0.417	0.99	
STELLIES33	Female	Coloured	Black	0.515	0.886	Coloured PP 0,459	Black Female	0.446	0.974	Coloured Female PP 0,349
STELLIES34	Female	Coloured	Coloured	0.49	0.767		White Female	0.499	0.81	
STELLIES35	Female	Coloured	Black	0.857	0.478		Black Female	0.591	0.336	
STELLIES36	Female	Coloured	White	0.599	0.237		White Female	0.468	0.529	
STELLIES37	Female	Coloured	Coloured	0.528	0.705	Black PP 0,435	Black Female	0.511	0.87	
STELLIES38	Female	Coloured	Coloured	0.586	0.503		Coloured Female	0.5	0.291	
STELLIES39	Female	Coloured	Coloured	0.519	0.657		Coloured Female	0.318	0.829	Black Female PP 0,245
STELLIES40	Female	Coloured	Black	0.624	0.664		Black Female	0.531	0.616	
STELLIES41	Female	Coloured	Coloured	0.558	0.955		Coloured Female	0.311	0.89	White Female PP 0.294
STELLIES42	Female	Coloured	Coloured	0.556	0.821		Coloured Female	0.411	0.711	
STELLIES43	Female	Coloured	Black	0.69	0.046		Black Female	0.742	0.016	
STELLIES44	Female	Coloured	White	0.488	0.599	Coloured PP 0.462	White Female	0.628	0.486	
STELLIES45	Female	Coloured	Coloured	0.461	0.675	Black PP 0.402	Coloured Female	0.547	0.453	
STELLIES46	Female	Coloured	White	0.551	0.185		White Female	0.477	0.14	
STELLIES47	Female	Coloured	Coloured	0.524	0.859	Black PP 0.448	Black Female	0.502	0.941	
STELLIES48	Female	Coloured	Black	0.673	0.686		Black Female	0.522	0.906	

FORDISC	RECORDED		ANCESTRY				SEX AND ANCESTRY			
Assigned no.	Record-ed sex	Recorded ancestry	Estimated ancestry	Posterior probability (PP)	Typicality F	Within 10% of correct PP	Estimated sex + ancestry	Posterior probability (PP)	Typicality F	Within 10% of correct PP
STELLIES49	Female	Coloured	Coloured	0.652	0.373		Black Female	0.373	0.408	Coloured Female PP 0.362
STELLIES50	Female	Coloured	Coloured	0.652	0.636		Black Female	0.582	0.909	
STELLIES51	Female	Coloured	Coloured	0.483	0.595	Black PP 0.469	Coloured Female	0.561	0.093	
STELLIES52	Female	Coloured	White	0.741	0.348		White Female	0.658	0.274	
STELLIES53	Female	Coloured	Black	0.727	0.203		Black Female	0.662	0.368	
STELLIES54	Female	Coloured	Coloured	0.556	0.534		Coloured Female	0.566	0.705	
STELLIES55	Female	Coloured	Coloured	0.463	0.015	Black PP 0.448	Coloured Female	0.49	0.489	
STELLIES56	Female	Coloured	Black	0.873	0.483		Black Female	0.485	0.468	
STELLIES57	Female	Coloured	Coloured	0.409	0.089	White PP 0.357; B PP 0.235	White Female	0.522	0.659	
STELLIES58	Female	Coloured	Black	0.675	0.873		Black Male	0.406	0.822	
STELLIES59	Female	Coloured	Black	0.472	0.056		Black Female	0.636	0.05	
STELLIES60	Female	Coloured	Black	0.422	0.507	Coloured PP 0.395	Coloured Female	0.352	0.467	Black Female PP 0.318
STELLIES61	Female	Coloured	White	0.915	0.316		White Male	0.482	0.593	
STELLIES62	Female	Coloured	White	0.722	0.936		Coloured Female	0.369	0.294	White Female PP 0.330
STELLIES63	Female	Coloured	Black	0.588	0.216		Black Female	0.692	0.036	
STELLIES64	Female	Coloured	Coloured	0.489	0.815		Coloured Female	0.303	0.648	Black Female PP 0.299
STELLIES65	Female	Coloured	Black	0.822	0.236		Black Male	0.813	0.209	
STELLIES66	Female	Coloured	Coloured	0.51	0.985		Coloured Female	0.453	0.969	
STELLIES67	Female	Coloured	Black	0.421	0.242	White PP 0.338	Coloured Male	0.256	0.711	White Male PP 0.219
STELLIES68	Female	Coloured	Coloured	0.49	0.901		Coloured Female	0.618	0.702	
STELLIES69	Female	Coloured	Coloured	0.595	0.101		Coloured Female	0.349	0.172	White Female PP 0.302
STELLIES70	Female	Coloured	Black	0.571	0.768		Coloured Female	0.51	0.937	Black Female PP 0.435
STELLIES71	Female	Coloured	Black	0.539	0.404	Coloured PP 0.449	Black Female	0.471	0.537	
STELLIES72	Female	Coloured	Black	0.494	0.302	Coloured PP 0.456	Black Female	0.482	0.504	

FORDISC	RECORDED		ANCESTRY				SEX AND ANCESTRY			
Assigned no.	Record-ed sex	Recorded ancestry	Estimated ancestry	Posterior probability (PP)	Typicality F	Within 10% of correct PP	Estimated sex + ancestry	Posterior probability (PP)	Typicality F	Within 10% of correct PP
STELLIES73	Female	Coloured	White	0.817	0.155		White Male	0.682	0.107	
STELLIES74	Female	Coloured	Coloured	0.539	0.445	Black PP 0.451	Coloured Male	0.298	0.334	Black Male PP 0.260
STELLIES75	Female	Coloured	Black	0.75	0.609		Black Female	0.703	0.392	
COLOURED MALES										
STELLIES76	Male	Coloured	Coloured	0.593	0.595		Coloured Male	0.548	0.164	
STELLIES77	Male	Coloured	White	0.535	0.896		White Female	0.426	0.975	
STELLIES78	Male	Coloured	Black	0.481	0.631	Coloured PP 0.459	Coloured Male	0.351	0.714	Black Male PP 0.337
STELLIES79	Male	Coloured	Black	0.687	0.558		Black Male	0.798	0.205	
STELLIES80	Male	Coloured	Black	0.677	0.38		Black Male	0.728	0.525	
STELLIES81	Male	Coloured	Black	0.477	0.596	Coloured PP 0.429	Coloured Male	0.337	0.823	
STELLIES82	Male	Coloured	Black	0.536	0.814		Black Male	0.364	0.805	
STELLIES83	Male	Coloured	Coloured	0.631	0.579		Coloured Male	0.292	0.689	White Female PP 0.223
STELLIES84	Male	Coloured	White	0.573	0.036		White Female	0.415	0.061	
STELLIES85	Male	Coloured	Black	0.699	0.631		Coloured Female	0.554	0.853	
STELLIES86	Male	Coloured	Black	0.974	0.076		Black Male	0.978	0.103	
STELLIES87	Male	Coloured	Black	0.532	0.533		Coloured Male	0.458	0.124	
STELLIES88	Male	Coloured	Coloured	0.492	0.883	White PP 0.423	Coloured Female	0.314	0.244	White Female PP 0.257
STELLIES89	Male	Coloured	Black	0.781	0.729		Black Male	0.452	0.608	
STELLIES90	Male	Coloured	Black	0.483	0.625	Coloured PP 0.444	Black Female	0.555	0.302	
STELLIES91	Male	Coloured	Coloured	0.396	0.419	White PP 0.364	White Male	0.297	0.175	White Female PP 0.255
STELLIES92	Male	Coloured	Black	0.396	0.244	Coloured PP 0.308	White Male	0.334	0.215	Black Male PP 0.326
STELLIES93	Male	Coloured	Coloured	0.527	0.306	Black PP 0.461	Coloured Female	0.488	0.025	
STELLIES94	Male	Coloured	Black	0.376	0.01		White Female	0.289	0.863	White Male PP 0.236
STELLIES95	Male	Coloured	Coloured	0.461	0.214		Coloured Male	0.45	0.488	

FORDISC	RECORDED		ANCESTRY				SEX AND ANCESTRY			
Assigned no.	Record-ed sex	Recorded ancestry	Estimated ancestry	Posterior probability (PP)	Typicality F	Within 10% of correct PP	Estimated sex + ancestry	Posterior probability (PP)	Typicality F	Within 10% of correct PP
STELLIES98	Male	Coloured	Black	0.74	0.504		Black Male	0.316	0.147	Black Female PP 0.284
STELLIES97	Male	Coloured	Black	0.373	0.892	Coloured PP 0.319	Coloured Male	0.326	0.975	White Male PP 0.246
STELLIES 98	Male	Coloured	Black	0.592	0.981		Black Female	0.249	0.958	Coloured Female PP 0.236
STELLIES99	Male	Coloured	Black	0.862	0.504		Black Male	0.482	0.899	
STELLIES100	Male	Coloured	Coloured	0.349	0.803	White PP 0.348	White Female	0.522	0.887	
STELLIES101	Male	Coloured	Black	0.979	0.039		Black Male	0.493	0.122	
STELLIES102	Male	Coloured	White	0.917	0.446		White Male	0.775	0.187	
STELLIES103	Male	Coloured	White	0.873	0.397		White Male	0.619	0.418	
STELLIES104	Male	Coloured	Black	0.536	0.552		Black Male	0.602	0.691	
STELLIES105	Male	Coloured	Black	0.821	0.594		Black Female	0.554	0.104	
STELLIES106	Male	Coloured	Black	0.714	0.113		Coloured Male	0.398	0.18	
STELLIES107	Male	Coloured	White	0.696	0.583		White Female	0.443	0.279	
STELLIES108	Male	Coloured	Coloured	0.483	0.899		Coloured Male	0.263	0.857	Coloured Female PP 0.248
STELLIES109	Male	Coloured	Coloured	0.371	0.97	Black PP 0.339	Coloured Female	0.272	0.986	Coloured Male PP 0.205
STELLIES110	Male	Coloured	White	0.37	0.658	Coloured PP 0.338	Coloured Male	0.294	0.38	White Male PP 0.279
STELLIES111	Male	Coloured	Coloured	0.507	0.557		Coloured Male	0.277	0.872	Coloured Female PP 0.266
STELLIES112	Male	Coloured	Coloured	0.527	0.306		Coloured Female	0.545	0.578	
STELLIES113	Male	Coloured	Black	0.719	0.308		Black Female	0.353	0.553	Coloured Female PP 0.295
STELLIES114	Male	Coloured	Coloured	0.491	0.085		White Female	0.379	0.583	White Male PP 0.311
STELLIES115	Male	Coloured	White	0.962	0.834		White Male	0.688	0.769	
STELLIES116	Male	Coloured	Black	0.464	0.988		Black Female	0.264	0.915	Black Male PP 0.202
STELLIES117	Male	Coloured	Coloured	0.569	0.612		Coloured Female	0.544	0.887	

FORDISC	RECORDED		ANCESTRY				SEX AND ANCESTRY			
Assigned no.	Record-ed sex	Recorded ancestry	Estimated ancestry	Posterior probability (PP)	Typicality F	Within 10% of correct PP	Estimated sex + ancestry	Posterior probability (PP)	Typicality F	Within 10% of correct PP
STELLIES118	Male	Coloured	Black	0.561	0.585		Coloured Male	0.358	0.679	Black Male PP 0.310
STELLIES119	Male	Coloured	Black	0.499	0.398		Coloured Male	0.415	0.962	
STELLIES120	Male	Coloured	Black	0.483	0.268	Coloured PP 0.402	Coloured Male	0.435	0.818	
STELLIES121	Male	Coloured	White	0.975	0.753		White Female	0.749	0.804	
STELLIES122	Male	Coloured	Black	0.487	0.306	White PP 0.272	Coloured Male	0.319	0.871	
STELLIES123	Male	Coloured	Black	0.507	0.433	Coloured PP 0.454	Coloured Female	0.279	0.525	Black Female PP 0.234
STELLIES124	Male	Coloured	White	0.834	0.975		White Male	0.512	0.886	
STELLIES125	Male	Coloured	Coloured	0.49	0.414		Coloured Male	0.512	0.378	
STELLIES126	Male	Coloured	Coloured	0.454	0.344		Black Female	0.323	0.519	Coloured Female PP 0.292
STELLIES127	Male	Coloured	Black	0.794	0.177		Black Male	0.909	0.181	
STELLIES128	Male	Coloured	White	0.618	0.998		White Female	0.378	0.575	
STELLIES129	Male	Coloured	Coloured	0.52	0.486		Black Male	0.326	0.313	
STELLIES130	Male	Coloured	White	0.38	0.028	Coloured PP 0.373	Coloured Female	0.754	0.219	
STELLIES131	Male	Coloured	Coloured	0.534	0.531		Coloured Female	0.256	0.827	Coloured Male PP 0.240
STELLIES132	Male	Coloured	Black	0.487	0.051	Coloured PP 0.411	Coloured Male	0.504	0.945	
STELLIES133	Male	Coloured	Coloured	0.467	0.47		Coloured Male	0.279	0.146	Coloured Female PP 0.229
STELLIES134	Male	Coloured	Black	0.596	0.943		Coloured Female	0.408	0.938	Black Female PP 0.308
STELLIES135	Male	Coloured	White	0.645	0.971		White Female	0.392	0.814	
STELLIES136	Male	Coloured	Black	0.809	0.363		Black Male	0.735	0.178	
STELLIES137	Male	Coloured	Black	0.842	0.308		Black Male	0.492	0.57	
STELLIES138	Male	Coloured	Black	0.543	0.536		Black Male	0.421	0.95	Coloured Male PP 0.359
STELLIES139	Male	Coloured	Black	0.699	0.607		Black Female	0.366	0.686	Coloured Female PP 0.304
STELLIES140	Male	Coloured	Coloured	0.608	0.661		Coloured Male	0.559	0.752	
STELLIES141	Male	Coloured	Black	0.809	0.309		Black Male	0.402	0.055	

FORDISC	RECORDED		ANCESTRY				SEX AND ANCESTRY			
Assigned no.	Record-ed sex	Recorded ancestry	Estimated ancestry	Posterior probability (PP)	Typicality F	Within 10% of correct PP	Estimated sex + ancestry	Posterior probability (PP)	Typicality F	Within 10% of correct PP
STELLIES142	Male	Coloured	Coloured	0.52	0.72	Black PP 0.466	Coloured Male	0.334	0.617	Black Male PP 0.269
STELLIES143	Male	Coloured	Black	0.614	0.695		Black Male	0.492	0.223	Coloured Male PP 0.418
STELLIES144	Male	Coloured	Black	0.502	0.457		Black Male	0.485	0.556	
STELLIES145	Male	Coloured	White	0.511	0.287		White Female	0.342	0.47	Coloured Female PP 0.341
STELLIES146	Male	Coloured	Black	0.532	0.66	Coloured PP 0.460	Black Female	0.657	0.186	
STELLIES147	Male	Coloured	White	0.953	0.476		White Male	0.527	0.199	
STELLIES148	Male	Coloured	Black	0.614	0.317		Coloured Male	0.417	0.911	
STELLIES149	Male	Coloured	White	0.415	0.182		White Female	0.291	0.143	Coloured Male PP 0.236
STELLIES150	Male	Coloured	Black	0.575	0.662		Black Female	0.437	0.423	
BLACK FEMALES										
WITS1	Female	Black	Black	0.573	0.501		Black Female	0.492	0.721	
WITS2	Female	Black	Black	0.647	0.98		Black Male	0.486	0.999	
WITS3	Female	Black	Coloured	0.51	0.866	Black PP 0.483	Black Female	0.485	0.877	Coloured Female PP 0.460
WITS4	Female	Black	Coloured	0.51	0.972	Black PP 0.478	Black Female	0.424	0.943	Coloured Female PP 0.339
WITS5	Female	Black	Black	0.555	0.81		Black Male	0.398	0.829	
WITS6	Female	Black	Black	0.544	0.359	Coloured PP 0.445	Coloured Female	0.525	0.548	
WITS7	Female	Black	Black	0.724	0.816		Black Female	0.354	0.82	Coloured Female PP 0.329
WITS8	Female	Black	Black	0.726	0.932		Black Female	0.379	0.777	Black Male PP 0.359
WITS9	Female	Black	Black	0.737	0.941		Black Female	0.596	0.998	
WITS10	Female	Black	Black	0.694	0.677		Black Female	0.469	0.99	Coloured Female PP 0.387
WITS11	Female	Black	Coloured	0.45	0.927	Black PP 0.435	Black Male	0.314	0.841	Coloured Female PP 0.228
WITS12	Female	Black	Black	0.803	0.674		Black Female	0.508	0.869	
WITS13	Female	Black	Black	0.497	0.118	Coloured PP 0.486	Black Female	0.471	0.156	

FORDISC	RECORDED		ANCESTRY				SEX AND ANCESTRY			
Assigned no.	Record-ed sex	Recorded ancestry	Estimated ancestry	Posterior probability (PP)	Typicality F	Within 10% of correct PP	Estimated sex + ancestry	Posterior probability (PP)	Typicality F	Within 10% of correct PP
WITS14	Female	Black	Black	0.529	0.675	Coloured PP 0.460	Black Female	0.386	0.661	
WITS15	Female	Black	Coloured	0.454	0.373		Coloured Female	0.676	0.46	
WITS16	Female	Black	Black	0.724	0.959		Black Male	0.635	0.944	
WITS17	Female	Black	Black	0.919	0.813		Black Male	0.678	0.665	
WITS18	Female	Black	Black	0.75	0.028		Black Female	0.491	0.013	
WITS19	Female	Black	Coloured	0.577	0.763		Coloured Female	0.511	0.713	
WITS20	Female	Black	Coloured	0.593	0.715		Coloured Female	0.454	0.7	
WITS21	Female	Black	Coloured	0.56	0.931		Coloured Female	0.443	0.912	
WITS22	Female	Black	Black	0.507	0.867	Coloured PP 0.480	Black Female	0.525	0.906	
WITS23	Female	Black	Black	0.884	0.821		Black Male	0.58	0.67	
WITS24	Female	Black	Black	0.643	0.795		Black Male	0.373	0.537	
WITS25	Female	Black	Black	0.839	0.014		Black Female	0.59	0.172	
WITS26	Female	Black	Black	0.861	0.133		Black Male	0.403	0.036	Coloured Female PP 0.335
WITS27	Female	Black	Black	0.621	0.761		Black Female	0.459	0.842	Coloured Female PP 0.408
WITS28	Female	Black	Black	0.862	0.211		Black Female	0.435	0.122	Black Male PP 0.390
WITS29	Female	Black	Black	0.654	0.97		Black Female	0.357	0.965	
WITS30	Female	Black	Black	0.64	0.978		Coloured Female	0.365	0.905	Black Female PP 0.352
WITS31	Female	Black	Coloured	0.66	0.842		Coloured Male	0.369	0.767	Black Male PP 0.279
WITS32	Female	Black	Black	0.7	0.632		Black Female	0.631	0.594	
WITS33	Female	Black	Black	0.632	0.19		Black Female	0.477	0.471	Coloured Female PP 0.387
WITS34	Female	Black	Coloured	0.642	0.323		Coloured Female	0.52	0.444	
WITS35	Female	Black	Black	0.518	0.271	Coloured PP 0.452	Black Female	0.394	0.497	
WITS36	Female	Black	Black	0.943	0.267		Black Female	0.713	0.209	
WITS37	Female	Black	Coloured	0.514	0.897	Black PP 0.437	Black Female	0.557	0.988	
WITS38	Female	Black	Black	0.8	0.699		Black Female	0.484	0.562	
WITS39	Female	Black	Coloured	0.698	0.931		Coloured Female	0.453	0.928	

FORDISC	RECORDED		ANCESTRY				SEX AND ANCESTRY			
Assigned no.	Record-ed sex	Recorded ancestry	Estimated ancestry	Posterior probability (PP)	Typicality F	Within 10% of correct PP	Estimated sex + ancestry	Posterior probability (PP)	Typicality F	Within 10% of correct PP
WITS40	Female	Black	White	0.65	0.764		White Female	0.586	0.696	
WITS41	Female	Black	Black	0.514	0.914	Coloured PP 0.452	Black Female	0.336	0.971	Coloured Female PP 0.275
WITS42	Female	Black	White	0.379	0.874	Coloured PP 0.373	White Female	0.402	0.937	
WITS43	Female	Black	Black	0.733	0.184		Coloured Female	0.535	0.256	
WITS44	Female	Black	Coloured	0.475	0.697	Black PP 0.386	Black Female	0.447	0.805	Coloured Female PP 0.382
WITS45	Female	Black	Coloured	0.612	0.34		Coloured Female	0.512	0.538	
WITS46	Female	Black	Black	0.814	0.636		Black Female	0.647	0.632	
WITS47	Female	Black	Coloured	0.47	0.952	White PP 0.382	White Female	0.409	0.963	Coloured Female PP 0.313
WITS48	Female	Black	Coloured	0.64	0.075		Black Female	0.526	0.326	
WITS49	Female	Black	Black	0.914	0.295		Black Female	0.444	0.345	Black Male PP 0.433
WITS50	Female	Black	Black	0.857	0.334		Black Male	0.835	0.532	
WITS51	Female	Black	Black	0.626	0.975		Black Female	0.492	0.901	
WITS52	Female	Black	Black	0.813	0.075		Coloured Female	0.503	0.043	Black Female PP 0.485
WITS53	Female	Black	Coloured	0.44	0.777	Black PP 0.355	Coloured Male	0.264	0.35	Black Female PP 0.189
WITS54	Female	Black	Coloured	0.426	0.671	Black PP 0.411	Coloured Female	0.378	0.858	Black Female PP 0.310
WITS55	Female	Black	Coloured	0.684	0.758		Coloured Female	0.655	0.664	
WITS56	Female	Black	Black	0.568	0.088		Black Female	0.489	0.106	Coloured Female PP 0.462
WITS57	Female	Black	White	0.942	0.327		White Female	0.706	0.223	
WITS58	Female	Black	Black	0.724	0.674		Black Female	0.665	0.585	
WITS59	Female	Black	White	0.495	0.055		White Female	0.501	0.066	
WITS60	Female	Black	Black	0.464	0.693	Coloured PP 0.416	Coloured Female	0.371	0.996	Black Female PP 0.352
WITS61	Female	Black	Coloured	0.637	0.155		Coloured Male	0.352	0.044	
WITS62	Female	Black	Black	0.475	0.809	Coloured PP 0.456	Coloured Female	0.364	0.887	Black Female PP 0.265

FORDISC	RECORDED		ANCESTRY				SEX AND ANCESTRY			
Assigned no.	Record-ed sex	Recorded ancestry	Estimated ancestry	Posterior probability (PP)	Typicality F	Within 10% of correct PP	Estimated sex + ancestry	Posterior probability (PP)	Typicality F	Within 10% of correct PP
WITS63	Female	Black	Black	0.781	0.435		Black Female	0.373	0.944	Coloured Female PP 0.339
WITS64	Female	Black	Coloured	0.409	0.65	White PP 0.313	White Female	0.262	0.75	Coloured Female PP 0.223
WITS65	Female	Black	White	0.465	0.378	Coloured PP 0.462	White Male	0.318	0.348	White Female PP 0.317
WITS66	Female	Black	Coloured	0.464	0.458		White Female	0.348	0.507	Black Female PP 0.346
WITS67	Female	Black	White	0.469	0.673	Coloured PP 0.437	White Female	0.398	0.96	
WITS68	Female	Black	Black	0.791	0.047		Black Female	0.556	0.032	
WITS69	Female	Black	Black	0.861	0.772		Black Female	0.503	0.669	
WITS70	Female	Black	Black	0.828	0.464		Black Female	0.683	0.546	
WITS71	Female	Black	Black	0.708	0.438		Black Female	0.597	0.31	
WITS72	Female	Black	Coloured	0.626	0.656		White Female	0.387	0.47	Coloured Female PP 0.379
WITS73	Female	Black	Black	0.917	0.415		Black Male	0.581	0.278	
WITS74	Female	Black	Black	0.596	0.701		Black Female	0.474	0.964	Coloured Female PP 0.415
WITS75	Female	Black	Black	0.559	0.318		Black Female	0.352	0.399	Coloured Female PP 0.327
BLACK MALES										
WITS76	Male	Black	Coloured	0.565	0.282		Black Male	0.715	0.666	
WITS77	Male	Black	Coloured	0.483	0.486	Black PP 0.401	Black Male	0.569	0.823	
WITS78	Male	Black	Black	0.852	0.406		Black Male	0.534	0.243	
WITS79	Male	Black	Coloured	0.47	0.961	Black PP 0.431	Coloured Male	0.503	0.997	
WITS80	Male	Black	Black	0.867	0.316		Black Male	0.792	0.225	
WITS81	Male	Black	Black	0.611	0.752		Coloured Male	0.481	0.559	Black Male PP 0.428
WITS82	Male	Black	White	0.786	0.793		White Male	0.584	0.835	
WITS83	Male	Black	Black	0.5	0.208	Coloured PP 0.416	Black Male	0.497	0.163	
WITS84	Male	Black	Coloured	0.468	0.611	Black PP 0.464	Coloured Male	0.507	0.482	
WITS85	Male	Black	Black	0.524	0.776	Coloured PP 0.466	Black Male	0.429	0.979	

FORDISC	RECORDED		ANCESTRY				SEX AND ANCESTRY			
Assigned no.	Record-ed sex	Recorded ancestry	Estimated ancestry	Posterior probability (PP)	Typicality F	Within 10% of correct PP	Estimated sex + ancestry	Posterior probability (PP)	Typicality F	Within 10% of correct PP
WITS86	Male	Black	Black	0.656	0.526		Black Male	0.456	0.444	Coloured Male PP 0.375
WITS87	Male	Black	Black	0.868	0.378		Black Male	0.506	0.916	
WITS88	Male	Black	Black	0.97	0.05		Black Male	0.677	0.023	
WITS89	Male	Black	Black	0.878	0.12		Black Male	0.818	0.712	
WITS90	Male	Black	Coloured	0.442	0.319		Coloured Male	0.284	0.617	Black Male PP 0.277
WITS91	Male	Black	Black	0.599	0.835		Coloured Female	0.338	0.574	
WITS92	Male	Black	Black	0.852	0.827		Black Male	0.843	0.684	
WITS93	Male	Black	Black	0.594	0.795		Black Male	0.504	0.955	
WITS94	Male	Black	Black	0.553	0.713		Black Male	0.556	0.725	
WITS95	Male	Black	Black	0.929	0.473		Black Male	0.671	0.283	
WITS96	Male	Black	Black	0.515	0.255	Coloured PP 0.434	Coloured Female	0.316	0.848	Black Female PP 0.2477
WITS97	Male	Black	Black	0.764	0.868		Black Female	0.504	0.632	
WITS98	Male	Black	Coloured	0.457	0.231	White PP 0.360	Coloured Male	0.328	0.196	White Male PP 0.264
WITS99	Male	Black	Black	0.901	0.59		Black Male	0.752	0.914	
WITS100	Male	Black	White	0.653	0.29		White Male	0.488	0.299	
WITS101	Male	Black	Coloured	0.579	0.257		Coloured Female	0.593	0.617	
WITS102	Male	Black	Black	0.479	0.898	Coloured PP 0.431	Black Male	0.443	0.966	Coloured Male PP 0.346
WITS103	Male	Black	Black	0.568	0.351		Black Male	0.511	0.286	
WITS104	Male	Black	Black	0.69	0.411		Black Male	0.493	0.683	
WITS105	Male	Black	Black	0.614	0.525		Black Male	0.406	0.441	
WITS106	Male	Black	Black	0.668	0.46		Black Male	0.638	0.219	
WITS107	Male	Black	Black	0.523	0.935	Coloured PP 0.456	Black Male	0.6	0.986	
WITS108	Male	Black	Coloured	0.549	0.965		Coloured Male	0.502	0.913	
WITS109	Male	Black	Coloured	0.521	0.823		Coloured Male	0.448	0.964	
WITS110	Male	Black	Coloured	0.543	0.859		Coloured Male	0.459	0.84	
WITS111	Male	Black	Coloured	0.462	0.741		Coloured Male	0.362	0.781	
WITS112	Male	Black	Black	0.855	0.26		Black Male	0.534	0.148	

FORDISC	RECORDED		ANCESTRY				SEX AND ANCESTRY			
Assigned no.	Record-ed sex	Recorded ancestry	Estimated ancestry	Posterior probability (PP)	Typicality F	Within 10% of correct PP	Estimated sex + ancestry	Posterior probability (PP)	Typicality F	Within 10% of correct PP
WITS113	Male	Black	Black	0.666	0.949		Black Male	0.475	0.977	
WITS114	Male	Black	Black	0.544	0.471		Coloured Female	0.38	0.565	Black Female PP 0.370
WITS115	Male	Black	Black	0.812	0.31		Black Male	0.81	0.537	
WITS116	Male	Black	Coloured	0.536	0.582	Black PP 0.440	Coloured Female	0.56	0.303	
WITS117	Male	Black	Coloured	0.442	0.994	White PP 0.390	White Female	0.258	0.994	Coloured Male PP 0.218
WITS118	Male	Black	Black	0.654	0.996		Black Male	0.526	0.988	
WITS119	Male	Black	Coloured	0.486	0.822		Coloured Male	0.351	0.679	
WITS120	Male	Black	Black	0.919	0.451		Black Male	0.4	0.43	Black Female PP 0.394
WITS121	Male	Black	Black	0.484	0.539		Black Male	0.449	0.668	
WITS122	Male	Black	Coloured	0.618	0.38		Coloured Female	0.438	0.272	
WITS123	Male	Black	Black	0.908	0.92		Black Male	0.842	0.951	
WITS124	Male	Black	Black	0.435	0.345	Coloured PP 0.402	Black Male	0.379	0.392	Coloured Male PP 0.300
WITS125	Male	Black	Black	0.838	0.706		Black Male	0.638	0.949	
WITS126	Male	Black	White	0.796	0.465		White Female	0.38	0.569	White Male PP 0.313
WITS127	Male	Black	Black	0.919	0.604		Black Male	0.853	0.439	
WITS128	Male	Black	White	0.531	0.107		White Male	0.639	0.141	
WITS129	Male	Black	Coloured	0.613	0.014		Black Male	0.649	0.016	
WITS130	Male	Black	Black	0.599	0.956		Black Male	0.552	0.97	
WITS131	Male	Black	Black	0.913	0.09		Black Male	0.683	0.226	
WITS132	Male	Black	White	0.787	0.53		White Male	0.779	0.586	
WITS133	Male	Black	Coloured	0.417	0.938	Black PP 0.336	Coloured Male	0.451	0.994	
WITS134	Male	Black	Black	0.673	0.7		Black Male	0.742	0.631	
WITS135	Male	Black	Black	0.437	0.319	Coloured PP 0.408	Black Male	0.375	0.347	Coloured Male PP 0.353
WITS136	Male	Black	Coloured	0.546	0.914		Black Male	0.449	0.859	Coloured Male PP 0.381
WITS137	Male	Black	Black	0.528	0.255		Black Female	0.392	0.295	

FORDISC	RECORDED		ANCESTRY				SEX AND ANCESTRY			
Assigned no.	Record-ed sex	Recorded ancestry	Estimated ancestry	Posterior probability (PP)	Typicality F	Within 10% of correct PP	Estimated sex + ancestry	Posterior probability (PP)	Typicality F	Within 10% of correct PP
WITS138	Male	Black	Black	0.856	0.843		Black Male	0.867	0.474	
WITS139	Male	Black	Black	0.651	0.773		Black Male	0.399	0.664	
WITS140	Male	Black	Black	0.824	0.1		Black Male	0.817	0.285	
WITS141	Male	Black	Coloured	0.558	0.165		Coloured Female	0.477	0.688	
WITS142	Male	Black	Coloured	0.498	0.2	Black PP 0.494	Black Male	0.472	0.561	
WITS143	Male	Black	Black	0.75	0.564		Black Male	0.754	0.796	
WITS144	Male	Black	Black	0.81	0.898		Black Male	0.802	0.955	
WITS145	Male	Black	Black	0.899	0.21		Black Female	0.767	0.501	
WITS146	Male	Black	Black	0.618	0.936		Coloured Female	0.419	0.767	Black Female PP 0.340
WITS147	Male	Black	Black	0.935	0.162		Black Male	0.866	0.286	
WITS148	Male	Black	Black	0.899	0.836		Black Male	0.382	0.588	Black Female PP 0.365
WITS149	Male	Black	White	0.531	0.218		Coloured Male	0.255	0.196	White Male PP 0.251
WITS150	Male	Black	Coloured	0.525	0.81		Coloured Male	0.368	0.962	
WHITE FEMALES										
STELLIES151	Females	White	White	0.967	0.52		White Female	0.872	0.455	
WITS151	Females	White	Coloured	0.478	0.05		Black Female	0.251	0.186	White Female PP 0.237
WITS152	Females	White	White	0.562	0.862		White Female	0.415	0.884	
WITS153	Females	White	White	0.633	0.395		White Female	0.741	0.746	
WITS154	Females	White	Coloured	0.442	0.487		Black Female	0.254	0.477	Coloured Female PP 0.218
WITS155	Females	White	White	0.544	0.043		White Female	0.57	0.108	
WITS156	Females	White	Coloured	0.437	0.673		Black Female	0.294	0.664	White Female PP 0.265
WITS157	Females	White	White	0.651	0.979		White Female	0.48	0.951	
WITS158	Females	White	White	0.968	0.8		White Female	0.811	0.809	
WITS159	Females	White	White	0.506	0.921		White Female	0.42	0.998	
WITS160	Females	White	White	0.983	0.105		White Female	0.515	0.338	White Male PP 0.456

FORDISC	RECORDED		ANCESTRY				SEX AND ANCESRTY			
Assigned no.	Record-ed sex	Recorded ancestry	Estimated ancestry	Posterior probability (PP)	Typicality F	Within 10% of correct PP	Estimated sex + ancestry	Posterior probability (PP)	Typicality F	Within 10% of correct PP
WITS161	Females	White	Coloured	0.519	0.241	Black PP 0.471	Black Female	0.603	0.843	
WITS162	Females	White	Black	0.586	0.941		Black Female	0.327	0.898	Coloured Female PP 0.251
WITS163	Females	White	Coloured	0.501	0.827		Coloured Female	0.343	0.797	White Female PP 0.322
WITS164	Females	White	Coloured	0.724	0.095		Coloured Female	0.599	0.402	
WITS165	Females	White	White	0.626	0.293		White Female	0.664	0.422	
WITS166	Females	White	White	0.45	0.082	Coloured PP 0.446	Coloured Female	0.368	0.072	Black Female PP 0.297
WITS167	Females	White	White	0.891	0.659		White Female	0.682	0.847	
WITS168	Females	White	White	0.653	0.353		White Female	0.739	0.658	
WITS169	Females	White	Coloured	0.6	0.761		Coloured Male	0.307	0.714	
WITS170	Females	White	White	0.818	0.958		White Female	0.418	0.756	White Male PP 0.334
WITS171	Females	White	Coloured	0.408	0.394	White PP 0.323	Coloured Female	0.289	0.275	White Female PP 0.209
WITS172	Females	White	Black	0.592	0.751		Coloured Female	0.294	0.515	Black Female PP 0.249
WITS173	Females	White	White	0.84	0.721		White Female	0.695	0.735	
WITS174	Females	White	White	0.716	0.13		White Female	0.684	0.258	
WITS175	Females	White	White	0.521	0.472	Coloured PP 0.431	White Female	0.709	0.544	
WITS176	Females	White	Coloured	0.571	0.079		White Female	0.491	0.649	
WITS177	Females	White	White	0.875	0.209		White Female	0.69	0.909	
WITS178	Females	White	White	0.951	0.239		White Female	0.81	0.178	
WITS179	Females	White	White	0.725	0.991		White Female	0.644	0.956	
WITS180	Females	White	Coloured	0.466	0.313		White Female	0.368	0.2	Black Female PP 0.318
WITS181	Females	White	Coloured	0.567	0.394		Coloured Female	0.494	0.559	
WITS182	Females	White	White	0.988	0.041		White Female	0.938	0.351	
WITS183	Females	White	White	0.516	0.682		White Female	0.569	0.8	
WITS184	Females	White	White	0.498	0.242		White Female	0.415	0.968	

FORDISC	RECORDED		ANCESTRY				SEX AND ANCESTRY			
Assigned no.	Record-ed sex	Recorded ancestry	Estimated ancestry	Posterior probability (PP)	Typicality F	Within 10% of correct PP	Estimated sex + ancestry	Posterior probability (PP)	Typicality F	Within 10% of correct PP
WITS185	Females	White	Coloured	0.593	0.976		Black Female	0.367	0.942	Coloured Female PP 0.363
WITS186	Females	White	White	0.856	0.075		White Female	0.82	0.297	
WITS187	Females	White	White	0.97	0.06		White Female	0.572	0.109	
WITS188	Females	White	White	0.717	0.736		White Female	0.66	0.945	
WITS189	Females	White	Coloured	0.56	0.896		Coloured Female	0.452	0.921	
WITS190	Females	White	White	0.486	0.536	Coloured PP 0.398	White Female	0.416	0.812	
WITS191	Females	White	White	0.915	0.648		White Female	0.665	0.816	
WITS192	Females	White	Coloured	0.403	0.989	White PP 0.343	White Female	0.253	0.898	Coloured Female PP 0.202
WITS193	Females	White	Coloured	0.686	0.992		Coloured Female	0.301	0.94	Black Female PP 0.226
WITS194	Females	White	White	0.606	0.71		White Female	0.572	0.607	
WITS195	Females	White	Black	0.678	0.52		Black Female	0.528	0.493	
WITS196	Females	White	Black	0.518	0.617	Coloured PP 0.424	Black Female	0.323	0.285	Black Male PP 0.291
WITS197	Females	White	White	0.911	0.069		White Female	0.804	0.74	
WITS198	Females	White	White	0.779	0.41		White Female	0.752	0.399	
WITS199	Females	White	White	0.513	0.346		White Female	0.433	0.942	
WITS200	Females	White	White	0.921	0.666		White Female	0.788	0.496	
WITS201	Females	White	White	0.858	0.102		White Female	0.823	0.312	
WITS202	Females	White	White	0.571	0.844		White Female	0.487	0.887	
WITS203	Females	White	White	0.994	0.626		White Female	0.716	0.477	
WITS204	Females	White	White	0.942	0.53		White Female	0.769	0.725	
WITS205	Females	White	Coloured	0.489	0.733		White Female	0.4	0.695	
WITS206	Females	White	White	0.684	0.431		White Female	0.602	0.28	
WITS207	Females	White	White	0.812	0.39		White Female	0.784	0.337	
WITS208	Females	White	White	0.869	0.63		White Female	0.592	0.748	
WITS209	Females	White	Coloured	0.776	0.393		Coloured Female	0.343	0.307	Black Female PP 0.273
WITS210	Females	White	Black	0.626	0.5		Coloured Female	0.499	0.457	
WITS211	Females	White	White	0.718	0.104		White Female	0.862	0.748	

FORDISC	RECORDED		ANCESTRY				SEX AND ANCESTRY			
Assigned no.	Record-ed sex	Recorded ancestry	Estimated ancestry	Posterior probability (PP)	Typicality F	Within 10% of correct PP	Estimated sex + ancestry	Posterior probability (PP)	Typicality F	Within 10% of correct PP
WITS212	Females	White	Coloured	0.486	0.767	White PP 0.455	White Female	0.413	0.782	
WITS213	Females	White	Coloured	0.448	0.788		Coloured Male	0.22	0.262	Black Female PP 0.187
WITS214	Females	White	Coloured	0.477	0.609		Coloured Female	0.408	0.568	White Female PP 0.360
WITS215	Females	White	White	0.948	0.823		White Female	0.785	0.913	
WITS216	Females	White	Black	0.607	0.914		Black Female	0.554	0.96	
WITS217	Females	White	White	0.852	0.669		White Female	0.785	0.745	
WITS218	Females	White	White	0.615	0.377		White Female	0.562	0.523	
WITS219	Females	White	White	0.977	0.638		White Female	0.784	0.412	
WITS220	Females	White	White	0.704	0.913		White Female	0.581	0.844	
WITS221	Females	White	Black	0.546	0.887		Black Female	0.448	0.987	Coloured Female PP 0.414
WITS222	Females	White	White	0.971	0.078		White Female	0.725	0.517	
WITS223	Females	White	Coloured	0.567	0.707		Coloured Male	0.289	0.498	Coloured Female PP 0.212
WITS224	Females	White	White	0.598	0.301		Coloured Male	0.226	0.6	White Female PP 0.184
WHITE MALES										
WITS225	Male	White	Black	0.807	0.846		Black Male	0.555	0.904	
WITS226	Male	White	Black	0.641	0.445		Black Male	0.375	0.983	Coloured Male PP 0.287
WITS227	Male	White	White	0.701	0.953		White Male	0.398	0.739	White Female PP 0.347
WITS228	Male	White	White	0.928	0.418		White Female	0.78	0.473	
WITS229	Male	White	Coloured	0.484	0.316		Coloured Female	0.329	0.699	
WITS230	Male	White	White	0.918	0.988		White Male	0.845	0.986	
WITS231	Male	White	Black	0.561	0.476		Black Male	0.473	0.202	
WITS232	Male	White	White	0.504	0.759		White Female	0.329	0.649	
WITS233	Male	White	Black	0.661	0.615		Coloured Female	0.452	0.623	Black Female PP 0.385

FORDISC	RECORDED		ANCESTRY				SEX AND ANCESTRY			
Assigned no.	Record-ed sex	Recorded ancestry	Estimated ancestry	Posterior probability (PP)	Typicality F	Within 10% of correct PP	Estimated sex + ancestry	Posterior probability (PP)	Typicality F	Within 10% of correct PP
WITS234	Male	White	Coloured	0.464	0.363		Coloured Female	0.313	0.31	White Female PP 0.272
WITS235	Male	White	Coloured	0.437	0.514		White Male	0.455	0.909	
WITS236	Male	White	White	0.976	0.783		White Female	0.532	0.914	
WITS237	Male	White	White	0.983	0.327		White Male	0.836	0.5	
WITS238	Male	White	White	0.583	0.88		White Female	0.282	0.699	White Male PP 0.201
WITS239	Male	White	White	0.424	0.659		White Male	0.354	0.629	Coloured Male PP 0.296
WITS240	Male	White	White	0.715	0.389		White Male	0.495	0.263	
WITS241	Male	White	Black	0.706	0.094		Black Male	0.579	0.146	
WITS242	Male	White	White	0.623	0.397		White Female	0.484	0.532	
WITS243	Male	White	White	0.659	0.395		White Female	0.331	0.784	
WITS244	Male	White	White	0.741	0.221		White Male	0.374	0.771	White Female PP 0.333
WITS245	Male	White	White	0.968	0.295		White Female	0.508	0.225	White Male PP 0.447
WITS246	Male	White	White	0.817	0.581		White Male	0.495	0.512	
WITS247	Male	White	Black	0.48	0.189	Coloured PP 0.381	Coloured Female	0.259	0.095	Black Male PP 0.246
WITS248	Male	White	Black	0.555	0.21		Black Male	0.39	0.387	Coloured Male PP 0.378
WITS249	Male	White	White	0.471	0.801	Coloured PP 0.401	White Male	0.387	0.699	Coloured Male PP 0.343
WITS250	Male	White	White	0.856	0.256		White Male	0.599	0.376	
WITS251	Male	White	White	0.442	0.995	Coloured PP 0.377	White Male	0.338	0.988	Coloured Male PP 0.272
WITS252	Male	White	White	0.524	0.669		White Male	0.469	0.655	Coloured Male PP 0.369
WITS253	Male	White	White	0.651	0.674		White Male	0.684	0.673	
WITS254	Male	White	White	0.894	0.186		White Male	0.815	0.256	
WITS255	Male	White	White	0.891	0.638		White Male	0.703	0.756	
WITS256	Male	White	Black	0.584	0.869		Black Male	0.379	0.477	

FORDISC	RECORDED		ANCESTRY				SEX AND ANCESTRY			
Assigned no.	Record-ed sex	Recorded ancestry	Estimated ancestry	Posterior probability (PP)	Typicality F	Within 10% of correct PP	Estimated sex + ancestry	Posterior probability (PP)	Typicality F	Within 10% of correct PP
WITS257	Male	White	Coloured	0.453	0.785	White PP 0.353	Coloured Male	0.354	0.852	
WITS258	Male	White	White	0.912	0.617		White Male	0.731	0.535	
WITS259	Male	White	White	0.917	0.947		White Male	0.373	0.879	White Female PP 0.360
WITS260	Male	White	White	0.893	0.695		White Female	0.541	0.796	
WITS261	Male	White	Coloured	0.543	0.527		Coloured Male	0.49	0.79	
WITS262	Male	White	White	0.867	0.686		White Male	0.819	0.712	
WITS263	Male	White	White	0.828	0.91		White Male	0.766	0.853	
WITS264	Male	White	Black	0.596	0.464		Black Male	0.513	0.636	
WITS265	Male	White	White	0.852	0.366		White Male	0.638	0.543	
WITS266	Male	White	White	0.847	0.883		White Male	0.804	0.925	
WITS267	Male	White	Coloured	0.491	0.503		Coloured Female	0.274	0.822	White Female PP 0.253
WITS268	Male	White	Coloured	0.493	0.365		Coloured Male	0.378	0.726	
WITS269	Male	White	White	0.413	0.947	Coloured PP 0.337	White Male	0.253	0.952	Coloured Male PP 0.2748
WITS270	Male	White	White	0.811	0.503		White Male	0.538	0.943	
WITS271	Male	White	White	0.894	0.116		White Male	0.647	0.633	
WITS272	Male	White	White	0.431	0.579	Coloured PP 0.365	White Female	0.324	0.417	
WITS273	Male	White	White	0.551	0.501		White Female	0.202	0.478	Black Male PP 0.197
WITS274	Male	White	White	0.708	0.92		White Male	0.478	0.828	
WITS275	Male	White	White	0.687	0.559		White Male	0.587	0.477	
WITS276	Male	White	White	0.68	0.181		White Male	0.424	0.188	White Female PP 0.358
WITS277	Male	White	Coloured	0.524	0.813		Coloured Male	0.288	0.688	Black Male PP 0.278
WITS278	Male	White	White	0.873	0.678		White Female	0.455	0.651	White Male PP 0.435
WITS279	Male	White	Coloured	0.391	0.699	Black PP 0.348	Black Male	0.252	0.86	Coloured Male PP 0.210

FORDISC	RECORDED		ANCESTRY				SEX AND ANCESTRY			
Assigned no.	Record-ed sex	Recorded ancestry	Estimated ancestry	Posterior probability (PP)	Typicality F	Within 10% of correct PP	Estimated sex + ancestry	Posterior probability (PP)	Typicality F	Within 10% of correct PP
WITS280	Male	White	White	0.599	0.683		White Male	0.37	0.874	Coloured Male PP 0.318
WITS281	Male	White	White	0.524	0.897		Coloured Male	0.298	0.68	White Male PP 0.274
WITS282	Male	White	White	0.65	0.037		White Female	0.628	0.557	
WITS283	Male	White	Black	0.451	0.134		Black Male	0.406	0.161	
WITS284	Male	White	White	0.635	0.12		White Male	0.489	0.021	
WITS285	Male	White	White	0.798	0.996		White Female	0.465	0.967	
WITS286	Male	White	Black	0.444	0.961	Coloured PP 0.402	Black Male	0.413	0.933	
WITS287	Male	White	White	0.891	0.855		White Male	0.658	0.913	
WITS288	Male	White	White	0.92	0.403		White Male	0.673	0.743	
WITS289	Male	White	Coloured	0.498	0.44	White PP 0.434	Coloured Female	0.397	0.559	
WITS290	Male	White	White	0.962	0.724		White Male	0.928	0.85	
WITS291	Male	White	White	0.611	0.843		White Male	0.411	0.937	
WITS292	Male	White	Coloured	0.453	0.423	White PP 0.392	Coloured Female	0.461	0.784	
WITS293	Male	White	White	0.956	0.133		White Female	0.606	0.9	
WITS294	Male	White	White	0.974	0.225		White Female	0.583	0.581	
WITS295	Male	White	Black	0.448	0.793		Black Male	0.425	0.811	
WITS296	Male	White	White	0.939	0.379		White Male	0.806	0.235	
WITS297	Male	White	White	0.679	0.652		White Male	0.442	0.689	
WITS298	Male	White	White	0.794	0.169		White Male	0.692	0.416	
WITS299	Male	White	White	0.632	0.609		White Male	0.384	0.345	