Facial Soft Tissue Thickness and Facial Recognition in Black South African Adults

by,

Keegan O. Meiring

Student number: 547058

A dissertation submitted to the School of Anatomical Sciences, Faculty of Health Science, The University of the Witwatersrand, South Africa, in fulfilment of the requirements for the degree **Master of Science in Medicine**

Johannesburg, South Africa

March 27th, 2019

Supervisor: Nanette Briers, Ph.D.

Declaration

I, Keegan Oliver Meiring, declare that this dissertation entitled “Facial Soft Tissue Thickness and Facial Recognition in Black South African Adults” is my own unaided work. It is being submitted for the Degree of Master of Science in Medicine at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at any other University.

Signed:  Date: March 27th, 2019

Presentations from this study

1. Oral presentation at the Anatomical Society of South Africa, 46th Annual Conference (April, 2018).
2. Poster presentation at 2018 Wits Research day & Postgraduate expo (September, 2018).

Abstract

Forensic facial approximations are used when conventional methods of identification provide insufficient results. The success of these facial approximations are founded on access to the applicable facial soft tissue thicknesses (FSTT), the culmination of muscle and connective tissue extending the distance from the outer surface of the bone to the external extreme of the skin. The aims of this study were to supplement existing South African FSTT data and investigate their reliability by utilizing facial recognition tests.

Sixty computed tomography (CT) scans were obtained retrospectively from the Wits Donald Gordon Medical Centre (Johannesburg) after obtaining ethical clearance. On each scan, FSTTs were measured at 10 midline and 10 bilateral landmarks. South African sex-specific & pooled data, as well as FSTT from the C-Table repository (available through http://www.CRANIOFACIALidentification.com) were utilized to perform several facial approximations using the Manchester method. Assessors were then asked to study the facial arrays and indicate which photographs of the 20 presented to them resembled the facial approximations.

The FSTT measurements were collected for 47 females (78.3%) and 13 males (21.7%). The range of the midline and lateral means were 2.4 to 11.3 mm and 6.5 to 29.8 mm respectively. The range of the midline and lateral shorths were 1.8 to 10.4 mm and 5.6 to 27.5 mm respectively. Moreover, the 75-shormaxes of midline and lateral landmarks ranged from 3.3 to 13.2 mm and from 8.4 to 37.6 mm respectively. Statistical significance between sexes for FSTT, was found to only exist at the mid-zygomatic landmark (*t*-test; $p<0.05$). One significant difference was accounted for at the supra M_2 pertaining to bilateral asymmetry. Age-related differences, mostly in the vicinity of the eyes and nose, were however, statistically significant at several midline and bilateral landmarks. The facial approximations

produced in this study displayed subtle changes depending on the data used to achieve them (i.e.: sex-specific or pooled data). Variations in the appearance of the jaw line, parotid region and mouth were the most noticeable; however, these did not influence recognisability. All categories of assessors exhibited above-chance success during the facial array tests. The female approximations were paired with correct photograph 30% more frequently than the male approximations; however, the difference was not statistically significant. White females with experience of anatomy and/or art displayed the highest success during the facial arrays and indicated that the bridge of the nose and breadth of the nostrils were the most telling features.

The results of this study have uncovered that the categorization of FSTTs does not significantly influence the recognition rates of facial approximation. Facial approximations produced using pooled data were identified more easily.

Furthermore, recognition biases associated with sex and ancestry were not significant in the current South African study.

Keywords

Facial arrays, Facial soft tissue thickness, Forensic facial approximation, Recognition, South Africa

Acknowledgements

I would like to thank the South African National Research Foundation (NRF) and the J.J.J. Smiezsek Fellowship for their sponsorship of this research. I would also like to acknowledge the Human Research Ethics Committee (HREC) of the University of the Witwatersrand for its approval to conduct this study.

I would like to thank my supervisor, Dr. N. Briers, for granting me this opportunity and for her continual advice and motivation throughout the duration of this project. Further thanks to Dr. M. Haagensen of the Department of Radiology of the Wits Donald Gordon Medical Centre (WDGMC) for granting this study the permission to collect and utilize CT scans from his department. Furthermore, thank you for having the time and patience to teach me how to navigate the software programs. Thank you to Dr T. Houlton of the University of the Witwatersrand for his efforts in the production of the forensic facial approximations used in this research. Thank you to Dr. J. Hemingway, also of the University of the Witwatersrand, for his assistance with the statistical analysis. I would also like to acknowledge the help of Ms. K. Els who performed the inter-observer reliability. Thank you for your interest and assistance.

Thank you to all of those who granted me permission to use their CT scans and photographs, as well as those who volunteered their time to participate in the facial array surveys. Your contribution has been invaluable towards this research.

A very special thank you goes to Lara for giving me strength and hope that I would complete this arduous task. Lastly, I would like to thank my mother, Heather and my sister Kirstin for their unwavering love and support throughout this project and my studies. Mom, I am all I am because of you.

Table of contents

Declaration **i**

Presentations from this study **ii**

Abstract **iii**

Acknowledgements **v**

Table of contents **vi**

List of figures **xi**

List of tables **xv**

List of abbreviations **xvii**

Chapter 1: Introduction

1.1. Introduction 1

1.2. Problem statement, aims & objectives 2

Chapter 2: Literature review

2.1 Introduction 4

2.2 Terminology 5

2.3 Identification of skeletal remains 5

2.4 The history of facial approximations 7

2.5 Forensic facial approximations 8

2.5.1 Methodologies 10

2.5.2 American method 10

2.5.3 Russian method 11

2.5.4 Manchester method 13

2.6 Anatomical landmarks and facial soft tissue thickness 14

2.6.1 FSTT landmarks 14

2.6.2 Collecting FSTT measurements 15

2.6.3 Analysis of FSTT values 18
2.7 Factors influencing FSTT 19
2.8 Facial soft tissue thickness in South African populations 22
2.9 Facial soft tissue thickness outside of Africa 23
2.9.1 European studies 23
2.9.2 Asian studies 26
2.9.3 American studies 27
2.9.4 Australian studies 31
2.10 Facial recognition 34
2.10.1 Own-population bias 35
2.10.2 Own-sex bias 37
2.10.3 Own-age bias 38
2.10.4 Recognition testing 38

Chapter 3: Methodology

3.1 Study design 40
3.1.1 Facial soft tissue thickness 40
3.1.2 Facial approximations and arrays 40
3.2 Site of study 40
3.2.1 Facial soft tissue thickness 40
3.2.2 Facial approximations and arrays 40
3.3 Study sample 41
3.4 Data collection 42
3.4.1 Facial soft tissue thickness 42
3.4.2 Facial approximations and arrays 50
3.5 Data analysis 54
3.6 Ethics 55

Chapter 4: Results

4.1 Introduction 57
4.2 Sample demographics 57
4.3 FSTT measurements of landmarks 58

4.4	Intra- and inter-observer error	87
4.5	Morphological analysis	93
4.5.1	Facial approximations	93
4.5.2	Facial arrays	97
4.5.3	Facial features	110

Chapter 5: Discussion

5.1	Introduction	115
5.2	Facial soft tissue thickness measurements	116
5.3	Facial approximation and facial arrays	123
5.4	Repeatability	126
5.5	Samples and protocols	128
5.6	Limitations	129
5.6.1	Sampling	129
5.6.2	Landmark identification	130
5.6.3	Photographs for facial arrays	132
5.7	Significance of this study	132

Chapter 6: Conclusion

Conclusion		135
Recommendations		137

References		138
Appendix 1	Scoring sheet	149
Appendix 2.1.1	Facial array presenting frontal view photographs of female participants and female forensic facial approximation (subject A) (in the centre) constructed utilizing sex-specific FSTT data	150
Appendix 2.1.2	Facial array presenting lateral view photographs of female participants and female forensic facial approximation (subject A) (in the centre) constructed utilizing sex-specific FSTT data	151

Appendix 2.1.3	Facial array presenting three-quarter view photographs of female participants and female forensic facial approximation (subject A) (in the centre) constructed utilizing sex-specific FSTT data	152
Appendix 2.2.1	Facial array presenting frontal view photographs of male participants and male forensic facial approximation (subject B) (in the centre) constructed utilizing sex-specific FSTT data	153
Appendix 2.2.2	Facial array presenting lateral view photographs of male participants and male forensic facial approximation (subject B) (in the centre) constructed utilizing sex-specific FSTT data	154
Appendix 2.2.3	Facial array presenting three-quarter view photographs of male participants and male forensic facial approximation (subject B) (in the centre) constructed utilizing sex-specific FSTT data	155
Appendix 2.3.1	Facial array presenting frontal view photographs of female participants and female forensic facial approximation (subject C) (in the centre) constructed utilizing pooled FSTT data	156
Appendix 2.3.2	Facial array presenting lateral view photographs of female participants and female forensic facial approximation (subject C) (in the centre) constructed utilizing pooled FSTT data	157
Appendix 2.3.3	Facial array presenting three-quarter view photographs of female participants and female forensic facial approximation (subject C) (in the centre) constructed utilizing pooled FSTT data	158
Appendix 2.4.1	Facial array presenting frontal view photographs of male participants and male forensic facial approximation (subject D) (in the centre) constructed utilizing pooled FSTT data	159
Appendix 2.4.2	Facial array presenting lateral view photographs of male participants and male forensic facial approximation (subject D) (in the centre) constructed utilizing pooled FSTT data	160
Appendix 2.4.3	Facial array presenting three-quarter view photographs of male participants and male forensic facial approximation (subject D) (in the centre) constructed utilizing pooled FSTT data	161
Appendix 3	INFORMATION SHEET for taking and using facial	162

	photographs	
Appendix 4	INFORMATION SHEET to act as an assessor for facial arrays	164
Appendix 5	INFORMATION SHEET for the use of CT scans	166
Appendix 6	INFORMED CONSENT FORM for the taking and using of facial photographs	168
Appendix 7	INFORMED CONSENT FORM to act as an assessor for facial arrays	169
Appendix 8	INFORMED CONSENT FORM for the use of CT scans	170
Appendix 9	Ethics clearance certificate	171

List of figures

Figure 3.1	Frontal and lateral views illustrating anatomical landmarks which were measured for FSTT	43
Figure 3.2.1	Photographs labelled A-H illustrates the method used for collecting the FSTT measurement of the rhinion (Rhi) from a topogram view of a CT scan	48
Figure 3.2.2	Photographs labelled A-H illustrated the method used for collecting the FSTT measurement of the mid-ramus from a horizontal slice view of a CT scan	49
Figure 3.3	The production process of a male facial approximation using pooled data by employing the Manchester method	52
Figure 4.1	Box and whisker plots of the FSTT measurements for midline landmarks	63
Figure 4.2	Box and whisker plots of the FSTT measurements at the mid-supraorbital bilaterally	64
Figure 4.3	Box and whisker plots of the FSTT measurements at the mid-infraorbital bilaterally	66
Figure 4.4	Box and whisker plots of the FSTT measurements of the mid-zygomatic bilaterally	67
Figure 4.5	Box and whisker plots of the FSTT measurements of the zygoma bilaterally	68
Figure 4.6	Box and whisker plots of the FSTT measurements of the mid-mandibular bilaterally	69
Figure 4.7	Box and whisker plots of the FSTT measurements of the mid-ramus bilaterally	71
Figure 4.8	Box and whisker plots of the FSTT measurements of the gonion bilaterally	72
Figure 4.9	Box and whisker plots of the FSTT measurements of the supra M_2 bilaterally	73
Figure 4.10	Box and whisker plots of the FSTT measurements of the Sub M_2 bilaterally	75

Figure 4.11	Box and whisker plots of the FSTT measurements of the mental tubercle bilaterally	76
Figure 4.12	Box and whisker plots of the FSTT measurements of lateral landmarks of the left hand side	77
Figure 4.13	Box and whisker plots of the FSTT measurements of the lateral landmarks on the right hand side	77
Figure 4.14	Artistic impressions of male facial approximations	95
Figure 4.15	Artistic impressions of female facial approximations	96
Figure 4.16	Comparison of female facial approximations (A: sex-specific data & C: pooled data) to the positive match photograph (B)	98
Figure 4.17	Comparison of male facial approximations (A: sex-specific data & C: pooled data) to the positive match photograph (B)	98
Figure 4.18	Histogram depicting the percentage of correct identification of the facial arrays in the entire sample	99
Figure 4.19	Histogram depicting the percentage of correct identification of the facial array in different populations	100
Figure 4.20	Histogram depicting the percentage of correct identification of the facial arrays by females and males	101
Figure 4.21	Histogram depicting the percentage of correct identification of the facial arrays among people whom had and did not have formal training in anatomy and/or art	102
Figure 4.22	Histogram depicting the recognition frequency of the facial array photographs against the facial approximation Subject A by black and white assessors	103
Figure 4.23	Histogram depicting the recognition frequency of the facial array photographs against the facial approximation Subject B by black and white assessors	103
Figure 4.24	Histogram depicting the recognition frequency of the facial array photographs against the facial approximation Subject C by black and white assessors	104
Figure 4.25	Histogram depicting the recognition frequency of the facial array photographs against the facial approximation Subject D by black and white assessors	104

Figure 4.26	Histogram depicting the recognition frequency of the facial array photographs against the facial approximation Subject A by female and male assessors	105
Figure 4.27	Histogram depicting the recognition frequency of the facial array photographs against the facial approximation Subject B by female and male assessors	106
Figure 4.28	Histogram depicting the recognition frequency of the facial array photographs against the facial approximation Subject C by female and male assessors	106
Figure 4.29	Histogram depicting the recognition frequency of the facial array photographs against the facial approximation Subject D by female and male assessors	107
Figure 4.30	Histogram depicting the recognition frequency of the facial array photographs against the facial approximation Subject A by assessors with and without training in anatomy and/or art	108
Figure 4.31	Histogram depicting the recognition frequency of the facial array photographs against the facial approximation Subject B by assessors with and without training in anatomy and/or art	108
Figure 4.32	Histogram depicting the recognition frequency of the facial array photographs against the facial approximation Subject C by assessors with and without training in anatomy and/or art	109
Figure 4.33	Histogram depicting the recognition frequency of the facial array photographs against the facial approximation Subject D by assessors with and without training in anatomy and/or art	109
Figure 4.34	Histogram depicting the selection frequency of facial features per facial approximation (A-D) by the entire sample of assessors of the facial arrays	111
Figure 4.35	Histogram depicting the selection frequency of facial features per facial approximation (A-D) by persons of different ancestry	112
Figure 4.36	Histogram depicting the selection frequency of facial features per facial approximation (A-D) by persons of different sex	113
Figure 4.37	Histogram depicting the selection frequency of facial features per facial approximation (A-D) amongst people whom had (x) and did	114

not have (y) formal training in anatomy and/or art

Figure 5.1 Photographs A-C are topogram views of CT scans demonstrating the varying amounts light scatter reflecting from metallic dental implants. 131

Figure 5.2 Photographs A-C are frontal (A) and lateral (B & C) depiction of a patient's skull. 131

List of tables

Table 3.1	Midline and bilateral landmarks, and its respective descriptions	44
Table 4.1	Basic descriptive statistics for the FSTT measurements of 60 black South African adults between the ages of 19 and 70 years	78
Table 4.2	Basic descriptive statistics for the FSTT measurements of 47 black South African adult females between the ages of 19 and 70 years	79
Table 4.3	Basic descriptive statistics for the FSTT measurements of 13 black South African adult males between the ages of 29 and 66 years	80
Table 4.4	Basic descriptive statistics for the FSTT measurements of 3 black South African adults between the ages of 19 and 29 years	81
Table 4.5	Basic descriptive statistics for the FSTT measurements of 11 black South African adults between the ages of 30 and 39 years	82
Table 4.6	Basic descriptive statistics for the FSTT measurements of 15 black South African adults between the ages of 40 and 49 years	83
Table 4.7	Basic descriptive statistics for the FSTT measurements of 23 black South African adults between the ages of 50 and 59 years	84
Table 4.8	Basic descriptive statistics for the FSTT measurements of 8 black South African adults aged 60 years and above	85
Table 4.9	Table 4.9.: Means, shorths and 75-shormaxes of FSTT data at midline and bilateral anatomical landmarks for the entire sample (females and males) of the current study.	86
Table 4.10	Intra-observer repeatability of midline landmarks expressed by the intra-class correlation coefficient (ICC), technical error of measurement (TEM) and the variable average value	89
Table 4.11	Intra-observer repeatability of bilateral landmarks expressed by the intra-class correlation coefficient (ICC), technical error of measurement (TEM) and the variable average value	89
Table 4.12:	Inter-observer repeatability of midline landmarks expressed by the intra-class correlation coefficient (ICC), technical error of measurement (TEM) and the variable average value	91
Table 4.13	Inter-observer repeatability of bilateral landmarks expressed by the	91

intra-class correlation coefficient (ICC), technical error of measurement (TEM) and the variable average value

Table 5.1 Comparison of midline FSTT data from the current study with Aulsebrook *et al*. (1996), Cavanagh & Steyn (2011) (C&S) and Stephan &Simpson (2008a) (S&S) 120

Table 5.2 Comparison of midline and left (L) lateral FSTT data from the current study (A) with Stephan & Priesler (2018) (B), Stephan & Sievwright (2018) (C) and Thiemann *et al*. (2017) (D) 121

List of abbreviations

BMI – Body mass index

CBCT – Cone beam computed tomography

CT – Computed tomography

DICOM – Digital Imaging and Communications in Medicine

EDA – Exploratory data analysis

FSTT – Facial soft tissue thickness

HREC – Human Research Ethics Committee

ICC – Intra-class correlation

MPR – Multi-planar reconstruction

MRI – Magnetic resonance imaging

MSCT – Multi-slice computed tomography

OAB – Own-age bias

OPB – Own-population bias

OSB – Own-sex bias

ROI – Region of interest

r-TEM – Relative technical errors of measurement

SD – Standard deviation

TDStats – Tissue depth statistics

TEM – Technical error of measurement

VAV – Variable average value

WDGMC – Wits Donald Gordon Medical Centre

Chapter 1 – Introduction

1.1. Introduction

When confronted with heavily decomposed or skeletonized remains, it is generally not possible to acquire DNA for the purposes of victim identification. When scientists are, however, able to collect viable tissue for analysis it may prove futile if there is no reference material for comparison (Chung *et al*., 2015). In this eventuality, the focus is shifted to the analysis of the skeleton in an attempt to compile a biological profile of the deceased. Since the early 20th century, anthropologists have engaged in the questionable practice of facial approximations (Christensen & Anderson, 2013). The process involves the resurfacing of a skull with artificial soft tissues in order to produce the possible ante mortem appearance of the deceased (Tedeschi-Oliveira *et al*., 2009). In order to do this, most practitioners rely on relevant facial soft tissue thickness (FSTT) datasets or repositories to assist them in rebuilding a face from bone alone. It has been a century since facial approximations were first utilized in the field of forensics, but many experts have yet to reach a consensus about the specific criteria necessary for its practice. For example, many modern day studies still refer to the topic as forensic reconstruction despite the controversy associated with the term (Stephan, 2015a). The methodologies pertaining to the collection of FSTT are also highly contentious. One major dispute pertains to the medical imaging equipment now available, and which of the wide variety of options are the best at measuring FSTTs (Domaracki & Stephan, 2006). The most debated part is whether facial approximations are actually worthwhile. It is argued that the subjectivity of the forensic artist will inevitably influence the facial approximation (Stephan & Henneberg, 2001). The field has been heavily criticized, especially as it is labelled by some as “scientific art” (Wilkinson, 2010). Despite their many short-comings,

forensic facial approximations have acted as last resorts when the more scientific approaches (e.g. DNA) are unable to yield any results (Wilkinson, 2004). Its primary purpose is not to necessarily establish a positive identification, but to instead serve as a trigger to obtain ante mortem information from the general public that in turn may assist in narrowing down the identity of the deceased.

1.2. Problem statement, aims & objectives

Unfavourable socio-economic circumstances are largely responsible for the high murder rate in South Africa (Bhorat *et al.*, 2017). Many of these deaths occur in rural areas and human remains can go unnoticed and can severely decompose if concealed well enough by the environment (Keyes *et al.*, 2016). As a result, there is a high demand for more forensic-orientated research to address the lack of information available on FSTT. There are currently only four FSTT studies (Aulsebrook *et al.*, 1996; Briers *et al.*, 2015; Cavanagh & Steyn, 2011; Phillips & Smuts, 1996) that have been conducted in South African black populations. This research is also intended to determine whether population-specific FSTT data are necessary to produce facial approximations that are recognisable within a South African context.

Considering this, the aims and objective of this research were to:

<u>Aim I</u>

Supplement the existing data used for facial approximations in a black South African population.

<u>Objectives</u>

a. Describe the sexual dimorphism of facial soft tissue thicknesses using midline and bilateral anatomical landmarks.

b. Determine whether a correlation exists between facial soft tissue thicknesses and age by utilizing midline and bilateral anatomical landmarks.
c. Ascertain the differences between descriptive statistics and the tissue depth statistics (TDStats) in terms of FSTTs.

Aim II

Ascertain the reliability of the data collected in this study with respect to facial array testing.

a. Determine whether pooled or subcategorized FSTT data resulted in a more recognisable facial approximation.
b. Describe the same- and cross-population and sex effects with respect to facial recognition.
c. Determine which facial features are recognized most frequently in same- & cross-population and sex faces.

Chapter 2: Literature review

2.1. Introduction

The ability to determine the identity of unknown, decomposed or skeletonized human remains is part of the domain of forensic anthropology. This is most commonly achieved by producing a possible ante-mortem appearance of the deceased or facial approximation - a technique that has become a pivotal feature of medico-legal investigations (Christensen & Anderson, 2013). Facial approximations are made possible by using facial soft tissue thicknesses (FSTTs) which are also helpful in the realms of plastic surgery and palaeoanthropology (Kurkcuoglu *et al*., 2011). When producing a facial approximation, it is generally ideal to consider the morphology of the individual skull in conjunction with the FSTTs. However, much debate exists regarding the validity of FSTTs. Usually measurements are collected from only a small proportion of a population and hence, they are not a true representation of the whole population (De Greef *et al*., 2009). Stephan & Sievwright (2018) stated that the values used for facial approximations are most commonly only the mean values and as such, ignore the natural variation amongst individuals.

This literature review will highlight the history of the practice of facial approximations, as well as the importance, reliability and utilization of FSTT. Current work performed within South African and on an international level will be emphasized, as well as the mainstream methodologies used for forensic facial approximations. Lastly, the psychology of facial recognition will discussed in terms of general trends and how it influences facial approximation recognition rates.

2.2. Terminology

The reproduction of a face from a dry skull provides a possible ante-mortem appearance, but the process has been heavily criticized, especially due to the methodologies incorporated and associated errors (Stephan, 2015a). The term 'facial approximation' is now deemed more suitable and less confusing than the term 'facial reconstruction' (Stephan, 2015a). Reconstruction implies that the result is an ante-mortem replica of the deceased, whereas, an approximation provides an idea of what the deceased may have looked like sometime during their life (Evision *et al*., 2016). The primary focus of forensic facial approximations is recognition, while, facial reconstructions function in determining identity (Wilkinson, 2005). Facial reconstructions are used to emphasize the importance of idiosyncrasies pertaining to an individual skull when developing the deceased's identity, whereas, facial approximations are more inclined to apply pre-determined FSTTs in accordance with sex, age and population affinity (Wilkinson, 2005). Moreover, Panenkova (2007) stated that features influential in recognitions cannot be predicted from cranial remains alone. In the current study, the process of restoring a face onto a dry skull will be referred to as facial approximation.

2.3. Identification of skeletal remains

Establishing the identity of unknown individuals is of legal and moral importance as it not only permits the family of the deceased to address matters concerning personal assets, but also provides a sense of closure for the bereaved (Leo *et al*., 2013). Identification can be classified as being tentative, presumptive or positive depending on the information available. Factors of individualization (e.g. tattoos, scars, skeletal irregularities) and analysis of biological material (e.g. DNA, dental records) are

generally utilized to ascertain the identity of unknown remains (Solla & Dominguez, 2018). However, on discovery of skeletal or severely decomposed remains, investigators and other members of the law enforcement fraternity are unable to conduct any molecular based testing due to the absence of viable biological soft tissue. In instances such as these, forensic anthropologists will attempt to compile a biological profile by estimating the sex, age, stature and ancestry of the deceased (Solla & Dominguez, 2018). Furthermore, the skills of a forensic artist can be incorporated to attempt to produce a facial approximation that can assist in the identification process (Herrera *et al*., 2017). Forensic facial approximations can be produced via two- or three-dimensional drawings, clay models or computer generated images (Herrera et al., 2017). These approximations are performed with the objective that the public will be able to assist law enforcement agencies in obtaining possibly relevant ante-mortem information (Herrera et al., 2017).

Producing a facial approximation relies on information pertaining to the facial soft tissue thicknesses (FSTTs) and corresponding bony landmarks on the skull (Herrera et al., 2017). Many researchers suggest that consideration should be given to the biological profile (e.g. sex, age, ancestry) of the deceased and that the appropriate category of FSTTs should be used (Randolph-Quinney *et al*., 2009). Conversely, other studies have indicated that data pooling is more effective as it increases reliability by eliminating the effects of small sample sizes and data noise (Stephan *et al.*, 2016).

Ultimately, the process of facial approximation provides investigators with insight into the appearance of their person of interest and hopefully facilitates positive identification.

2.4. The history of facial approximations

Since the Neolithic Age, there has been evidence of restoring faces of the dead for the purposes of funerary, ancestor-worship (Evison *et al*., 2016) or scientific resolve (Verzé, 2009). In 1953, skulls with built up faces made of plaster, believed to have been buried between 7000 – 10000 years ago, were excavated beneath the floors of houses in the city of Jericho, Jordan (Verzé, 2009). This discovery is known to be the earliest evidence of facial approximtion and functioned in ancestral worship (Verzé, 2009). Another early example of facial approximations includes Egyptian mummification, dating back to 1370 B.C. (Wilkinson, 2004). Before the end of the 19th century, the only purpose of facial approximations was for creating death masks and not for recognition testing (Verzé, 2009).

The first work in the field of facial approximation and recognition was performed and recorded by a German anatomist, Wilheim His, in 1895. He produced a facial reconstruction of Johann Sebastian Bach, the famous 17th century German composer. Wilheim His achieved this by using FSTT depths provided by another German anatomist and physiologist, Hermann Welcker (Stephan, 2015a). Welcker established his dataset by inserting blades into the soft tissue of cadavers, which His further refined by using rubber-ended needles that reduced the distortion of the facial soft tissue (Vanezis & Vanezis, 2000). In 1898, Kollman and Buchly published their research pertaining to a face they had produced from a skull of an unknown individual (Stephan, 2015a). Stephan (2015a) argues that the work of Kollman and Buchly should be considered the first scientific attempt at face prediction as the approximation of Bach was based on ante-mortem portraits together with a skull presumed to belong to Bach. Welcker was the first to record facial soft tissue thickness measurements, which he subsequently used to perform approximations of

other celebrated individuals of the time (e.g. Schiller, Kant and Dante) (Omstead, 2002). Towards the end of the 19th century, Kollman & Buchly investigated the FSTTs of white males and females (Vanezis & Vanezis, 2000). Their results were found to correspond with those of His and are still considered as appropriate norms for the restoration of faces for people categorised as white (Vanezis & Vanezis, 2000). Throughout the 20th century many researchers pursued the practice of facial prediction. These included Mikhal Gerasimov, Wilton M. Krogman and Richard Neave. Gerasimov was a Russian archaeologist and inventor of the anatomy-based Russian method of facial approximation (Stephan, 2015a). Krogman was an American anthropologist and inventor of the measurement-based American method (Stephan, 2015a). Neave is a British forensic artist and collaborator of the Manchester or combination method (Stephan, 2015a). The principles of all these methods are widely used in modern day forensic settings as well as in the sphere of palaeoanthropology (Verzé, 2009).

2.5. Forensic facial approximation

Forensic craniofacial approximation is the process whereby cranial remains are utilized to estimate the identity of an unknown individual (Gupta *et al*., 2015). It is dependent on the correlations between the facial features, subcutaneous soft tissues and the underlying bony structure of the skull (Cavanagh & Steyn, 2011). Alternatively, craniofacial approximation is also useful for paleontological and archaeological studies that assist us in perceiving the appearance of our ancestors or other historic Figures (Lee *et al*., 2014). During the growth and developmental stages of life, small variations in the bony structures and FSTTs will arise and are responsible for the large amount of facial variation within humankind (Wilkinson, 2010). Lodha *et al*. (2016) stated that the variation in the facial appearances of

humans is the result of small variations in muscular and cranial idiosyncrasies. Wilkinson (2004) emphasized that facial approximations serve as a final resort when conventional methods of identification are inadequate, but their success is heavily dependent on publicity and social awareness campaigns.

Currently, facial approximations are conducted by using two- or three- dimensional techniques dependent on the expertise, technology and resources available (Basu & Nag, 2015). These techniques include the use of facial soft tissue thicknesses, the position and shape of facial muscles, or both (Gupta *et al*., 2015). These facial approximations are ultimately presented as sketches, clay sculptures or three-dimensional computer images (Gupta *et al*., 2015). Factors such as an individual's biological profile (e.g. sex, age, stature and ancestry) and their personal skeletal variation are influential when attempting a facial approximation (Lodha *et al*., 2016). Generally, a facial approximation can only provide an estimated ante-mortem appearance of the deceased. Hair styles, hair textures, facial hair in males, and factors of individualization (e.g. tattoo, scars, wrinkles) are documented to have a crucial impact on an assessor whilst considering if a facial approximation correlates with the deceased (Starbuck & Ward, 2007). The Rule of Thumbs, a set of principles formulated by Krogmann in 1946, highlights the necessity for the correct positioning of the eyes, ear, nose and mouth when producing facial approximations (De Greef *et al*., 2009). The implementation of Krogman's principles facilitates the process of recognition (Herrera *et al*., 2017) and allows for the quantification of results and subsequently, enables repeatability (De Greef *et al*., 2009).

2.5.1. Methodologies

There are three methods that can be utilized to produce three dimensional forensic facial approximations: the American method (also known as the morphometric method); the Russian method (also known as the morphoscopic or Gerasimov's method); and the Manchester Method (also known as the British or Melbourne or Combination method) (Herrera *et al*., 2017).

Despite being developed in different parts of the world, the three techniques are noticeable similar to each other and not as differentiated as some would claim (Stephan, 2006). Both the American and Russian methods consider FSTT and facial anatomy and hence, are mixtures of the two (Stephan, 2006). Stephan (2006) suggests that the term 'Combination (Manchester)' method should be discarded as it introduces further confusion by suggesting it differs from the American and Russian methods. Despite the overwhelming evidence to conform to the recommendations provided by Stephan (2006), the three methods will be addressed individually so as to offer insight into the history of each.

2.5.2. The American method

The American method is completely dependent on reproducing the soft tissue of the face with the use of FSTT at multiple anatomical landmarks (Verzé, 2009). In 1946, Krogmann was the first to utilize FSTTs for the purposes of forensic identification, the practice underlying the American method (Herrera *et al*., 2017). Although Krogmann was the first to be credited for the inception of utilizing FSTTs for facial approximation, Stephan (2006) states that FSTT methods utilized for facial approximation had indeed begun in Germany. The practice of facial approximation in the United States of America was already in practice several decades earlier, but was

centred on archaeological practices rather than forensics (Starbuck & Ward, 2007). In 1915, McGregor used the technique to recreate faces of early hominids, although Wilder is believed to have brought this technique from Europe in order to recreate the faces of Native Americans in 1912 (Verzé, 2009). The method involves positioning the skull in the Frankfurt Horizontal plane and assigning rubber cylinders of corresponding FSTT heights at specific landmarks (Wilkinson, 2004). The FSTT is an aggregation of all soft tissue (e.g. muscle, fat, glands and connective tissue) between the hard surface of the bone and the outermost layer of the skin (Wilkinson, 2004). The American method (unlike the Russian method) is commonly reported to place less importance on the anatomy of the bony skull and that tissue absent between each landmark is interpolated (Claes *et al*., 2010). This is an unfavourable practice as the cranial morphology is integral in the shaping of facial musculature (Cavanagh & Steyn, 2011). This distinction pertaining to the homogenous nature of the American method is false, as Krogmann himself had emphasized the need to consider the architecture of a skull and that anatomical prowess was of great importance when performing the reconstruction (Stephan, 2006).

2.5.3. The Russian method

The Russian method was developed by the Russian anatomist Mikhail Gerasimov, and involves reproducing a face by replacing the muscles of mastication and facial expression onto a subject's skull with modelling materials primarily made of bees wax (Ullrich & Stephan, 2016). The literature on forensic facial approximation generally emphasizes that the Russian method does not incorporate the use of facial soft tissue depths, and that its practice only encompasses the shaping of the facial soft tissue in accordance to the shape of the skull (Ullrich & Stephan, 2016). However, it was documented that Gerasimov collected and recorded FSTTs from 71 fresh

cadavers, and used these FSTTs for his facial approximations (Ullrich & Stephan, 2016). The common misconception is that Gerasimov did not utilize FSTT, even though he had published FSTT at 19 landmarks and continued to supplement his dataset over time (Ullrich & Stephan, 2016). Gerasimov only created partial approximations and never included the muscles of facial expression as he thought it would be inaccurate to attempt to determine the precise attachments (Ullrich & Stephan, 2011). Instead, he only ever constructed the temporalis and masseter muscles (Ullrich & Stephan, 2011). Despite the misconceptions pertaining to Gerasimov's methods, he emphasized that great effort should be taken with the representation of the facial and neck musculature as these were highly recognisable features (Omstead, 2002). These inconsistencies regarding Gerasimov's method are attributed to the poor translation of his work from Russian into English (Ullrich & Stephan, 2016).

When the three-dimensional method was employed by Gerasimov, he positioned the skull in the Frankfurt Horizontal plane and restored half the face at a time, using the other half as a bony reference (Ullrich & Stephan, 2016). Gerasimov was also believed to have destroyed his first approximation and perform the approximation for a second time, for the purposes of testing accuracy (Ullrich & Stephan, 2016). Verzé (2009) explains that Gerasimov stated that although individual muscles differ in shape and size amongst all people, he could accurately rebuild a face by anticipating what musculature had previously existed via the remnants that persisted on a skull. In the beginning of Gerasimov's work, he included pyramidal shaped markers made of modelling mastic at various anatomical landmarks, but he began to omit this step as he had become familiar with the FSTTs and found it unnecessary to include it (Ullrich & Stephan, 2016).

In summary, the Russian or Anatomical method is an adaption of methods first used in Germany and shares similar foundations to the American method. Furthermore, the muscles of mastication were documented to have first been moulded by someone other than Gerasimov (Ullrich & Stephan, 2016). Gerasimov highlighted that specificities regarding the facial appendages and the eyes could be deducted with careful inspection of the respective bony areas of the skull (Wilkinson, 2004).

2.5.4. The Manchester method

Wilkinson (2004) described the Manchester method as an amalgamation of expertise ranging from facial anatomy and anthropology to artistry and expressionism. The method was developed by Richard Neave in 1977, a former forensic artist at the University of Manchester. The method incorporates the principles of FSTT and facial anatomy (Gupta *et al*., 2015), the foundations of the American and Russian methods respectively. This method is generally divided into three stages, namely: the preparation of the skull, moulding of the shape of the face, and sculpting of the facial features (Prag and Neave, 1997). Firstly, a cast of the skull is produced so as to preserve the original (Prag and Neave, 1997). Secondly, the skull is positioned in the Frankfurt Horizontal plane and pegs are placed on the surface of the skull to represent the FSTTs for the subsequent application of musculature made of clay. Finally, the facial features as described by Krogmann (1946) are sculptured. Wilkinson (2004) recommends the combination method as it provides more than satisfactory accuracy rates that are a result of remaining in the parameters of the subject's skull. Furthermore, the combination method limits an artist's subjectivity during the production of the facial approximation (Wilkinson, 2004).

2.6. Anatomical landmarks and facial soft tissue thickness

Craniofacial identification techniques such as facial approximation and facial superimposition rely upon facial soft tissue thickness measurements to provide data about the tissue depths at various anatomical landmarks and evaluate the contours of a face respectively (Domaracki & Stephan, 2006). Before any FSTT measurements can be collected, an investigator should seek to correctly identify the selected anatomical landmarks. The section below discusses the approach used by Caple & Stephan (2016) in terms of landmark selection. Caple & Stephan (2016) emphasized a need to standardize the nomenclature of all anatomical landmarks in order to prevent ambiguity amongst researchers. This is deemed necessary to increase the reliability of FSTT results and to standardize the practice of collecting FSTT for the purposes of facial approximations as the incorrect identification of landmarks will have adverse effects on the reliability of the FSTT measurements (Caple & Stephan, 2016).

2.6.1. FSTT landmarks

In the field of biological anthropology there has been consensus for the need to standardize methodologies and nomenclatures used to identify anatomical landmarks (Caple & Stephan, 2016). The selection of anatomical landmarks for the collection of FSTTs should be carefully considered and sourced. There are three categories of landmarks.Type I landmarks are defined as points where several different tissues intersect (Campomanes-Álvarez *et al*., 2015) and are naturally homogenous in humans (Caple & Stephan, 2016). Type II landmarks are identified as points of extreme curvature (Campomanes-Álvarez *et al*., 2015), and located in accordance to geometric standards (Caple & Stephan, 2016). Type III landmarks are typically located on a curved surface (Campomanes-Álvarez *et al*., 2015) and determined using

instrumentation (Caple & Stephan, 2016). The ability to ascertain the precise location of bony landmarks, as well as their corresponding soft tissue landmarks, is vital in the facial approximation process. Caple & Stephan (2016) identified inconsistencies in the nomenclature such as several terms being associated with a certain landmark, as well as landmarks being defined using ambiguous lay terms. Campomanes-Álvarez *et al*. (2015) tested the precision of forensic experts by issuing them with photographs for the purpose of correctly identifying anatomical landmarks commonly used in craniofacial superimposition. Their results indicated that the participants were able to identify type I and II landmarks with a much higher frequency than landmarks classified as type III (Campomanes-Álvarez *et al*., 2015). It was also reported that the midline type III landmark used in their study were identified with a considerably high accuracy compared to the lateral type III landmarks. FSTTs are not only adversely affected by incorrectly locating the bony landmarks, but are also affected by inconsistent and misleading definitions regarding the corresponding soft tissue landmarks (Caple & Stephan, 2016). The standardization of bony and soft tissue landmarks are necessary for ensuring that the directionality of the tangent measuring line is correct. Ultimately, these standardizations are highly influential towards better repeatability.

2.6.2. Collecting facial soft tissue thickness measurements

Before any facial approximation can be performed, the methodology for collecting the FSTT must be considered. The needle puncture method, radiography, ultrasounds, computed tomography (CT) and magnetic resonance imaging (MRI) comprise the techniques available for measuring the FSTT. There are advantages and disadvantages associated with each of these methods.

The needle-puncture method is the oldest of all available and currently used methods. It is commonly performed using cadavers and is advantageous as the equipment needed is inexpensive. The disadvantage is, however, that movement of the needles are a common occurrence. Radiography, ultrasounds, CT and MRI rival the needle puncture technique as all of these can be performed on the living, eliminating the effects of post-mortem dehydration (De Greef *et al*., 2006). Furthermore, all these techniques are non-contact (with the exception of ultrasounds), thereby eradicating measurement error associated with the manipulation of the soft tissue as experienced with the needle puncture method (Stephan & Simpson, 2008a). Despite the overwhelming advantages connected to the modern FSTT measuring techniques, factors such as high cost, the volunteers' exposure to radiation and the presence of imaging artefacts (e.g. light scatter as a result of metallic dental implants) cannot be overcome (Stephan & Simpson, 2008a). Of all the imaging equipment currently available for the analysis of FSTTs, MRI and CT scans have been deemed the best (Domaracki & Stephan, 2006). CT scans assists a researcher in measuring the FSTT more accurately as they are non-contact and provides better resolution (Hwang *et al*., 2012). However due to effects of gravity, CT scans should ideally be taken in the upright position (Munn & Stephan, 2018). The effect of gravity can be overcome with cone beam computed tomography (CBCT), a scanner predominately used for orthodontics and orthognathic surgery (Hwang *et al*., 2012). The images are taken while patients are upright and furthermore, radiation exposure is less than that of multi-slice CT (Hwang *et al*., 2012). An additional constraint associated with CT and CBCT is the need for a qualified radiographer to operate the machinery. Researchers may not have easy access to qualified radiographers nor may they have the finances to employ the services of one. However, Stephan & Preisler (2018) emphasized that

measuring FSTTs with ultrasound has become increasingly prevalent as it is relatively inexpensive, omits much less radiation and can be done in the upright position unlike many of the other technological methods. In the study performed by Munn & Stephan (2018), the effect of gravity was investigated by measuring FSTTs in both the upright and supine position with the use of B-mode ultrasound. This type of ultrasound is also known as two-dimensional mode (2D mode) and is the most commonly used type of ultrasound as it interprets the intensity of the echoes by displaying an image (Cobbold, 2006). It was found that the FSTTs in the vicinity of the eyes and cheeks varied up to four millimetres on average when taken in the supine position compared to the upright position (Munn & Stephan, 2018). Consequentially, the study postulated correction values that could be added or subtracted to several midline and lateral landmarks. These correction values, however, were cautiously recommended as the method used to collect the FSTT (B-mode ultrasound) yields considerably large measurement errors (Munn & Stephan, 2018).

With the advent of imaging equipment, these methods have been favoured over the needle-puncture method. Although these radiological methods are generally more time efficient, researchers tend to prefer conducting studies which involve the living based on the premise that FSTTs collected from cadavers are less accurate (Stephan & Simpson, 2008a). The aforementioned view was later discredited when Stephen & Preisler (2018) found that the differences in FSTT collected from the dead and living are insignificant and are most likely due to errors associated with measurements and study samples.

2.6.3. Analysis of FSTT values

FSTTs have generally been analysed using descriptive statistics (e.g. arithmetic means), a method believed to be erroneous due its inability to sufficiently describe central tendency (Stephan *et al*, 2013). Furthermore, the mean is noted to be limiting as it fails to recognise the natural variation between individuals and simply places them in one category (Stephan & Sievwright, 2018). A FSTT dataset is generally not normally distributed and displays skewness due to the presence of outliers (Stephan *et al*, 2013). Further considerations include small sample sizes (less than 40), which are then further sub-categorized according to sex, age, ancestry and body mass index (BMI) (Stephan *et al*., 2015b). These practices, in conjunction with using inappropriate methods of statistical analysis, are responsible for the significant statistical differences reported in many current FSTT studies (Stephan, 2014). Therefore, a statistical package known as Tissue Depth Statistics (TDStats) was developed for the purpose of introducing a more efficient and valid method of interpreting FSTTs (Stephan, 2018). The shorth and 75-shormax are two examples of TDStats capable of interpreting skewed FSTT data (Stephan, 2018). The shorth is described as the mean of the densest half of the data and the 75-shormax is the 75^{th} percentile between the shorth and the maximum value of a dataset. The shorth provides a point approximation of the apex of a FSTT dataset and the 75-shormax provides an estimate of central tendency within the right tail of a bell curve, a portion that could be utilized for obese individuals (Stephan *et al.*, 2016).

In addition to the shorth and 75-shormax, TDSstats performs exploratory data analysis (EDA). This type of analysis, unlike mere statistical significance tests, acknowledges the role of sampling errors and sam0ple sizes. The primary function of

EDA is to provide visual representation of the data to draw scientific inferences (Stephan, 2018).

2.7. Factors influencing FSTT

Sex, age, ancestry, and BMI are critical factors when creating an accurate facial approximation (Jeelani *et al*., 2017). However, Stephan (2017) suggested that the categorization of FSTT is superfluous as differences attributed to these categorizations are not statistical significant and are ultimately misleading. Establishing these biological features from decomposed or skeletonized remains are also rarely accurate and using subcategorized FSTT data is flawed from the onset (De Greef *et al*., 2009).

The influence of sex on FSTT was tested in a study by Stephan *et al*. (2016), wherein FSTTs were collected from 52 Australian adults. Their sample ranged in age from 18 to 30 years and the tissue thicknesses were measured using B-mode ultrasound (Stephan *et al*., 2016). It was argued that the stratification of FSTT data according to sex introduces errors associated with small sample sizes that are later misconstrued as illustrating significant differences (Stephan *et al*., 2016). By normalising BMI, Stephan *et al*. (2016) reported that females had larger FSTTs than males by a factor of 2.7. This supports the findings of an Egyptian study of 204 adults performed by El-Mehallawi and Soliman (2001).

Sexual dimorphism was found to characterize several landmarks during a study that investigated Pakistani sub-adults aged 9-18 years (Jeelani *et al*., 2017). However, these sexual dimorphic differences were only noticeable in children undergoing pubertal changes (ages 13-18 years) (Jeelani *et al*., 2017). The results also indicated a two millimetre difference with regards to age differences between their sample and

studies performed in Japan (Utsuno *et al*., 2010) and a white Canadian population (Garlie & Saunders, 1999). The two millimetre differences were, however, likely due to the natural variation within the samples of the aforementioned studies and not indicative of differences amongst the various population groups. A review conducted by Stephan & Simpson (2008b) investigated the effects of age in relation to FSTT from five sub-adult studies. The sample consisted of 830 individuals in total, but there was an uneven distribution of measurements per landmark. The review found that tissue thickness increased with age, but that studies should consider only categorizing sub-adults in two cohorts (Stephan & Simpson, 2008b). A threshold of 11 years of age was selected as the FSTT means of those sub-adults between 12 and 17 years differed by less than one millimetre compared to FSTTs of adults (Stephan & Simpson, 2008a). However, Stephan & Simpson (2008a) highlighted that there was poor repeatability amongst FSTT data for sub-adult populations (in excess of two millimetres). These issues can only be addressed in future studies if the methodologies utilized for both adult and sub-adults FSTT studies are standardized in terms of sample sizes and measuring equipment (Stephan & Simpson, 2008b).

Research pertaining to bilateral asymmetry reported that 50% of lateral landmarks were statistically different, but the absolute differences were all less than one millimetre (De Greef *et al*., 2006). These authors stated that these differences would be meaningless when constructing a facial approximation and that the significant differences observed were definitely due to their excessively large sample size of almost one thousand (De Greef *et al*., 2006). A more recent study investigating 320 German adults (Thiemann *et al*., 2017) concurred with the results of De Greef *et al*. (2006). Statistically significant differences were also found at multiple bilateral landmarks, however, the differences between right and left sided FSTT means only

ranged between 0.5 to 1.8 millimetres. These small millimetre differences were not considered large enough to alter the recognition of a facial approximation (Thiemann *et al*., 2017).

Stephan *et al*. (2015b) strongly disagreed with the categorization of FSTT and subsequent reporting of statistically significant differences at landmarks among categories of persons. Their study researched the impact of data noise and how researchers justified this noise as meaningful deductions by relying on p-values. The authors (Stephan *et al*., 2015b) stated that observed differences are most often a result of noise and not indicative of biological differences. Furthermore, the primary reasons for the occurrences of significant differences are problems associated with sample sizes, the statistical analyses and sampling errors (Stephan *et al*., 2015b). Small sample sizes were noted to be highly sensitive to data noise, whereas very large samples delivered multiple significant differences (Stephan *et al*., 2015b). Non-randomised sampling results in errors with regard to the reliability and accuracy of the results of a study (Stephan *et al*., 2015b). The authors concluded that minute differences in FSTTs, regardless of their significance, should not be given any emphasis as it is most probably due to one or more of the aforementioned reasons (Stephan *et al*., 2015b). As such, data pooling is stated to be an effective approach to alleviating the misleading effects of small sample size and measuring errors (Stephan *et al*., 2016). More so, statistical errors associated with large samples are managed with the incorporation of the values like the shorth and 75-shormax, statistics of central tendency (Stephan *et al*., 2016).

2.8. Facial soft tissue thickness in South African populations

Previous literature suggested the need for population specific FSTT data (Bulut *et al*., 2014). This idea possibly exists due to the genetic variability within and amongst populations of different regions of the world (Bulut *et al*., 2014). Wilkinson (2002) indicated that FSTT measurements of European white populations are applicable estimations for the white South African population. However, FSTTs obtained from black Americans cannot be utilized for a black South African population (Aulsebrook *et al*. 1996). The reason for the aforementioned conclusion is that the South African black population has relatively little genetic admixture, unlike the American black population (Aulsebrook *et al*., 1996). As such, it is necessary to assess whether there is a need for the categorization of FSTTs according to population groups in a South African context.

Stephan & Simpson (2008a) disagreed with the practice of FSTT categorization and stated that the miniscule differences among FSTTs of several populations have little applied importance in terms of facial approximation and recognition. Furthermore, these differences are identifiably due to error and data uncertainty and should be alluded to with the use of multivariate analysis (Stephan & Simpson, 2008a). At present, studies performed by Aulsebrook *et al*. (1996) and Cavanagh & Steyn (2011) are the sole sources of information pertaining to FSTTs in black South African adults. Aulsebrook *et al*. (1996) investigated the FSST of adult Zulu males and Cavanagh & Steyn (2011) examined the FSTT of black females. Cavanagh & Steyn (2011) and Philips & Smuts (1996) employed computed tomography scans to collect the FSTT measurements. In contrast, Briers *et al*. (2015) and Aulsebrook *et al*. (1996) collected their data utilizing cephalometric radiography.

The research performed by Briers *et al*. (2015) and Phillips & Smuts (1996) were on Coloured children and adults respectively. Therefore, due to the small number of studies currently available in South Africa, data pooling must be considered to increase the sample size and hence, the validity of the data. The pooling of data has been advocated by Stephan *et al*. (2016), who highlighted that it is the safest route in avoiding interpreting data noise due to ancestry.

2.9. Facial soft tissue thickness outside of Africa

Studies pertaining to FSTT have been performed around the world and variation in tissue depths have been especially noted amongst populations that are geographically separated. Therefore, it is necessary to conduct more research as not one population specific data set can be utilized in a different population (Lodha *et al*., 2016). This need is convoluted when populations cannot be clearly defined in terms of ethnicity, like those found in Brazil (Tedeschi-Oliveira *et al*., 2009); or in South Africa, due to the large degree of genetic mixture propagated by generations of international immigration (Philips & Smuts, 1996). With this in mind, it is necessary to ascertain how practical and feasible the current FSTT datasets are and what influence recommendations like data pooling have on future research.

2.9.1. European studies

Numerous literature pertaining FSTT extend across Europe. Studies extending from the Mediterranean to Eastern Europe provide the necessary insight into the effects of geographical dispersion. Panenkova *et al*. (2012) was the first to investigate the FSTT of a Central European population. Their study consisted of CT scans of 160 Slovak adults with an equal distribution of males and females comprised the sample and 14 anatomical landmarks were used (six midline and eight bilateral). The landmarks used

by Panenkova *et al*. (2012) did not include any from the mandible, but mid-face landmarks used were stated to be the most important in facial recognition. The results of Panenkova *et al*. (2012) were centred on determining difference amongst sexes and several age groups. Males were reported to have higher FSTT than females at landmarks around the mouth, but not on the upper face. In terms of age categories, females displayed significant difference at double the number of landmarks compared to the males of the sample (Panenkova *et al*., 2012). These trends amongst sex and age were argued to be negligible as the FSTTs differed by no more than one millimetre (Panenkova *et al*., 2012). A similar methodology was incorporated by Drgáčová *et al*. (2016) in a modern day Czech population. The FSTT at forty landmarks were examined from CT scans of a 102 patients. Again, age and sex formed part of the study as well as bilateral asymmetry. The FSTTs were reported to be larger in males than females. However, females displayed a considerable amount of age significant differences, almost double compared to the males. The researchers subjected the results to discriminant function analysis and confirmed the validity of their correlations. Drgáčová *et al.* (2016) then drew comparisons between their and other current cross-continental studies which revealed that their FSTT results were larger and possibly attributed to great prevalence of obesity within the Czech population. Pertaining to bilateral asymmetry, the FSTTs recorded on the right side was larger than on the left. These asymmetries were all located at landmarks associated with the teeth (Drgáčová *et al*., 2016). No definitive reasons were provided for these asymmetries reported by both Drgáčová *et al*. (2016) and Panenkova *et al*. (2012).

Unlike the previous to studies, a Portuguese study by Codinha (2009) collected FSTT from cadavers using the needle puncture method. The method had however, been

slightly modified to exclude the conventional rubber stopper as it was believed to alter the measurement on removal of the rubber stop from the needle. Instead a forceps was used to handle the needle. The results of this study calculated minimal technical error of measurement and not exceeding 0.5 mm but the author stipulated that this was most probably a result of taking the second round of measurements within two hours of the first round and not forgetting that the author was still able to identify the first puncture site (Codinha, 2009). Their results demonstrated that males generally had thicker facial soft tissue than female, particularly around the mouth and chin region. Females, however, had larger FSTTs on and in the vicinity of the cheeks. The study determined the Portuguese sample showed significantly larger FSTTs (up to 5.0 mm) compared to a Spanish study performed by Valencia in 2007 (Codinha, 2009). The study concluded that the before-mentioned illustrated that despite the geographical proximity between different populations, FSTT datasets should not be considered interchangeable between countries (Codinha, 2009).

Other studies include those performed in Turkish (Bulut *et al*., 2014) and German (Thiemann *et al*., 2017) samples. The two studies had equal large sample sizes of 320 individuals which were further equally distributed between males and females. Both studies collected their data from retrospective CT scan, though Thiemann *et al.* (2017) used multi-slice computed tomography (MSCT). The German study reported numerous significant differences with respect to sex, BMI and asymmetry. Similarly, Bulut *et al*. (2014) indicated that sex and age were highly influential to the FSTT in the Turkish population. More interesting though is that when this sample was compared to a Korean and Belgian sample, the results showed that the Turkish sample had larger FSTTs than white Europeans (Belgians), but smaller FSTTs compared to Koreans. The differences were reported to have been significant (Bulut *et al*., 2014).

This sentiment was not shared by Thiemann *et al*. (2017) who concluded that their inter-population analysis indicated no need for population specific data. This finding by Thiemann *et al*. (2017) is in agreement with Stephan *et al*. (2015b).

2.9.2. Asian studies

A study performed in Pakistan involved measuring FSTTs at 11 midline landmarks using lateral cephalograms of 231 children for the purposes of exploring the effects of adolescent growth (Jeelani *et al*., 2017). Sex differences were calculated to be insignificant using Bonferonni corrections for multiple comparisons ($p < 0.005$). Analysis of age differences found significant differences at the nasal landmarks. Compared to other international sub-adult FSTT studies, differences in excess of two millimetres were identified at several landmarks in the vicinity of the mouth (Jeelani *et al*., 2017). The authors reasoned that these are primarily due to the variable morphology of maxilo-mandibular region which is continually changing as a result of growth and development (Jeelani *et al*., 2017).

FSTTs studies conducted for the purposes of ascertaining the effects of sexual dimorphism were performed utilizing CT scans in Indian Gujarati (n = 489) (Lodha *et al*., 2016) and Taiwanese (n = 193) (Chung *et al*., 2015) populations. Chung *et al*. (2015) found that males possessed a thicker FSTT on average compared to its female subjects and attributed this difference to increased collagen production found in females compared to males. Age differences indicated decreased FSTT around the forehead, nose and mouth with increased age. This is consistent the results reported by De Greef *et al*. (2009) despite statements made in Wilkinson (2004) that age-related FSTT changes are highly variable. Similar trends in sex were observed by Lodha *et al*. (2016). However, many more of these differences were calculated to be

significant ($p < 0.05$). Age and BMI differences were significant at several midline (e.g. glabella, nasion, rhinion, pogonion, menton) and lateral (e.g. supra M_2, sub M_2, supraorbital, infraorbital, gonion) landmarks (Lodha *et al*., 2016).

Similarly, CBCT scan were utilized to test for sexual dimorphism with respect to FSTTs in Korean adults (n = 100) (Hwang *et al*., 2012) and Iranian adults (n = 250) (Masoume *et al*., 2015). CBCT has the advantages of it exposing patients to lower doses of radiation in conjunction to the scans being taken in the upright position compared to the supine position seen in multi-slice CT which lessens the effects of gravity (Hwang *et al*., 2012). However, in order to reduce the effects of noise and pixilation, Hwang *et al*. (2012) commented that slice thickness and possibly the radiation dosage would need to be increased. In the Korean sample, males had larger FSTT compared to females at 60 and 50 % of midline and bilateral landmarks respectively (Hwang *et al*., 2012). The zygoma was the only landmark where females were observed to have a statistically significant larger FSTT than males and no significant differences were detected in terms of bilateral asymmetry on analysis of the entire sample (Hwang *et al*., 2012). In the study of Masoume *et al*. (2014), BMI in their Iranian sample was witnessed to be directly proportional to FSTT, yielding significant differences at all the investigated landmarks except the inferior malar.

2.9.3. American studies

Facial soft tissue thickness studies have be conducted in several populations and countries across North and South America. These include Brazil (Tedeschi-Oliveira *et al*., 2009; de Almeida *et al*., 2013), Columbia (Ruiz, 2013), Canada (Huculak & Peckman, 2012; Peckman *et al*., 2015) and the United States of America (U.S.A.) (Parks *et al*., 2014) to name a few. Needle-puncture, cone beam computed

tomography (CBCT) and ultrasound are some of the methodologies used to collect FSTTs from cadaveric and *in vivo* individuals. Of all the aforementioned methods, recent studies have suggested that ultrasound is the most preferable as it is the most cost-effective and safest in terms of radiation poisoning (Stephan & Prieler, 2018).

Huculak & Peckman (2012) and Peckman *et al*. (2015) researched the FSTT of African Nova Scotian children (African Canadian population) and Mi'kmaq adults (people indigenous to Canada's Atlantic Provinces). In both studies FSTT measurements were taken from the right handed side of the face using ultrasound whilst participants sat in an upright position. Typically, FSTT measurements are obtained from the left side of the face. However, neither Huculak & Peckman (2012) nor Peckman *et al*. (2015) provided their reasons for collecting the FSTTs from the right side of the face. Peckman *et al*. (2015) reported that older adults had larger FSTTs than their younger counterparts. The presence of this result was stated to exist due to ageing process like wrinkling, but furthermore, may be due to genetics, level of hydration and even pregnancy (Peckman *et al*., 2015). Huculak & Peckman (2012) had a sample of sub-adults ranging in age from three to 18 years of age and reported that individuals aged nine to 13 years demonstrated the highest variation in FSTT which was possible due to the effects of puberty. Furthermore, their study determined significant differences to exist between sub-adult males and females in different age categories. It was thus suggested that sex-specific FSTT data be separated with respect to the approximate onset of puberty due to commencement of sexual dimorphic changes (Huculak & Peckman, 2012). Regarding sexual dimorphism, Peckman *et al.* (2015) acknowledged the statements shared by Stephan & Simpson (2008a) regarding the small and impractical difference between male and female FSTT data, but stated that there were very evident differences in their study especially

in the vicinity of the cheeks and jaw. Furthermore, Peckman *et al*. (2015) indicated that no attention was paid to the BMI of their volunteers, but in a study by Stephan *et al*. (2016) it was stated that females generally have a larger FSTT than males when BMI is standardized. The results of Stephan *et al*. (2016) pertaining to effects of BMI when evaluating sexual dimorphism were similar to those of earlier studies conducted in South American populations.

The Brazilian population is highly heterogeneous in nature and molecular markers have indicated that its population is a combination of European and African ancestry (de Almeida *et al*., 2013). With the addition of Asian ancestry, the Columbian population is genetically similar to the Brazilian population (Ruiz, 2013). The study of de Almeida *et al*. (2013) was conducted using the needle puncture method on 100 cadavers, whereas the FSTT of 30 living persons were collected using CBCT in Ruiz (2013). Several significant differences were reported in the study of Ruiz (2013) with respect to sex (when BMI was standardized) and varying BMI categories especially around the masseteric region. When this dataset was compared to European studies it was reported that large differences at certain landmarks were noted despite the Spanish genetic influence in the Columbian population. The results of Ruiz (2013) were more in line with a Portuguese population (Codinha, 2009) although the method used differed (i.e. CBCT *v*. needle puncture). A previous Columbian study performed using radiographs conducted by Bermudez & Mora in 2000, indicated much larger FSTTs (differences ranging from 0.0 to 8.6 mm) compared to the sample of Ruiz (2013). These differences were stated to be due to the difference in the methodologies used to collect the FSTT (Ruiz, 2013). CBCT scans, used by Ruiz (2013), have better resolution and are more reliable than radiographic imaging. Although a difference in methodology may contribute towards the differences in FSTTs between several

studies, the state of an individual's health has also been reported to contribute towards the variation of FSTT measurements (Tedeschi-Oliveira *et al*., 2009).

The nutritional status of an individual had been reported to be influence the FSTT of the buccal, infraorbital and zygoma regions of the face (Tedeschi-Oliveira *et al*., 2009). These researchers suggested that consideration should be given to producing facial approximations with varying nutritional states. The FSTTs of American whites and American blacks were found to be considerably different to Brazilian white and Brazilian black populations respectively (Tedeschi-Oliveira *et al*., 2009). This finding advocates the need for population-specific data as cranial morphology is diverse in terms of geographical location, despite there being no evidence for distinct racial groupings among humans (Tedeschi-Oliveria *et al*., 2009). The findings of Tedeschi-Oliveira *et al*. (2009) however, negate the effects of the dissimilarity in diet between different regions of the world; a factor possibly more responsible for FSTT variation than cranial morphology.

FSTTs collected from a modern *in vivo* American population included 388 CT scans, ranging in age from 18 to 62 years belonging to the Federal Bureau of Investigation (FBI) (Parks *et al*., 2014). One of the focuses of this study was to compare FSTTs of their study to a previous American study conducted by Rhine & Moore in 1984. The results of the latter are still currently utilized by the FBI, although being 30 years old (Parks *et al*., 2014). Their study determined that the FSTTs differed most in the labial-buccal region when compared to the results of Rhine & Moore of 1984. The outcome, however, was attributed to the doubling in the rate of obesity since the year 1980 (Parks *et al*., 2014). The incidence of obesity had increased by 30% between the time of the modern American sample and the sample of obese individuals belonging to Rhine & Moore (Parks *et al*., 2014). Parks *et al.* (2014) continued by highlighting that

the significance of these differences could possibly have been perpetuated do large difference in sample size. This trend of large sample sizes were repeated on analysis of FSTTs collected in Parks *et al*. (2014) and those in Stephan & Simpson (2008a), with sample sizes totalling 388 and more than 5700 respectively. Stephan & Simpson (2008a) explained that significances are under and over-estimated when samples are too small or too large respectively.

2.9.4. Australian studies

Most FSTT studies performed in Australian have been conducted using cadavers and recently one in living individuals. The first study was conducted by Sutton in 1969, and the sample consisted of 104, but measurements were only collected at the lateral landmark, the zygion (Stephan & Preisler, 2018). The sample sizes of other cadaveric studies ranged from nine to 150 and FSTTs were measured at least 20 landmarks using the needle puncture method. All the cadavers were classified as white and of European ancestry, ranging in age from 13 to 101 years. The last published cadaveric study was conducted by Domaracki & Stephan (2006). Their study consisted of 33 cadavers and measurements were collected at seven midline and six bilateral landmarks. Instead of the conventional needle-puncture method, the authors selected the sooted-pin method which unlike the former does not require any alteration of the measuring tool. It was also reported that this method resulted in higher intra- and inter-observer accuracy due to the persistence of soot on the surface of the skin (Domaracki & Stephan, 2006). Collecting FSTTs from the dead could result in up to 30% measurement errors due to post-mortem changes, the positioning of the body (supine *v.* upright) and/or the incorrect identification of landmarks (Stephan & Preisler, 2018). These measurements errors can, however, be negated with the use of modern imaging equipment such as the ultrasound (Stephan *et al*., 2016).

Currently, there have been three in *vivo* Australian studies all using B-mode ultrasound to collect FSTTs (Stephan *et al*., 2016; Stephan & Preisler, 2018; and Stephan & Sievwright, 2018). The aims of these three studies included determining whether sexual dimorphism influenced FSTT when BMI was normalized (Stephan *et al.*, 2016), what were the effects of posture and the consequences of using a particular study sample (living *v*. dead) (Stephan & Prielser, 2018), and if craniometrics and linear regression models was more useful than FSTT arithmetic means (Stephan & Sievwright, 2018).

Stephan *et al*. (2016) measured FSTTs at 14 landmarks in 52 living adults for the purpose of discrediting the use of data categorization according to sex. The authors stated that BMI is a determining factor in FSTT difference which should be normalized before comments pertaining to sexual dimorphism could be made conclusively. The results of their study determined that females have larger FSTTs compared to males (Stephan *et al*., 2016). It was recommended that despite the influence of BMI, it should not be sub-categorized as body mass cannot be determined from skeletal remains. Instead, data pooling should always be preferred. By implementing a central tendency statistic such as the 75-shormax, facial approximations could be produced for larger individuals (Stephan *et al*., 2016).

A study investigating FSTTs in *in vivo* Australians and consisting of 64 individuals was conducted using ultrasound in both the supine and upright position (Stephan & Preisler, 2018). Their results indicated that very few differences were recognized between measurements taken in the supine and upright regardless of sex or age (Stephan & Preisler, 2018). When the pooled results of Stephan & Preisler (2018) were compared against the pooled data of cadaveric Australian studies it was determined that the largest difference (4.1 mm) was between the measurements taken

at the gonion. However, Stephan & Preisler (2018) explained that the differences must be considered negligible due to sampling biases and the effects of data noise caused by inconsistencies in the measuring process. The study by Stephan & Preisler (2018) highlights that no differentiation can be made between FSTTs collected from the living and dead because it would be premature to associated significance with the recorded differences. Factors such as sampling errors (with respect to both the selected landmarks and the individuals comprising the sample) and measurement errors were acknowledged as being responsible for FSTT differences amongst studies (Stephan & Preisler, 2018). Lastly, the authors expressed that in order to collect more FSTTs of both adults and sub-adults, future research should incorporate ultrasound techniques. The recommendation was made as ultrasound does not expose participants to radiation and facilitates more random sampling.

A more recent study performed by Stephan & Sievwright (2018) included 71 white Australian adults and followed the methodological guidelines provided by Simpson & Henneberg (2001). These guidelines dictated that large enough sample sizes and using ultrasound in living persons yield the most reliable FSTT data (Stephan & Sievwright, 2018). The purpose of their study was to determine the degree of variation in FSTTs across all previous Australian studies (cadaveric and *in vivo*) compared to theirs (Stephan & Sievwright, 2018). The FSTT means of Stephan & Sievwright (2018) were found to fall within the ranges (0.0 to 1.1 mm) of a previous ultrasound study conducted by Stephan & Preisler (2018). The results of Stephan & Sievwright (2018) were also determined to be similar to that of cadaveric Australian studies with differences of 0.1 to 3.6 mm, except at the gonion which differed by 6.6 mm. The possible justifications for these differences included the discrepancy in the number of studies wherein ultrasound was used instead of the needle-puncture method in

cadavers (Stephan & Seivwright, 2018). The cadaveric studies was also said to be susceptible to FSTT changes due to a combination of post-mortem dehydration and by the embalming process (Stephan & Sievwright, 2018). Lastly, the sample sizes of Stephan & Prieler (2018) and Stephan & Sievwright (2018) were two to three times smaller compared to the average cadaveric sample which is a concern as it is less representative of the population (Stephan & Sievwright, 2018).

2.10. Facial recognition

The human face and associated facial expressions provides signals necessary for interaction (Wallis *et al*., 2012) and the ability to recognise and remember faces is crucial in order to function optimally in society (Barzut & Zdravkovic, 2013). Wilkinson (2004) stated that the face is the primary indicator of a person's personality.

In a forensic setting, understanding the technicalities of facial recognition is highly imperative as eyewitness identification can be exploited in the judicial system (Shriver *et al*., 2008). An incorrect identification could result in the unjust incarceration of an individual or the assigning of an incorrect identity to skeletal remains (Towler *et al.*, 2017). Therefore, the accuracy and reliability of facial approximations in terms of resemblance is of great importance in order to assess its usefulness for medico-legal purposes (Lee & Wilkinson, 2016). Several studies have found that the ability to recognise and correctly identify a face is dependent on the individual biases of the viewer (Shriver *et al*., 2008). Previous research has also demonstrated that there is a remarkable ability to recognise a face categorically similar to the observer (Rhodes & Anastasi, 2012). This recognition ability was quantitatively demonstrated when there was increased activity in a specialized region

of the brain when observers were presented with photographs of faces which were similar in race (Barzut & Zdrakovic, 2013). Photography and videography have permitted organisations to develop facial matching systems which are now commonly used in judicial proceedings (Towler *et al*., 2017).

There are two theories which attempt to elucidate the biases seen in the field of facial recognition. The first, and more traditional of the two, is Perpetual Expertise Model which hypothesizes that population biases are correlated to the amount of (quality) interaction one has with people of other population group (Lee & Wilkinson, 2016). The second model, known as the Social-Cognitive Model, includes theories not attributed to biases arising from the amount of time the observer spent with people of different population groups. Instead, observers will cognitively categorize people based societal classifications in order to determine whether more attention and energy should be expending (e.g. age, sex, population and attractiveness) (Lee & Wilkinson, 2016). The above-mentioned models are used extensively to understand and describe the results of studies which have investigated own-population, own-sex and own-age biases.

2.10.1. Own-population bias

Described as a deficiency to correctly recognize faces of different populations, own-population biases (OPB) can commonly also result in the observer merely categorising a face and not taking the adequate time to consider the stimulus (Sporer, 2001). Sporer (2001) noted that other-population faces were often misidentified due to lack to attention used to process the face as it was unfamiliar. More so, Meissner *et al*. (2005) explained that there is a greater relationship between the confidence and accuracy of an assessor when the face being observed is of the same population

group. Sangrigoli *et al.* (2005) tested OPB by evaluating facial recognition in a group of Korean born children who were adopted and grew up in France and a group of Korean adults who temporarily lived in France. Their results showed that the former had a greater affinity to recognise Caucasian faces than Korean faces whereas the inverse was true for group of Korean adults who had grown up in Korea. This finding is in line with the teachings of the Contact Hypothesis which stipulates that the more meaningful interactions one has with a certain group of individuals, the higher the probability of successful recognition (Barzut & Zdrakovic, 2013). Other studies suggest that it is not the quantity, but rather the quality of other-population exposure (Lavrakas *et al*., 1976). The discussion between quantity and quality was further disputed when it was debated that the ability to remember and recognise a face is a determinant of inherent attitudes towards people of different ancestral groups (Lavrakas *et al*., 1976). OPB between white and black individuals found that people classified as white exhibited higher OPB (Lavrakas *et al*., 1976). A study conducted by Blais *et al*. (2008) involved tracking the eye movements of a white European and Asian population discovered that different populations analyse different regions of the face when recognising and categorising faces. However, significant visual training coupled with requesting that observers concentrate on specific features (e.g. lips, eyes, nose, etc.) has a substantial positive effect in facial recognition across population groups (Lavrakas *et al*., 1976). Another factor of influence, known as field dependence, is the ability to not rely on environmental or external stimuli to influence their cognitive ability and instead, use their analytic capability to access a face (Nozari & Siamian, 2015). These results display the need to consider theories of both the Perpetual Expertise Model and the Social-Cognitive Model when evaluating OCB in facial recognition studies.

2.10.2. Own-sex bias

Studies performed to test for own-sex bias (OSB) have demonstrated that females have a better capacity to recognise faces than men. Herlitz & Loven (2013) discussed that this phenomena is possible due to fact that females practice eye contact more often than males. Females are said to be better equipped in recognising and discriminating facial expressions due to their attentive demeanour (Herlitz & Loven, 2013). Majority of studies suggests that OSB is predominately found amongst females (Barzut & Zdravkovic, 2013). Wright & Sladden (2003) found the OSB was common amongst both sexes and continued by stating that the current OSB trends are due to the uneven distribution of studies performed amongst the sexes. These differences across studies may be attributed to the Perpetual Expertise Model, but Wright & Sladden (2003) were of the opinion that humans may have evolved in such a way that each sex was more in line with identifying potential threats instead of potential mating partners. These authors also stated that their subjects remembered the faces presented to them approximately 10 times more often when hair was present or shown in the photographs and suggested that the judiciary should take this into account during eye-witness testimonies (Wright & Sladden, 2003).

2.10.3. Own-age bias

Growth generally tends to conclude after adolescence, although changes in skin texture, hair colours, and increased robustness of the ears and nose for examples are cues of advancing age (Rhodes & Anastasi, 2012). Infants and young children can more readily be identified, but not necessarily recognised, due to their stature and disproportionate features (e.g. larger forehead and eyes) (Rhodes & Anastasi, 2012). Unlike most studies related to OSB, perpetual expertise is more influential in own-age

bias (OAB) (Hills & Lewis, 2011). An experiment conducted by Schaich *et al*. (2016) consisting of presenting photographs of younger and older adults for varying durations of time to two different age groups of assessors. The results found that each group of adults more readily identified and recognised the faces of the respective age groups and attributed this finding to the fact that they had more numerous meaningful interactions with persons closer to their age (Schaich *et al*., 2016). Compared to own population or own sex biases, own age bias is not affected as severely to the underlying theories of the social-cognitive model. This incidence of this observation is as such because humans will continually progress throughout different age categories and have experience in each, but will remain within a particular population and sex group throughout their lifetime (Rhodes & Anastasi, 2012).

2.10.4. Recognition testing

A facial array test assesses the frequency a target individual can be positively matched to a facial approximation (Herrera *et al*., 2016). This method of identification has been used extensively in mass disasters such as with migrants whom have died attempting to cross the Mediterranean Sea (Caplova *et al*., 2017). A series of facial array tests were conducted by Caplova *et al.* (2017) wherein art students and forensic professionals were instructed to select a matching photograph of the several presented to them. Their study found that the professionals displayed a greater success rate than the students (80% and 78% respectively), and the nose was recorded as the most useful feature by both categories of participants (Caplova *et al*., 2017). A similar study used four facial approximations that were assessed by people with varying degrees of knowledge of forensic dentistry (Herrera *et al.*, 2016). The results of Herrera *et al.* (2016) indicated that their female assessors did not have significantly greater success rates compared to the male assessors ($p < 0.05$). Furthermore,

experience and knowledge of facial dentistry did not make a significant difference in the success rate of assessors; a finding similar to that made by Caplova *et al*. (2017). The most frequently selected features of influence in Herrera *et al*. (2016) were the nose, face shape and chin.

Chapter 3: Methodology

3.1. Study design

3.1.1. *Facial soft tissue thickness*

The first part of this study was a retrospective, observational and cross-sectional description of facial soft tissue thicknesses (FSTTs) obtained from computed tomography (CT) scans. These FSTT measurements were then analysed in order determine the influence of sex, age and ancestry.

3.1.2. *Facial approximations and arrays*

The second part of this study was prospective and entailed utilizing facial approximations to test for recognition. The recognition tests were conducted using facial arrays. The participants of the facial arrays did so voluntarily and were invited to participate based on the sex, ancestry and level of education in anatomy and/or art.

3.2. Site of study

3.2.1. *Facial soft tissue thickness measurements*

The first part of the study involved the collection of CT scans from the Department of Radiology located at the Wits Donald Gordon Medical Centre (WDGMC). The FSTT measurements were taken from the CT scans and recorded at the School of Anatomical Sciences of the University of the Witwatersrand.

3.2.2. *Facial approximations and arrays*

Facial approximations were conducted by a forensic artist in the School of Anatomical Sciences of the University of the Witwatersrand. The photographs for the facial arrays were taken of patients from WDGMC as well as participants from inside

and outside the premises of the university. Facial arrays tests were conducted inside and outside of the university premises dependent on the location of the participants.

3.3. Study sample

The first part of this study was conducted using CT scans that were originally taken for the purposes of diagnosis and treatment of patients whom were under the care of the WDGMC. Only black adult (≥ 18 years of age) South African males and females were considered. All patients who exhibited any bone pathologies or abnormalities of the face were excluded due to the possibility of FSTT alterations. These criteria were put in place as very little data pertaining to FSTT exists for the black South African population. FSTT measurements were taken from 60 patients (47 females and 13 males) whose ages ranged from 19 to 70 years.

For the second part of this study, the CT scans of two patients (one male and one female) from the WDGMC were used for the facial approximations. These two patients also had their photographs taken for the purposes of positive matches or controls for the facial array tests. A further 18 participants (nine females and nine males) had their photographs taken to serve as negative matches in the facial arrays. The selected 18 participants were similar in age, sex and general appearance to the two positive matches. The people who volunteered to have their photographs taken included persons associated to the University of the Witwatersrand, as well as lay-persons not associated to the university. The facial arrays tests were completed by 80 people (who will further be known as 'assessors'). The assessors were selected based on their ancestry, sex and whether they had formal education in anatomical sciences and/or facial artistry.

3.4. Data collection

3.4.1. *Facial soft tissue thickness*

An experienced radiographer employed at the WDGMC selected the scans of patients who fulfilled the inclusion and exclusion criteria so as to ensure the anonymity of the patients. The names, personal details and medical information associated with each of the scans were not disclosed, preventing any breach of confidentiality. The only information that accompanied each scan was that of the sex and the age of the patient which was necessary to test several objectives of this study.

All the scans used in this study were obtained by a Philips 64 Slice CT Brilliance between October 2011 and April 2017. Digital Imaging and Communications in Medicine (DICOM) software converted the CT scans into three dimensional images that were subsequently uploaded to Osirix Lite Version 9.0 software. The facial soft tissue thicknesses were measured at approximately 90 degrees from the three-dimensional images to one decimal place at 10 midline and 10 bilateral anatomical landmarks (see Figure 3.1 and Table 3.1 for detailed description of landmarks).

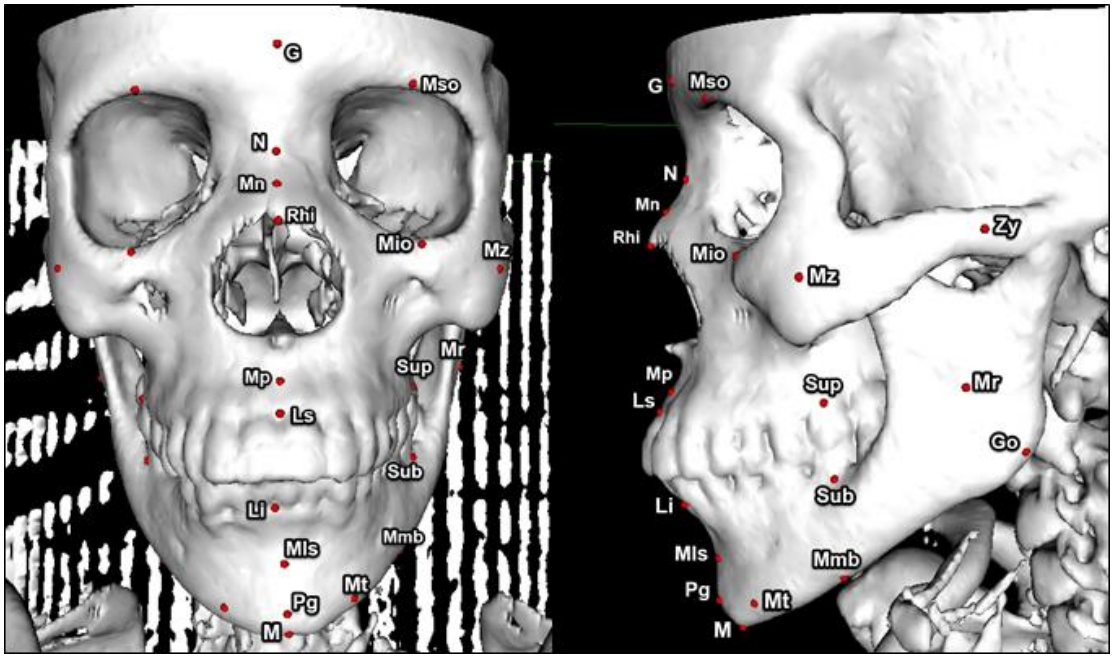


Figure 3.1: Frontal and lateral views illustrating anatomical landmarks that were measured for FSTT. Table 3.1 provides the full names and definitions of the abbreviations.

Table 3.1: Midline and bilateral landmarks, and its respective descriptions [Adapted from Briers *et al*., 2015; Cavanagh & Steyn, 2011; Stephan & Simpson, 2008a; Thiemann *et al*., 2017].

Landmarks	Abbreviation	Description
Midline		
Glabella	G	The most prominent point between the supraorbital ridges in the midsagittal plane
Nasion	N	The midpoint of the suture between the frontal and the two nasal bones
Mid-nasal	Mn	Point on the internasal suture midway between the nasion and rhino
Rhinion	Rhi	Midline point at the inferior free end of the internasal suture
Midphiltrum	Mp	The deepest midline point on the indentation between the nasal spine and the supradentale
Labiale superius	Ls	The apex of the alveolus in the midline between the maxillary central incisors
Labiale inferius	Li	The apex of the alveolus in the midline between the mandibular central incisors
Labiomental sulcus	Lms	Deepest midline point in the groove superior to the mental eminence
Pogonion	Pg	The most anterior point in the midline on the mental protuberance
Menton	M	Most inferior midline point at the mental symphysis of the mandible

Landmarks	Abbreviation	Description
Bilateral		
Mid-supraorbital	Mso	Point on the supraorbital rim at the mid-sagittal plane of the orbit
Mid-infraorbital	Mio	Point on the infraorbital rim at the mid-sagittal plane of the orbit
Mid-zygomatic	Mz	Lined up with the lateral border of the orbit on the centre of the zygomatic process
Zygion	Zy	Point of maximum lateral extent of the lateral surface of the zygomatic arch
Mid-mandibular border	Mmb	Point on the inferior border of the corpse of the mandible midway between the pogonion and gonion
Mid-ramus	Mr	Point at the centre of the mandibular ramus
Gonion	Go	The most lateral point on the mandibular angle
Supra M2	Sup	Above the second maxillary molar
Sub M2	Sub	Below the second maxillary molar
Mental tubercle	Mt	Most prominent point on the lateral bulge of the chin mound

The FSTT located at the midline and bilateral anatomical landmarks were measured from a three-dimensional multi-planar reconstruction (3D MPR) topogram and horizontal view of the CT scans respectively (Figures 3.2.1.B & 3.2.2.B). First, each scan was set to a brightness:contrast ratio of WL: 40/WW: 350 and at one millimetre slice thickness. These specifications were selected on the advice of a radiologist (Dr. M. Haagensen) employed at the WDGMC because it provided the best degree of clarity and accuracy when determining the positions of the landmarks. To ensure the highest accuracy when identifying an anatomical landmark, the '3D surface rendering' option was selected, which presented the bony surface of the CT scan (Figures 3.2.1.A & 3.2.2.A). A region of interest (ROI) was selected and the CT scan would be reverted to a 3D MPR which then indicated the position of the landmark, facilitating the measurement of FSTT (Figures 3.2.1.B-C & 3.2.2.B-C). Measurements at each landmark were measured using an angle tool that outlined the ROI and immediate surrounding bone (Figures 3.2.1.D & 3.2.2.D). Then the angle it made was subtracted from 180 degrees (a straight line) and the difference was divided by a factor of two. On one side of the angle, the angle would be increased by half of the difference calculated in the previous step (Figures 3.2.1.E & 3.2.2.E) and the other side of the angle would then be adjusted until it was at 89.9 – 90 degrees of the landmark (Figures 3.2.1.F & 3.2.2.F). Finally the measuring tool overlapped the angle tool to ensure the FSTT measurement was taken at the correct angle (Figures 3.2.1.G-H & 3.2.2.G-H). With respect to the bilateral landmarks, the measurements were taken on both the left and right sides. However, only the FSTT measurements from the left side were used by the forensic artist to perform the facial approximations. Current literature utilizes only left-sided FSTT data and therefore, this study complied with that norm. The right-sided FSTT measurements were collected for the purposes

of testing bilateral asymmetries and whether the data indicated statistically significant differences. For the purposes of continuum and eventual comparison, the cranial anatomical landmarks that were utilized in this study were adapted from previous South African studies (See Table 3.1).

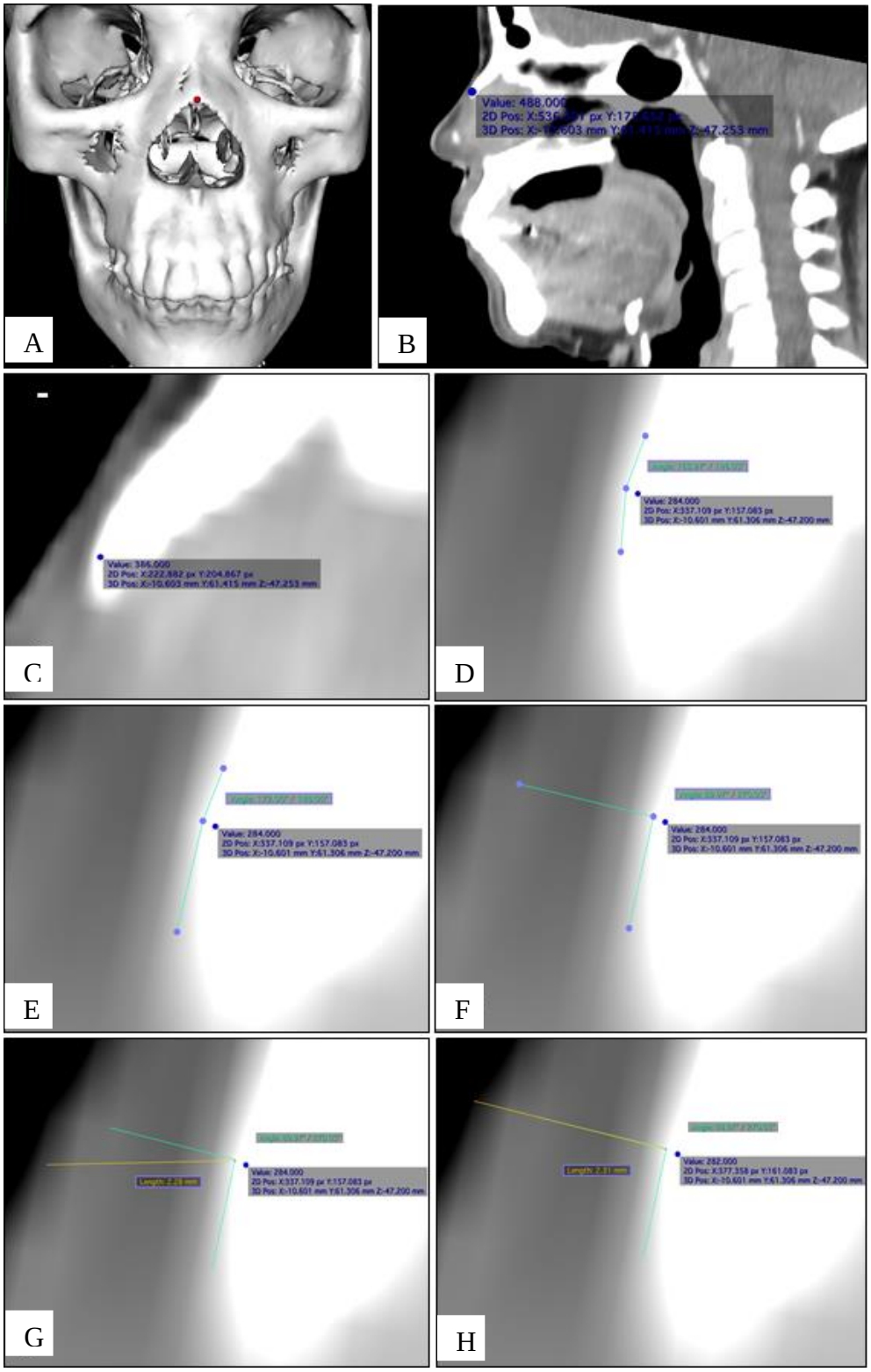


Figure 3.2.1.: Photographs labelled A-H illustrates the method used for collecting the FSTT measurement of the rhinion (Rhi) from a topogram view of a CT scan.

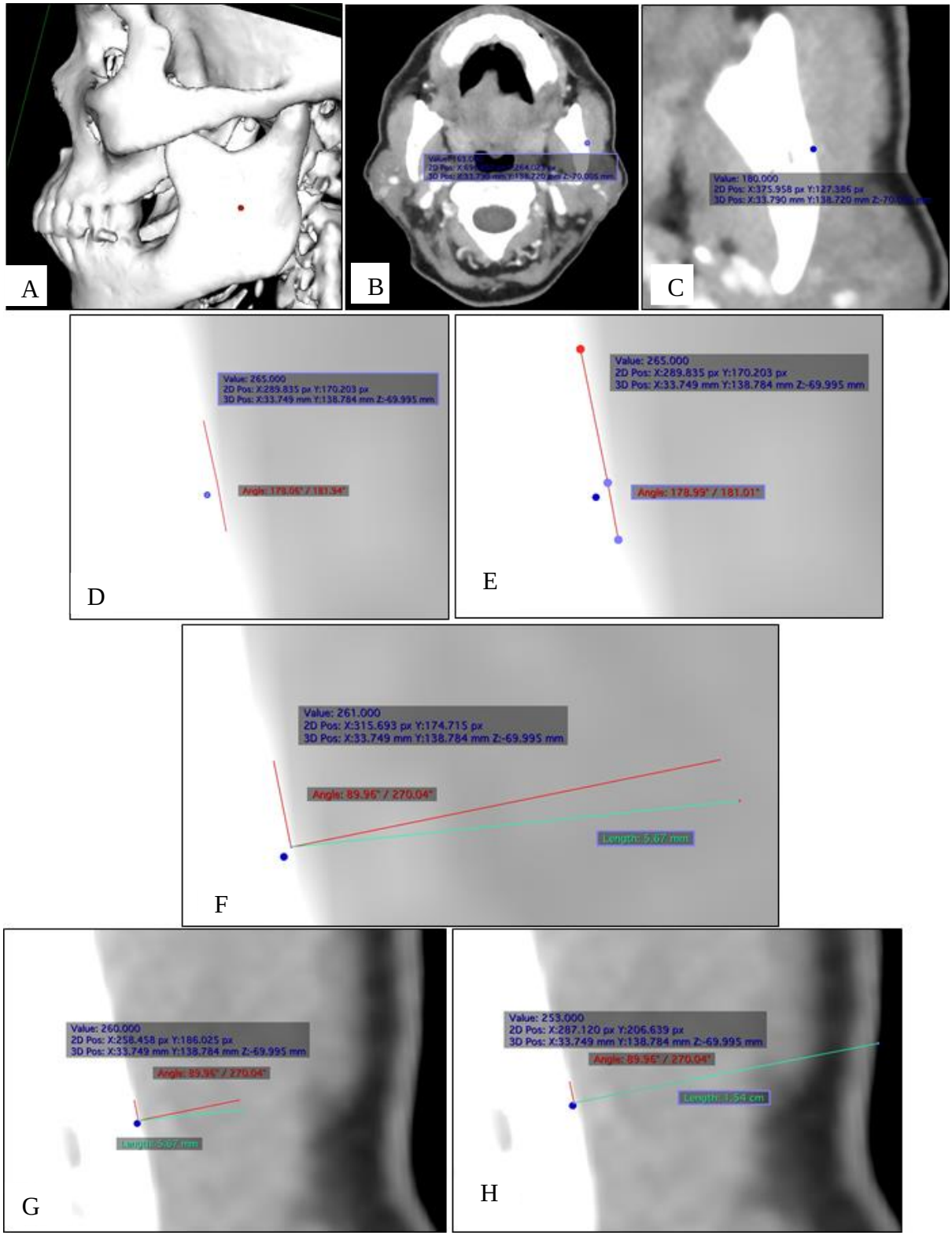


Figure 3.2.2.: Photographs labelled A-H illustrated the method used for collecting the FSTT measurement of the mid-ramus from a horizontal slice view of a CT scan.

3.4.2. *Facial approximations and arrays*

In order to perform the facial approximations required for this study, the CT scans of two patients from the WDGMC were prospectively selected and utilized as templates. The two scans consisted of one male and one female and were selected in accordance to the same inclusion and exclusion criteria that the other 60 CT scans were subjected to. As the participation of these patients was two-fold, they were given information sheets and letters of informed consent allowing the principle investigator to obtain their CT scans (Appendices 5 & 8) from the Department of Radiology of the WDGMC as well as having their photographs taken for the purposes of the facial array test (Appendices 3 & 6).

Facial approximations were conducted by a forensic artist from the School of Anatomical Sciences (University of the Witwatersrand) by extrapolating the FSTT data from part 1 of this study onto the skulls of the two CT scans collected prospectively. The mean FSTT values collected in part 1 of this study were pooled and categorized according to sex in order to produce a total of four facial approximations. Two female approximations were produced using sex-specific and sex-pooled data respectively; and two male approximations were produced using the same method. In addition to the four facial approximations produced using the data of the current study, a further two (one male and one female) facial approximations were produced using the C-Table repository available on http://www.craniofacial identification.com. The C-Table includes adult and sub-adult data reported by age, sex, ancestry, height and body weight. Developed by Professor Carl Stephan (Queesnsland University, Australia), the repository holds data for >1,700 individuals, collected at up to 25 standardized craniofacial landmarks.

The purpose of using sex-specific and pooled data was to test the premise made by Stephan *et al.* (2015) wherein it was stated that the separation of data was unnecessary as the differences in FSTT was negligible. Stephan *et al*. (2015b) explained that these differences in FSTT, despite it being statistically significant at times, did not change the appearance of the approximation enough so as to influence its recognition. The forensic artist was blinded and thus, not shown the photographs of the faces he had reconstructed. Geomagic FreeForm Plus software (from 3D Systems Inc., USA), which utilizes a phantom haptic feedback device was used to construct the facial approximations *via* the Manchester Method. The facial approximations were subsequently textured using an image editing software; Adobe Photoshop CS. After a three-dimensional scan of the skull (A) was uploaded to the software, the eyeballs (B) and tissue depth markers (C) are added to the skull. The musculature (D) and skin (E) of the face was then added. Finally, the approximation was textured and shaded to produce a more lifelike appearance (F). The Manchester method, also known as the combination method, was selected because it is the most reliable method available as it takes into account the morphology of the skull in conjunction to FSTTs. These approximations were subsequently used to conduct facial arrays to test for facial recognition and resemblance between it and the positive matches.

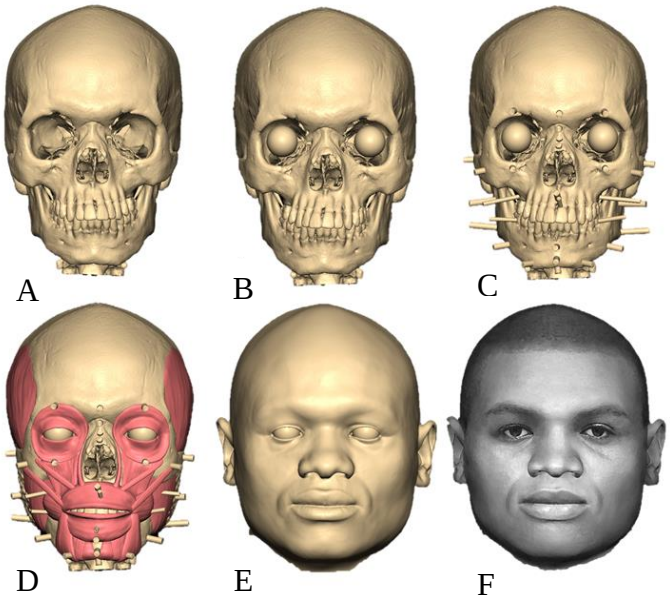


Figure 3.3.A-F: The production process of a male facial approximation using pooled data by employing the Manchester method.

The facial arrays were conducted as follows: the photographs of the male and female patients who were the positive matches were positioned amongst photographs of other males and females (negative matches). In total, 18 negative matches (nine males and nine female) were included in the facial array tests. These candidates for the facial arrays were selected based on their sex, age, ancestry and general appearance compared to the positive matches. This was necessary to avoid easy identification of the positive matches by the assessors. All the participants who had their photographs taken for the facial arrays were issued with an information sheet about the study (appendix 3) and a letter of informed consent (appendix 6) prior to having their photographs taken. The facial arrays were portrayed in Microsoft Power Point presentations which were subsequently printed and presented to assessors together with an information sheet and letter of consent (See appendices 4 & 7 respectively). The facial array tests consisted of frontal, lateral and three-quarter angled photographs. In each of the facial array tests, a facial approximation was positioned in

the centre of the page and 10 photographs consisting of the positive match and nine negative matches were placed around the facial approximation in a circular fashion (Appendix 2). This method of presenting the photographs on the perimeter of the page around the facial approximation was to attempt to equalise the assessors' focus on each photograph. In contrast, photographs in closer proximity to the facial approximation are generally given higher preference when photographs are simply placed in rows beneath a facial approximation. The participants' heads were positioned in the Frankfurt Horizontal Plane whilst maintaining a neutral facial expression. These criteria were necessary in order to maintain uniformity and avoid certain photographs from attracting attention and possibly being selected due to aesthetics. All the photographs were taken at a distance of 1.2 meters using a Nikon D3200 camera.

The assessors of the facial arrays consisted of student and staff volunteers of the University of the Witwatersrand and lay-persons not associated with the university (e.g. family members, friends, random members of the public). The assessors consisted of any person who fulfilled the selection criteria and was interested in participating. The assessors' participation was voluntary and they were required to read information sheets and complete consent forms (see appendices 5 & 8 respectively). A scoring sheet was handed to each assessor wherein they were instructed to indicate their ancestry (e.g. black or white), sex (e.g. male or female) and if they had any formal training in anatomy and/or facial artistry. Assessors were required to identify which of the photographs presented to them most closely resemble the facial approximation and indicate which facial features were the most influential and similar between the facial approximation and photograph of their choice (Appendix 1). In total there were 80 assessors which were separated into eight

groups of ten. The groups were classified as follows: 1) white and 2) black individuals who had formal training in anatomical sciences/facial artistry; 3) white and 4) black individuals who had no formal training in anatomical sciences/facial artistry; 5) males and 6) females who had formal training in anatomical sciences/facial artistry; and, 7) males and 8) females who had no formal training in anatomical sciences/facial artistry. These distinctions enabled the investigator to approximate the effects of same-ancestral and same-sex biases in facial recognition in a general population.

3.5. Data analysis

The FSTT data were recorded in a Microsoft Excel spread sheet and were categorized according to sex and age. The age threshold was that of persons below the age of 50 years and those above the age of 50 as the sample was approximately equally distributed. Descriptive statistics such as: means and standard deviations were recorded for the measurements of each anatomical landmark. Student t-tests were conducted to test for significant difference at 95% confidence intervals between sexes, age groups and bilateral measurements. Tissue depth statistics (TDStats) were used to calculate the shorth and 75-shormax.

In order to test for intra- and inter-observer error the measurements of 19 scans were repeated post the initial data collection by the principle investigator and by an independent observer respectively. The independent observer was deemed capable of performing the necessary duties as they had undergraduate knowledge of anatomy and physical anthropology. The selection of the 19 scans in question was not randomised, but chosen as these had a complete set of measurements. In contrast, the other scans that had not been selected to test for inter- and intra-observer error had one or several missing measurements missing due to issues that will be addressed later in this study.

The intra-observer and inter-observer repeatability was calculated as the intra-cluster correlation and inter-rater agreement respectively. A value of 1 was used as the upper limit and indicative of high repeatability. Furthermore, the technical error or measurements and variable average values were determined. The former was expressed as both an absolute value and a percentage using the equations adapted from Perini *et al.* (2005).

The results obtained from the facial arrays were analysed using Student's t-tests to test for statistical significance ($p < 0.05$) for correlations between the recognition ability for same-population and cross-population effects; as well as, own-sex and cross-sex effects. As there were 10 photographs per facial array there was a 10% chance that any photograph could have been selected. The results for the facial arrays were assessed against a 10% chance factor. Dashed lines were placed on each graph pertaining to the results of the facial arrays; so as to indicate the frequency above chance each photograph was selected. Furthermore, the facial features documented by the assessors as being influential when making their decisions during the facial arrays were subjected to descriptive statistically analyses.

3.6. Ethics

Ethical approval to conduct this study was granted by the Human Research Ethics Committee of the University of the Witwatersrand (Clearance number M170331) (Appendix 9). Permission to access patient CT scans was issued by Dr. Mark Haagensen of the Wits Donald Gordon Medical Centre (WDGMC). The ethical considerations included the protection of patient confidentiality.

The radiographer (employed at WDGMC) who selected the scans used in this study, de-identified all the scans and did not disclose any personal patient information to the

principle investigator. The patients whose CT scans were collected prospectively, those who volunteered to have their photographs taken and those who participated as assessors for the facial array were not asked to disclose any personal information.

Furthermore, letters of informed consent as well as information sheets which outlined the details of the study and the purpose of their participation were given to all the volunteers (Appendices 3 - 8). In addition to these measures, the information of the all the participants were uploaded to a password protected file to ensure the maximum level of anonymity and confidentiality possible.

All CT scans, photographs and facial array scoring sheets will be kept for a minimum of five years post publication of this research.

Chapter 4: Results

4.1. Introduction

In this study, both metric and morphological analyses were performed. The metric analysis entailed the facial soft tissue thicknesses (FSTTs) that were measured from the retrospectively collected CT scans. The morphological component of this study included the utilization of the FSTT to construct several facial approximations which were tested for its recognisability *via* facial array tests.

4.2. Sample demographic

The sample consisted of 60 patients of whom there were 47 females (78.3%) and 13 males (21.7%). Ideally, an equal distribution of males and females would have been preferred for the purposes of statistical analysis. The mean age of the sample was 49.1 with a standard deviation (SD) of 11.1 years; whereas, in the female and male samples the mean age was 47 (SD = 11.5) years and 49.6 (SD = 10.3) years respectively. The sample was also categorized according to age and the majority of patients were in the age groups of 40-49 and 50-59 constituting 25% and 38.3% of the sample respectively. Furthermore, the persons with an age less than 50 years and above the age 50 years consisted 48.3% and 51.7% of the sample respectively. Box and whisker diagrams were constructed for each measurement (Figures 4.1 - 4.13). Figure 4.1 displays the distribution of FSTT at the midline landmarks and Figures 4.2 – 4.13 illustrate the same at the bilateral landmarks. Descriptive statistics were provided for the entire sample (Table 4.1), the female sample (Table 4.2), the male sample (Table 4.3), persons aged 19 – 29 years (Table 4.4), 30 – 39 years (Table 4.5), 40 – 49 years (Table 4.6), 50 – 59 years (Table 4.7) and 60 or older (Table 4.8).

Lastly, the tissue depth statistics (TDStats), including the shorths and 75-shormaxes ,were tabulated against the means for each landmark in Table 4.9.

4.3. FSTT measurements of landmarks

A – *Glabella*

Identified as the most anterior prominence on the forehead and lying between the supraorbital ridges, this measurement was recorded for 51 (85.0%) individuals. The mean value was 4.4 mm (SD = 1.1; range = 4.9 mm) (Figure 4.1). The tissue depth statistics at the glabella indicated that the shorth was 3.6 mm and the 75-shormax was 5.5 mm (Table 4.9). The FSTT mean at this landmark in females was 4.3mm (SD = 1.1) and for males it was 4.6 mm (SD = 1.3). Age groups 40-49, 50-59 and 60+ had means closest to that of the overall sample ranging from 4.3 – 4.7 mm. However, the age group 60+ reported the largest standard deviation across all five age groups with 2.0 mm. The smallest measurement was 2.6 mm and the largest measurement was 7.5 mm found in the 30-39 years and above 60 years categories respectively. There was no significant difference between the sexes; however there was a significant difference between persons above and below the age of 50 years (t-test, $p < 0.05$).

B – *Nasion*

Located at the junction of the frontal bone and inter-nasal suture, this measurement was recorded for 59 (98.3%) individuals. The mean value was 4.6 mm (SD = 1.4; range = 6.2 mm) (Figure 4.1). The tissue depth statistics at the nasion indicated that the shorth was 4.5 mm and the 75-shormax was 6.4 mm (Table 4.9). The FSTT mean was 4.4 mm (SD = 1.3) and 5.1 mm (SD = 1.6) in females and males respectively. The age group with the most similar mean with 4.2 mm was the 40-49 year olds. The

smallest measurement (2.2 mm) and largest measurements (8.4 mm) in this sample were obtained from the age categories 30-39 years old and above 60 years. There was no significant difference between the sexes; however there was a significant difference between persons above and below the age of 50 years (t-test, $p < 0.05$).

C – *Mid-nasal*

Identified as the midpoint of the inter-nasal suture, this measurement was recorded in 53 (88.3%) individuals. The mean value was 2.6 mm (SD = 1.1; range = 5.4 mm) (Figure 4.1). The tissue depth statistics at the mid-nasal indicated that the shorth was 2.1 mm and the 75-shormax was 3.6 mm (Table 4.9). The FSTT mean at the mid-nasal in females was 2.6 (SD =1.2). The smallest (0.7 mm) and largest (6.1 mm) values were recorded from the age group 50-59 years of age. In males, 12 of 13 measurements were recorded and the mean was 2.5 mm (SD =0.8). The mean values across all age groups ranged between 2.1 – 2.7 mm. The standard deviation was the largest amongst patients aged 60 and above, despite it only being 0.2 millimetres more than the standard deviation of the entire sample. There was no significant difference between the sexes; however there was a significant difference between persons above and below the age of 50 years (t-test, $p < 0.05$).

D – *Rhinion*

Identified as the most anterior point at the distal end of the nasal bone, a total of 59 (98.3%) measurements were taken that yielded a mean value of 2.4 mm (SD = 1.0; range = 4.5 mm) (Figure 4.1). The tissue depth statistics at the rhinion indicated that the shorth was 1.8 mm and the 75-shormax was 3.3 mm (Table 4.9). In 46 (97.8%) females the mean value, standard deviation, and range was identical to that of the entire sample. The minimum mean value of 0.6 mm and the maximum mean value of

5.1 mm were measured in the age categories 40-49 years and 50-59 years respectively. The mean value for the 13 males was 2.5 (SD = 0.8) and across all the age groups the mean values ranged between 2.0 and 2.9 mm. There was no significant difference between the sexes; however there was a significant difference between persons above and below the age of 50 years (t-test, $p < 0.05$).

E – *Midphiltrum*

The midphiltrum was measured in 54 (90%) individuals and the mean value recorded was 9.0 mm (SD = 1.4; range = 6.3 mm) (Figure 4.1). The tissue depth statistics at the mid-philturm indicated that the shorth was 9.0 mm and the 75-shormax was 10.6 mm (Table 4.9). The mean was slightly smaller in females (8.9 mm, SD = 1.4) but larger in males (9.7 mm, SD = 1.4). The minimum (6.2 mm) and maximum (12.5 mm) mean values were recorded in age groups 50 -59 years and 60 year or above respectively. There were no significant differences with respect to sex or age (t-test, $p < 0.05$).

F – *Labiale superius*

This was the least recorded midline measurements with a total of 47 (78.3%) as it was the most difficult to locate due to light scatter caused by metallic dental implants/restorations. The mean value for the entire sample was 9.6 mm (SD = 1.8; range = 6.8 mm) (Figure 4.1) compared to 9.5 mm (SD =1.9) in females and 10.0 mm (SD = 1.1) in males. The minimum, maximum and the range were identical between the sample on a whole and the female sample specifically. In the male sample, the range was 3.4 mm. The tissue depth statistics at the labiale superius indicated that the shorth was 9.7 mm and the 75-shormax was 12.1 mm (Table 4.9). The mean FSTT value tended to be inversely proportional to age, i.e.: the largest mean was 10.9 mm and the smallest mean was 8.4 mm which were recorded from persons aged 19-29

years and 60 years or over respectively. There were no significant differences with respect to sex or age (t-test, $p < 0.05$).

G – *Labiale inferius*

This measurement was recorded in 54 (90%) of all patients, the FSTT mean value at this landmark was 10.2 mm (SD = 2.2; range = 9.9 mm) (Figure 4.1). The mean value in females was slightly smaller (10.1 mm ± 2.4) whilst, larger in the male sample (10.6 mm ± 1.7). The tissue depth statistics at the labiale inferius indicated that the shorth was 9.8 mm and the 75-shormax was 12.6 mm (Table 4.9). Both the minimum (6.0 mm) and maximum (16.1 mm) were obtained from females aged 60 years or above. There was no significant difference between the sexes. There was a significant difference, however, between persons above and below the age of 50 years (t-test, $p < 0.05$).

H – *Labiomental sulcus*

This measurement was recorded in 58 (96.7%) of instances and yielded a mean value of 11.3 mm (SD = 1.9; range = 8.9 mm) (Figure 4.1). Females had a smaller mean of 11.2 mm (SD = 1.9), but males had a larger mean of 11.6 mm (SD = 1.8). The tissue depth statistics at the labiomental sulcus indicated that the shorth was 10.4 mm and the 75-shormax was 13.2 mm (Table 4.9). The sample minimum of 6.5 mm and maximum of 15.4 mm were logged from the age categories 60 years or above and the 40-49 years respectively. There was no significant difference between the sexes; however there was a significant difference between persons above and below the age of 50 years (t-test, $p < 0.05$).

I – *Pogonion*

This measurement was recorded from 58 (96.7%) scans. The mean was calculated as 10.7 mm (SD = 2.4; range = 13.6 mm) (Figure 4.1). In the female sample the mean was 10.6 mm (SD = 2.6) and 10.9 mm (SD = 1.4) in males. The tissue depth statistics at the pogonium indicated that the shorth was 10.2 mm and the 75-shormax was 13.1 mm (Table 4.9). The minimum (2.3 mm) and maximum (15.9 mm) values were gathered from females in the categories 50-59 years and 30-39 years respectively. The range amongst males was only 4.3 mm. There were no significant differences with respect to sex or age (t-test, $p < 0.05$).

J – *Menton*

Measured in 59 (98.3%) individuals, the FSTT mean value at this landmark was 5.0 mm (SD = 2.7: range = 17.9 mm) (Figure 4.1). The tissue depth statistics at the menton indicated that the shorth was 4.6 mm and the 75-shormax was 7.6 mm (Table 4.9). The minimum value of 1.1 mm and maximum value of 19.0 mm were found in the 50-59 years group. The mean and standard deviation of females (5.1 mm ± 2.9) was larger than both the collective sample and the sample of males (4.7 mm ± 1.7). There were no significant differences with respect to sex or age (t-test, $p < 0.05$).

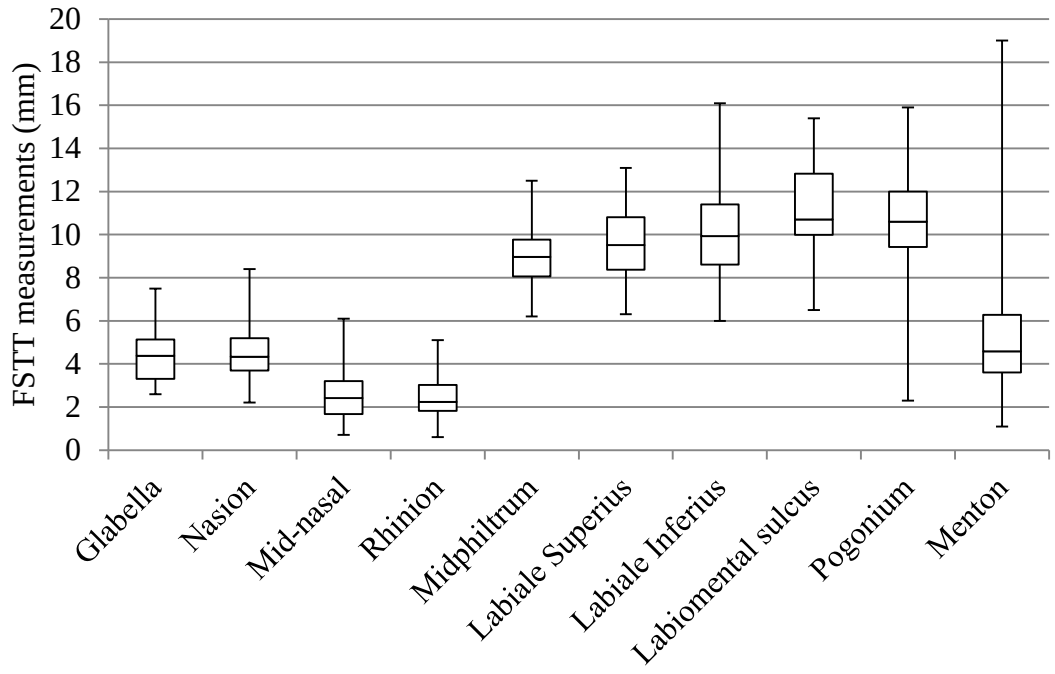


Figure 4.1.: Box and whisker plots of the FSTT measurements for midline landmarks. The horizontal line in the centre of the box represents the median value and the whiskers represent the minimum and maximum values.

K – *Mid-supraorbital*

On the left side, 58 (96.7%) measurements were collected and the mean value was calculated as 7.0 mm (SD =1.7; range = 10.2 mm). The mean values for the female and male samples were 7.1 mm (SD = 1.7) and 6.9 mm (1.8) respectively. The minimum value was 3.5 mm whereas the maximum value was 13.7 mm. These were retrieved from the age groups 50-59 years and 60 years or above respectively. The tissue depth statistics for the left mid-supraorbital indicated that the shorth was 7.3 mm and the 75-shormax was 8.7 mm (Table 4.9). There were no significant differences between sexes but a significant difference did exist at this landmark between persons aged below 50 years and aged above 50 years (t-test, $p < 0.05$).

On the right side, 58 (96.7%) measurements were taken from a possible 60 cases. The data for this measurement are as follows: mean = 7.1 mm; SD = 1.7; and range = 9.1 mm – quite similar when compared to the same measurement on the left side. Regarding the sex-specific samples, the mean value was 7.2 mm (SD = 1.8) for females and 6.6 mm (SD = 1.6) for males. The tissue depth statistics for the right mid-supraorbital indicated that the shorth was 6.7 mm and the 75-shormax was 8.7 mm (Table 4.9). There were no significant differences between sexes and age groups (t-test, $p < 0.05$). Furthermore, bilateral asymmetry was not significant at the mid-supraorbital (t-test, $p < 0.05$).

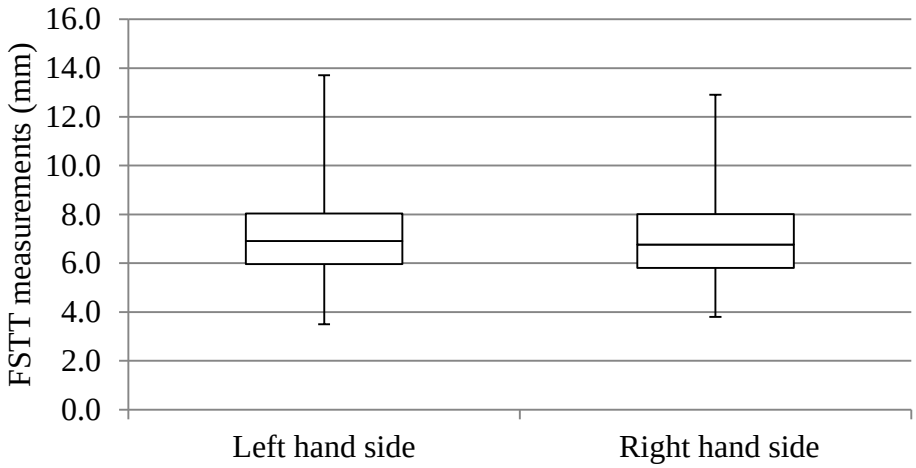


Figure 4.2.: Box and whisker plots of the FSTT measurements at the mid-supraorbital bilaterally. The horizontal line in the centre of the box represents the median value and the whiskers represent the minimum and maximum values.

L – *Mid-infraorbital*

On the left side, this measurement was recorded 59 out of 60 times (98.3%). The mean value at this site for the entire sample was 6.6 mm (SD = 2.5; range = 10.6 mm). The minimum value for this sample was 2.6 mm and the maximum was 13.2 mm and was derived from the 40 - 49-year olds male and the 60 years or older groups

respectively. The mean values for the female and male samples were 6.8 mm (SD = 2.5) and 5.5 mm (SD = 2.4) respectively. Similar to FSTT of the mid-supraorbital, the mean FSTT collected per age category had increase with the progress of age and ranged between 5.1 -7.8 mm. The tissue depth statistics for the left mid-infraorbital indicated that the shorth was 4.6 mm and the 75-shormax was 8.4 mm (Table 4.9). A significant difference was absent in the sexes but present between persons above 50 years and below 50 years.

On the right side, the same number of measurements was taken as on the left side. The collective sample means was also 6.6 mm (SD = 2.4; range = 9.2 mm). The means for the female and male samples were the same as on the left side – 6.8 mm (SD = 2.3) and 5.5 mm (SD = 2.4) respectively. Once again, FSTT increased in response to age as the mean values per age categories ranged from 5.0mm to 7.5 mm. The minimum value was 2.9 mm, found in the age category 40 -49 years. The maximum value was 12.1 mm and was found in the age category 50 – 59 years. The tissue depth statistics for the right mid-infraorbital indicated that the shorth was 5.6 mm and the 75-shormax was 8.9 mm (Table 4.9). Sex differences were insignificant but age yielded a significant difference below 50 years *vs*. above 50 years (t-test, $p < 0.05$). Furthermore, bilateral asymmetry was not significant at the mid-supraorbital (t-test, $p < 0.05$).

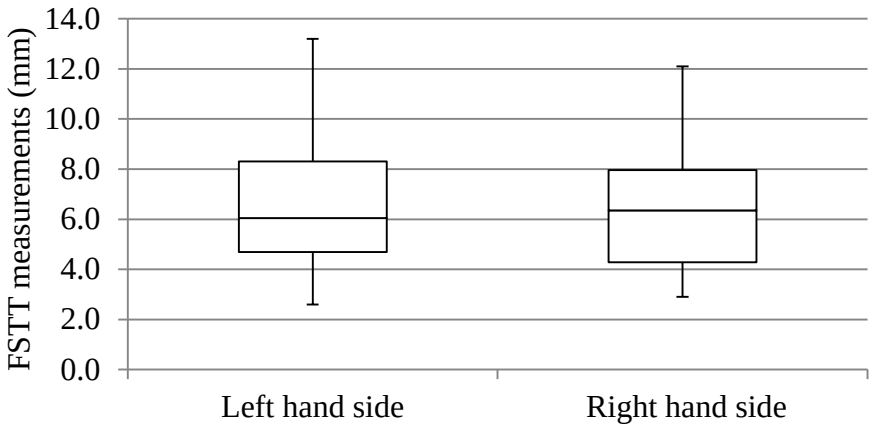


Figure 4.3.: Box and whisker plots of the FSTT measurements at the mid-infraorbital bilaterally. The horizontal line in the centre of the box represents the median value and the whiskers represent the minimum and maximum values.

M – *Mid-zygomatic*

On the left side, measurements were taken from 59 (98.3%) scans and the mean value was calculated to be 9.5 mm (SD = 2.3; range = 10.9 mm). The age group 30 -39 years provided the minimum value of 4.4 mm, and the maximum value of 15.3 mm was collected from the above 60 years category. The FSTT mean specifically for females was 10.0 mm (SD = 2.2) and 7.3 mm in the sample of males. The tissue depth statistics for the left mid-zygomatic indicated that the shorth was 8.7 mm and the 75-shormax was 11.9 mm (Table 4.9). FSTT was significantly different between the sexes, however not for age (t-test, $p < 0.05$).

On the right side, the same number of measurements was taken like that on the left side. The mean value and range had decreased to 9.0 mm (SD = 2.3; range = 10.1 mm). The minimum (4.5 mm) and maximum (14.6 mm) values were recorded from persons aged 50 – 59 years and above 60 years respectively. Females yielded a mean

of 9.4 mm whilst the mean of the males sample was considerably smaller (7.0 mm). The tissue depth statistics for the right mid-zygomatic indicated that the shorth was 9.5 mm and the 75-shormax was 11.5 mm (Table 4.9). FSTT was significantly difference between the sexes, however not for age (t-test, $p < 0.05$). Furthermore, bilateral asymmetry was not significant at the mid-zygomatic (t-test, $p < 0.05$).

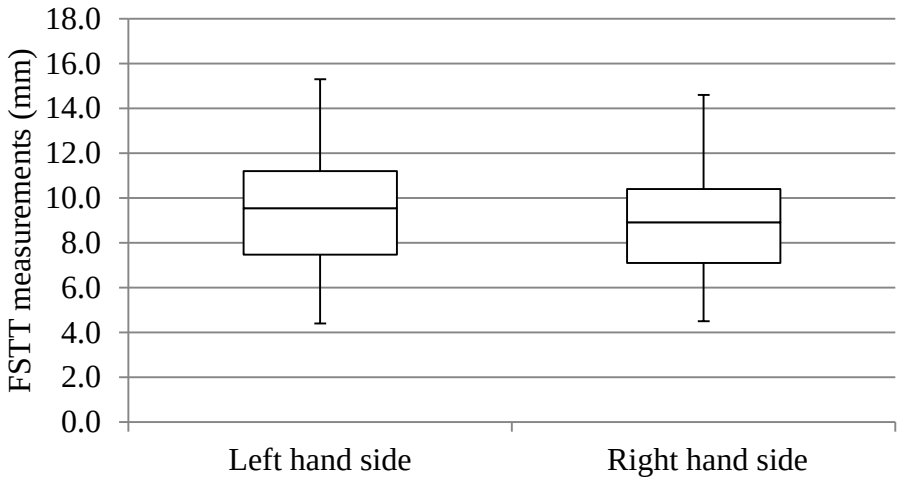


Figure 4.4.: Box and whisker plots of the FSTT measurements of the mid-zygomatic bilaterally. The horizontal line in the centre of the box represents the median value and the whiskers represent the minimum and maximum values.

N – *Zygion*

On the left side, the sample comprised of 59 (98.3%) measurements and the mean value was determined as 8.9 mm (SD = 3.0; range = 15.1 mm). The mean value for females was 9.2 mm ± 2.9 and 7.4 mm ± 3.2 for males. The age group 50-59 presented an identical mean as the whole sample, whereas the least similar mean was calculated as 10.7 mm and was represented in the 60-years and older age group. The minimum value was 3.1 mm and the maximum was 18.2 mm. These were obtained from the 50 – 59 years groups and 60 years or above group respectively. The tissue

depth statistics for the left zygion indicated that the shorth was 8.0 mm and the 75-shormax was 11.9 mm (Table 4.9). Sex and age differences were insignificant (t-test, $p < 0.05$).

On the right side, despite the same number of measurements being taken as on the left side, the mean had decreased to 8.5 mm (SD = 2.8; range = 13.3 mm). The female and male samples produced means of 8.8 mm and 7.1 mm respectively, and were smaller than the means for the same landmark on the opposite side of the face. The minimum and maximum values for this landmark were 2.8 mm and 16.1 mm respectively. The minimum value was collected from the 40-49 years category and the maximum value from the above 60 years category. The tissue depth statistics for the right zygion indicated that the shorth was 8.1 mm and the 75-shormax was 11.4 mm (Table 4.9). Sex and age differences were insignificant (t-test, $p < 0.05$). Furthermore, bilateral asymmetry was not significant at the zygion (t-test, $p < 0.05$).

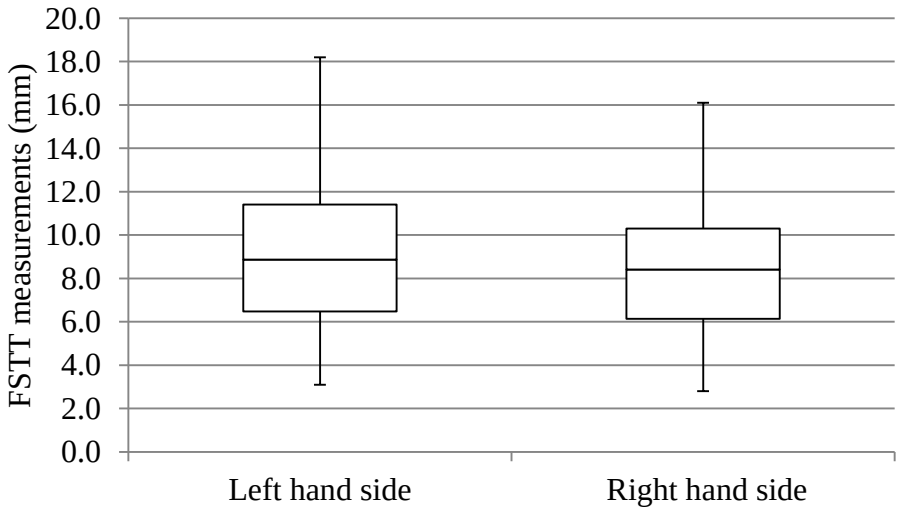


Figure 4.5.: Box and whisker plots of the FSTT measurements of the zygoma bilaterally. The horizontal line in the centre of the box represents the median value and the whiskers represent the minimum and maximum values.

O – *Mid-mandibular border*

On the left side, this landmark was measured 58 (96.7%) times and rendered a mean value of 12.4 (SD = 5.0; range = 23.6 mm) with minimum and maximum values of 2.5 mm and 26.1 mm respectively. The minimum value was cited from persons aged 40 – 49 years and the maximum was cited from persons aged 50 – 59 years. The mean was slightly larger (12.5 mm ± 5.0) in the female sample however, slightly smaller (12.2 mm ± 5.4) in the male sample. The tissue depth statistics for the left mid-mandibular indicated that the shorth was 11.5 mm and the 75-shormax was 18.7 mm (Table 4.9). Sex and age were not statistically significant (t-test, $p < 0.05$).

On the right side, the measurement was recorded 59 (98.7%) times and specified the mean as 12.6 mm (SD = 5.8; range = 29.5 mm). The minimum value was 2.9 mm whereas the maximum was 32.4 mm and were measured in persons aged 30 – 39 years and 50 – 59 years respectively. The mean value was 12.6 mm for the females and less in the male sample (12.2 mm). The tissue depth statistics for the right mid-mandibular indicated that the shorth was 11.9 mm and the 75-shormax was 18.0 mm (Table 4.9). Sex and age were not statistically significant (t-test, $p < 0.05$). Furthermore, bilateral asymmetry was not significant (t-test, $p < 0.05$).

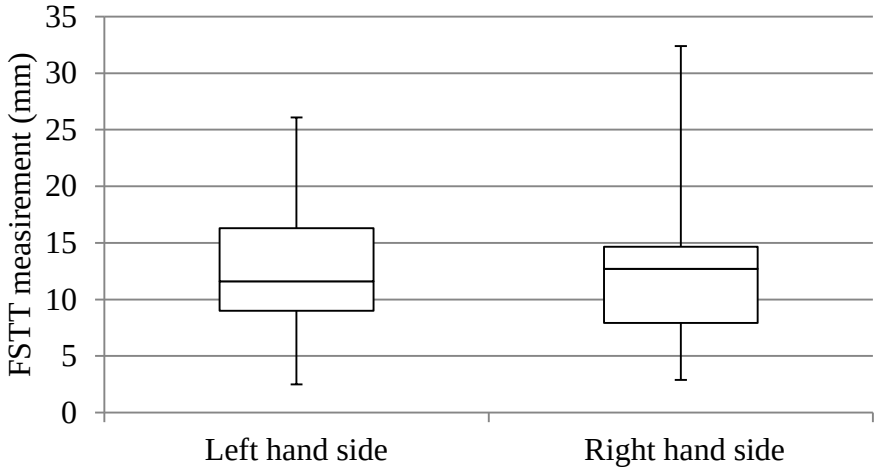


Figure 4.6.: Box and whisker plots of the FSTT measurements of the mid-mandibular bilaterally. The horizontal line in the centre of the box represents the median value and the whiskers represent the minimum and maximum values.

P – *Mid-ramus*

On the left side, 59 measurements were taken and provided a mean value of 21.9 mm (SD = 4.2; range = 18.9 mm). The minimum (14.0 mm) and maximum (33.1 mm) values were both collected from people aged 50 – 59 years and above 60 years respectively. The mean value for the males (22.0 mm) exceeded both the female mean (21.7 mm) and the mean for the collective sample. The tissue depth statistics for the left mid-ramus indicated that the shorth was 21.5 mm and the 75-shormax was 27.1 mm (Table 4.9). Sex and age were not statistically significant (t-test, $p < 0.05$).

On the right side, there were also 59 measurements providing a mean value of 21.1 mm (SD = 4.4; range =19.7 mm). The minimum value (13.4 mm) and maximum value (33.1 mm) were measured in people aged 50-59 years. The mean values calculated per sex was 20.6 mm (SD = 4.0) for females and 22.3 mm (SD = 5.7) for males. The tissue depth statistics for the right mid-ramus indicated that the shorth was 21.7 mm and the 75-shormax was 26.6 mm (Table 4.9). Sex and age were not statistically significant (t-test, $p < 0.05$). Furthermore, bilateral asymmetry was not significant (t-test, $p < 0.05$).

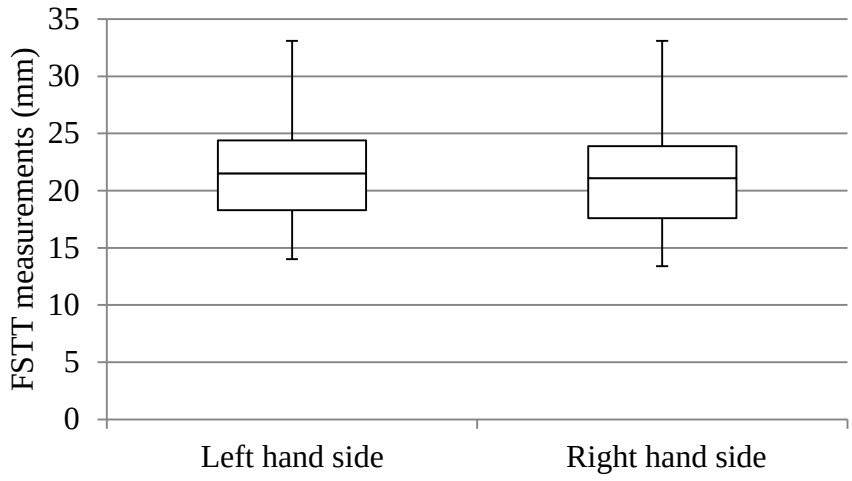


Figure 4.7.: Box and whisker plots of the FSTT measurements of the mid-ramus bilaterally. The horizontal line in the centre of the box represents the median value and the whiskers represent the minimum and maximum values.

Q – *Gonion*

The landmark was measured in all 60 scans on both sides of the face. On the left side, the statistics are as follows: mean = 22.0 mm; standard deviation = 6.9; and range = 29.8 mm. The minimum value of 7.2 mm belonged to a female aged between 40 – 49 years. The maximum value was 3.0 mm and was collected from a male aged between 50 - 59-years. The mean in females was 21.7 mm and 22.6 mm in males. The tissue depth statistics for the left gonion indicated that the shorth was 19.8 mm and the 75-shormax was 29.7 mm (Table 4.9). There was a significant difference in FSTT between persons aged below 50 years and those above 50 years. Sex was not statistically significant (t-test, $p < 0.05$).

On the right hand, there was a mean of 21.5 mm (SD = 6.9; range = 29.2 mm). The smallest value of 9.7 mm was scored from a female between the ages of 40 and 47 years. The largest value of 38.9 mm was measured from a male between the ages of

50 and 59 years. Means for females and males were 21.3 mm and 21.8 mm respectively. The tissue depth statistics for the right gonion indicated that the shorth was 22.0 mm and the 75-shormax was 28.5 mm (Table 4.9). Sex and age displayed no significant differences (t-test, $p < 0.05$). Furthermore, bilateral asymmetry was not significant (t-test, $p < 0.05$).

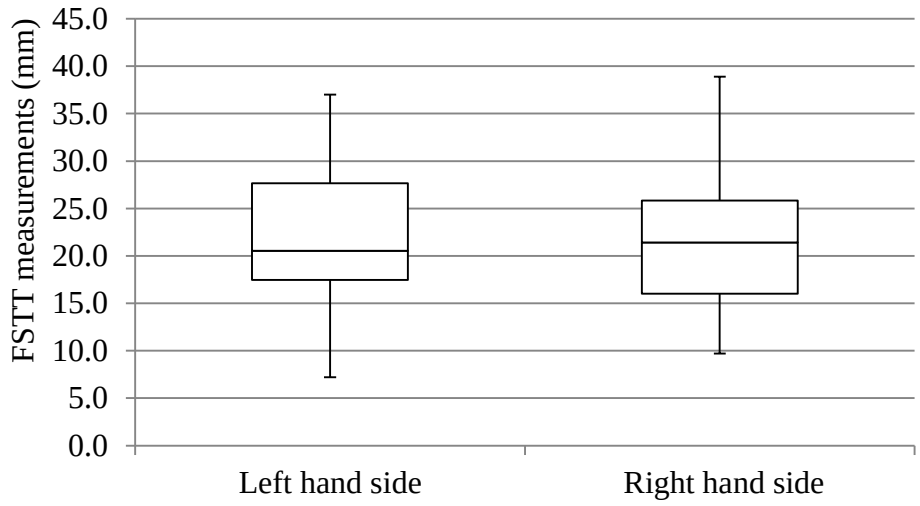


Figure 4.8.: Box and whisker plots of the FSTT measurements of the gonion bilaterally. The horizontal line in the centre of the box represents the median value and the whiskers represent the minimum and maximum values.

R – *Supra M_2*

This measurement was only recorded 43 (71.6%) and 42 (70.0%) times on the left and right side respectively.

On the left the mean was 32.2 mm (SD = 4.7; range = 21.8 mm). The minimum and maximum values at this landmark were 19.4 mm and 41.2 mm respectively. The minimum value came from female aged 50 - 59-years and the maximum from a female above the age of 60 years. Sex-specific data depicts that the mean value amongst females was 32.3 mm, slightly larger than the 32.0 mm mean found in males.

The tissue depth statistics for the left supra M_2 indicated that the shorth was 30.7 mm and the 75-shormax was 37.6 mm (Table 4.9). The FSTTs at this landmark were not statistically significant for either sex or age (t-test, $p < 0.05$).

On the right the mean was 29.8 mm (SD = 4.9; range = 18.4 mm). The minimum and maximum values at this landmark were 21.2 mm and 39.6 mm respectively. The former was collected from a male between the age of 40 and 49; and the latter from a female between the ages of 50 – 59 years. Sex-specific data depicts that the mean value amongst females was 30.2 mm compared to the 28.2 mm in males. The tissue depth statistics for the right supra M_2 indicated that the shorth was 27.5 mm and the 75-shormax was 35.7 mm (Table 4.9) The FSTTs at this landmark were not statistically significant for either sex or age (t-test, $p < 0.05$). The supra M_2 was the only landmark to indicate significant bilateral asymmetry (t-test, $p < 0.05$).

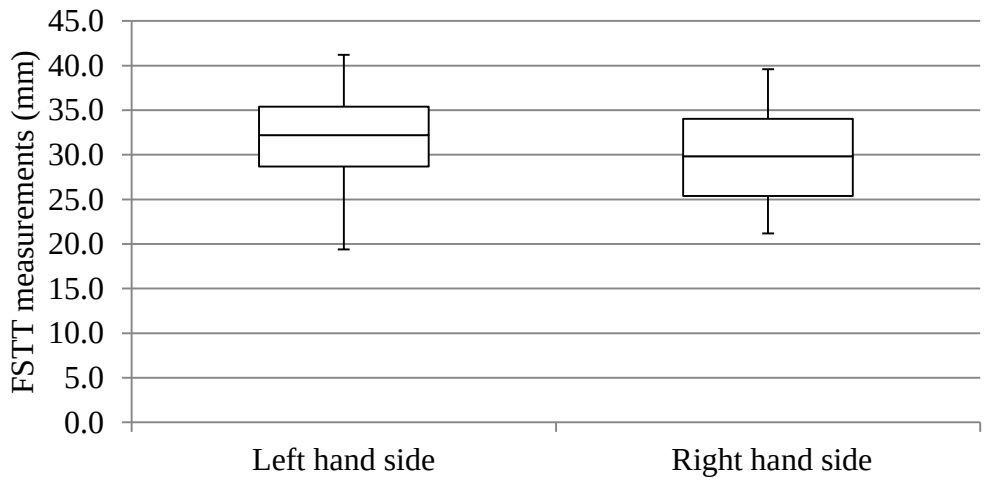


Figure 4.9.: Box and whisker plots of the FSTT measurements of the supra M_2 bilaterally. The horizontal line in the centre of the box represents the median value and the whiskers represent the minimum and maximum values.

S – *Sub M_2*

The least number of measurements were recorded at the sub M_2. There were only 35 (58.3%) and 32 (53.3%) measurements taken on the left and right side respectively.

The mean on the left side was 22.2 mm (SD = 3.8; range = 14.7 mm). The minimum (15.3 mm) and the maximum (30.0 mm) were scored from females aged 30 – 39 years and 50 - 56-years respectively. The mean value in females was 22.4 mm and 21.4 mm in males. The tissue depth statistics for the left sub M_2 indicated that the shorth was 21.2 mm and the 75-shormax was 26.5 mm (Table 4.9). Sex and age differences were insignificant (t-test, $p < 0.05$).

The mean on the right side was 21.6 mm (SD = 3.6; range = 16.0 mm) and fell between the mean values of females (21.4) and males (22.2 mm). The minimum (13.4 mm) and maximum (29.4 mm) values were collected from females in the age categories 40 – 49 years and 50- 59 years respectively. The tissue depth statistics for the right sub M_2 indicated that the shorth was 19.7 mm and the 75-shormax was 26.0 mm (Table 4.9). Sex and age differences were insignificant (t-test, $p < 0.05$). Furthermore, bilateral asymmetry was not significant (t-test, $p < 0.05$).

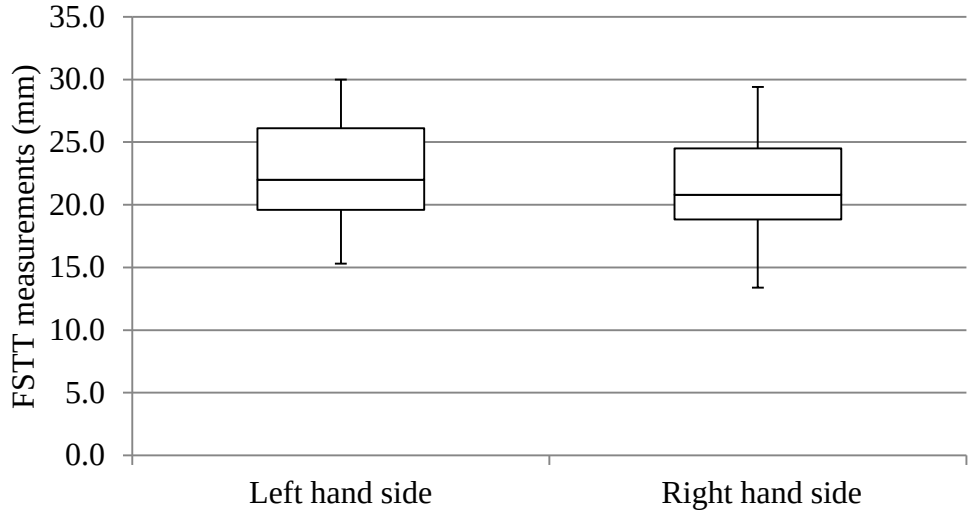


Figure 4.10.: Box and whisker plots of the FSTT measurements of the Sub M_2 bilaterally. The horizontal line in the centre of the box represents the median value and the whiskers represent the minimum and maximum values.

T – *Mental tubercle*

Measured 59 (98.3%) times on both the left and right side, this landmark is identified as the most inferior point under the chin.

The mean value on the left was 9.2 mm (SD = 1.9; range = 10.0 mm), with a minimum of 5.5 mm and a maximum of 15.5 mm found in people aged 40 -49 years and above 60 years respectively. The female sample and entire sample displayed the same mean, standard deviation, minimum, maximum and range. In the male sample the mean was 8.9 mm (SD = 1.7). The tissue depth statistics for the left mental tubercle indicated that the shorth was 8.9 mm and the 75-shormax was 10.9 mm (Table 4.9). This landmark displayed a significant difference between persons younger than 50 years and those older than 50 years. Sex differences were insignificant (t-test, $p < 0.05$).

On the right side, the mean was 9.3 mm (SD = 1.9; range = 9.1 mm), with a minimum of 5.7 mm and a maximum of 14.8 mm collected from people who were 40 – 49 years old and above 60 years respectively. Once again, females had the same mean, minimum, maximum and range as the entire sample. Males had a smaller mean of 8.9 mm (SD = 1.6; range = 5.5 mm); a value identical to the mean calculated for males on the left side. The tissue depth statistics for the left mental tubercle indicated that the shorth was 8.5 mm and the 75-shormax was 11.5 mm (Table 4.9). No statistically significant differences were determined with respect to sex and age. Furthermore, bilateral asymmetry was not significant (t-test, $p < 0.05$).

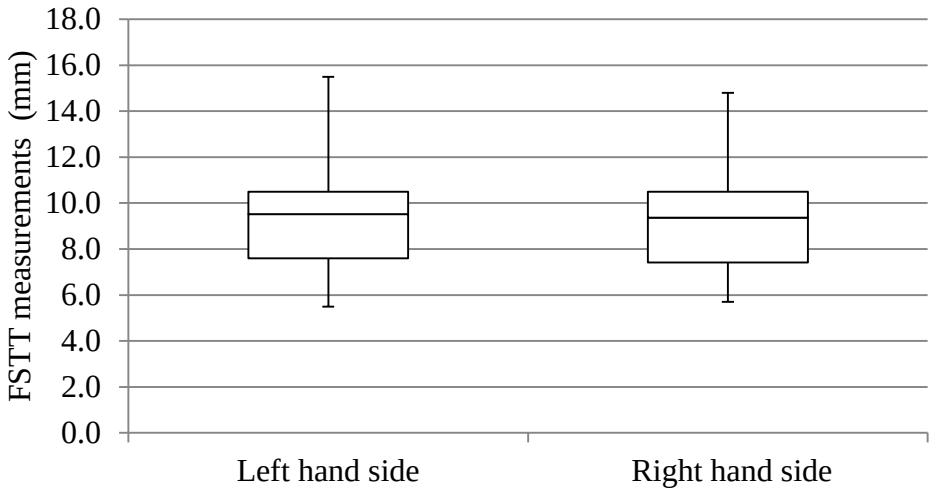


Figure 4.11.: Box and whisker plots of the FSTT measurements of the mental tubercle bilaterally. The horizontal line in the centre of the box represents the median value and the whiskers represent the minimum and maximum values.

A summary of the left and right handed FSTTs are illustrated in Figures 4.12 and 4.13. The FSTTs at gonion displayed the largest interquartile range on both the left (10.2 mm) and right (9.9 mm) hand sides. The smallest interquartile range was recorded at the mid-supraorbital (left = 2.0 mm and right = 2.2 mm). The largest and

smallest median values were recorded at supra M_2 (left = 32.2 mm and right = 29.9 mm) and the mid-infraorbital (left = 6.0 mm and right = 6.4 mm) respectively.

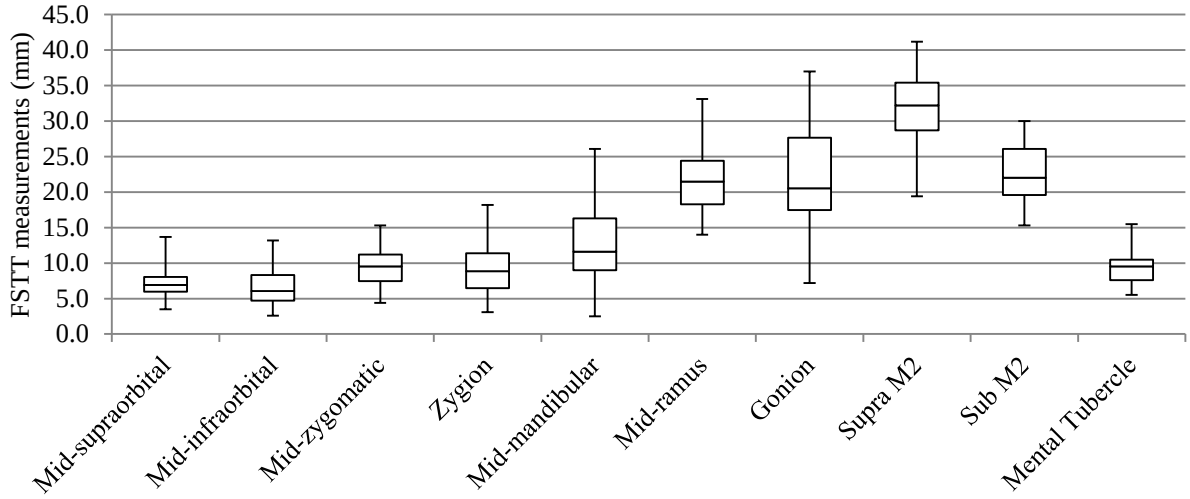


Figure 4.12.: Box and whisker plots of the FSTT measurements of lateral landmarks of the left side. The horizontal line in the centre of the box represents the median value and the whiskers represent the minimum and maximum values.

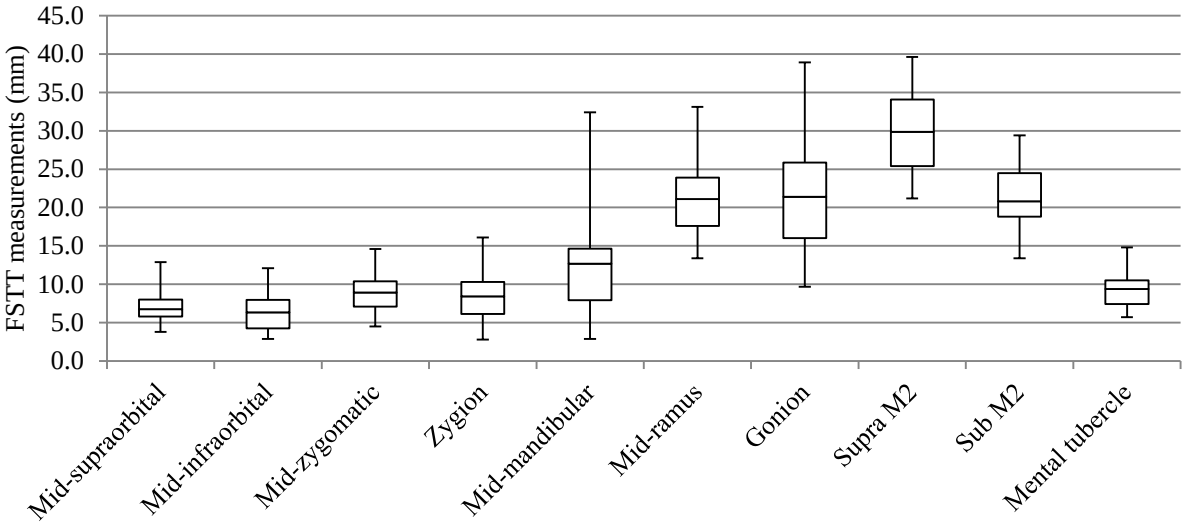


Figure 4.13.: Box and whisker plots of the FSTT measurements of the lateral landmarks on the right side. The horizontal line in the centre of the box represents the median value and the whiskers represent the minimum and maximum values.

Table 4.1.: Basic descriptive statistics for the FSTT measurements of 60 black South African adults between the ages of 19 and 70 years (average age: 49.1), (n = sample size for each measurement; S.D. = standard deviation; min = minimum; max = maximum).

Facial soft tissue thickness in mm (entire sample)						
Landmark	**N**	**Mean**	**S.D.**	**Mode**	**Min**	**Max**
Midline						
Glabella	51	4.4	1.1	3.1	2.6	7.5
Nasion	59	4.6	1.4	2.4	2.2	8.4
Mid-nasal	53	2.6	1.1	3.1	0.7	6.1
Rhinion	59	2.4	1.0	1.8	0.6	5.1
Midphiltrum	54	9.0	1.4	10.6	6.2	12.5
Labiale superius	47	9.6	1.8	10.5	6.3	13.1
Labiale inferius	54	10.2	2.2	10.8	6.0	16.1
Labiomental sulcus	58	11.3	1.9	10.6	6.5	15.4
Pogonion	58	10.7	2.4	10.6	2.3	15.9
Menton	59	5.0	2.7	8.4	1.1	19.0
Bilateral						
Mid-supraorbital (LHS)	58	7.0	1.7	5.0	3.5	13.7
Mid-supraorbital (RHS)	58	7.1	1.7	5.3	3.8	12.9
Mid-infraorbital (LHS)	59	6.6	2.5	N/A	2.6	13.2
Mid-infraorbital (RHS)	59	6.6	2.4	4.3	2.9	12.1
Mid-zygomatic (LHS)	59	9.5	2.3	10.3	4.4	15.3
Mid-zygomatic (RHS)	59	9.0	2.3	10.8	4.5	14.6
Zygion (LHS)	59	8.9	3.0	12.4	3.1	18.2
Zygion (RHS)	59	8.5	2.8	12.3	2.8	16.1
Mid-mandibular (LHS)	58	12.4	5.0	18.7	2.5	26.1
Mid-mandibular (RHS)	59	12.6	5.8	13.4	2.9	32.4
Mid-ramus (LHS)	59	21.9	4.2	24.1	14.0	33.1
Mid-ramus (RHS)	59	21.1	4.4	27.4	13.4	33.1
Gonion (LHS)	60	22.0	6.9	30.5	7.2	37.0
Gonion (RHS)	60	21.5	6.9	25.3	9.7	38.9
Supra M_2 (LHS)	43	32.2	4.7	33.8	19.4	41.2
Supra M_2 (RHS)	42	29.8	4.9	26.5	21.2	39.6
Sub M_2 (LHS)	35	22.2	3.8	19.6	15.3	30.0
Sub M_2 (RHS)	32	21.6	3.6	17.9	13.4	29.4
Mental tubercle (LHS)	59	9.2	1.9	10.9	5.5	15.5
Mental tubercle (RHS)	59	9.3	1.9	11.9	5.7	14.8

Table 4.2.: Basic descriptive statistics for the FSTT measurements of 47 black South African adult females between the ages of 19 and 70 years (average age: 48.9), (n = sample size for each measurement; S.D. = standard deviation; min = minimum; max = maximum).

Facial soft tissue thickness (females)						
Landmark	**n**	**Mean**	**S.D.**	**Mode**	**Min**	**Max**
Midline						
Glabella	39	4.3	1.1	5.1	2.6	6.4
Nasion	46	4.4	1.3	2.4	2.2	8.0
Mid-nasal	41	2.6	1.2	3.2	0.7	6.1
Rhinion	46	2.4	1.0	2.0	0.6	5.1
Midphiltrum	44	8.9	1.4	10.6	6.2	12.5
Labiale superius	39	9.5	1.9	12.1	6.3	13.1
Labiale inferius	43	10.1	2.4	9.4	6.0	16.1
Labiomental sulcus	46	11.2	1.9	10.6	6.5	14.6
Pogonion	46	10.6	2.6	10.9	2.3	15.9
Menton	47	5.1	2.9	8.4	1.1	19.0
Bilateral						
Mid-supraorbital (LHS)	46	7.1	1.7	8.3	4.6	13.7
Mid-supraorbital (RHS)	46	7.2	1.8	6.5	4.5	12.9
Mid-infraorbital (LHS)	47	6.8	2.5	N/A	2.8	13.2
Mid-infraorbital (RHS)	47	6.8	2.3	7.6	3.2	12.1
Mid-zygomatic (LHS)	47	10.0	2.2	10.3	5.2	15.3
Mid-zygomatic (RHS)	47	9.4	2.2	10.8	4.5	14.6
Zygion (LHS)	47	9.2	2.9	12.4	3.1	18.2
Zygion (RHS)	47	8.8	2.7	11.3	3.3	16.1
Mid-mandibular (LHS)	47	12.5	5.0	18.7	2.5	25.2
Mid-mandibular (RHS)	46	12.6	5.4	12.7	2.9	28.8
Mid-ramus (LHS)	47	21.7	4.4	24.4	14.0	33.1
Mid-ramus (RHS)	47	20.6	4.0	27.4	13.4	27.4
Gonion (LHS)	47	21.7	7.1	30.5	7.2	36.1
Gonion (RHS)	47	21.3	6.4	25.3	9.7	37.2
Supra M_2 (LHS)	35	32.3	4.8	28.5	19.4	41.2
Supra M_2 (RHS)	34	30.2	4.7	26.5	21.8	39.6
Sub M_2 (LHS)	29	22.4	3.9	26.5	15.3	30.0
Sub M_2 (RHS)	24	21.4	3.8	17.9	13.4	29.4
Mental tubercle (LHS)	47	9.2	1.9	10.2	5.5	15.5
Mental tubercle (RHS)	47	9.3	2.0	10.5	5.7	14.8

Table 4.3.: Basic descriptive statistics for the FSTT measurements of 13 black South African adult males between the ages of 29 and 66 years (average age: 49.6), (n = sample size for each measurement; S.D. = standard deviation; min = minimum; max = maximum).

Facial soft tissue thickness (males)						
Landmark	**n**	**Mean**	**S.D.**	**Mode**	**Min**	**Max**
Midline						
Glabella	12	4.6	1.3	N/A	3.0	7.5
Nasion	13	5.1	1.6	N/A	3.1	8.4
Mid-nasal	12	2.5	0.8	N/A	1.0	3.9
Rhinion	13	2.5	1.0	N/A	1.2	4.1
Midphiltrum	10	9.7	1.4	N/A	7.1	12.1
Labiale superius	8	10.0	1.1	10.5	7.6	11.0
Labiale inferius	11	10.6	1.7	11.7	8.1	13.9
Labiomental sulcus	12	11.6	1.8	10.6	9.9	15.4
Pogonion	12	10.9	1.4	10.6	9.0	13.2
Menton	12	4.7	1.7	N/A	1.9	7.4
Bilateral						
Mid-supraorbital (LHS)	12	6.9	1.8	6.1	3.5	9.6
Mid-supraorbital (RHS)	12	6.6	1.6	5.3	3.8	9.0
Mid-infraorbital (LHS)	12	5.5	2.4	N/A	2.6	10.9
Mid-infraorbital (RHS)	12	5.5	2.4	N/A	2.9	11.2
Mid-zygomatic (LHS)	12	7.3	1.5	N/A	4.4	10.3
Mid-zygomatic (RHS)	12	7.0	1.3	N/A	4.6	8.9
Zygion (LHS)	12	7.4	3.2	N/A	3.2	14.1
Zygion (RHS)	12	7.1	3.1	N/A	2.8	13.0
Mid-mandibular (LHS)	12	12.2	5.4	N/A	6.0	26.1
Mid-mandibular (RHS)	12	12.7	7.2	12.1	5.5	32.4
Mid-ramus (LHS)	12	22.0	3.7	24.1	17.0	27.2
Mid-ramus (RHS)	12	22.3	5.7	N/A	14.0	33.1
Gonion (LHS)	13	22.6	6.4	N/A	15.6	37.0
Gonion (RHS)	13	21.8	8.8	N/A	11.3	38.9
Supra M_2 (LHS)	8	32.0	4.8	N/A	24.9	38.6
Supra M_2 (RHS)	8	28.2	5.8	N/A	21.2	38.5
Sub M_2 (LHS)	6	21.4	3.7	19.6	16.1	26.1
Sub M_2 (RHS)	8	22.2	3.5	N/A	17.4	28.0
Mental tubercle (LHS)	12	8.9	1.7	N/A	6.6	11.2
Mental tubercle (RHS)	12	8.9	1.6	N/A	6.4	11.9

Table 4.4.: Basic descriptive statistics for the FSTT measurements of 3 black South African adults between the ages of 19 and 29 years (average age: 23.7), (n = sample size for each measurement; S.D. = standard deviation; min = minimum; max = maximum).

Facial soft tissue thicknesses of persons in mm (aged 19-29 years)						
	n	**Mean**	**SD**	**Mode**	**Min**	**Max**
Midline						
Glabella	3	3.6	0.6	N/A	3.1	4.3
Nasion	3	4.1	0.3	N/A	3.9	4.4
Mid-nasal	3	2.3	0.5	N/A	1.9	2.9
Rhinion	3	10.0	1.7	N/A	8.5	11.8
Midphiltrum	3	10.9	2.0	N/A	9.1	13.1
Labiale superius	3	8.0	0.5	N/A	7.5	8.5
Labiale inferius	3	11.0	1.6	N/A	10.0	12.8
Labiomental sulcus	3	9.7	1.6	N/A	7.9	10.7
Pogonion	3	4.8	2.2	N/A	3.4	7.4
Menton	3	2.5	0.7	N/A	1.6	2.9
Bilateral						
Mid-Supraorbital (LHS)	3	5.9	0.8	N/A	5.1	6.6
Mid-Supraorbital (RHS)	3	6.0	0.9	N/A	5.0	6.6
Mid-Infraorbital (LHS)	3	5.1	0.4	N/A	4.6	5.4
Mid-Infraorbital (RHS)	3	5.0	0.8	N/A	4.3	5.9
Mid-zygomatic (LHS)	3	7.9	0.5	N/A	7.6	8.5
Mid-zygomatic (RHS)	3	7.6	1.1	N/A	6.8	8.8
Zygion (LHS)	3	7.9	0.4	N/A	7.5	8.4
Zygion (RHS)	3	8.0	0.4	N/A	7.5	8.3
Mid-mandibular (LHS)	3	10.8	1.3	N/A	9.8	12.2
Mid-mandibular (RHS)	3	12.5	3.8	N/A	8.4	15.8
Mid-ramus (LHS)	3	19.8	2.1	N/A	17.9	22.0
Mid-ramus (RHS)	3	20.8	2.8	N/A	18.7	24.0
Gonion (LHS)	3	18.7	6.1	N/A	11.8	23.3
Gonion (RHS)	3	19.7	5.6	N/A	13.3	23.8
Supra M_2 (LHS)	3	32.9	4.6	N/A	29.3	38.1
Supra M_2 (RHS)	3	33.3	5.3	N/A	27.9	38.5
Sub M_2 (LHS)	2	23.7	1.5	N/A	22.6	24.7
Sub M_2 (RHS)	3	23.9	2.3	N/A	21.4	26.0
Mental tubercle (LHS)	3	7.7	0.4	N/A	7.3	8.1
Mental tubercle (RHS)	3	7.0	0.3	N/A	6.8	7.4

Table 4.5.: Basic descriptive statistics for the FSTT measurements of 11 black South African adults between the ages of 30 and 39 years (average age: 35.8), (n = sample size for each measurement; S.D. = standard deviation; min = minimum; max = maximum).

Facial soft tissue thicknesses of persons in mm (aged 30-39 years)						
	n	**Mean**	**SD**	**Mode**	**Min**	**Max**
Midline						
Glabella	9	3.7	1.0	N/A	2.6	5.8
Nasion	11	3.8	1.2	N/A	2.2	5.7
Mid-nasal	9	2.1	0.9	N/A	1.0	3.7
Rhinion	11	2.1	0.5	1.8	1.5	3.0
Midphiltrum	11	9.0	1.0	N/A	6.6	10.4
Labiale superius	9	9.7	1.8	N/A	6.6	11.8
Labiale inferius	11	9.4	2.1	N/A	7.0	14.2
Labiomental sulcus	11	10.5	1.4	N/A	8.4	13.2
Pogonion	11	10.6	2.2	11	7.9	15.9
Menton	11	4.6	2.5	N/A	1.5	8.4
Bilateral						
Mid-Supraorbital (LHS)	11	6.5	1.3	N/A	4.6	8.2
Mid-Supraorbital (RHS)	10	6.6	1.0	N/A	4.8	7.6
Mid-Infraorbital (LHS)	11	5.5	2.2	N/A	2.8	9.5
Mid-Infraorbital (RHS)	11	5.6	2.1	N/A	3.1	10.2
Mid-zygomatic (LHS)	11	9.7	3.1	10.3	4.4	13.6
Mid-zygomatic (RHS)	11	9.5	3.1	N/A	4.6	13.5
Zygion (LHS)	11	8.7	3.0	N/A	3.9	12.5
Zygion (RHS)	11	8.6	2.7	N/A	4.3	12.2
Mid-mandibular (LHS)	11	10.5	4.3	13.8	3.4	17.2
Mid-mandibular (RHS)	10	11.8	5.0	13.4	2.9	17.8
Mid-ramus (LHS)	11	22.2	2.4	N/A	18.2	25.9
Mid-ramus (RHS)	11	21.2	3.2	24.4	15.4	25.5
Gonion (LHS)	11	17.4	6.3	N/A	9.8	30.5
Gonion (RHS)	11	18.4	6.9	21.4	10.7	27.5
Supra M_2 (LHS)	9	31.0	4.7	28.5	24.9	37.5
Supra M_2 (RHS)	9	29.4	5.6	24.7	23.7	36.6
Sub M_2 (LHS)	9	20.9	4.3	N/A	15.3	27.2
Sub M_2 (RHS)	6	19.9	5.0	N/A	13.4	26.3
Mental tubercle (LHS)	11	8.8	1.7	N/A	6.8	11.1
Mental tubercle (RHS)	11	9.5	2.2	N/A	7.0	12,4

Table 4.6.: Basic descriptive statistics for the FSTT measurements of 15 black South African adults between the ages of 40 and 49 years (average age: 45.8), (n = sample size for each measurement; S.D. = standard deviation; min = minimum; max = maximum).

Facial soft tissue thicknesses of persons in mm (aged 40-49 years)						
	n	**Mean**	**SD**	**Mode**	**Min**	**Max**
Midline						
Glabella	15	4.3	1.2	3.0	3.0	6.4
Nasion	15	4.2	0.8	N/A	2.4	5.8
Mid-nasal	12	2.3	0.9	N/A	1.5	4.1
Rhinion	14	2.0	0.7	N/A	0.6	3.5
Midphiltrum	15	9.3	1.3	N/A	7.5	12.1
Labiale superius	12	9.8	1.3	9.6	7.0	12.6
Labiale inferius	12	9.4	2.1	N/A	6.9	13.9
Labiomental sulcus	15	10.7	2.0	N/A	7.9	15.4
Pogonion	15	10.7	2.1	N/A	7.3	15.0
Menton	15	4.8	1.9	N/A	1.9	8.4
Bilateral						
Mid-Supraorbital (LHS)	15	6.5	1.4	N/A	4.6	8.6
Mid-Supraorbital (RHS)	15	6.8	1.4	5.3	4.5	8.9
Mid-Infraorbital (LHS)	15	5.8	2.4	N/A	2.6	10.6
Mid-Infraorbital (RHS)	15	5.8	2.2	N/A	2.9	9.8
Mid-zygomatic (LHS)	15	9.0	2.2	10.7	6.4	13.4
Mid-zygomatic (RHS)	15	8.3	2.2	10.4	5.7	11.7
Zygion (LHS)	15	7.9	3.0	N/A	3.2	13.2
Zygion (RHS)	15	7.5	2.9	12.3	2.8	12.3
Mid-mandibular (LHS)	15	11.9	4.9	16.3	2.5	18.7
Mid-mandibular (RHS)	15	10.8	5.2	12.7	3.1	23.5
Mid-ramus (LHS)	15	21.9	4.2	24.1	15.5	28.9
Mid-ramus (RHS)	15	21.4	4.0	27.4	14.7	27.4
Gonion (LHS)	15	19.9	5.7	16.8	7.2	28.8
Gonion (RHS)	15	20.8	5.9	17.1	9.7	30.0
Supra M_2 (LHS)	10	32.9	5.4	N/A	25.8	41.2
Supra M_2 (RHS)	11	29.1	5.5	26.5	21.2	38.2
Sub M_2 (LHS)	9	23.0	3.5	19.6	19.6	27.1
Sub M_2 (RHS)	8	22.5	3.5	24.5	17.9	29.4
Mental tubercle (LHS)	15	8.6	1.9	N/A	5.5	11.6
Mental tubercle (RHS)	15	8.9	1.7	10.0	5.7	12.1

Table 4.7.: Basic descriptive statistics for the FSTT measurements of 23 black South African adults between the ages of 50 and 59 years (average age: 55.4), (n = sample size for each measurement; S.D. = standard deviation; min = minimum; max = maximum).

Facial soft tissue thicknesses of persons in mm (aged 50-59 years)						
	n	**Mean**	**SD**	**Mode**	**Min**	**Max**
Midline						
Glabella	19	4.7	0.9	N/A	3.1	6.3
Nasion	22	5.2	1.5	N/A	2.4	8.0
Mid-nasal	22	2.9	1.2	N/A	0.7	6.1
Rhinion	23	2.9	1.0	N/A	1.1	5.1
Midphiltrum	20	8.8	1.5	N/A	6.2	11.8
Labiale superius	20	9.4	2.0	12.1	6.4	12.7
Labiale inferius	22	11.2	1.7	10.8	9.0	15.6
Labiomental sulcus	21	11.9	1.3	10.6	10.2	14.2
Pogonion	21	10.5	2.6	11.3	2.3	14.7
Menton	22	5.4	3.6	4.4	1.1	19.0
Bilateral						
Mid-Supraorbital (LHS)	21	7.5	1.5	N/A	3.5	10.9
Mid-Supraorbital (RHS)	22	7.4	1.9	5.8	3.8	12.3
Mid-Infraorbital (LHS)	22	7.3	2.5	N/A	3.7	13.1
Mid-Infraorbital (RHS)	22	7.3	2.5	N/A	3.2	12.1
Mid-zygomatic (LHS)	22	9.4	2.0	N/A	5.2	13.9
Mid-zygomatic (RHS)	22	9.0	1.9	10.7	4.5	12.2
Zygion (LHS)	22	8.9	2.9	12.0	3.1	13.0
Zygion (RHS)	22	8.5	2.8	N/A	3.3	14.4
Mid-mandibular (LHS)	22	13.9	5.9	18.7	4.7	26.1
Mid-mandibular (RHS)	22	14.1	7.0	N/A	5.5	32.4
Mid-ramus (LHS)	22	21.4	4.8	27.6	14.0	29.2
Mid-ramus (RHS)	22	20.3	4.9	14.0	13.4	33.1
Gonion (LHS)	23	25.8	7.1	N/A	13.1	37.0
Gonion (RHS)	23	23.6	7.5	25.3	10.4	38.9
Supra M_2 (LHS)	17	31.7	4.9	N/A	19.4	39.1
Supra M_2 (RHS)	17	29.5	4.3	25.1	21.8	39.6
Sub M_2 (LHS)	13	22.3	4.0	N/A	15.5	30.0
Sub M_2 (RHS)	14	21.0	3.1	N/A	17.9	28.0
Mental tubercle (LHS)	22	9.3	1.7	10.2	6.5	12.4
Mental tubercle (RHS)	22	9.1	1.6	10.6	6.3	11.9

Table 4.8.: Basic descriptive statistics for the FSTT measurements of 8 black South African adults aged 60 years and above (average age: 64.5), (n = sample size for each measurement; S.D. = standard deviation; min = minimum; max = maximum).

Facial soft tissue thicknesses of persons in mm (aged 60+ years)						
	n	**Mean**	**SD**	**Mode**	**Min**	**Max**
Midline						
Glabella	5	4.7	2.0	N/A	3.2	7.5
Nasion	8	5.1	1.8	N/A	2.9	8.4
Mid-nasal	7	2.7	1.3	N/A	1.4	5.3
Rhinion	8	2.5	1.4	N/A	1.0	4.4
Midphiltrum	5	9.0	2.2	N/A	6.6	12.5
Labiale superius	3	8.4	1.8	N/A	6.3	9.5
Labiale inferius	6	10.4	3.3	N/A	6.0	16.1
Labiomental sulcus	8	11.8	3.0	14.2	6.5	14.6
Pogonion	8	11.3	3.0	N/A	6.7	15.8
Menton	8	5.4	1.1	5.6	4.5	7.9
Bilateral						
Mid-Supraorbital (LHS)	8	7.9	2.7	N/A	5.6	13.7
Mid-Supraorbital (RHS)	8	7.6	2.4	N/A	5.7	12.9
Mid-Infraorbital (LHS)	8	7.8	2.7	N/A	4.2	13.2
Mid-Infraorbital (RHS)	8	7.5	2.2	7.6	3.8	11.8
Mid-zygomatic (LHS)	8	10.6	2.6	N/A	7.3	15.3
Mid-zygomatic (RHS)	8	9.8	2.7	10.8	5.6	14.6
Zygion (LHS)	8	10.7	3.8	N/A	6.0	18.2
Zygion (RHS)	8	10.0	3.1	N/A	6.1	16.1
Mid-mandibular (LHS)	8	12.5	4.0	N/A	7.3	19.6
Mid-mandibular (RHS)	8	12.5	4.2	13.2	6.8	18.7
Mid-ramus (LHS)	8	22.6	5.5	N/A	16.1	33.1
Mid-ramus (RHS)	8	21.7	5.9	N/A	14.2	32.4
Gonion (LHS)	8	21.8	3.8	N/A	18.4	29.8
Gonion (RHS)	8	20.7	7.1	N/A	14.1	37.0
Supra M_2 (LHS)	4	35.1	3.1	N/A	32.9	39.6
Supra M_2 (RHS)	2	33.4	5.4	N/A	29.6	37.2
Sub M_2 (LHS)	2	22.9	5.2	N/A	19.2	26.6
Sub M_2 (RHS)	1	26.4	N/A	N/A	26.4	26.4
Mental tubercle (LHS)	8	10.8	1.9	10.5	9.5	15.5
Mental tubercle (RHS)	8	10.7	2.1	10.4	7.7	14.8

Table 4.9.: Means, shorths and 75-shormaxes of FSTT data at midline and bilateral anatomical landmarks for the entire sample (females and males) of the current study.

Facial soft tissue thicknesses in mm (entire sample)			
Landmarks	**Mean (SD)**	**Shorth**	**75-shormax**
Midline			
Glabella	4.4 (1.1)	3.6	5.5
Nasion	4.6 (1.4)	4.5	6.4
Mid-nasal	2.6 (1.1)	2.1	3.6
Rhinion	2.4 (1.0)	1.8	3.3
Midphiltrum	9.0 (1.4)	9.0	10.6
Labiale superius	9.6 (1.8)	9.7	12.1
Labiale inferius	10.2 (2.2)	9.8	12.6
Labiomental sulcus	11.3 (1.9)	10.4	13.2
Pogonium	10.7 (2.4)	10.2	13.1
Menton	5.0 (2.7)	4.6	7.6
Bilateral			
Mid-supraorbital (LHS)	7.0 (1.7)	7.3	8.7
Mid-supraorbital (RHS)	7.1 (1.7)	6.7	8.7
Mid-infraorbital (LHS)	6.6 (2.5)	4.6	8.4
Mid-infraorbital (RHS)	6.6 (2.4)	5.6	8.9
Mid-zygomatic (LHS)	9.5 (2.3)	8.7	11.9
Mid-zygomatic (RHS)	9.0 (2.3)	9.5	11.5
Zygion (LHS)	8.9 (3.0)	8.0	11.9
Zygion (RHS)	8.5 (2.8)	8.1	11.4
Mid-mandibular (LHS)	12.4 (5.0)	11.5	18.7
Mid-mandibular (RHS)	12.6 (5.8)	11.9	18.0
Mid-ramus (LHS)	21.9 (4.2)	21.5	27.1
Mid-ramus (RHS)	21.1 (4.4)	21.7	26.6
Gonion (LHS)	22.0 (6.9)	19.8	29.7
Gonion (RHS)	21.5 (6.9)	22.0	28.5
Supra M_2 (LHS)	32.2 (4.7)	30.7	37.6
Supra M2 (RHS)	29.8 (4.9)	27.5	35.7
Sub M_2 (LHS)	22.2 (3.8)	21.2	26.5
Sub M_2 (RHS)	21.6 (3.6)	19.7	26.0
Mental tubercle (LHS)	9.2 (1.9)	8.9	10.9
Mental tubercle (RHS)	9.3 (1.9)	8.5	11.5

4.4. Intra- and inter-observer error rates

Nineteen scans were re-measured to test for intra- and inter-observer reliability and calculated using intra-class correlation (ICC). Technical error of measurement (TEM) and the variable average value (VAV) at each landmark were also calculated. Tables 4.9 – 4.12 presents these statistics for the intra- and inter-observer reliability.

In total there were three measurements per landmark. The first reading represents the original data collected by the investigator and second reading represents the data taken the second time by the investigator necessary to calculate the intra-observer reliability. A person external to the study and graduate student of anatomy with knowledge of facial landmarks conducted the inter-observer reliability test. The measurements collected by the external graduate student were labeled as the third reading.

The correlation coefficient (r) is an inferential statistic that describes the strength in correlation between pairs of observations (Fancourt & Stephan, 2018). This value of r has an upper limit of 1 which indicates perfect correlation and hence, there would be an insignificant statistical difference. Furthermore, the measurements would be found to reliable and indicating good repeatability.

Technical error of measurement functions as an index of accuracy and is calculated by determining the standard deviation between repeated measurements (Goto, 2007). This can be presented as an absolute value or a percentage. The variable average value (also known as the average of averages) is a useful statistic between samples of equal size for the comparison to the mean values in each sample (Perini *et al*., 2005).

For the purposes of this study, an accepTable ICC would range between 0.75 – 1.00 as cited in Cicchetti (1994). The intra-observer reliability at all the midline and bilateral landmarks were found to well within this range. The landmark with lowest

ICC-value of 0.953 was the pogonion. All the landmarks investigated in this study the relative technical errors of measurement (r-TEM) between readings 1 and 2 were less than 5 %. The lowest r-TEM was 0.72% at the right mid-ramus and the highest r-TEM was 4.58% at the mid-nasal.

The inter-observer reliability in this study tended to indicate a greater variation of ICC values in comparison to the intra-observer reliability; as only 17 of the 30 landmarks exhibited an ICC of 0.75 or more. The highest ICC was determined to exist at the right mid-ramus (0.991). The landmarks which displayed the worst measures of repeatability whereby the ICCs were below 0.5 included the labial inferius (0.469) and menton (0.481). These differences may be consequence to the fact that the principle investigator was more proficient in using the software and locating the landmarks than the external observer. Nonetheless, the largest difference between the means of readings 1 and 3 did not exceed 2.2 millimeters. The r-TEM (%) for the inter-observer reliability indicated much larger value compared to the intra-observer reliability. Only 5 of the 30 landmarks yielded r-TEMs less than 5%. The lowest r-TEMs were found at the left and right mid-rami (3.42 & 1.97 respectively). The highest r-TEMs were 28.2 and 36.8 at the mid-nasal and menton respectively. The highest r-TEMs for bilateral FSTT measurements included regions around the jaw (e.g. zygion, mid-mandibular and gonion). Poor interpretation of the definitions for these landmarks in conjunction to the external observer's inexperience with imaging software may have resulted in the misidentification of the landmarks and the subsequent large difference in r-TEMS between the original and inter-observer measurements. Furthermore, no statistical differences were found to exist between readings 1 and 2, and readings 1 and 3 (student t-tests, $p < 0.05$).

Table 4.10: Intra-observer repeatability of midline landmarks expressed by the intra-class correlation coefficient (ICC), technical error of measurement (TEM) and the variable average value (VAV).

Measurement	Reading	Mean (SD)	ICC	Absolute TEM	VAV	Relative TEM %
Glabella	1	4.2 (0.9)	0.974	0.2	4.3	3.8
	2	4.3 (1.0)				
Nasion	1	4.5 (1.3)	0.991	0.1	4.5	2.9
	2	4.6 (1.4)				
Mid-nasal	1	2.5 (1.1)	0.989	0.1	2.5	4.6
	2	2.5 (1.0)				
Rhinion	1	2.5 (0.9)	0.993	0.1	2.5	3.0
	2	2.5 (0.9)				
Mid-philtrum	1	9.0 (1.3)	0.993	0.1	9.0	1.3
	2	9.0 (1.4)				
Labiale Superius	1	10.0 (1.7)	0.967	0.3	10.0	3.1
	2	10.1 (1.7)				
Labiale Inferius	1	10.2 (2.2)	0.991	0.2	10.3	2.1
	2	10.4 (2.2)				
Labiomental sulcus	1	11.3 (1.5)	0.993	0.1	11.3	1.1
	2	11.3 (1.5)				
Pogonion	1	10.6 (1.8)	0.953	0.3	10.7	2.3
	2	10.8 (1.9)				
Menton	1	5.5 (3.7)	0.999	0.1	5.5	1.8
	2	5.4 (3.6)				

Table 4.11: Intra-observer repeatability of bilateral landmarks expressed by the intra-class correlation coefficient (ICC), technical error of measurement (TEM) and the variable average value (VAV).

Measurement	Reading	Mean (SD)	ICC	Absolute TEM	VAV	Relative TEM %
Mid-supraorbital (LHS)	1	7.1 (1.6)	0.997	0.1	7.1	1.3
	2	7.1 (1.7)				
Mid-supraorbital (RHS)	1	7.0 (1.6)	0.997	0.1	7.0	1.4
	2	7.0 (1.6)				
Mid-infraorbital (LHS)	1	7.0 (2.4)	0.998	0.1	7.0	1.4

	2	7.0 (2.4)				
Mid-infraorbital (RHS)	1	6.7 (2.2)	0.998	0.1	6.8	1.3
	2	6.8 (2.2)				
Mid-zygomatic (LHS)	1	9.6 (2.5)	0.999	0.1	9.6	1.0
	2	9.6 (2.6)				
Mid-zygomatic (RHS)	1	9.3 (2.4)	0.997	0.1	9.3	1.3
	2	9.4 (2.4)				
Zygion (LHS)	1	9.3 (2.6)	0.999	0.1	9.3	1.0
	2	9.4 (2.8)				
Zygion (RHS)	1	9.2 (2.8)	0.998	0.1	9.2	1.5
	2	9.2 (2.8)				
Mid-mandibular (LHS)	1	13.9 (5.7)	0.997	2.3	13.9	2.4
	2	14.0 (6.1)				
Mid-mandibular (RHS)	1	14.5 (7.0)	0.997	0.4	14.4	2.9
	2	14.3 (6.8)				
Mid-ramus (LHS)	1	22.9 (3.9)	0.998	0.8	22.8	3.4
	2	22.9 (4.0)				
Mid-ramus (RHS)	1	22.2 (4.5)	0.999	0.2	22.2	0.7
	2	22.1 (4.5)				
Gonion (LHS)	1	23.8 (7.3)	0.999	0.2	23.7	1.0
	2	23.6 (7.4)				
Gonion (RHS)	1	24.8 (7.0)	0.994	0.6	24.6	2.3
	2	24.4 (7.3)				
Supra M_2 (LHS)	1	33.0 (4.8)	0.953	1.1	32.6	3.3
	2	32.2 (5.0)				
Supra M_2 (RHS)	1	30.7 (5.2)	0.996	0.3	30.6	1.1
	2	30.6 (5.4)				
Sub M_2 (LHS)	1	22.6 (3.9)	0.987	0.4	22.7	2.0
	2	22.9 (3.9)				
Sub M_2 (RHS)	1	22.0 (3.6)	0.982	0.5	22.1	2.2
	2	22.1 (3.6)				
Mental tubercle (LHS)	1	9.2 (1.5)	0.992	0.1	9.3	1.5
	2	9.3 (1.5)				
Mental tubercle (RHS)	1	9.4 (1.5)	0.991	0.2	9.4	1.6
	2	9.4 (1.6)				

Table 4.12: Inter-observer repeatability of midline landmarks expressed by the intra-class correlation coefficient (ICC), technical error of measurement (TEM) and the variable average value (VAV)

Measurement	Reading	Mean (SD)	ICC	Absolute TEM	VAV	Relative TEM %
Glabella	1	4.2 (0.9)	0.931	0.2	3.7	4.3
	3	4.3 (1.1)				
Nasion	1	4.5 (1.3)	0.841	0.6	4.3	12.9
	3	4.1 (1.3)				
Mid-nasal	1	2.5 (1.1)	0.669	0.7	2.6	28.2
	3	2.6 (1.3)				
Rhinion	1	2.5 (0.9)	0.884	0.3	2.5	13.1
	3	2.6 (1.0)				
Mid-philtrum	1	9.0 (1.3)	0.825	0.6	9.2	6.4
	3	9.4 (1.3)				
Labiale Superius	1	10.0 (1.7)	0.854	0.7	10.1	6.5
	3	10.1 (1.7)				
Labiale Inferius	1	10.2 (2.2)	0.469	1.2	10.0	11.9
	3	9.5 (1.2)				
Labiomental sulcus	1	11.3 (1.5)	0.703	0.9	11.4	7.7
	3	11.5 (1.6)				
Pogonion	1	10.6 (1.8)	0.746	1.0	10.2	9.8
	3	9.8 (1.9)				
Menton	1	5.5 (3.7)	0.481	2.2	6.0	36.7
	3	6.6 (1.9)				

Table 4.13: Inter-observer repeatability of bilateral landmarks expressed by the intra-class correlation coefficient (ICC), technical error of measurement (TEM) and the variable average value (VAV)

Measurement	Reading	Mean (SD)	ICC	Absolute TEM	VAV	Relative TEM %
Mid-supraorbital (LHS)	1	7.1 (1.6)	0.962	0.3	7.1	4.2
	3	7.2 (1.4)				
Mid-supraorbital (RHS)	1	7.0 (1.6)	0.906	0.4	7.0	6.3
	3	7.1 (1.2)				
Mid-infraorbital (LHS)	1	7.0 (2.4)	0.959	0.5	6.9	7.1
	3	6.8 (2.3)				
Mid-infraorbital (RHS)	1	6.7 (2.2)	0.971	0.4	6.6	5.8

	3	6.5 (2.2)				
Mid-zygomatic (LHS)	1	9.6 (2.5)	0.977	0.4	9.4	4.1
	3	9.3 (2.4)				
Mid-zygomatic (RHS)	1	9.3 (2.4)	0.965	0.5	9.2	4.9
	3	9.2 (2.4)				
Zygion (LHS)	1	9.3 (2.6)	0.633	1.6	9.7	16.9
	3	10.0 (2.6)				
Zygion (RHS)	1	9.2 (2.8)	0.638	1.7	9.5	18.1
	3	9.8 (2.8)				
Mid-mandibular (LHS)	1	13.9 (5.7)	0.854	2.3	14.3	16.1
	3	14.8 (6.0)				
Mid-mandibular (RHS)	1	14.5 (7.0)	0.834	3.0	15.3	19.2
	3	16.0 (6.9)				
Mid-ramus (LHS)	1	22.9 (3.9)	0.966	0.8	22.8	3.4
	3	22.6 (4.4)				
Mid-ramus (RHS)	1	22.2 (4.5)	0.991	0.4	22.1	2.0
	3	22.0 (4.3)				
Gonion (LHS)	1	23.8 (7.3)	0.693	4.1	23.0	17.9
	3	22.2 (7.1)				
Gonion (RHS)	1	24.8 (7.0)	0.785	3.0	24.6	12.3
	3	24.4 (5.7)				
Supra M_2 (LHS)	1	33.0 (4.8)	0.686	3.1	31.9	9.8
	3	30.8 (5.7)				
Supra M_2 (RHS)	1	30.7 (5.2)	0.626	3.4	30.2	11.3
	3	29.7 (5.7)				
Sub M_2 (LHS)	1	22.6 (3.9)	0.651	2.9	23.1	12.5
	3	23.7 (5.5)				
Sub M_2 (RHS)	1	22.0 (3.6)	0.646	2.6	22.2	11.7
	3	22.3 (4.9)				
Mental tubercle (LHS)	1	9.2 (1.5)	0.821	0.8	9.1	8.6
	3	8.8 (2.0)				
Mental tubercle (RHS)	1	9.4 (1.5)	0.715	1.0	9.1	11.3
	3	8.8 (2.2)				

4.5. Morphological analysis

4.5.1. Facial approximations

The FSTTs recorded in the first part of this study were used to perform four facial approximations. An additional two approximations were performed using raw pooled data from the C-Table repository sourced from http://www.craniofacial identification.com. The purpose for producing approximations using the FSTT data from the C-Table was to assess the difference in appearance of the approximations using the C-Table and the approximations using FSTT from this study. All the approximations were photographed and presented in frontal, oblique and lateral views.

According to the forensic facial artist who prepared the facial approximations; the principles proposed by Krogmann in 1946, related to the position of the eyeball to the orbit, the positioning and dimensions of the ear, the shape of the tip if the nose and the mouth width remained constant across all the facial approximations (Gupta *et al*., 2015). This was necessary in order to test for the influence of the varying FSTTs with regards to recognisability.

Figures 4.14 (A-C) & Figures 4.14 (D-F) represents male facial approximations achieved using sex-specific and pooled FSTT data respectively which were collected in this study. Figures 4.14 (G-I) are that of males facial approximations using the C-Table repository. Similarly, Figures 4.15 (A-C) & Figures 4.15 (D-F) are female facial approximations using sex-specific and pooled FSTT data respectively which collected in this study. Figures 4.15 (G-I) were created using the C-Table repository.

By visual inspection of the frontal view of the male approximations there is a slight, but noticeable difference in the shape of the jaw. This is most apparent when comparing the approximations of this study to the approximation achieved using the

C-Table's raw data. The jaw line in Figures 4.14 (A & D) was broad but narrower and more gracile in Figure 4.14 (G). In the oblique views of the male approximations the parotid area is more pronounced in Figures 4.14 (B & E) and bulges inferiorly in comparison to the same area in Figure 4.14 (H). This area overlies the mid-ramus landmark. The mean value in this study, however, does fall within the range for the mid-ramus as recorded in Stephan & Simpson (2008). In the lateral views of the male approximations, the mouth is more prognathic when the C-Table data (Figure 4.14.I) was employed compared to FSTT collected in this study (Figures 4.14 C & F). Alternatively, the labiomental sulcus was more concave in appearance in the approximations whereby the FSTTs of this study (Figures 4.14 C & F) were applied.

On analysis of the female approximations, similar differences as noted in the male approximations are found to exist, but to a lesser degree. The jaw line is broader in the approximations created from sex-specific (Figure 4.15.A) and pooled (Figure 4.15.D) FSTT data collected in this study compared to the C-Table approximation (Figure 4.15.G). When observing the facial approximation from an oblique angle, the parotid area is slightly more enlarged in Figures 4.15 B & E and is noticeable as more of the neck muscle (sternocleidomastoid) is visible in Figure 4.15.H. Once again, the approximation created using the C-Table raw data (Figure 4.15.I) presents a more prognathic profile compared to the approximations created using this study's FSTT data (Figures 4.15 C & F).

The appearance of the male and female facial approximations tends to become more gracile as the FSTT sample size increases (Figures 4.14 & 4.15 A-F *vs*. Figures 4.14 & 4.15 G-I). However, resemblance and recognition of the facial approximations appears to be uncompromised.

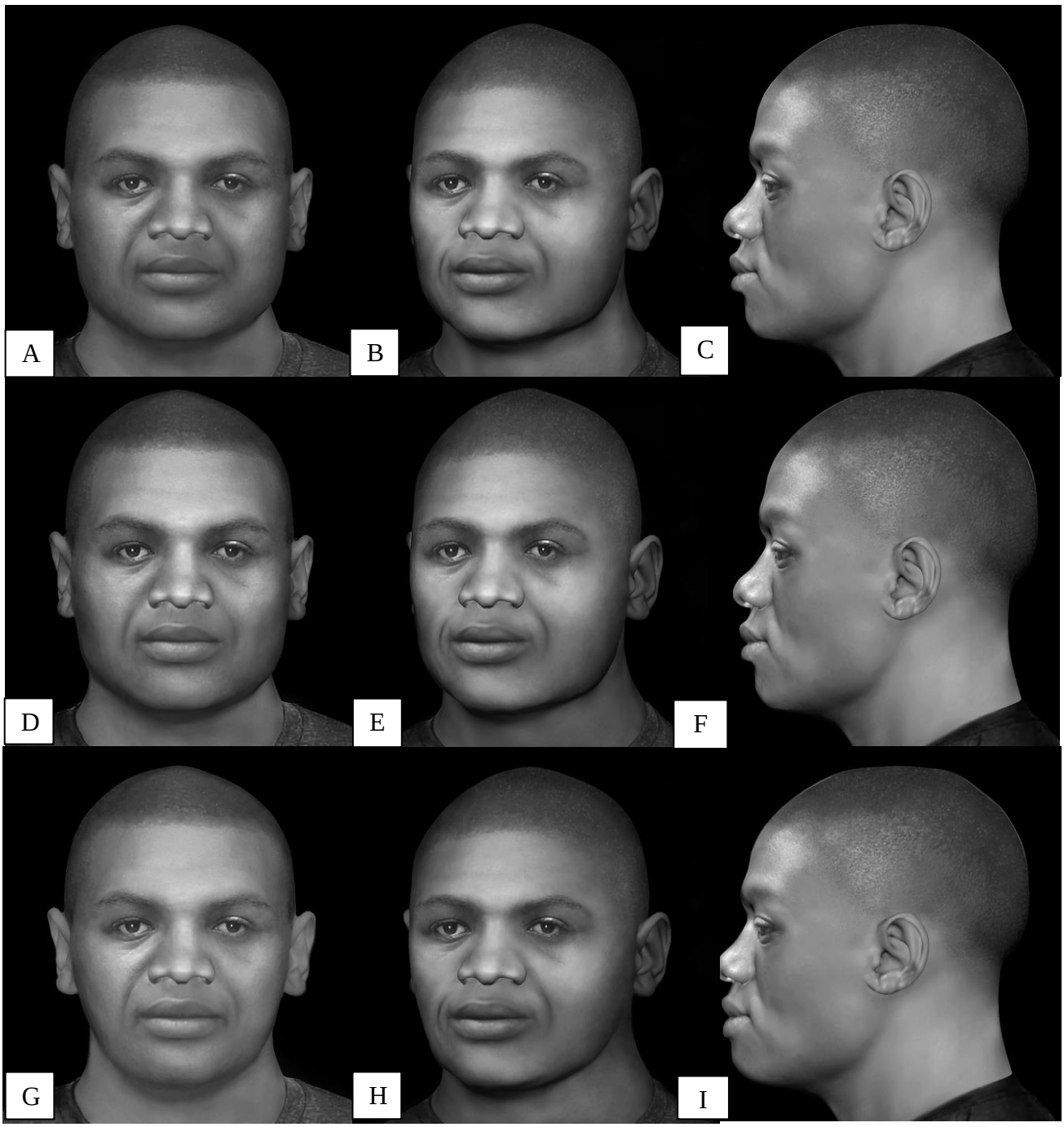


Figure 4.14: Frontal, oblique and lateral artistic impressions of males facial approximations utilitizing sex-specific FSTT (A-C) and pooled FSTT data (D-F) from the current study; and raw pooled data (G-I) from the C-Table repository (http://www.craniofacialidentification.com).

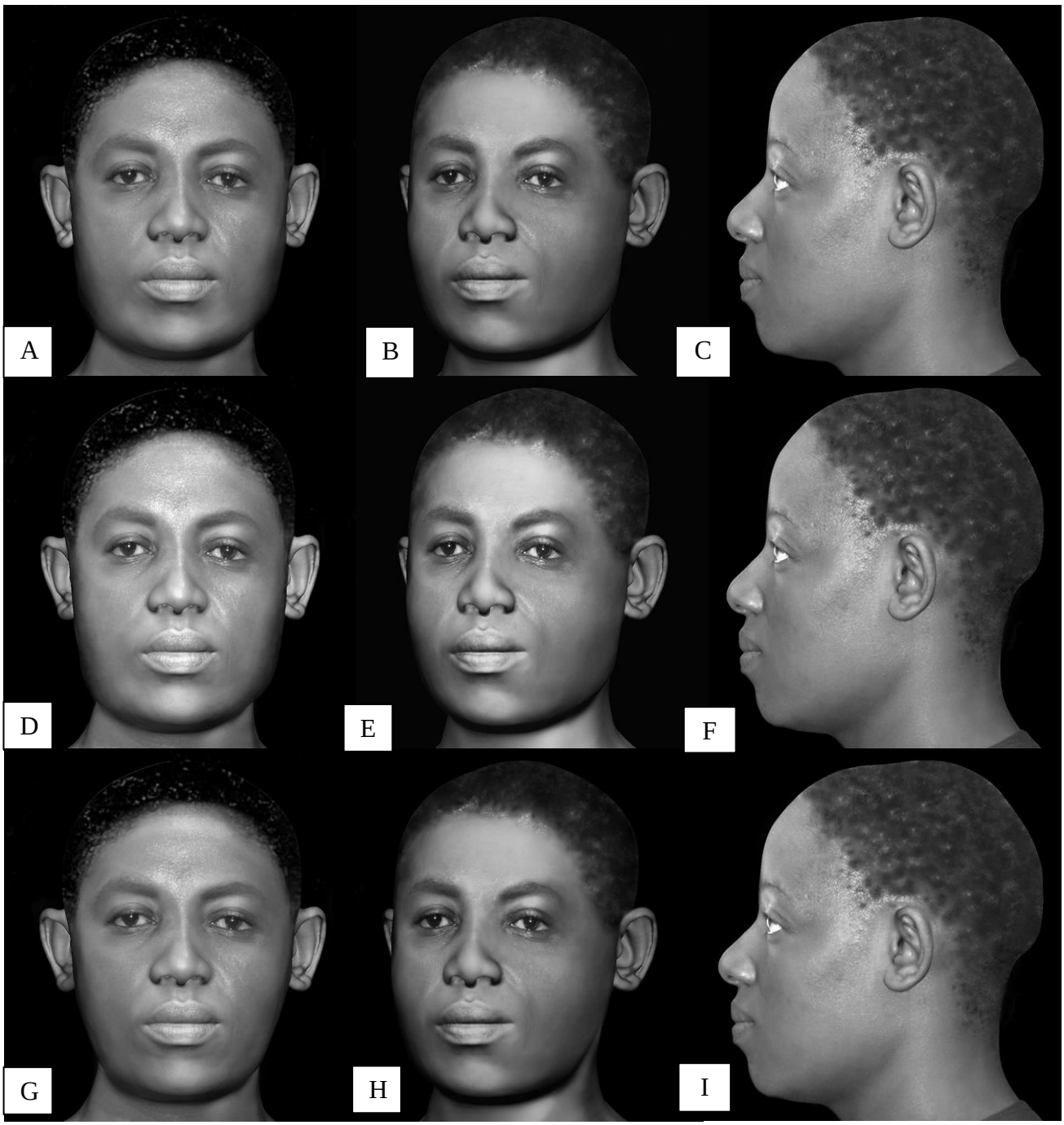


Figure 4.15: Frontal, oblique and lateral artistic impressions of females facial approximations utilitizing sex-specific FSTT (A-C) and pooled FSTT data (D-F) from the current study; and raw pooled data (G-I) from the C-Table repository (http://www.craniofacialidentification.com)

4.5.2. Facial arrays

Eighty volunteers participated as assessors in the facial arrays and their opinions were categorized according to sex, ancestry and experience in the fields of anatomy and/or art associated with facial proportions. The assessors were categorized according to sex, population affinity and education. As such, there were 40 males and 40 females, 40 white individuals and 40 black individuals, and 40 individuals with training in anatomy and/or art and 40 individuals with no training. Furthermore, there were eight sub-categories consisting of 10 individuals each; namely, 1) white females with training in anatomy/art, 2) white females with no training in anatomy/art, 3) white males with training in anatomy/art, 4) white males with no training n anatomy/art, 5) black females with training in anatomy/art, 6) black females with no training in anatomy/art, 7) black males with training in anatomy/art, and 8) black males with no training in anatomy/art. Figures 4.16 and 4.17 illustrate the matching photographs for their respective facial approximations. The graphs presented in Figures 4.18 – 4.21, illustrate the frequencies in which correct matches were made in accordance to different categories of assessors. Alternatively, Figures 4.22 to 4.33 demonstrate the selection frequency not only for correct photograph but all which were displayed in the facial arrays. There were four facial approximations consisting of two male and two female approximations. These were further subcategorized depending on the data used to produce them (i.e. sex-specific and pooled). Thus subject A was a female approximation created from sex-specific (female) data. Subject B was a male approximation created from sex-specific (male) data. Subject C and D were created from pooled data and comprised a female and male approximation respectively. Subject A and C correspond to photograph number five and, subject B and D correspond to photograph number seven.

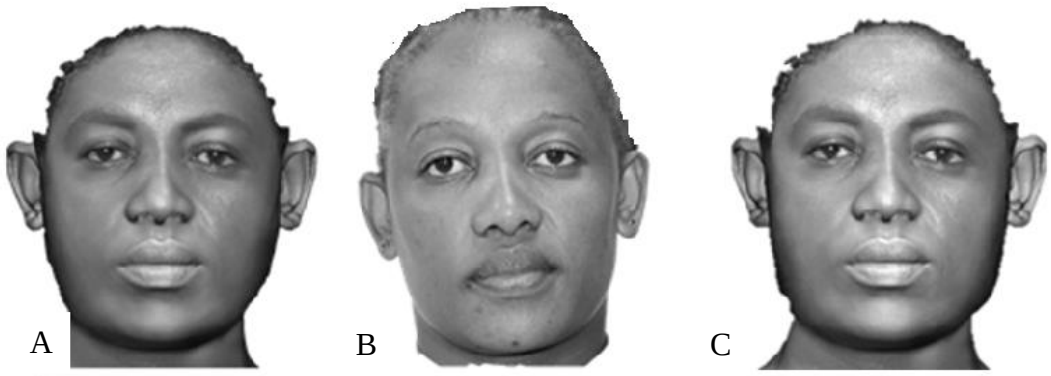


Figure 4.16: Comparison of female facial approximations (A: sex-specific data & C: pooled data) to the positive match photograph (B).

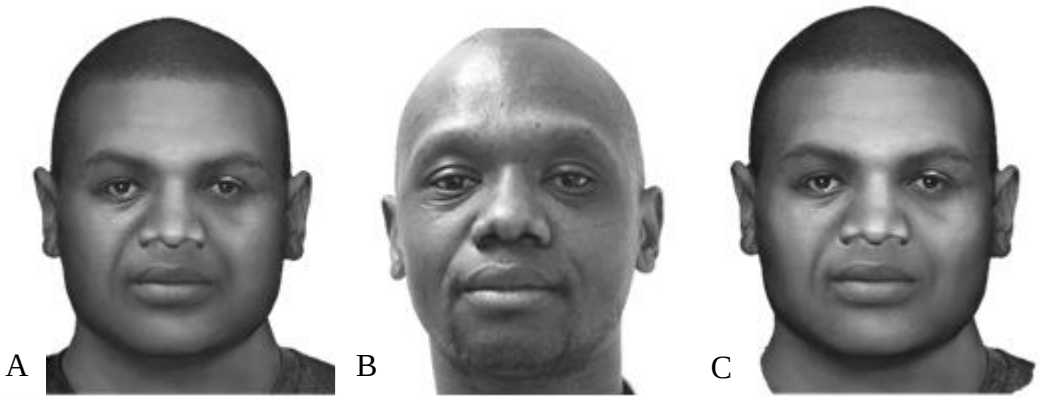


Figure 4.17: Comparison of male facial approximations (A: sex-specific data & C: pooled data) to the positive match photograph (B).

Finally, a horizontal dotted line was placed on each graph. The line demarcates the 10% chance of selecting each photograph and elucidates how often each category of assessors had selected a certain photograph above or below chance.

Analysis of the overall sample (Figure 4.18) indicated no observable significant differences (t-test, $p < 0.05$) between the frequencies in which each approximation was correctly matched. However, the correct photograph was matched with the females approximations approximately 30% above chance compared to the male facial approximations. Subject C (the approximation produced using pooled data) was

correctly matched 5% more often than subject A (the approximation produced using sex-specific data). However, all the approximations were matched correctly 25 – 60% above chance (indicated by the black dotted line located at 10%).

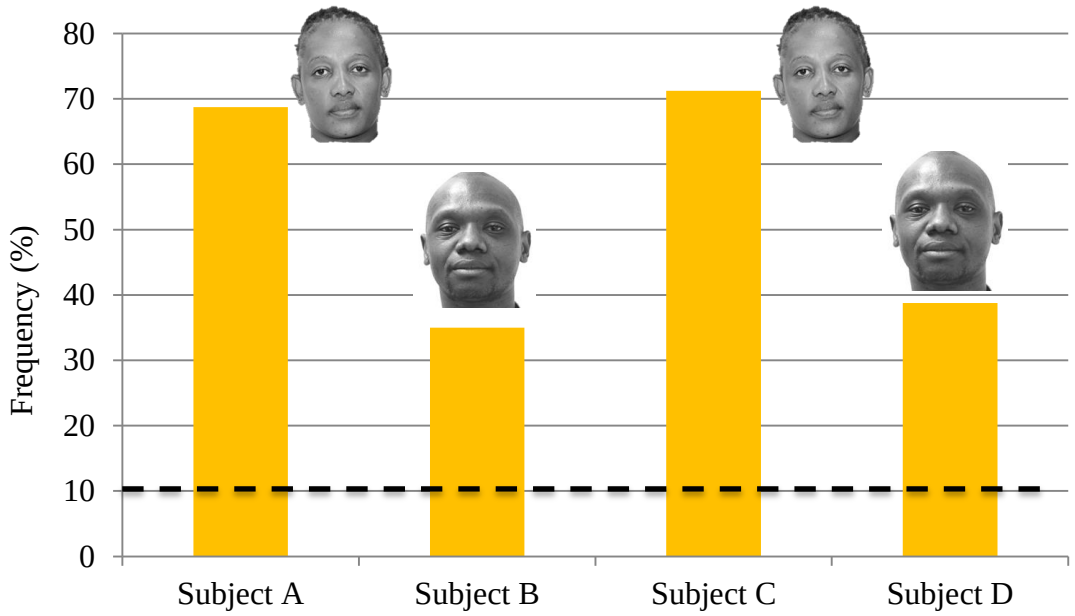


Figure 4.18: Histogram depicting the percentage of correct identification of the facial arrays in the entire sample. The dashed line represents the 10% chance of selecting the correct photograph per facial approximation.

With respect to ancestry, both groups (black and white) were able to identify the correct photograph corresponding to the respective facial approximation above chance. The dotted line positioned on the 10% in Figure 4.19 demarcates the level of chance and the correct pairing was made 10 – 65% above chance. A significant difference (t-test, $p < 0.05$) was found to exist for only one of the facial approximations (subject B).

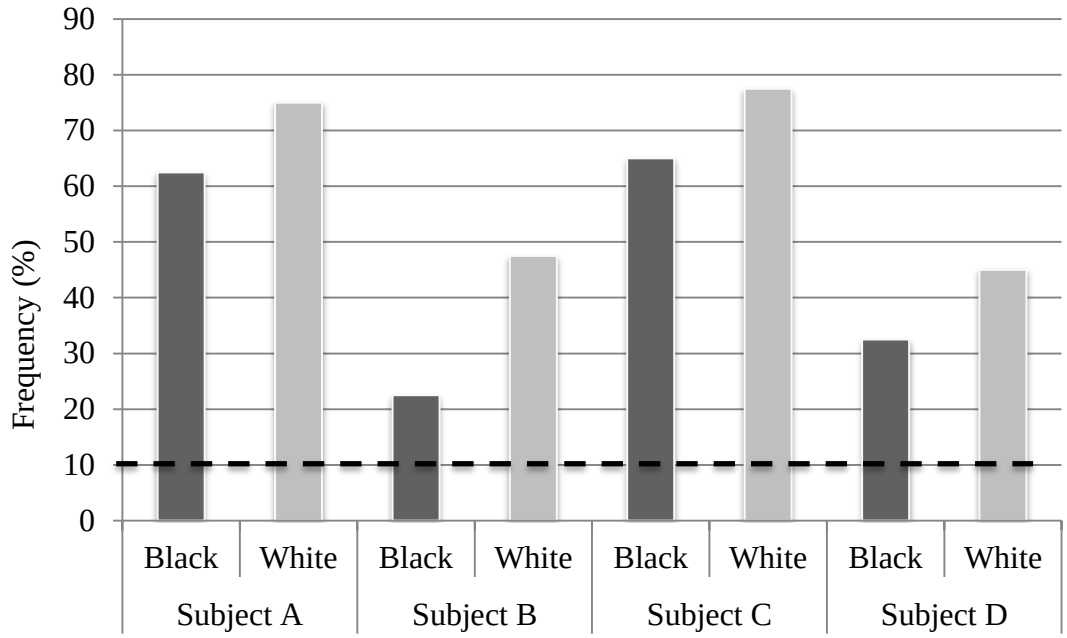


Figure 4.19: Histogram depicting the percentage of correct identification of the facial array in different populations. The dashed line represents the 10% chance of selecting the correct photograph per facial approximation.

Females generally performed better at the facial arrays compared to males. This discrepancy was most noticeable with regards to the male approximations (subject B & D), as female assessors selected the correct photograph matching these facial approximations more often than the male assessors (t-test, $p < 0.05$). The black dotted line at the 10% mark in Figure 4.20 indicates the cut off for chance.

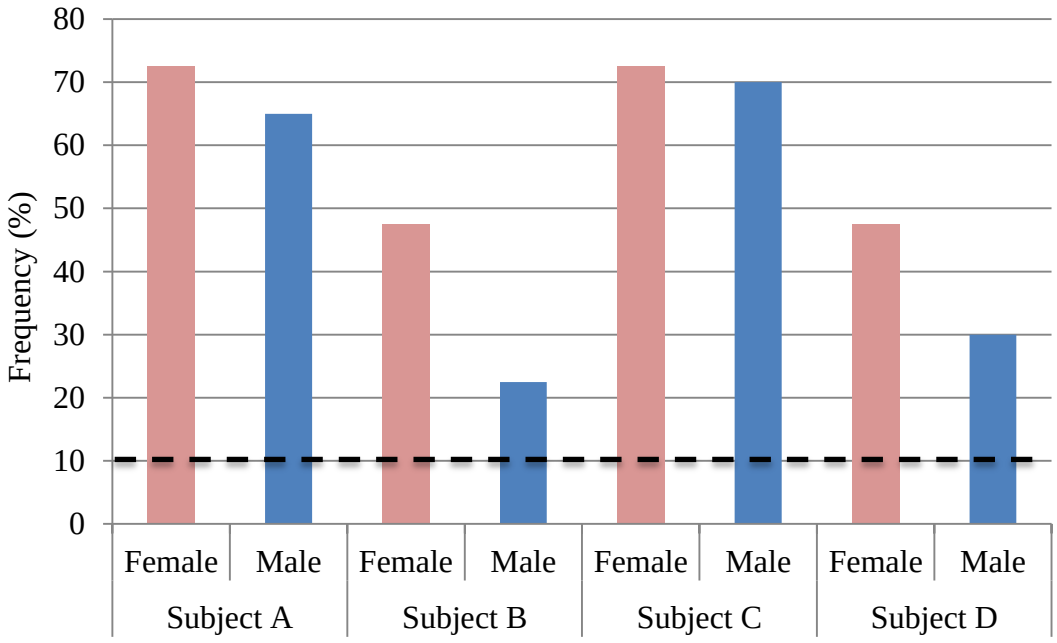


Figure 4.20: Histogram depicting the percentage of correct identification of the facial arrays by females and males. The dashed line represents the 10% chance of selecting the correct photograph per facial approximation.

Persons who had some formal training in anatomy and/or art did better in three of the four approximations. However, the frequencies for correct identification between the two categories of assessors were calculated to be insignificant (t-test, $p < 0.05$). Once again the sex-specific male facial approximation was identified the least (42.5% by females and 27.5% by males). The sex-specific male facial approximation was, however, correctly matched with the corresponding photographs 32.5% and 17.5% above chance by those assessors who had formal training in anatomy and/or art and those who stated they did not, respectively (Figure 4.21). Overall, the correct photographs were more frequently selected when the facial approximations being assessed were derived from the pooled FSTT data of this study.

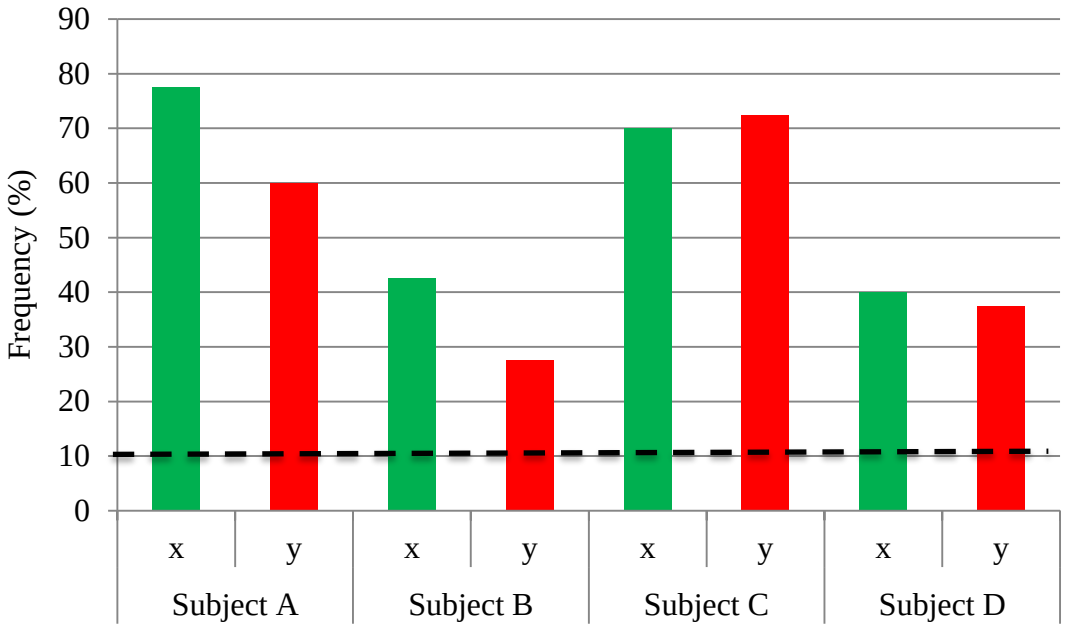


Figure 4.21: Histogram depicting the percentage of correct identification of the facial arrays among people whom had (x), and did not have formal training in anatomy and/or art (y).

Figure 4.22 to Figure 4.25 represents the selection frequency of all the photographs included in the facial arrays for subjects A – D amongst the assessors who identified as either black or white. Photograph number five was the positive match for both the facial approximations named subject A and C. The graphs of subject A (Figure 4.22) and subject C (Figure 4.24) demonstrates the high level of confidence in which the correct photograph was selected. There were no significant differences between the population groups when analysing selection for the correct photographs for subjects A – D. The correct selection was however, not made as frequently in the facial arrays of facial approximations B – D. Photographs number seven was the positive match to both subject B and D. Unlike in Figures 4.22 and 4.24, Figures 4.23 and 4.25 depicts that several of the negative matches were selected two to three the times above chance.

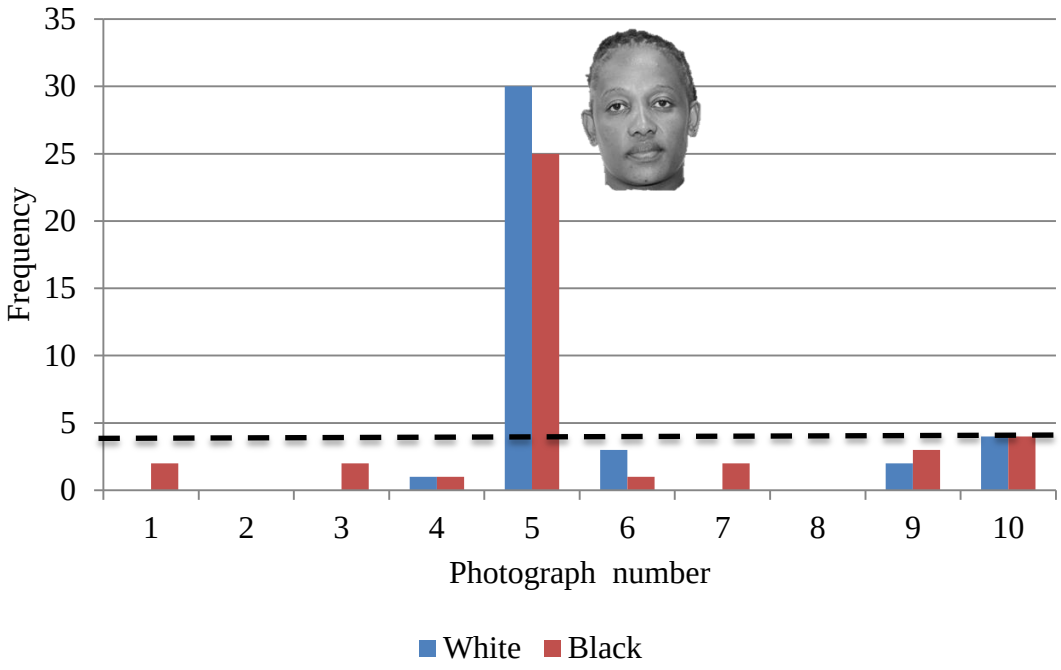


Figure 4.22: Histogram depicting the recognition frequency of the facial array photographs against the facial approximation Subject A by black and white assessors

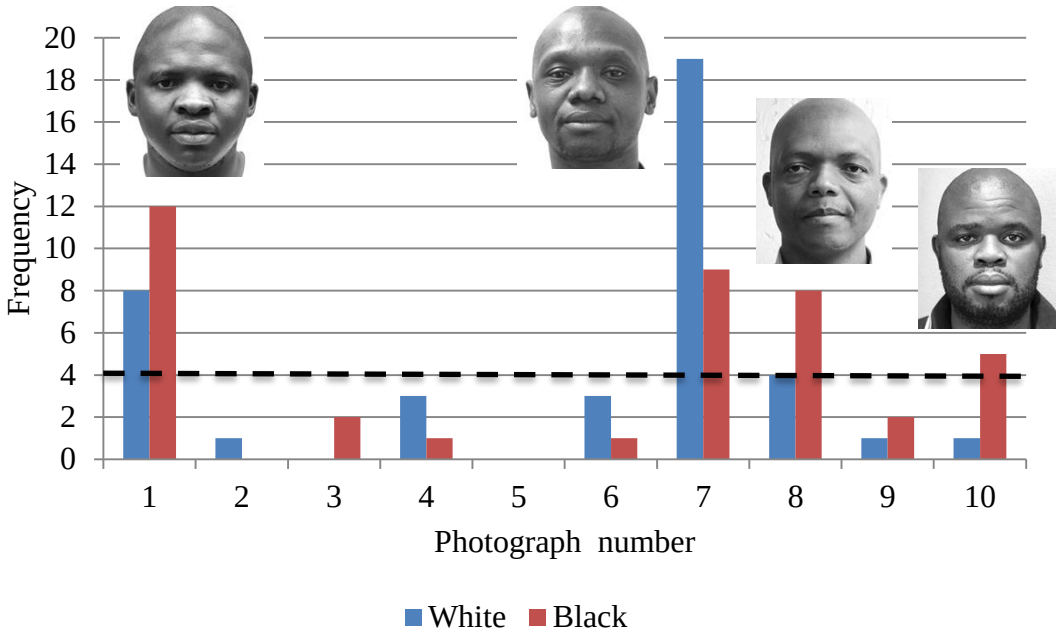


Figure 4.23: Histogram depicting the recognition frequency of the facial array photographs against the facial approximation Subject B by black and white assessors

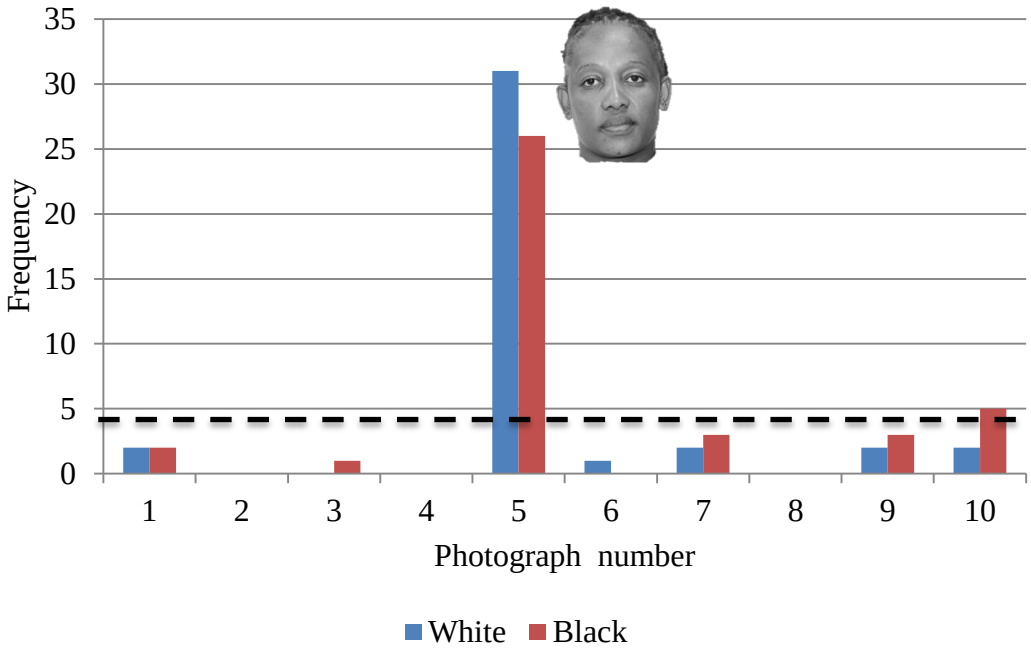


Figure 4.24: Histogram depicting the recognition frequency of the facial array photographs against the facial approximation Subject C by black and white assessors

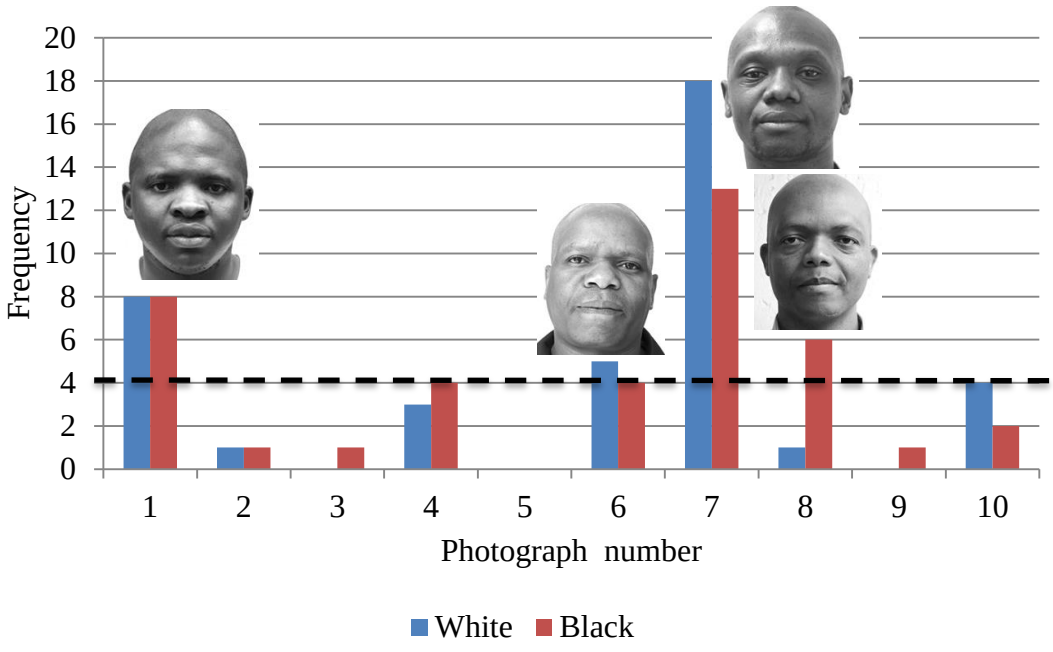


Figure 4.25: Histogram depicting the recognition frequency of the facial array photographs against the facial approximation Subject D by black and white assessors

A similar trend in frequency was evident when analysing the results across the sexes (Figures 4.26 – 4.29). Both males and females correctly matched the female approximations (subject A and C) to the corresponding photographs more frequently compared to the facial arrays pertaining to male approximations (subject B and D). Females however, performed better than the males even in the male approximation facial arrays (subject B and D). The correct matches were made above chance in both sexes for all four arrays but there was greater dispersion of selections amongst the male assessors. Furthermore, females performed significantly better than males for both male facial array tests (subject B and D) (t-test, $p < 0.05$).

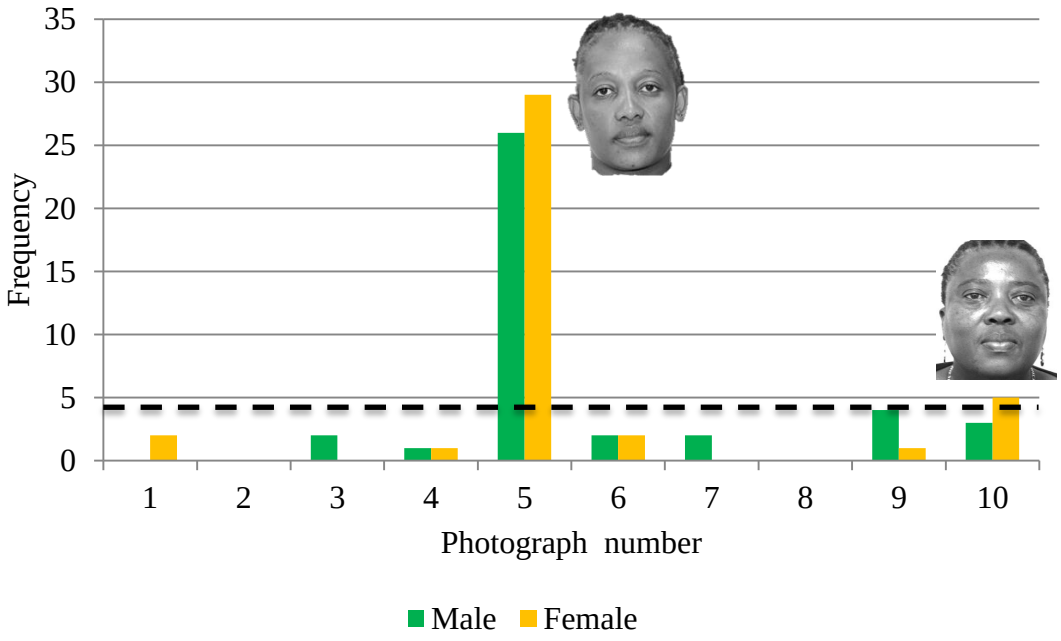


Figure 4.26: Histogram depicting the recognition frequency of the facial array photographs against the facial approximation Subject A by female and male assessors

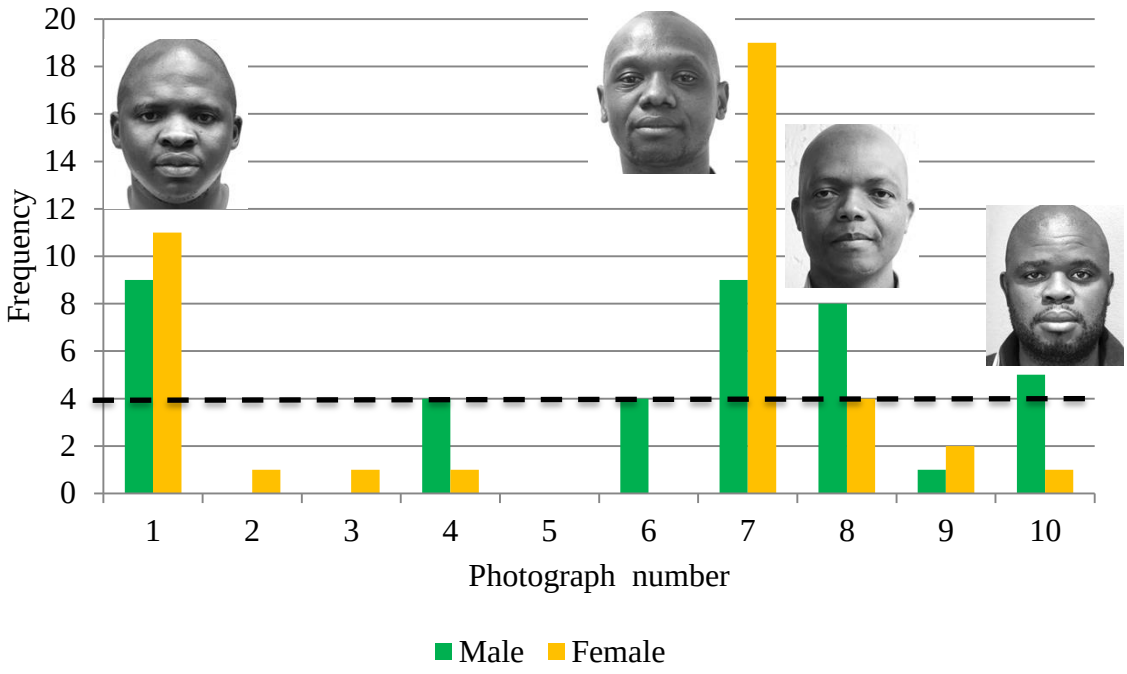


Figure 4.27: Histogram depicting the recognition frequency of the facial array photographs against the facial approximation Subject B by female and male assessors

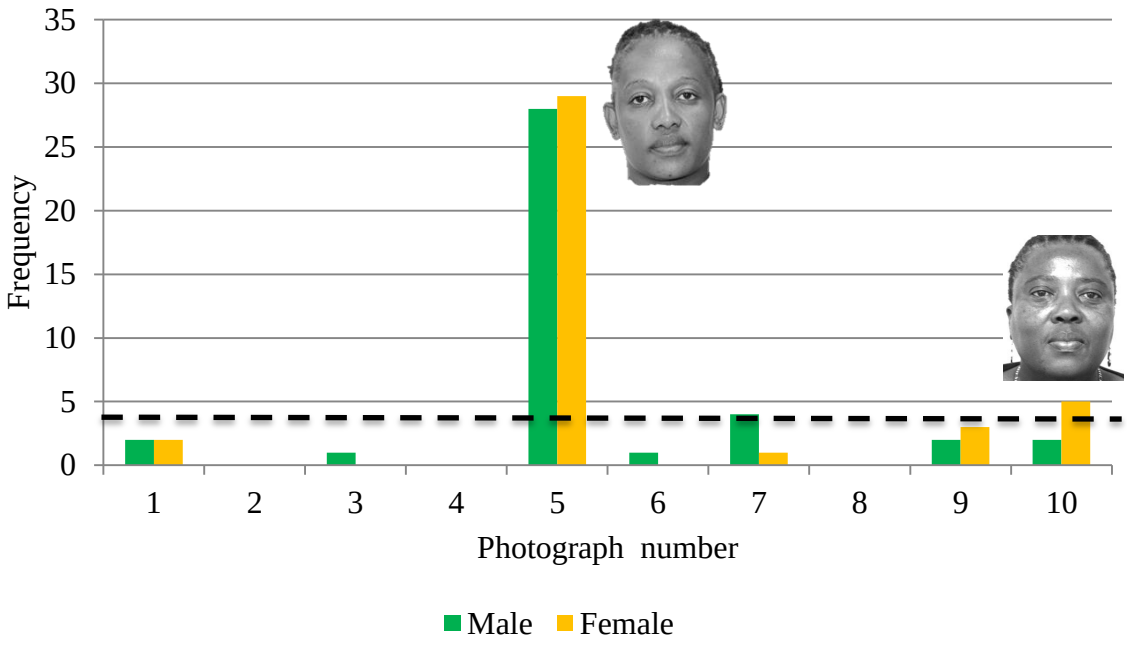


Figure 4.28: Histogram depicting the recognition frequency of the facial array photographs against the facial approximation Subject C by female and male assessors

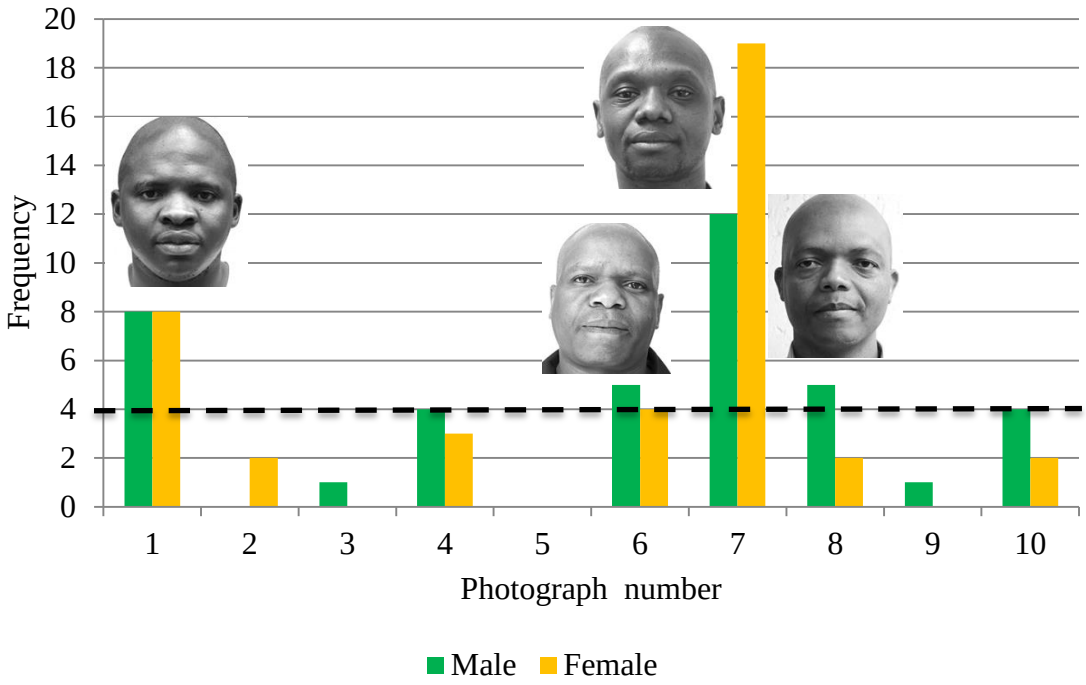


Figure 4.29: Histogram depicting the recognition frequency of the facial array photographs against the facial approximation Subject D by female and male assessors

The assessors that had formal training in anatomy and/or art performed better at the facial arrays than their counterparts but not significantly different (t-test, $p < 0.05$). Figures 4.30 – 4.33 illustrates the frequencies and dispersions of the selections made. In the facial arrays for the female approximations (subject A and C) both groups of assessors selected the correct photograph 50 – 70% above chance. With regards to the facial arrays which comprised the male approximations (subject B and D), both groups selected it above chance, but there is indication that another face was also very favourable amongst the assessors (photograph number 1).

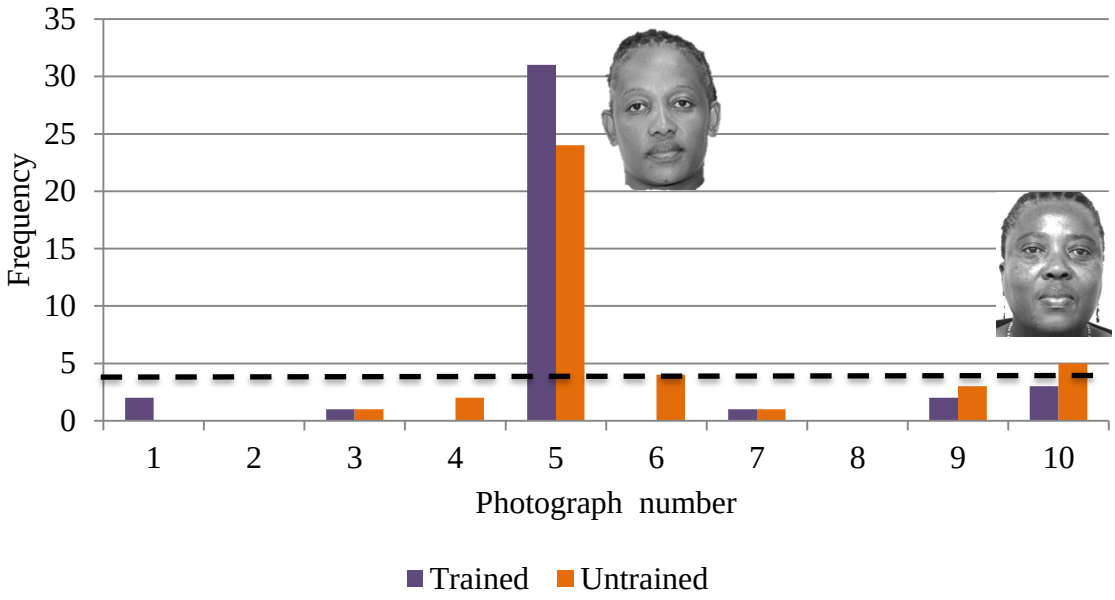


Figure 4.30: Histogram depicting the recognition frequency of the facial array photographs against the facial approximation Subject A by assessors with and without training in anatomy and/or art

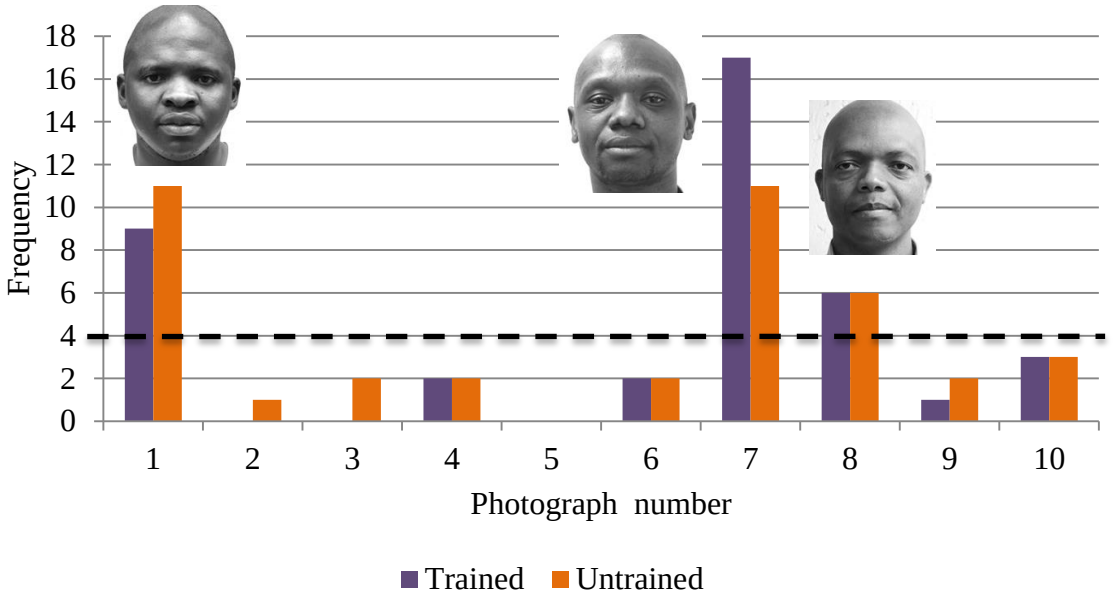


Figure 4.31: Histogram depicting the recognition frequency of the facial array photographs against the facial approximation Subject B by assessors with and without training in anatomy and/or art

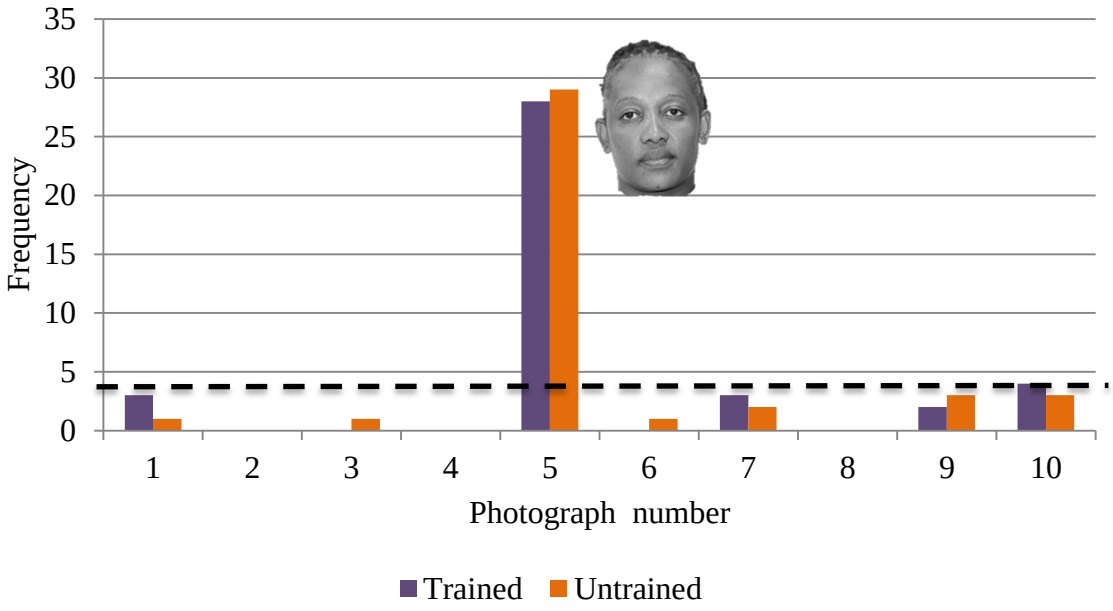


Figure 4.32: Histogram depicting the recognition frequency of the facial array photographs against the facial approximation Subject C by assessors with and without training in anatomy and/or art

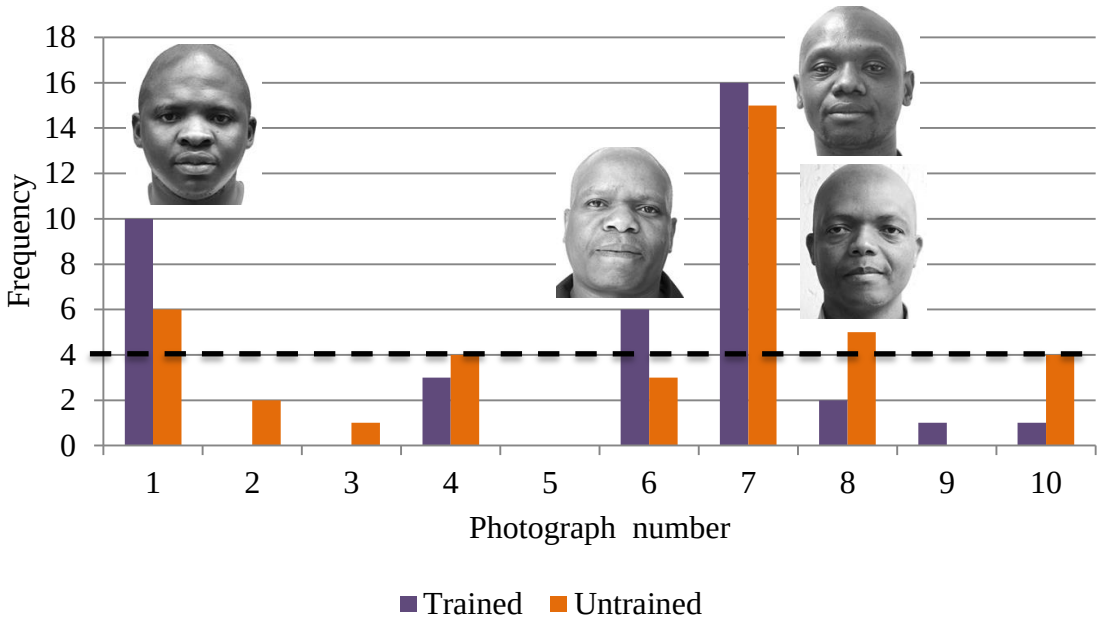


Figure 4.33: Histogram depicting the recognition frequency of the facial array photographs against the facial approximation Subject D by assessors with and without training in anatomy and/or art

4.5.3. Facial features

The second part of the facial arrays required the assessors to select facial feature(s) they thought to be influential in their decision making process. As not every assessor had matched the correct photographs to the facial approximations, this section is only concerned with the comments made by the assessors who paired the facial approximations with the correct photographs. The assessors whom had made the incorrect selections had however selected features similar to those discussed below. None of the assessors had commented on the FSTTs of the facial approximations. The features selected by the all assessors comprised those mentioned in Krogmann's Rule of Thumbs.

Overall, the most frequently indicated features amongst the assessors who correctly paired the positive photographs to the facial approximations were the eyes, lips, nose, jawline and ears (Figure 4.34). The eyes were the most frequently indicated facial feature in all four of the facial arrays (subject A to D). This feature was listed 32.7% in facial array A (appendices 2.1.1 – 2.1.3), 32.1% in facial array B (appendices 2.2.1 – 2.2.3), 28.1% in facial array C (appendices 2.3.1 – 2.3.3), and 35.5% in facial array D (appendices 2.4.1 – 2.4.3). Features listed as other included the hairline and the cheek bones but comprised only 4.1% of the results.

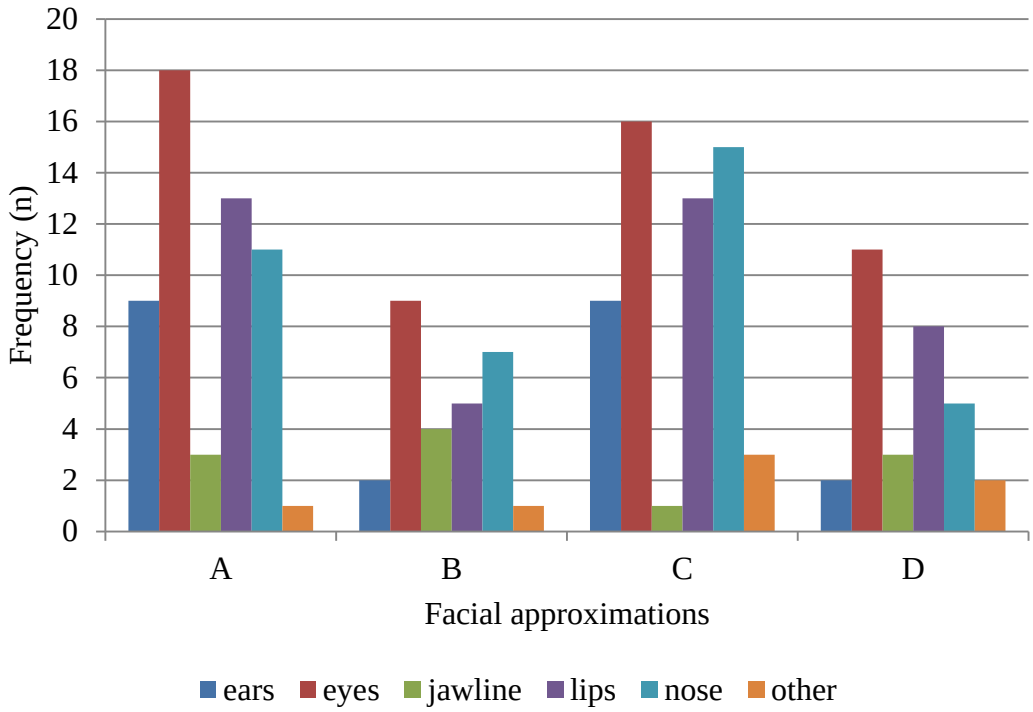


Figure 4.34: Histogram depicting the selection frequency of facial features per facial approximation (A-D) by the entire sample of assessors of the facial arrays

When the identified features were categorized according to population groups (Figure 4.35) the eyes were most frequently selected. In facial array A, the black population listed the eyes 44% of the time, whereas, the white population most frequently selected the nose (33%) as being more influential in their decision making process. In facial array B, the jawline (33%) and eyes (42.1%) were listed most often by the black and white populations respectively. In facial array C, black and white populations favoured the nose (26.9%) and eyes (32.3%). And, in facial array D, the eyes were selected by both populations (black: 23.1%, white: 44.4%). Additionally, the black population listed the lips at the same frequency as they did the eyes.

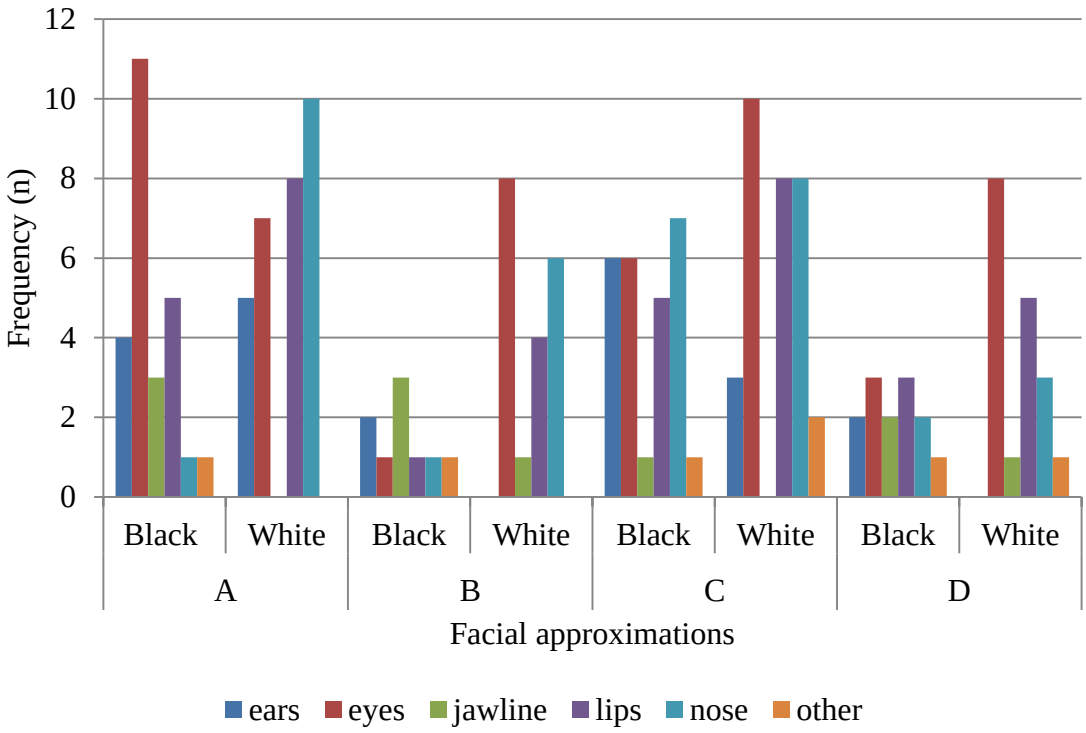


Figure 4.35: Histogram depicting the selection frequency of facial features per facial approximation (A-D) by persons of different population groups

Amongst the sexes (Figure 4.36), the eyes were listed most regularly in all the four of the facial arrays. The opinions of the female assessors stated this with frequencies of 37.9%, 36.8%, 27.6%, and 36.8% for facial arrays A to D respectively.

Similarly, the eyes were listed by the male assessors with frequencies of 26.9%, 22.2%, 28.6% and 33.3% for facial arrays A to D respectively. Other frequently mentioned features made by the male assessors included the lips and nose (both 22.2%) in facial array B. More so, males recognised the nose as significant with the identical frequency as the eyes in facial array C.

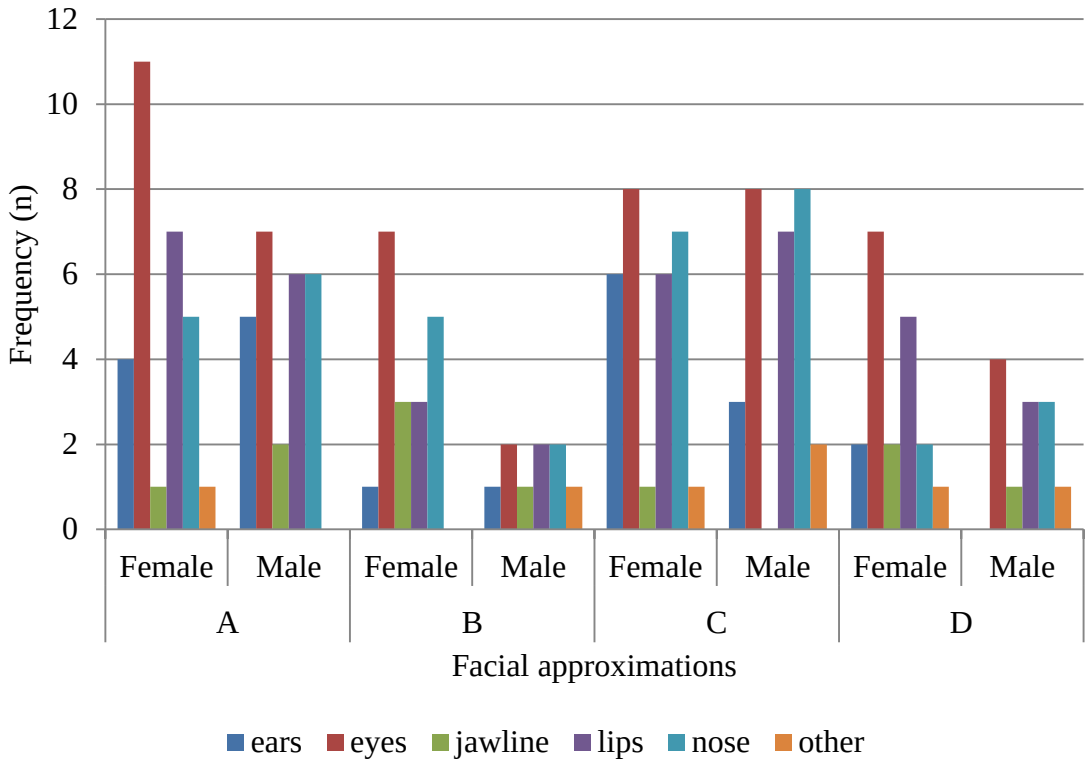


Figure 4.36: Histogram depicting the selection frequency of facial features per facial approximation (A-D) by persons of different sex

When consideration was given to the comments according to the educational background of the assessors (Figure 4.37), once again the eyes were most commonly recorded. Those assessors whom had stated they had received formal education in anatomy and/or art listed the eyes with frequencies of 32.3%, 29.4%, 32.1% and 31.3% in facial arrays A to D respectively. However, in facial array B, the nose was more frequent at 29.4%.

The other group of assessors, those with no formal training in anatomy and/or art also listed the eyes and nose most regularly. In facial arrays A, B and D the eyes yielded frequencies of 33.3%, 33.3% and 40% respectively. In facial array C, the nose was listed 24.1% of the time.

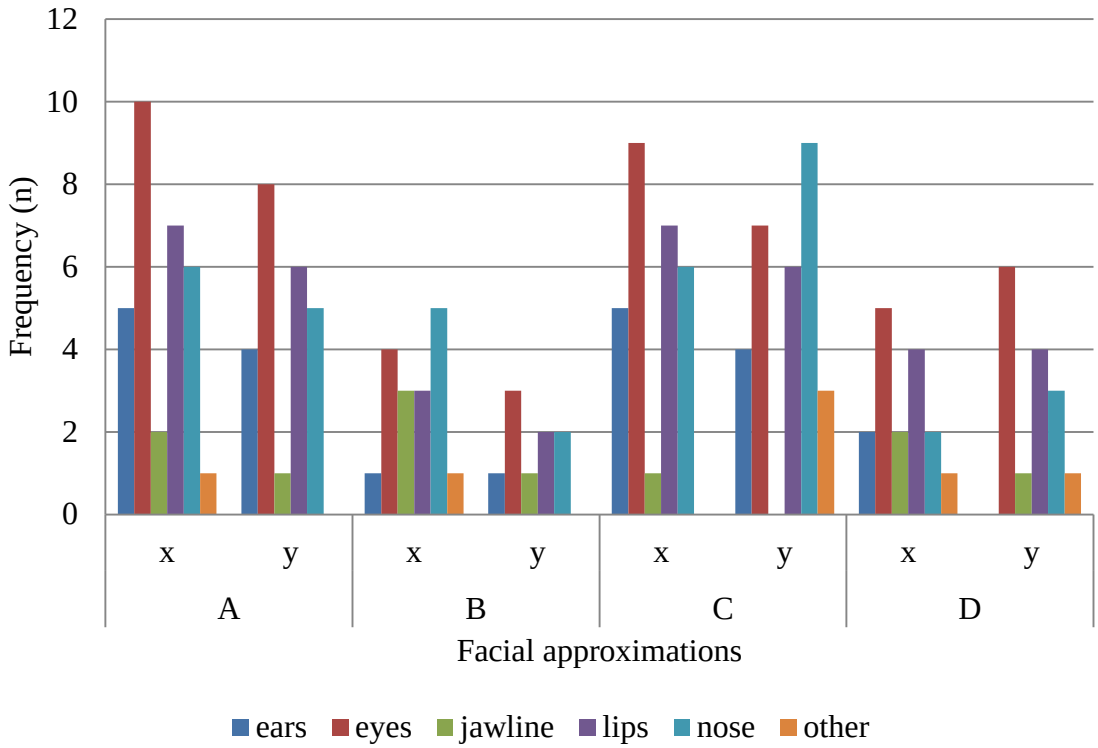


Figure 4.37: Histogram depicting the selection frequency of facial features per facial approximation (A-D) amongst people whom had (x) and did not have (y) formal training in anatomy and/or art

The results of the current facial arrays illustrated that the assessors relied heavily on features unrelated to FSTTs. Features such as the eyes, lips and nose were listed as the most similar features between the selected photographs and facial approximations. The aforementioned was true for the assessors who made the correct and incorrect selections. Therefore, the practice of categorizing FSTTs is not required as stated by Stephan *et al*. (2016). Moreover, the pooling of FSTTs prevents the misinterpretation of significant differences due to sampling and methodological errors. Furthermore, the facial approximations produced utilizing the pooled FSTT data were identified correctly more frequently than the facial approximations produced *via* sex-specific FSTT data.

Chapter 5: Discussion

5.1. Introduction

This study had to two main purposes. Firstly, facial soft tissue thicknesses (FSTTs) were measured at 30 anatomical landmarks (10 midline and 10 bilateral) from 60 retrospectively collected computed tomography (CT) scans of black South African adults in order to establish FSTT Tables according to categories of sex and age.

The second aspect of this study was the production of four facial approximations and recognition by means of facial array tests. The Manchester Method was chosen for this study as the artist who performed the four approximations was trained using this methodology. Four facial approximations were produced on two additional CT scans (one females and one male) that did not form part of the FSTT dataset of the current study. The approximations were performed using the current study's pooled FSTT and sex-specific FSTT data. There were 80 assessors who participated in the facial arrays. All the assessors were invited to participate based on their sex, ancestry and previous training or education in anatomy and/or facial artistry. Additionally, two more facial approximations (one female and one male) were performed using the C-Table raw data repository developed by Professor Carl Stephan of the University of Queensland, Australia (http://www.craniofacialidentification.com). The purpose of this was to provide a visual representation in order to qualitatively discuss the perceived differences between the approximations of this study and that of the approximations achieved using all the currently available pooled FSTT data collected from around the world. In this chapter, the results of the FSTTs and facial arrays of the current study will be compared to previously published literature. Finally, the challenges encountered in all the facets of this study will be focused upon.

5.2. Facial soft tissue thickness measurements

The FSTT data presented in this study were retrospectively collected from CT scans of patients at a Johannesburg hospital. The sample consisted of 60 living black adult South African adults.

The midline landmarks that revealed the smallest FSTT included the glabella, nasion, mid-nasal, rhinion, and menton. These are all midline line landmarks that are in close relation to the underlying bone. These landmarks did not display significance differences between males and females, but FSTTs at three midline landmarks (i.e. glabella, nasion and rhinion) were found to be statistically different with regards to progressive age. Furthermore, significant age-related differences were found at the labial inferius and labiomental sulcus. The bilateral landmarks that showed the smallest FSTT included the right and left mid-supraorbitals and infraorbitals, both of which were statistically significant with respect to age, but not sex.

The midline landmarks that showed the largest FSTT included the labiomental sulcus and the pogonion, with the former illustrating significant differences with increased age. With respect to the bilateral landmarks, the supra M_2 and sub M_2 on either side yielded the largest FSTTs. Neither of which were significant for sex or age changes but rather with regards to bilateral asymmetry at the supra M_2 only. The mean value of left supra M_2 was 2.4 millimetres larger than the mean on the right side. The FSTTs measured at the left and right mid-zygomatic were 2.4 to 2.7 millimetres larger in females compared to males, and was the only sex-related significance difference. Additionally, age-related differences were identified at the gonion and mental tubercle, both of which increased with the advancement of age.

When compared to the FSTTs of other studies, numerous significant differences (t-test, $p < 0.05$) were found (Table 5.1). Firstly, the results of the entire sample from the current study were compared to results reported by a South African study on black adult females (Cavanagh & Steyn, 2011). The FSTTs measured in the current study illustrated that the midline measurements were smaller than those reported by Cavanagh & Steyn (2011), ranging from 0.3 to 4.6 millimetres. The lateral FSTTs were larger in the current study with an upper range of 3.8 millimetres. Therefore, significant differences were present across 77.8% of midline and 37.5% of lateral landmarks (e.g. glabella, nasion, mid-philtrum, labial superius, labiale inferius, labiomental sulcus, menton, gonion, supra M_2, the mental tubercle). Only comparisons could be made between the left sided data as Cavanagh & Steyn (2011) had not investigated the FSTT on both sides of the face.

The second comparison study was that of Aulsebrook *et al*. (1995). The comparisons were only calculated between midline landmarks due to unclear definitions of the landmarks taken from their oblique radiographs. Their sample investigated the FSTTs in black South African adult males. The FSTTs in the current study were smaller compared to those in Aulsebook *et al*. (1995) and nine of ten midline measurements were statistically different (Table 5.1). The differences in means between the current study and those reported by Auslsebrook *et al*. (1996) ranged from 0.6 mm to 6.1 mm (Table 5.1). These differences in FSTT means could possibly be attributed to Aulsebrook *et al*. (1996) utilizing cephalometric radiography. Cephalograms are lower in resolution compared to CT scans (Hwang *et al*., 2012) and it is possible that the measurements in Aulsebrook *et al*. (1996) were over-estimated.

The FSTTs of the current study was also compared to a review study of published adult data by Stephan & Simpson (2008a). Their review calculated the pooled mean

values from numerous studies across the world regardless of sex-specificity, ancestry or population group, age or BMI.

Eight of the ten midline FSTT means calculated in the current study were significantly different from the pooled means found in Stephan & Simpson (2008a). The FSTT means of the current study were generally smaller than those in Stephan & Simpson (2008a), and the differences between the means had an upper limit of 2.8 millimetres (Table 5.1). However, as the sample size of Stephan & Simpson (2008a) was approximately 10 times larger than that of the current study (Table 5.1), the p-value has been grossly over-estimated. Similarly, the FSTTs were significantly different at 80% of the landmarks between the current study and the study conducted by Cavanagh & Steyn (2011), despite CT scans being utilized in both studies. The sample size of Cavanagh & Steyn (2011) was larger than that of the current study, which possibly resulted in the over-estimation of p-value. The differences in means ranged from 0.1 to 4.5 mm (Table 5.1). More so, the study sample of Cavanagh and Steyn (2011) comprised of females only, whereas the current study contained males and females. Research conducted by Stephan *et al*. (2016) reported that females have larger FSTTs than males. The results of the current study illustrated that females had larger FSTTs than males; therefore, the FSTT means of the entire sample was lowered due to the inclusion of males. As such, the significant differences reported between Cavanagh & Steyn (2011) and the current study may also have been due to the difference the demographics of the samples.

Unlike the studies conducted by Stephan & Simpson (2008a) and Cavanagh & Steyn (2011), the sample size of Aulsebrook *et al* (1996) was similar (n = 55) to that of the current study (n = 60). However, Aulsebrook *et al.* (1996) collected FSTT measurements using cephalometric radiography, a method that has been documented

to over-estimate FSTTs (Hwang *et al*., 2012). This over-estimation is due to worse resolution compared to CT scans utilized in the current study. The differences in means between the current study and those of Aulsebrook *et al*. (1996) ranged from 0.6 to 6.1 mm.

The significant differences reported between the current study and the aforementioned studies of Aulsebrook *et al*. (1996), Cavanagh & Steyn (2011) and Stephan & Simpson (2008a) were a result of differences in sample sizes and/or the methodologies used to collect FSTTs. Furthermore, the current study as well as the aforementioned studies did not consider the effects of BMI which has been documented to influence FSTT (Stephan *et al*., 2016).

Table 5.1.: Comparison of midline FSTT data from the current study with Aulsebrook *et al*. (1996), Cavanagh & Steyn (2011) and Stephan & Simpson (2008a)

Landmarks	Study	n	Mean	SD	p-value
Glabella	Current	51	4.4	1.1	-
	Aulsebrook *et al*.	55	5.8	0.9	**< 0.05**
	Cavanagh & Steyn	152	6.3	1.3	**< 0.05**
	Stephan & Simpson	5791	5.5	1.0	**< 0.05**
Nasion	Current	59	4.6	1.4	-
	Aulsebrook *et al*.	55	7.0	1.1	**< 0.05**
	Cavanagh & Steyn	141	6.0	1.6	**< 0.05**
	Stephan & Simpson	6159	6.5	1.5	**< 0.05**
Mid-nasal	Current	53	2.6	1.1	-
	Aulsebrook *et al*.	55	4.8	1.0	**< 0.05**
	Cavanagh & Steyn	NA	NA	NA	NA
	Stephan & Simpson	1272	4.0	1.0	**< 0.05**
Rhinion	Current	59	2.4	1.0	-
	Aulsebrook *et al*.	55	3.0	0.6	< 0.05
	Cavanagh & Steyn	132	2.7	1.0	> 0.05
	Stephan & Simpson	5511	3.0	1.0	**< 0.05**
Mid-philtrum	Current	54	9.0	1.4	-
	Aulsebrook *et al*.	55	12.1	1.6	**< 0.05**
	Cavanagh & Steyn	138	10.9	1.4	**< 0.05**
	Stephan & Simpson	5508	11.5	2.5	**< 0.05**
Labiale superius	Current	47	9.6	1.8	-
	Aulsebrook *et al*.	55	14.6	2.2	**< 0.05**
	Cavanagh & Steyn	121	13.3	1.8	**< 0.05**
	Stephan & Simpson	5106	11.5	3.0	**< 0.05**
Labiale inferius	Current	54	10.2	2.2	-
	Aulsebrook *et al*.	55	16.3	2.0	**< 0.05**
	Cavanagh & Steyn	95	14.7	1.9	**< 0.05**
	Stephan & Simpson	4886	13.0	2.5	**< 0.05**
Labiomental sulcus	Current	58	11.3	1.9	-
	Aulsebrook *et al*.	55	12.9	1.7	**< 0.05**
	Cavanagh & Steyn	91	12.2	2.0	**< 0.05**
	Stephan & Simpson	5792	11.0	2.0	> 0.05
Pogonion	Current	58	10.7	2.4	-
	Aulsebrook *et al*.	55	11.7	1.8	> 0.05
	Cavanagh & Steyn	64	10.6	1.9	> 0.05
	Stephan & Simpson	6786	11.5	2.5	> 0.05
Menton	Current	59	5.0	2.7	-
	Aulsebrook *et al*.	55	7.3	2.0	**< 0.05**
	Cavanagh & Steyn	42	6.7	1.5	**< 0.05**
	Stephan & Simpson	4475	7.0	2.5	**< 0.05**

Table 5.2.: Comparison of midline and left (L) lateral FSTT data from the current study (A) with Stephan & Preisler (2018) (B), Stephan & Sievwright (2018) (C) and Thiemann *et al*. (2017) (D)

Landmarks	A	B	C	D	A	B	C	D	A	B	C	D
Midline	Mean				Shorth				75-Shormax			
G	4.4	5.6	5.5	5.5	3.6	5.6	5.3	5.1	5.5	6.7	6.2	6.1
N	4.6	5.6	7.4	8.4	4.5	7.1	7.7	8.3	6.4	8.5	8.7	9.6
Mn	2.6	NA	5.4	6.0	2.1	NA	5.6	6.3	3.6	NA	6.5	7.6
Rhi	2.4	2.3	2.3	3.6	1.8	2.4	2.3	3.6	3.3	3.1	2.9	4.2
Mp	9.0	11.3	10.6	14.5	9.0	11.2	10.5	14.0	10.6	12.9	12.7	16.2
Ls	9.6	NA	10.1	NA	9.7	NA	9.6	NA	12.1	NA	11.6	NA
Li	10.2	NA	11.0	NA	9.8	NA	11.0	NA	12.6	NA	12.5	NA
Lms	11.3	11.3	10.1	NA	10.4	11.5	9.2	NA	13.2	13.2	12.0	NA
Pg	10.7	8.5	11.0	NA	10.2	8.3	10.3	NA	13.1	10.6	12.8	NA
M	5.0	7.3	7.4	NA	4.6	7.1	7.0	NA	7.6	8.9	8.9	NA
Lateral	Mean				Shorth				75-Shormax			
L Mso	7.1	7.8	6.8	7.5	6.7	7.9	6.7	7.5	8.7	8.9	7.7	9.2
L Mio	6.5	5.4	4.8	6.0	5.6	5.4	4.6	6.6	8.9	6.4	6.1	7.2
L Mz	8.9	NA	NA	7.9	9.5	NA	NA	7.4	11.5	NA	NA	8.8
L Zy	8.4	8.3	9.0	7.6	8.1	7.8	8.2	7.2	11.4	9.6	10.7	9.4
L Mmb	12.6	8.7	10.1	NA	11.9	8.2	9.9	NA	18.0	10.9	13.6	NA
L Mr	21.0	20.1	20.2	NA	21.7	22.1	20.0	NA	26.6	24.4	23.4	NA
L Go	21.4	10.3	7.8	NA	22.0	8.5	6.7	NA	28.5	13.0	10.3	NA
L Sup	29.8	NA	NA	30.7	27.5	NA	NA	32.2	35.7	NA	NA	34.0
L Sub	21.6	NA	18.3	NA	19.7	NA	18.2	NA	26.0	NA	21.4	NA
L Mt	9.2	NA	NA	NA	9.5	NA	NA	NA	11.5	NA	NA	NA

Stephan (2017) stated that due to general misinterpretation of means and p-values, TDStats such as the shorth and 75-shormax should be incorporated into FSTT studies. The current study calculated these values and tabulated against those of three other recent studies (Table 5.2). These statistics of central tendency are better suited to

FSTTs due to the non-normalized nature of tissue depth data (Stephan, 2018). The shorths for the midline measurements of the current study differed to those of the comparison studies by 0.1 – 5.0 mm. A difference of 0.1 mm was calculated between the current study and Stephan & Sievwright (2018) at the labiale superius, and a difference of 5.0 mm between the current study and Thiemann *et al*. (2017) at the mid-philtrum. There was also a larger range of 0.2 to 15.3 mm in terms of the left bilateral landmarks. The smallest difference (0.2 mm) was determined at the left mid-infraorbital between this study and Stephan & Preisler (2018), and the largest difference of 15.3 mm was calculated at the left gonion between the current study and Stephan & Sievwright (2018). The FSTT measurement at the left gonion in the current study was also larger than the shorth of Stephan & Preisler (2018) by 13.5 mm.

Possible reasons for these large differences in shorths, especially at the gonion, could be due to differences in BMI as well as the methodologies used to measure the FSTTs. BMI was a controlled variable in Thiemann *et al*. (2017), and B-mode ultrasound was used instead of CT scans by Stephan & Preisler (2018) and Stephan & Sievwright (2018). BMI has been reported to influence FSTTs, especially at landmarks in the vicinity of the lower face (e.g. gonion) (Stephan *et al*. 2016). The sample used by Thiemann *et al*. (2017) consisted only of persons categorized as having normal BMI, whereas the current study did not as this information was not recorded by the WDGMC radiology department. More so, Thiemann *et al*. (2017) utilized multislice CT scans, which unlike conventional CT scans, provide better resolution. B-mode ultrasound has been stated as the preferred method for measuring FSTTs as measurements can be taken in the upright position (Munn & Stephan, 2018). Measurements taken in upright position provides a true reflection of FSTTs *in*

situ as it eliminates the positional effects of gravity as seen with conventional CT scans (Munn & Stephan, 2018).

Considering the differences reported between the current study and all the aforementioned studies, it is imperative that the sample sizes, the demographics of the sample and methodologies utilized in future FSTT studies should be standardized in order to determine the validity of significant difference between different studies (Stephan *et al*., 2015b).

5.3. Facial approximations and facial arrays

A total of six approximations were created for this study. Four approximations were made utilizing the current FSTT data set and their recognition tested using facial arrays. There were two male and two female approximations that were further subcategorized by the data used to created it (i.e., sex-specific or pooled). The reason for this was to test whether the categorization of influences the rate of correct match pairing of facial approximation to photograph. The other two approximations were performed using the freely available C-Table raw data repository (http://www.craniofacial identification.com). The purpose of this was to assess the morphological differences between facial approximations that were produced utilizing different data-sets (i.e. sex-specific data and pooled data).

Differences in the appearances of the male and female approximations indicated varying jaw shapes, thicknesses in the parotid region and degrees of prognathism. These differences were noticeable due to several available orientations of the approximations.

With analysis, it was evident that the jaw line appeared more gracile, displayed a less pronounced bucco-parotid area and revealed more prognathism as the FSTT sample sizes increased. These changes were most apparent whilst inspecting the facial approximations compiled using the current study's sex specific data and the C-Table repository. Furthermore, the lips were more protrusive due to a more pronounced labiomental sulcus in facial approximations based on smaller samples.

The assessors of the facial arrays in the current study were able to correctly match the positive match photographs to the facial approximations 25% to 60% above chance. These results were considerably better compared to a Brazilian study performed by Herrera *et al*. (2016) which indicated a success rate of only 10 to 20% from their 120 assessors. Higher success rates of 30 to 80% above chance were achieved by Caplova *et al*. (2017) in an Italian population of students (n = 41) and forensic professionals (n = 5). The success rates of the facial arrays were 80% and 78% for the forensic professionals and students respectively, a statistically insignificant result (Caplova *et al*., 2017). In the current study, both sexes (i.e. females = 40 & males = 40) performed well, but not significantly better, when presented with the female approximations and the approximations derived from pooled data. A similar outcome was noted by Herrera *et al*. (2016), with the exception that the male assessors of their study performed significantly better compared to the female assessors included in their study. In the current study, the effects of the assessors' ancestry was mostly insignificant except in one instance when assessors classified as white performed 25% better than those assessors classified as black (t-test, $p < 0.05$). Surprisingly, assessors whom had had formal training in anatomy and/or art displayed no clear advantage over their counterparts. The former had perform slightly better (but not significantly) than the latter in three of four facial arrays. Similarly, the studies by Herrera *et al*.

(2016) and Caplova *et al*. (2017) contained assessors with knowledge of anatomy and/or experience in the field of forensics. Both their results, like those in this study, indicated performance based education or work experience.

In the current study, the feature that was noted as being the most similar significant in terms of identification (between a respective facial approximation and positive match photograph) were the eyes (31.6%). Other frequently selected features included the lips (22.8%) and the nose (22.2%). In Caplova *et al*. (2017), forensic professionals and art students had selected the nose with frequencies of 29% and 36.9% respectively. It was also commented that features such as teeth, scars and discoloration were infrequently indicated as similar between the facial approximations and photographs, but only by the forensic professionals (Caplova *et al.*, 2017). Additionally, the presence of facial hair in several of the photographs shown to assessors of Caplova *et al*. (2017) was indicated as problematic that possibly led to lower success rates in one of their facial array tests. Therefore, the current study did not include photographs of males with excessive facial hair so as prevent the implications highlighted by Caplova *et al*. (2017). Furthermore, the current study utilized black and white photographs so as to negate the effects of colour.

The results of the facial arrays also cound that negative matches were selected more often when assessors were presented with the facial approximations based on sex-specific data. This observation was especially noticeable with regards to the male facial approximations. The assessors of the facial arrays also mentioned that the appearance of sex-specific and pooled data facial approximations varied slightly. This observational difference might have been due to the difference in the gross number of FSTT measurements per landmark between the sex-specific and pooled databases. Naturally, the sex-specific FSTT data always consisted of fewer cases than the pooled

data. The FSTT means of the female sample (n = 47) closely resembled that of the entire sample (n = 60), with differences between mean values ranging from 0.0 to 0.5 mm. The FSTT means of the male sample (n = 13) differed by 0.1 to 2.2 mm compared to the entire sample (n = 60) of the current study. The recognition and resemblance of the facial approximations were not significantly impacted despite these millimetre differences (Figure 4.14 & 4.15). Another possible reason for this observation was that the facial approximations based on the sex-specific FSTT data was always presented first and the facial approximations based on the pooled data presented second. It is possible that the assessors were more attentive and had a better understanding of the facial arrays when assessing the facial approximations that were produced using the pooled FSTT data.

5.4. Repeatability

The intra- and inter-observer repeatability tests were conducted after the initial measurement of all the CT scans. The 19 CT scans used for inter- and intra-repeatability were selected on the basis of having measurements recorded at all 30 anatomical landmarks. In order to prevent any bias, all the CT scans were deleted from the OsiriX software program and then re-uploaded before the intra-observer measurements were collected. With regards to the inter-observer measurements, the same 19 CT scans were re-measured by a graduate student familiar with anatomy and landmark placement.

Comparisons between the intra- and inter-observer reliability were determined by computing the intra-class correlation (ICC), the absolute technical error of measurements (TEM), the relative-TEM and the variable average value. The original data (reading 1) *vs*. the second round measurements (reading 2) indicated high

repeatability. The ICCs calculated between these two sets of measurements were all in excess of 0.95. More so, the r-TEMs did not exceed 5%, a very accepTable and pleasing result. TEM has been characterized as the best predictor for error as it analyses the raw data instead of the difference between means as seen in Student t-test (Fancourt & Stephan, 2018). The r-TEM in the current study ranged from 0.7 (right mid-ramus) to 4.6 (mid-nasal) and smaller than the r-TEM range of 1-6% reported in Stephan & Sievwright (2018).

The inter-observer error displayed a greater degree of variation and was demonstrated with analysis of the ICC and r-TEM. The low ICC values of the labiale inferius (0.469) and menton (0.481) are especially alarming. Cicchetti (1994) categorized ICC values of 0.75 – 1.00 as excellent, 0.60 – 0.74 as good, 0.40 – 0.59 as fair, values less than 0.40 as poor. Fortunately, there were no instances of poor ICC value. A similar trend in terms of r-TEM was noticeable in the inter-observer reliability. The r-TEM exceeded 10% at the nasion, rhinion, labiale inferius, left and right zygion, left and right mid-mandibular, left and right gonion, left and right sub M_2, right supra M_2 and right mental tubercle. Furthermore, the mid-nasal and menton had r-TEMs in excess of 20% and 30% respectively. These poor r-TEMs of 20% and 30% may possibly be due to the several factors such as the poor interpretation of the landmarks' definition, the misidentification of these landmarks, or errors that occurred during the measurement process. The clarity of the skin or bone surface may also have been distorted as a result of increasing the focus. Despite these large r-TEM values, similar results were recorded by Stephan & Sievwright (2018). Their study had r-TEMs ranging from 4 to 35%. The inter-observer error r-TEM range in the current study was smaller (2.0 to 28.2%) but not statistically significant (t-test, $p < 0.05$).

Even though the repeatability between readings 1 and 3 was not at the same standard as the repeatability between readings 1 and 2, no significant differences were found to exist (t-test, $p < 0.05$). This could be misleading as the sample consisted of only 19 CT scans and p-values tend to under-represent probable significances in small samples (Stephan *et al*., 2015b). Alternatively, the differences between means had never exceeded 1.5 millimetres. Fancourt and Stephan (2018) stated that the reliability of p-values should only been classified as significant when a difference of 20 millimetres is recorded between measurements.

5.5. Samples and protocols

As previously mentioned, there are currently only four studies that have investigated South African populations and only two (Aulsebrook *et al*., 1995 and Cavanagh & Steyn, 2011) that have investigated the FSTTs in black adults. The current study's sample consisted of 60 individuals. Aulsebrook *et al.* (1995) studied a sample of 55 black adult males whilst a sample of a 154 black adult females was examined in Cavanagh & Steyn (2011). Domaracki & Stephan (2006) noted that the larger a sample the more representative it is of a population. However, larger samples tend to over-estimate statistical significances, and erroneous inferences are made regarding the relevance of small differences between FSTT measurements (Fancourt & Stephan, 2018). Due the non-normalized nature of FSTT datasets, mean values are not ideal for estimating central tendency (Domaracki & Stephan, 2006). The sample size used in the current study ($n = 60$) is by no means large, but adheres to the suggestion made by Stephan *et al*. (2015b) that a sample size of 40 or more is less vulnerable to data noise.

5.6. Limitations

5.6.1. Sampling

As this study entailed interpreting CT scans that were taken retrospectively, many difficulties were encountered throughout the study. Firstly, there was a much smaller percentage of males compared to females despite that the current sample of CT scans were taken over a six year span extending to the year 2011. This cut-off date was implemented as the Radiology Department of the Wits Donald Gordon Medical Centre (WDGMC) had used a different CT scanner prior to 2011. The department generally receives patients who are or were at some point diagnosed with cancer, most of whom are females afflicted with breast cancer. Originally, the sample was intended to consist of 30 males and 30 females for the purposes of achieving statistically significance, however, the final sample consisted of 47 females and 13 males. Although both samples (females & males) were subjected to significance testing (t-test, $p < 0.05$); no accurate and definite inferences could be made regarding sex differences due to the unequal distribution of males to females in the current sample. The adjustment to the sample was obligatory due to the factors mentioned above in conjunction to the inclusion and exclusion criteria selected for this study.

Finally, as the CT scans used in this study were retrieved exclusively from a hospital in Johannesburg (Gauteng), it may be a good indicator of South Africa's demographics due to the heterogeneous population which inhabits it.

5.6.2. Landmark identification

The landmarks used in this study were adapted from Briers *et al*. (2015), Cavanagh & Steyn (2011), Stephan & Simpson (2008) and Thiemann *et al*. (2017). The selection of landmarks was based on how reliable they were accurately located in previous studies. These landmarks were also selected due to the fact that the definitions of each were the least ambiguous and could be identified with relative ease. As the CT scans used in the current study were collected retrospectively, several landmarks were unobtainable as many scans were only taken from eye level down to the level of the chin. As a result, no measurements could be recorded for landmarks situated on the forehead. In one instance, the rhinion (Rhi) could not be located for measurement because the scan did not include the nasal region.

The average age for this sample was 49.1 (SD = 11.1) years, which is a possible explanation as to why a considerable proportion had metallic dental fillings. These fillings produced varying amounts of light scatter when the CT scan was taken (Figure 5.1.A-C); thus certain landmarks could not be identified and measured. This effect was especially predominant for the landmarks supra M_2 and sub M_2 on both the left and right sides (Figure 5.2.A & B). Furthermore, supra M_2 and sub M_2 could not be measured from partially or completely edentulous patients (Figure 5.2.C). This issue resulted in considerably smaller sample size, but still within the acceptable statistical ranges (Stephan & Simpson, 2008).

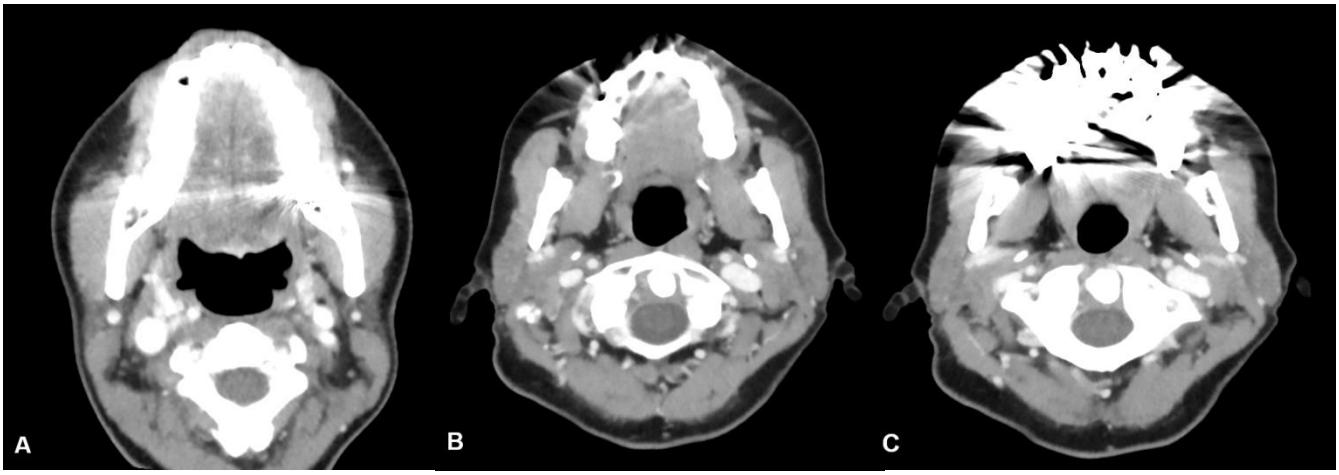


Figure 5.1: Photographs A-C are topogram views of CT scans demonstrating the varying amounts light scatter reflecting from metallic dental implants.

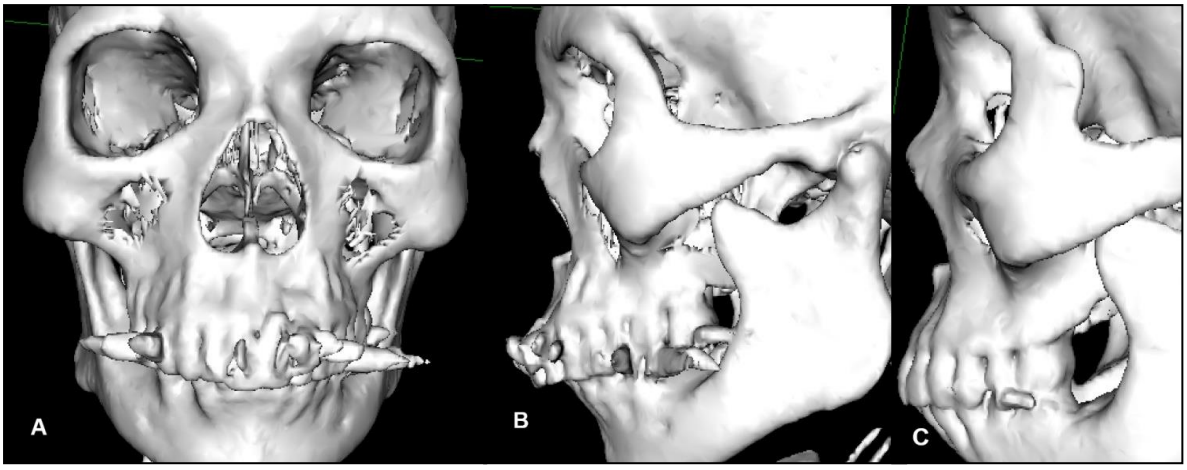


Figure 5.2: Photographs A-C are frontal (A) and lateral (B & C) depiction of a patient's skull.

In Figure 5.2 there are spike-like objects projecting perpendicular to the surface of the teeth (A-C) as a result of light scatter. In several scans, patients were also found to be edentulous as seen in C and FSTTs at the supra and sub M_2 could not be measured.

A further problem included very few scans being taken in the Frankfurt Horizontal Plane. This was problematic because the scans were set to one-millimetre thicknesses, allowing very little room for error whilst locating the landmarks (especially those

situated on the mid-line). Even though the OsiriX software allowed the user to alter the axes, the process was labourious and very time consuming.

Due to the ethical implications pertaining to exposing people to radiation in cases not related to diagnosis and/or treatment, the current study could not recruit more appropriate candidates. Additionally, such a process would be expensive to finance independently.

5.6.3. Photographs for facial arrays

Since the second part of this study was conducted prospectively, it was possible to capture photographs of the patients who volunteered their CT scans for the facial approximations. These photographs were kept as positive matches for the imminent facial array tests. Thereafter, negative matches similar in sex, age, population group and appearance to that of the positive matches were gathered. Individuals fitting this description were approached by the investigator and issued with an information sheet explaining their role in the study. If the individuals were willing, they were then asked to complete a letter of informed consent. These participants included university students and staff, as well as patients from a local hospital. Many of the people who were approached to participate declined for reasons such feeling uncomfortable being photographed in fear of the wrongful use of their photographs despite ethical clearance given by the HREC. Difficulties associated with those who had agreed to participate included these individuals not maintaining the Frankfurt position and a neutral position. In these instances the photographs were discarded. The photographs which were ultimately selected for the facial arrays required alteration in terms of lighting and contrast in order to standardize their appearances so no photograph drew

unwarranted attention. For these reasons, all the photographs were displayed in grey-scale in the facial array tests.

5.7. Significance of this study

This study is the first to simultaneously investigate FSTTs in adult males and females whom are classified as black South Africans. It was also the first study in South Africa to investigate and report on FSTT bilateral asymmetry. The results of this study provide forensic facial artists with a new population-specific data set that could supplement existing South African FSTT standards. These FSTT mean values are only to be used as guidelines as it would be misleading to believe that an accurate ante-mortem appearance could be achieved without access to ante-mortem data of the deceased. Examples included whether the deceased was emaciated, obese or of normal weight; the presence of scarring or facial abnormities; and hair styles. A possible manner of managing the appearance of a facial approximation in terms of its health status would be to produce additional approximation by incorporating the shorths and 75-shormaxes of FSTTs at each anatomical landmark (Stephan *et al*., 2016). These variations of the same approximation should subsequently be presented together and not as individual identities. The results of the current study further support the pooling of FSTT data as suggested by Stephan *et al*. (2015b). Despite the categorization of the FSTT data in this study, the facial approximations produced utilizing sex-specific and pooled data are not highly differentiated. The results of the current study indicated higher recognition rates of facial approximations when pooled FSTT data was applied. Furthermore, the assessors of the facial arrays had indicated that features not associated with FSTT (e.g. eyes, lips, and nose) were the most influential during the identification process. The aforementioned finding further reinforces the proposed notion of moving away from FSTT data sub-categorisation.

As with any FSTT dataset, forensic facial artists should continue to practice their professional discretion, as well as account for the individual peculiarities associated with each individual skull.

Chapter 6: Conclusion

The facial soft tissue thickness data (FSTTs) produced in the current study were extracted from the CT scans of 60 black South African adults. The purpose of this study was to ascertain whether there is a need for population-specific FSTT data as well as the influence thereof with respect to the recognition of facial approximations. The results of this study revealed good repeatability and should be given consideration for the purposes of facial approximations.

- The FSTT results displayed significant differences according to sex, age and bilateral asymmetry. However, the metric difference between these means amounted to only a few millimetres. The effects of these millimetre differences did influence the appearance of the approximations, but did not alter their recognition.
- Population-specific data are important, however, the over-emphasis of its importance related to recognition is unwarranted, as metric differences may be a result of data noise or inappropriate statistical testing.
- The results of the facial arrays indicated that assessors paid more attention to the eye, lips and nose of the photographs and facial approximations. The FSTT differences between the facial approximations did not seem to influence the assessors during the arrays, regardless of whether they selected the correct or incorrect photograph.
- Inherent biases associated with sex and population affinity were also found to be absent within the current group of facial array assessors. Certain facial features were more frequently recorded amongst certain categories of assessors. These differences, however, were calculated to be insignificant.

- The data retrieved after the facial arrays indicated that people assess faces differently and may remember faces based on certain features. This observation should be investigated further as it may perhaps assist forensic facial artists in isolating which features require more attention during the facial approximation process.
- Significant differences were calculated at several landmarks for the FSTT of the current study; however, analysis of the facial arrays indicated no noteworthy effects on the recognition of the facial approximations.
- The results of this study are in accordance with those that promote the pooling of FSTT data instead of its continual sub-categorization. Although, subtle millimetre differences will inevitably exist across and within populations, the morphology of the skull is believed to be more significant.

The results of the current study have provided greater insights into the sex and age related FSTT differences. This was the first South African study to report on bilateral asymmetry. This study was also the first South African study to use its own FSTT data to produce three-dimensional computer generated facial approximations. Furthermore, no previous study in the country has considered testing the recognition of its facial approximations in terms of biases associated to sex, population group and education (i.e.; anatomy and or facial artistry). The current study also attempted to test the idea that recognition rates are not dependent on the category of FSTT data used to produce facial approximations (i.e. sex-specific vs. pooled), specifically in the context of South Africa. Until researchers begin to implement a standardized methodology with appropriate samples sizes, no declaration can be made regarding the need for categorised or population-specific data.

Recommendations

The purpose of FSTT data is for producing recognisable facial approximations. The current study has shown that the sub-categorization of FSTT data is unnecessary and hence, future research should adopt the pooling of FSTT data. Researchers should also be aware that assessors of the facial arrays in the current study placed more emphasis on facial features (e.g., the eyes, nose, lips) compared to tissue thicknesses whilst analysing the facial approximations. Therefore, future research should attempt to investigate the relationship between the frequencies of positive identification and the facial features of facial approximations.

References

Aulsebrook, W.A., Becker, P.J. and İşcan, M.Y., 1996. Facial soft-tissue thicknesses in the adult male Zulu. *Forensic Science International*, *79*(2), pp.83-102.

Barzut, V. and Zdravković, S., 2013. Discrimination of faces of the same and other race and gender modulated by familiarity. *Psihologija*, *46*(1), pp.45-59.

Basu, M.N. and Nag, M.S., 2015. Facial Reconstruction -A Review. *International Education and Research Journal*, *1*(5), pp.34-36.

Bhorat, H., Lilenstein, A., Monnakgotla, J., Thornton, A. and van der Zee, K. (2017). "The SocioEconomic Determinants of Crime in South Africa: an Empirical Assessment". Development Policy Research Unit Working Paper 201704. DPRU, University of Cape Town.

Blais, C., Jack, R.E., Scheepers, C., Fiset, D. and Caldara, R., 2008. Culture shapes how we look at faces. *PloS one*, *3*(8), p.e3022.

Briers, N., Briers, T.M., Becker, P.J. and Steyn, M., 2015. Soft tissue thickness values for black and coloured South African children aged 6–13 years. *Forensic Science International*, *252*, pp.188-e1.

Bulut, O., Sipahioglu, S. and Hekimoglu, B., 2014. Facial soft tissue thickness database for craniofacial reconstruction in the Turkish adult population. *Forensic Science International*, *242*, pp.44-61.

Campomanes-Álvarez, B.R., Ibáñez, O., Navarro, F., Alemán, I., Cordón, O. and Damas, S., 2015. Dispersion assessment in the location of facial landmarks on photographs. *International Journal of Legal Medicine*, *129*(1), pp.227-236.

Caple, J. and Stephan, C.N., 2016. A standardized nomenclature for craniofacial and facial anthropometry. *International Journal of Legal Medicine*, *130*(3), pp.863-879.

Caplova, Z., Obertova, Z., Gibelli, D.M., Mazzarelli, D., Fracasso, T., Vanezis, P., Sforza, C. and Cattaneo, C., 2017. The reliability of facial recognition of deceased persons on photographs. *Journal of Forensic Sciences*, *62*(5), pp.1286-1291.

Cavanagh, D. and Steyn, M., 2011. Facial reconstruction: soft tissue thickness values for South African black females. *Forensic Science International*, *206*(1-3), pp.215-e1.

Christensen, A.M. and Anderson, B.E., 2013. Methods of personal identification. In: Tersigni-Tarrant, M.A. & Shirley, N.R. (Eds.), *Forensic Anthropology: An Introduction*. CRC Press, pp.397-420.

Chung, J.H., Chen, H.T., Hsu, W.Y., Huang, G.S. and Shaw, K.P., 2015. A CT-scan database for the facial soft tissue thickness of Taiwan adults. *Forensic Science International*, *253*, pp.132-e1.

Cicchetti, D.V., 1994. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, *6*(4), p.284.

Claes, P., Vandermeulen, D., De Greef, S., Willems, G., Clement, J.G. and Suetens, P., 2010. Computerized craniofacial reconstruction: conceptual framework and review. *Forensic Science International*, *201*(1-3), pp.138-145.

Cobbold, R.S., 2006. *Foundations of biomedical ultrasound*. Oxford University Press.

Codinha, S., 2009. Facial soft tissue thicknesses for the Portuguese adult population. *Forensic Science International*, *184*(1-3), pp.80-e1.

de Almeida, N.H., Michel-Crosato, E., de Paiva, L.A.S. and Biazevic, M.G.H., 2013. Facial soft tissue thickness in the Brazilian population: New reference data and anatomical landmarks. *Forensic Science International*, *231*(1-3), pp.404-e1.

De Greef, S., Claes, P., Vandermeulen, D., Mollemans, W., Suetens, P. and Willems, G., 2006. Large-scale in-vivo Caucasian facial soft tissue thickness database for craniofacial reconstruction. *Forensic Science International*, *159*, pp.S126-S146.

De Greef, S., Vandermeulen, D., Claes, P., Suetens, P. and Willems, G., 2009. The influence of sex, age and body mass index on facial soft tissue depths. *Forensic Science, Medicine, and Pathology*, *5*(2), pp.60-65.

Domaracki, M. and Stephan, C.N., 2006. Facial soft tissue thicknesses in Australian adult cadavers. *Journal of Forensic Sciences*, *51*(1), pp.5-10.

Drgáčová, A., Dupej, J. and Velemínská, J., 2016. Facial soft tissue thicknesses in the present Czech population. *Forensic Science International*, *260*, pp.106-e1.

El-Mehallawi, I.H. and Soliman, E.M., 2001. Ultrasonic assessment of facial soft tissue thicknesses in adult Egyptians. *Forensic Science International*, *117*(1-2), pp.99-107.

Evison, M.P., Iwamura, E. S. M. and Guimarães, M.A.G. (2016). Forensic facial reconstruction and its contribution to identification in missing person cases. In Morewitz, S.J. and Sturdy Colls, C. (eds.), Handbook of Missing Persons, New York: Springer, pp. 427-41.

Fancourt, H.S. and Stephan, C.N., 2018. Error measurement in craniometrics: The comparative performance of four popular assessment methods using 2000 simulated cranial length datasets (g-op). *Forensic Science International*, *285*, pp.162-171.

Garlie, T.N. and Saunders, S.R., 1999. Midline facial tissue thicknesses of subadults from a longitudinal radiographic study. *Journal of Forensic Science*, *44*(1), pp.61-67.

Goto, R., 2007. Precision of measurement as a component of human variation. *Journal of Physiological Anthropology*, *26*(2), pp.253-256.

Gupta, S., Gupta, V., Vij, H., Vij, R. and Tyagi, N., 2015. Forensic facial reconstruction: The final frontier. *Journal of Clinical and Diagnostic Research: JCDR*, *9*(9), pp.26.

Herlitz, A. and Lovén, J., 2013. Sex differences and the own-gender bias in face recognition: a meta-analytic review. *Visual Cognition*, *21*(9-10), pp.1306-1336.

Herrera, L.M., Strapasson, R.A.P., da Silva, J.V.L. and Melani, R.F.H., 2016. Forensic facial approximation assessment: can application of different average facial tissue depth data facilitate recognition and establish accepTable level of resemblance?. *Forensic Science International*, *266*, pp.311-319.

Herrera, L.M., Strapasson, R.A.P., Zanin, A.A., da Silva, J.V.L. and Melani, R.F.H., 2017. Comparison among manual facial approximations conducted by two methodological approaches of face prediction. *Journal of Forensic Sciences*, *62*(5), pp.1279-1285.

Hills, P.J. and Lewis, M.B., 2011. Rapid communication: the own-age face recognition bias in children and adults. *Quarterly Journal of Experimental Psychology*, *64*(1), pp.17-23.

Huculak, M.A. and Peckmann, T.R., 2012. *In vivo* facial tissue depth study of African Nova Scotian children. *Canadian Society of Forensic Science Journal*, *45*(3), pp.126-142.

Hwang, H.S., Park, M.K., Lee, W.J., Cho, J.H., Kim, B.K. and Wilkinson, C.M., 2012. Facial soft tissue thickness database for craniofacial reconstruction in Korean adults. *Journal of Forensic Sciences*, *57*(6), pp.1442-1447.

Jeelani, W., Fida, M. and Shaikh, A., 2017. Age and sex-related variations in facial soft tissue thickness in a sample of Pakistani children. *Australian Journal of Forensic Sciences*, *49*(1), pp.45-58.

Keyes, C.A., Hill, L. and Gordon, G.M., 2016. An Appraisal of Decomposition Cases Received at the Johannesburg Forensic Pathology Service Medico- legal Mortuary During 2010–2011. *Journal of Forensic Sciences*, *61*(2), pp.452-457.

Kurkcuoglu, A., Pelin, C., Ozener, B., Zagyapan, R. and Sahinoglu, Z., 2011. Facial soft tissue thickness in individuals with different occlusion patterns in adult Turkish subjects. *HOMO-Journal of Comparative Human Biology*, *62*(4), pp.288-297.

Lavrakas, P.J., Buri, J.R. and Mayzner, M.S., 1976. A perspective on the recognition of other-race faces. *Perception & Psychophysics*, *20*(6), pp.475-481.

Lee, W.J., Yoon, A.Y., Song, M.K., Wilkinson, C.M. and Shin, D.H., 2014. The archaeological contribution of forensic craniofacial reconstruction to a portrait drawing of a Korean historical Figure. *Journal of Archaeological Science*, *49*, pp.228-236.

Lee, W.J. and Wilkinson, C.M., 2016. The unfamiliar face effect on forensic craniofacial reconstruction and recognition. *Forensic Science International*, *269*, pp.21-30.

Leo, C., O'Connor, J.E. and McNulty, J.P., 2013. Combined radiographic and anthropological approaches to victim identification of partially decomposed or skeletal remains. *Radiography*, *19*(4), pp.353-362.

Lodha, A., Mehta, M., Patel, M.N. and Menon, S.K., 2016. Facial soft tissue thickness database of Gujarati population for forensic craniofacial reconstruction. *Egyptian Journal of Forensic Sciences*, *6*(2), pp.126-134.

Meissner, C.A., Tredoux, C.G., Parker, J.F. and MacLin, O.H., 2005. Eyewitness decisions in simultaneous and sequential lineups: A dual-process signal detection theory analysis. *Memory & Cognition*, *33*(5), pp.783-792.

Masoume, J., Farzad, E. and Mohaddeseh, M.M., 2015. Facial soft tissue thickness in North-West of Iran. *Advances in Biosciences & Clinical Medicine*, *3*(1), p.29.

Munn, L. and Stephan, C.N., 2018. Changes in face topography from supine-to-upright position—And soft tissue correction values for craniofacial identification. *Forensic Science International*, *289*, pp.40-50.

Nozari, A.Y. and Siamian, H., 2015. The relationship between field dependent-independent cognitive style and understanding of English text reading and academic success. *Materia Socio-Medica*, *27*(1), p.39.

Omstead, J., 2002. Facial reconstruction. *Totem: The University of Western Ontario Journal of Anthropology*, *10*(1), pp.7.

Parks, C.L., Richard, A.H. and Monson, K.L., 2014. Preliminary assessment of facial soft tissue thickness utilizing three-dimensional computed tomography models of living individuals. *Forensic Science International*, *237*, pp.146-e1.

Panenková, P., 2007. Face approximation and information about facial soft tissue thickness. *EAA Summer School eBook*, *1*, pp.233-9.

Panenková, P., Beňuš, R., Masnicová, S., Obertova, Z. and Grunt, J., 2012. Facial soft tissue thicknesses of the mid-face for Slovak population. *Forensic Science International*, *220*(1-3), pp.293-e1.

Peckmann, T.R., Harris, M., Huculak, M., Pringle, A. and Fournier, M., 2015. *In vivo* facial tissue depth for Canadian Mi'kmaq adults: A case study from Nova Scotia, Canada. *Journal of Forensic and Legal Medicine*, *29*, pp.43-53.

Perini, T.A., Oliveira, G.L.D., Ornellas, J.D.S. and Oliveira, F.P.D., 2005. Technical error of measurement in anthropometry. *Revista Brasileira de Medicina do Esporte*, *11*(1), pp.81-85.

Phillips, V.M. and Smuts, N.A., 1996. Facial reconstruction: utilization of computerized tomography to measure facial tissue thickness in a mixed racial population. *Forensic Science International*, *83*(1), pp.51-59.

Prag, J, & Neave, R., 1997. *Making faces: using forensic and archaeological evidence*. London: British Museum Press.

Randolph-Quinney, P.S., Mallett, X. and Black, S.M., 2009. Forensic Anthropology: In Jamieson, A. and Moenssens, A. (Eds.), *Wiley Encyclopedia of Forensic Science*, pp. 1-27.

Rhodes, M.G. and Anastasi, J.S., 2012. The own-age bias in face recognition: a meta-analytic and theoretical review. *Psychological Bulletin*, *138*(1), pp.146.

Ruiz, N.A.P., 2013. Facial soft tissue thickness of Colombian adults. *Forensic Science International*, *229*(1-3), pp.160-e1.

Sangrigoli, S., Pallier, C., Argenti, A.M., Ventureyra, V.A.G. and de Schonen, S., 2005. Reversibility of the other-race effect in face recognition during childhood. *Psychological Science*, *16*(6), pp.440-444.

Schaich, A., Obermeyer, S., Kolling, T. and Knopf, M., 2016. An own-age bias in recognizing faces with horizontal information. *Frontiers in Aging Neuroscience*, *8*, pp.264.

Shriver, E.R., Young, S.G., Hugenberg, K., Bernstein, M.J. and Lanter, J.R., 2008. Class, race, and the face: Social context modulates the cross-race effect in face recognition. *Personality and Social Psychology Bulletin*, *34*(2), pp.260-274.

Solla, H.E. and Domínguez, Z., 2018. Skeletal remains of Ramón Roberto Peré: A victim of a dictatorial regime in Uruguay. *Paripex-Indian Journal of Research*, pp. *7*(3).

Sporer, S.L., 2001. Recognizing faces of other ethnic groups: An integration of theories. *Psychology, Public Policy, and Law*, *7*(1), pp.36.

Starbuck, J.M. and Ward, R.E., 2007. The affect of tissue depth variation on craniofacial reconstructions. *Forensic Science International*, *172*(2-3), pp.130-136.

Stephan, C.N. and Henneberg, M., 2001. Building faces from dry skulls: are they recognized above chance rates?. *Journal of Forensic Science*, *46*(3), pp.432-440.

Stephan, C.N., 2006. Beyond the sphere of the English facial approximation literature: ramifications of German papers on western method concepts. *Journal of Forensic Sciences*, *51*(4), pp.736-739.

Stephan, C.N. and Simpson, E.K., 2008a. Facial soft tissue depths in craniofacial identification (Part I): an analytical review of the published adult data. *Journal of Forensic Sciences*, *53*(6), pp.1257-1272.

Stephan, C.N. and Simpson, E.K., 2008b. Facial soft tissue depths in craniofacial identification (Part II): An Analytical Review of the Published Sub- Adult Data. *Journal of Forensic Sciences*, *53*(6), pp.1273-1279.

Stephan, C.N., Simpson, E.K. and Byrd, J.E., 2013. Facial soft tissue depth statistics and enhanced point estimators for craniofacial identification: the debut of the shorth and the 75- shormax. *Journal of Forensic Sciences*, *58*(6), pp.1439-1457.

Stephan, C.N., 2014. The application of the central limit theorem and the law of large numbers to facial soft tissue depths: T- Table robustness and trends since 2008. *Journal of Forensic Sciences*, *59*(2), pp.454-462.

Stephan, C.N., 2015a. Facial approximation—from facial reconstruction synonym to face prediction paradigm. *Journal of Forensic Sciences*, *60*(3), pp.566-571.

Stephan, C.N., Munn, L. and Caple, J., 2015b. Facial soft tissue thicknesses: noise, signal, and P. *Forensic Science International*, *257*, pp.114-122.

Stephan, C.N., Preisler, R., Bulut, O. and Bennett, M., 2016. Turning the Tables of sex distinction in craniofacial identification: Why females possess thicker facial soft tissues than males, not vice versa. *American Journal of Physical Anthropology*, *161*(2), pp.283-295.

Stephan, C.N., 2017. 2018 tallied facial soft tissue thicknesses for adults and sub-adults. *Forensic Science International*, *280*, pp.113-123.

Stephan, C.N., 2018. TDStats—A fast standardized capability for facial soft tissue thickness analysis in R. *Forensic Science International*, *28*. pp 304-309

Stephan, C.N. and Preisler, R., 2018. *In vivo* facial soft tissue thicknesses of adult Australians. *Forensic Science International*, *282*, pp.220-e1.

Stephan, C.N. and Sievwright, E., 2018. Facial soft tissue thickness (FSTT) estimation models—And the strength of correlations between craniometric dimensions and FSTTs. *Forensic Science International*, *286*, pp.128-140.

Tedeschi-Oliveira, S.V., Melani, R.F.H., de Almeida, N.H. and de Paiva, L.A.S., 2009. Facial soft tissue thickness of Brazilian adults. *Forensic Science International*, *193*(1-3), pp.127-e1.

Thiemann, N., Keil, V. and Roy, U., 2017. *In vivo* facial soft tissue depths of a modern adult population from Germany. *International Journal of Legal Medicine*, *131*(5), pp.1455-1488.

Towler, A., Kemp, R. I., & White, D. (2017). Unfamiliar face matching systems in applied settings. In Face Processing: Systems, Disorders and Cultural Difference, Bindemann, M. & Megreya, A.M. (Eds.), Nova Science.

Ullrich, H. and Stephan, C.N., 2011. On Gerasimov's plastic facial reconstruction technique: new insights to facilitate repeatability. *Journal of Forensic Sciences*, *56*(2), pp.470-474.

Ullrich, H. and Stephan, C.N., 2016. Mikhail Mikhaylovich Gerasimov's authentic approach to plastic facial reconstruction. *Anthropologie*, *54*(2), pp.97.

Utsuno, H., Kageyama, T., Uchida, K., Yoshino, M., Miyazawa, H. and Inoue, K., 2010. Facial soft tissue thickness in Japanese children. *Forensic Science International*, *199*(1-3), pp.109-e1.

Vanezis, M. and Vanezis, P., 2000. Cranio-facial reconstruction in forensic identification—historical development and a review of current practice. *Medicine, Science and the Law*, *40*(3), pp.197-205.

Verzé, L., 2009. History of facial reconstruction. *Acta Bio Medica Atenei Parmensis*, *80*(1), pp.5-12.

Wallis, J., Lipp, O.V. and Vanman, E.J., 2012. Face age and sex modulate the other-race effect in face recognition. *Attention, Perception, & Psychophysics*, *74*(8), pp.1712-1721.

Wilkinson, C.M., 2002. In vivo facial tissue depth measurements for white British children. *Journal of Forensic Science*, *47*(3), pp.459-465.

Wilkinson, C., 2004. *Forensic Facial Reconstruction*. Cambridge University Press.

Wilkinson, C., 2005. Computerized forensic facial reconstruction. *Forensic Science, Medicine, and Pathology*, *1*(3), pp.173-177.

Wilkinson, C., 2010. Facial reconstruction–anatomical art or artistic anatomy?. *Journal of Anatomy*, *216*(2), pp.235-250.

Wright, D.B. and Sladden, B., 2003. An own gender bias and the importance of hair in face recognition. *Acta Psychologica*, *114*(1), pp.101-114.

Appendix 1

Scoring Sheet

Study Title: Facial soft tissue thickness and facial recognition rates of black South African adults.

Details of Assessor: White | Black | Male | Female

Do you have any knowledge of anatomy, facial artistry or have an occupation which involves understanding facial proportions? Yes | No

If yes, please specify: _____

Instructions: Study the provided photographs in conjunction with the facial approximations. Select which photograph most closely resembles the facial approximation and enter it next to the relevant option. Subsequently, enter the facial feature (e.g. nose, mouth, eyes, brow ridge, shape of face, etc.) which was most influential in the decision previously made. Please try your best not to guess and note that a decision must be made. Thank you for your participation.

Facial Approximation	Corresponding Photograph	Facial Feature
Subject A		
Subject B		
Subject C		
Subject D		

Why did you select this facial feature?

A: _____

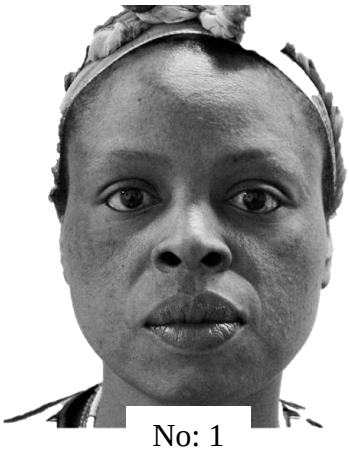
B: _____

C: _____

D: _____

Appendix 2.1.1: Facial array presenting frontal view photographs of female participants and female forensic facial approximation (subject A) (in the centre) constructed utilizing sex-specific FSTT data

No: 10

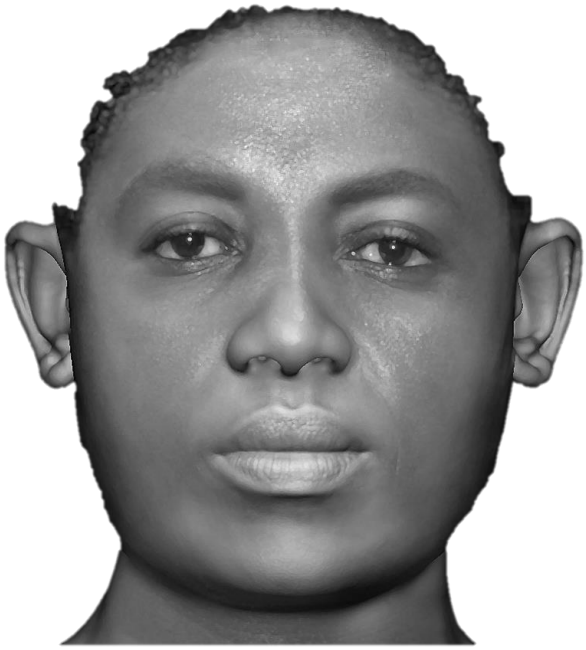


No: 1

No: 2

No: 3

No: 9



No: 4

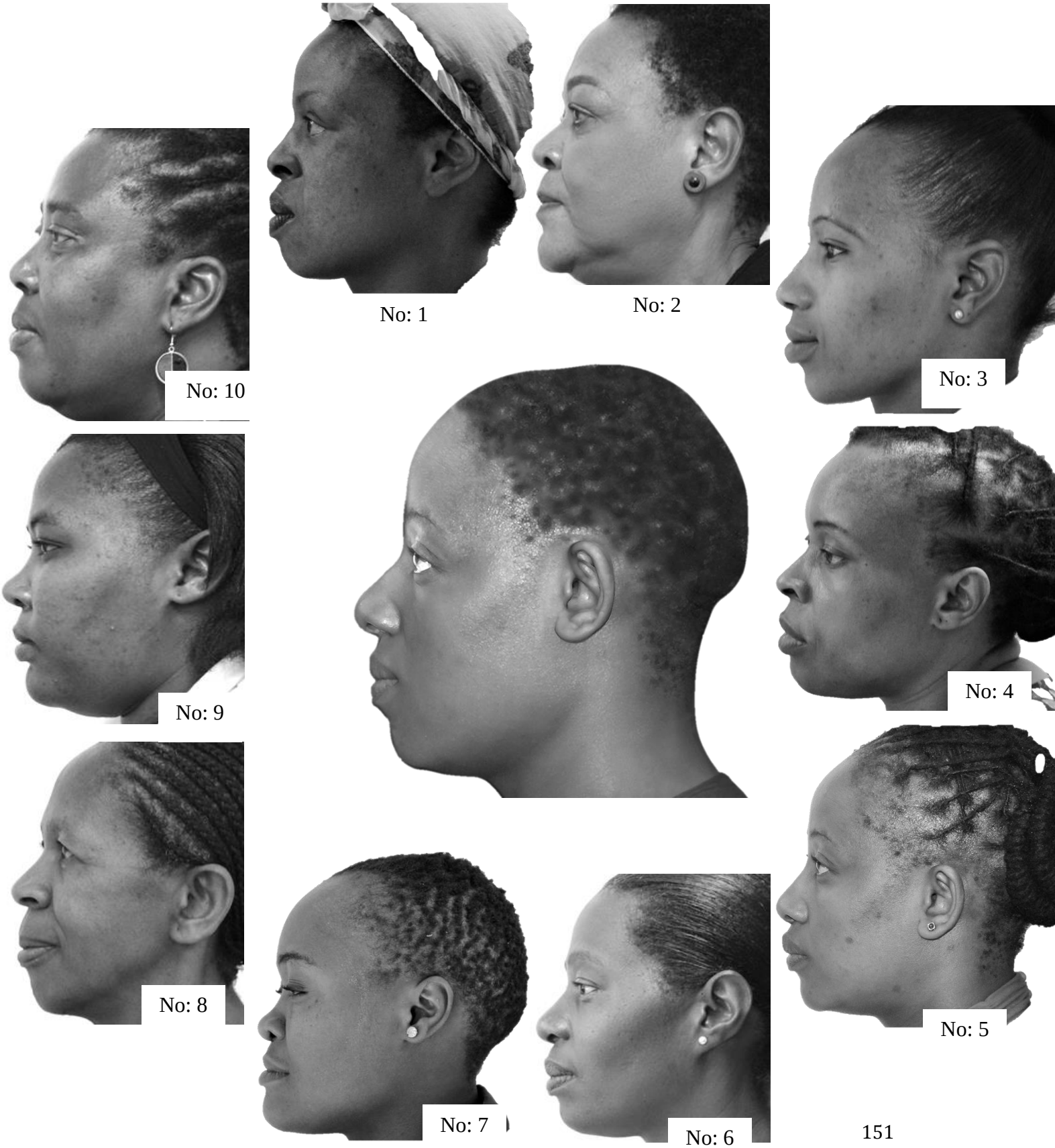
No: 8

No: 7

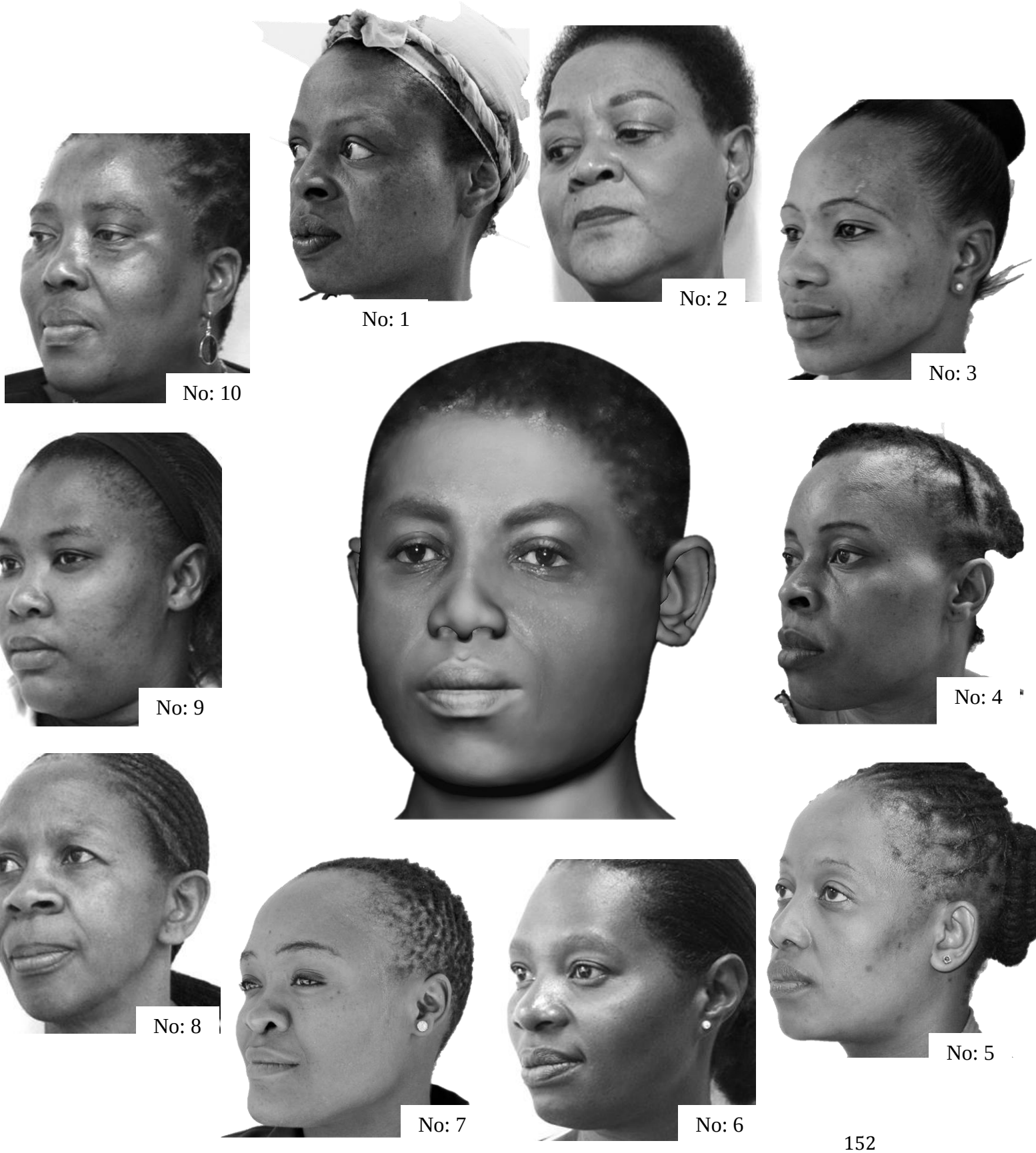
No: 6

No: 5

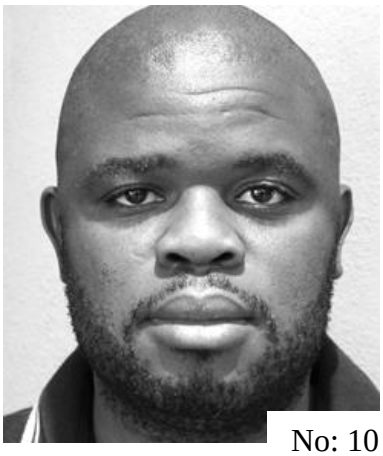
Appendix 2.1.2: Facial array presenting lateral view photographs of female participants and female forensic facial approximation (subject A) (in the centre) constructed utilizing sex-specific FSTT data



Appendix 2.1.3: Facial array presenting three-quarter view photographs of female participants and female forensic facial approximation (subject A) (in the centre) constructed utilizing sex-specific FSTT data



Appendix 2.2.1: Facial array presenting frontal view photographs of male participants and male forensic facial approximation (subject B) (in the centre) constructed utilizing sex-specific FSTT data



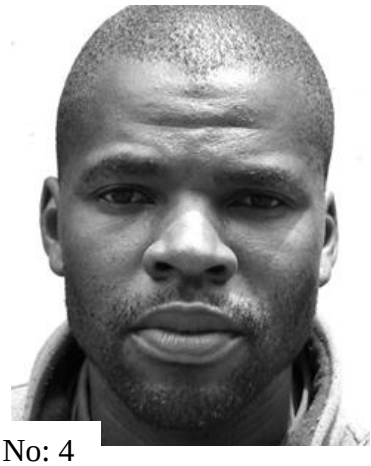
No: 10

No: 1

No: 2

No: 3

No: 9



No: 4

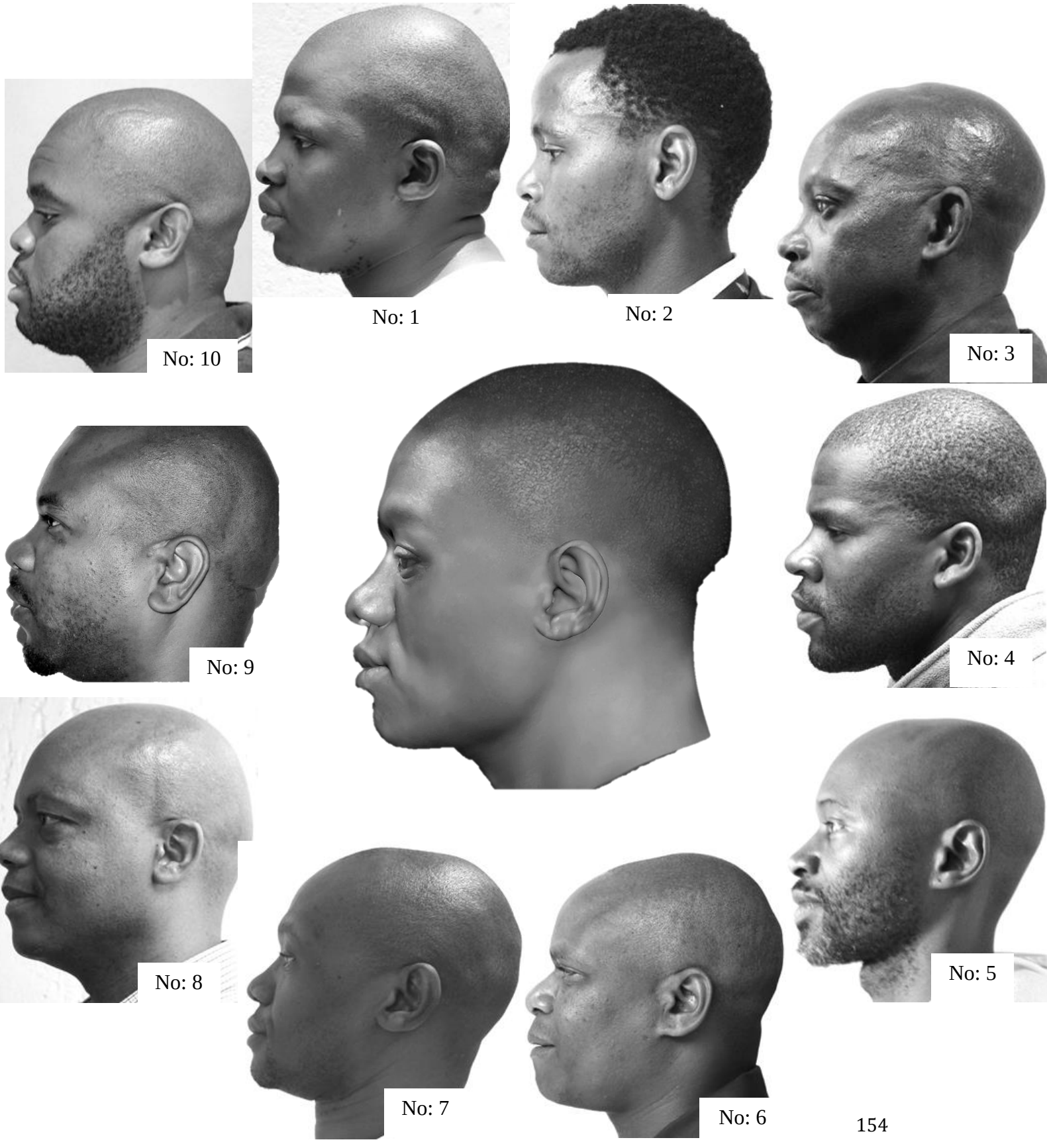
No: 8

No: 7

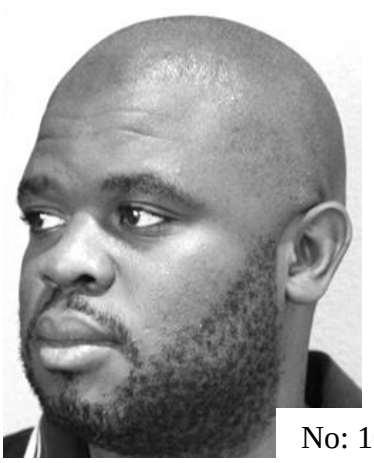
No: 6

No: 5

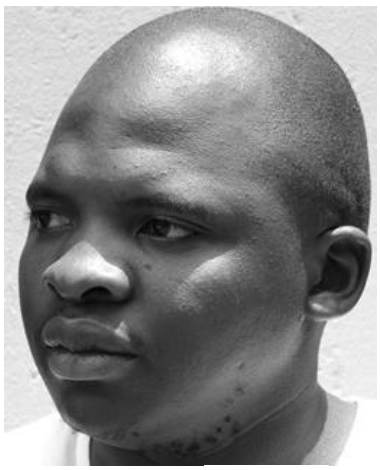
Appendix 2.2.2: Facial array presenting lateral view photographs of male participants and male forensic facial approximation (subject B) (in the centre) constructed utilizing sex-specific FSTT data



Appendix 2.2.3: Facial array presenting three-quarter view photographs of male participants and male forensic facial approximation (subject B) (in the centre) constructed utilizing sex-specific FSTT data



No: 10



No: 1

No: 2

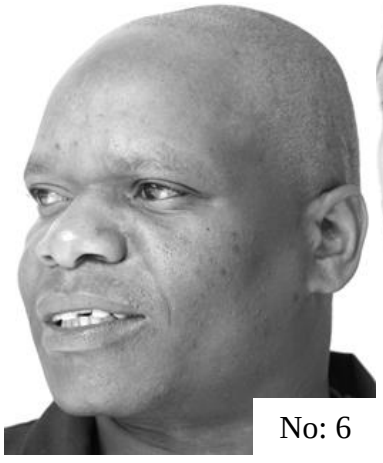
No: 3

No: 9

No: 4

No: 8

No: 7

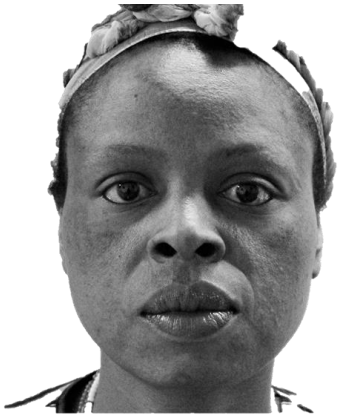


No: 6

No: 5

Appendix 2.3.1: Facial array presenting frontal view photographs of female participants and female forensic facial approximation (subject C) (in the centre) constructed utilizing pooled FSTT data

No: 10

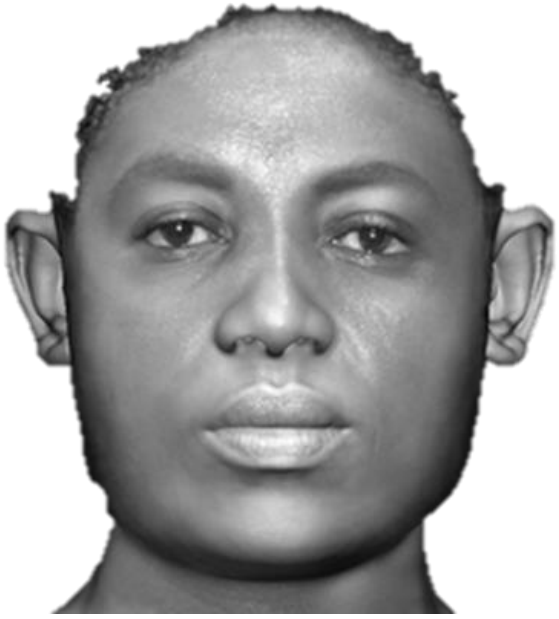


No: 1

No: 2

No: 3

No: 9



No: 4

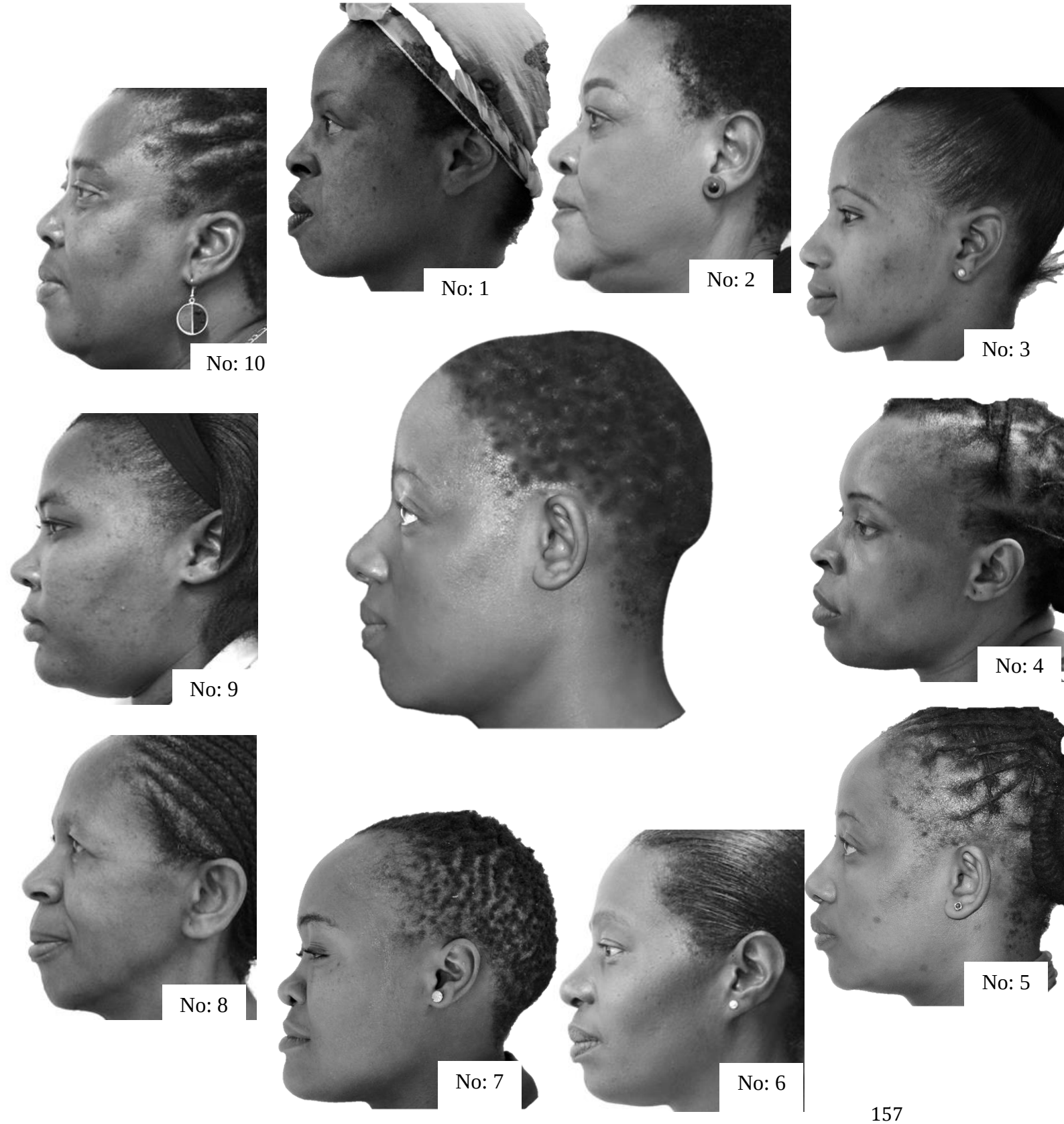
No: 8

No: 7

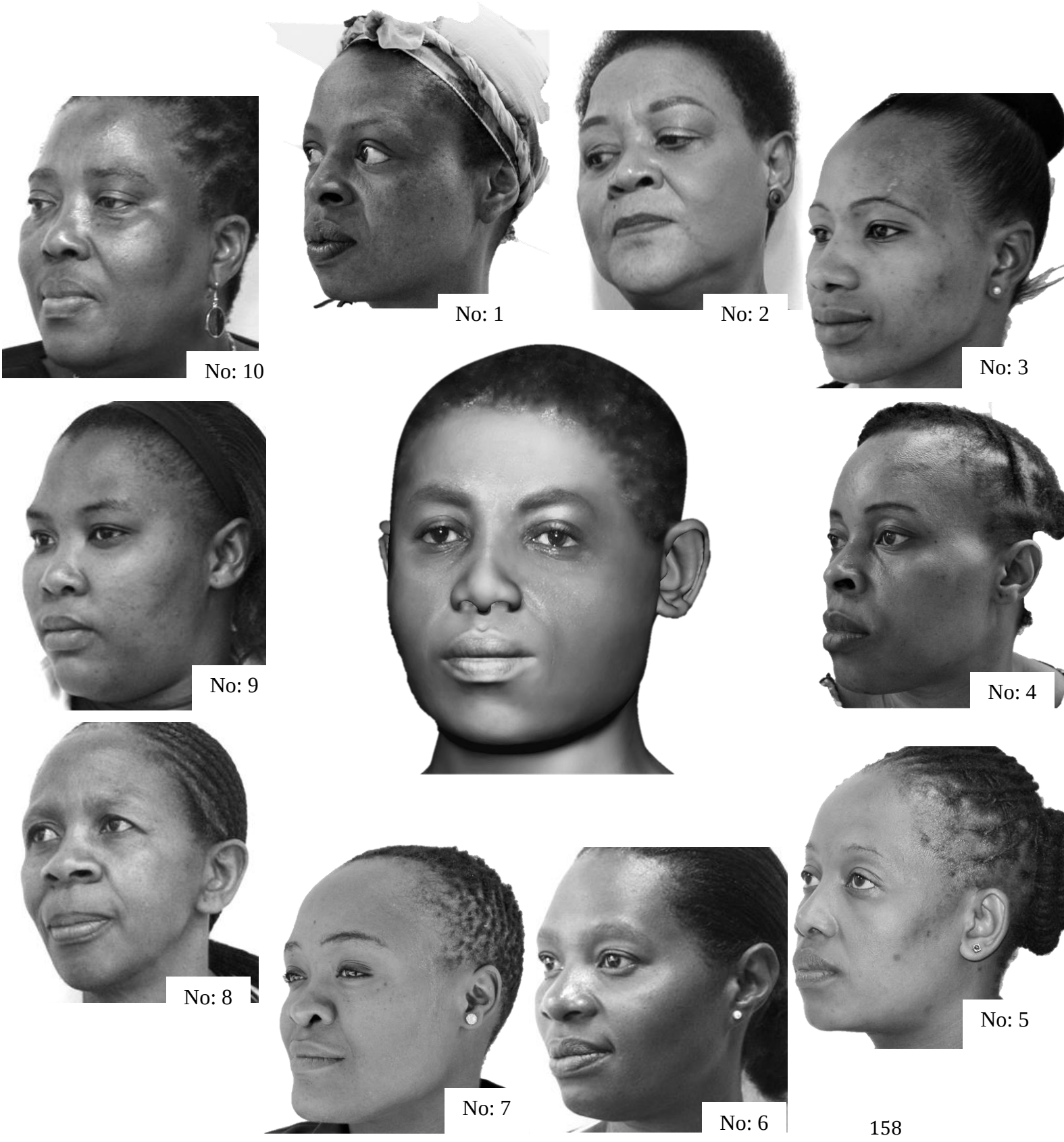
No: 6

No: 5

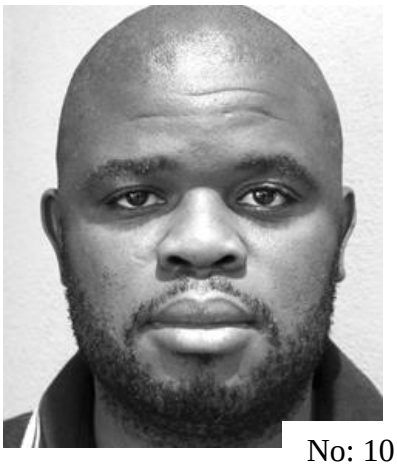
Appendix 2.3.2: Facial array presenting lateral view photographs of female participants and female forensic facial approximation (subject C) (in the centre) constructed utilizing pooled FSTT data



Appendix 2.3.3: Facial array presenting three-quarter view photographs of female participants and female forensic facial approximation (subject C) (in the centre) constructed utilizing pooled FSTT data



Appendix 2.4.1: Facial array presenting frontal view photographs of male participants and male forensic facial approximation (subject D) (in the centre) constructed utilizing pooled FSTT data



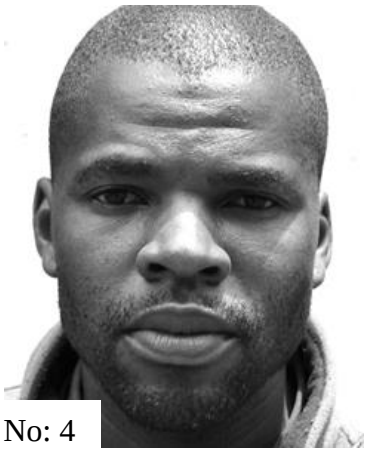
No: 10

No: 1

No: 2

No: 3

No: 9



No: 4

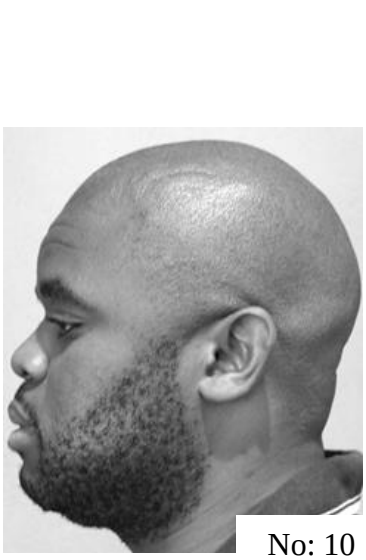
No: 8

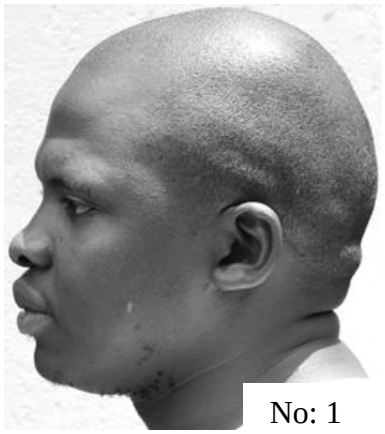
No: 7

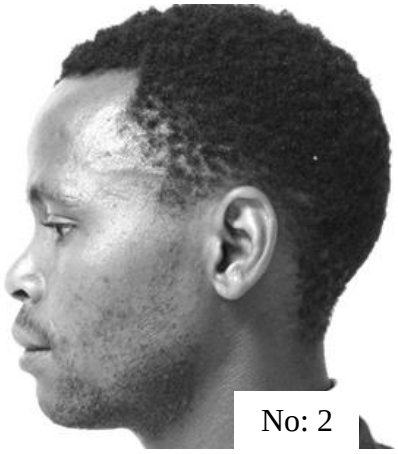
No: 6

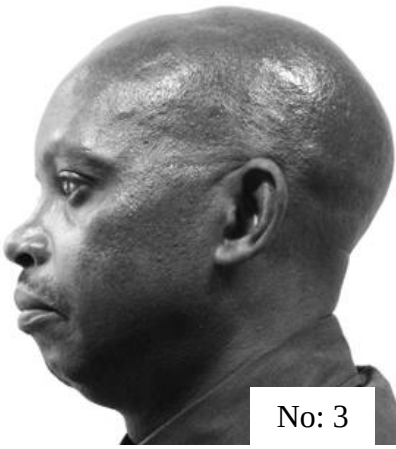
No: 5

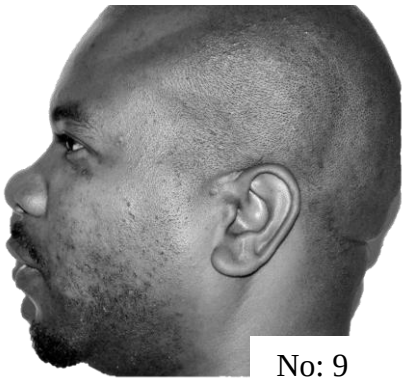
Appendix 2.4.2: Facial array presenting lateral view photographs of male participants and male forensic facial approximation (subject D) (in the centre) constructed utilizing pooled FSTT data

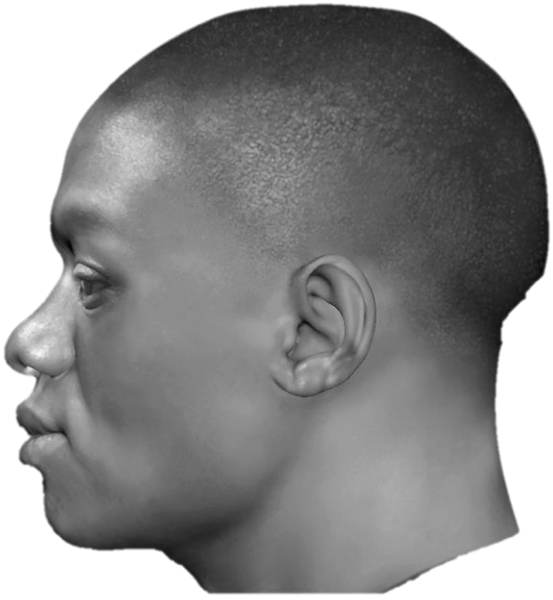

No: 10

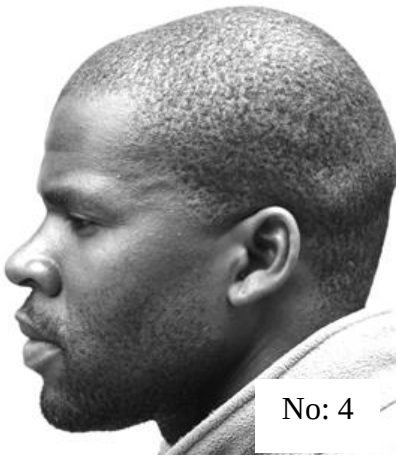

No: 1

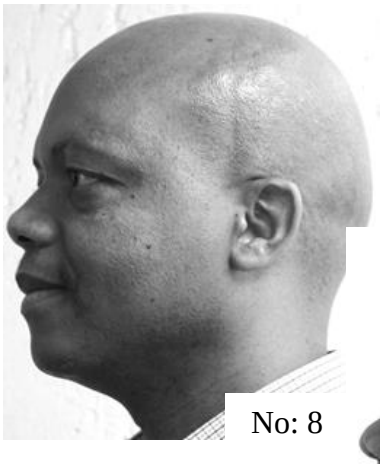

No: 2

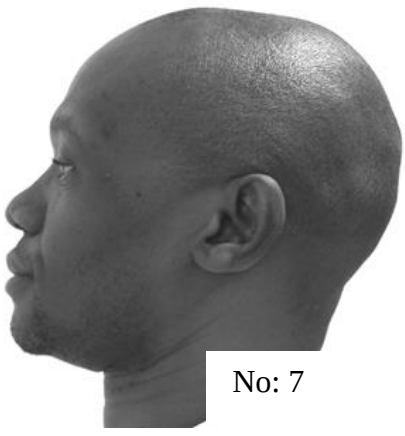

No: 3

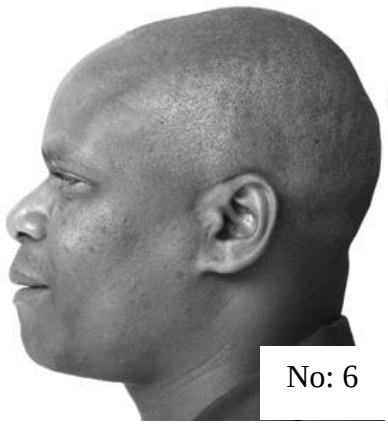

No: 9

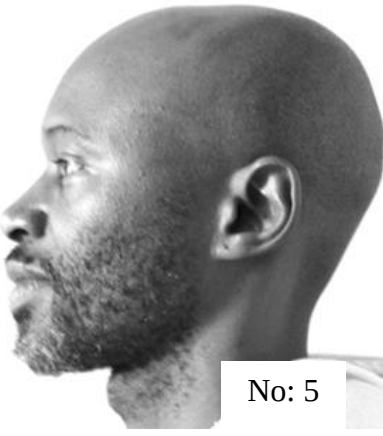



No: 4

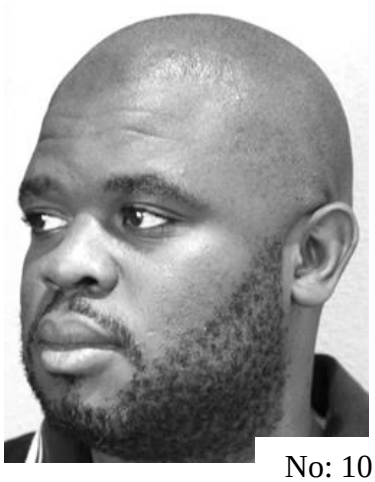

No: 8

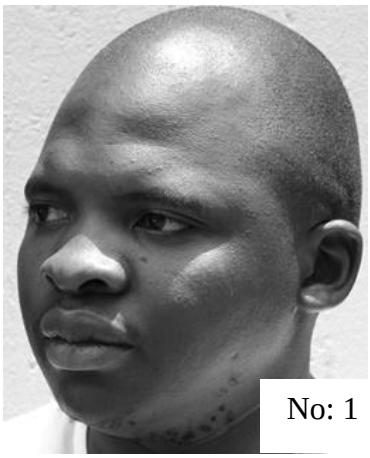

No: 7


No: 6


No: 5

Appendix 2.4.3: Facial array presenting three-quarter view photographs of male participants and male forensic facial approximation (subject D) (in the centre) constructed utilizing pooled FSTT data

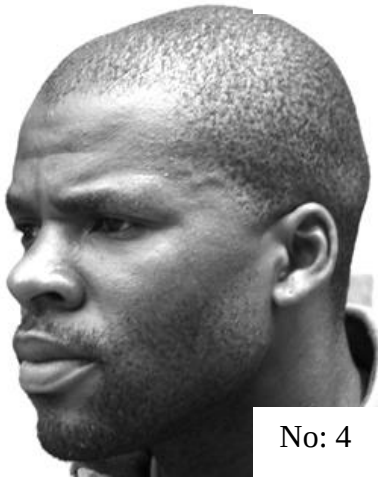

No: 10


No: 1


No: 2

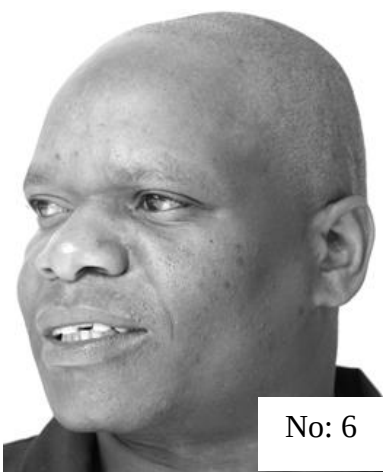

No: 3


No: 9


No: 4


No: 8


No: 7


No: 6


No: 5

Appendix 3 - INFORMATION SHEET for taking and using facial photographs

INFORMATION SHEET for taking and using facial photographs

Title: Facial Soft Tissue Thickness and Facial Recognition in Black South African Adults

Introduction

Good day, my name is Keegan Meiring. I would like to invite you to take part in this study. I am required to conduct research as part of my studies in Masters of Science in Medicine Degree (Biological Anthropology) at the University of the Witwatersrand, Faculty of Health Sciences, and Department of Anatomical Sciences.

What is the purpose of this Research Study and why have I been chosen?

The purpose of the study is to possibly assist in missing persons' cases when little to no information regarding the person is available. In South Africa very little has been done with regards to facial reconstruction and recognition, leaving a gap within an ever-growing database. You will be requested to allow the researcher to photograph your face for purpose of facial recognition testing. You will not be required to tell the research any personal information such as: your name & surname, contact details or address. On a whole, this study will collect facial soft tissue thickness measurements from special x-rays (CT scans) for picturing the bones in head to compare with photographs of the face. We wish to invite individuals of African descent, both males and females, to participate in our study. We wish to invite individuals of African ancestry, both males and females, to participate in our study.

What will happen to me if I take part in this study and what do I have to do?

If you wish to participate in this study your photographs will used for surveys (facial array tests). You will be asked to sit with your head upright and face the camera directly whilst your photograph is taken. This will take less than five minutes; thereafter, your participation will be complete.

What are the side effects or risks of taking part?

There is no physical, emotional or financial harm if you wish to participate in this study. We will only require 5 minutes of your time.

What are the Benefits for participating in the study?

There are no direct benefits if you wish you to participate in this study. Participation, however, can help this and future research which can assist with missing persons' cases.

Is there reimbursement for study participation?

There will be no reimbursement.

Do I have to take part?

Your participation in this study is entirely voluntary. It is up to you to decide whether to take part. If you do decide to take part you will be given this Information sheet to keep and will be asked to sign an informed consent form. Should you wish to discontinue participation, you may do so freely at any time and without giving a reason. If you choose not to participate or to withdraw from the study there will be no negative consequences.

Will my taking part in this study be kept confidential?

All information obtained during this study will be kept safe and confidential. Your photographs will only be used for the purposes of this research and will only be accessible to Dr. N. Briers & Mr. K. Meiring. Data may be reported in scientific journals and will not include any information that identifies you as a participant in this study. Data will be kept for five years if published or ten years if not published, after this period the will be destroyed. In the event of data sharing with other researchers for academic purposes written permission will be sought from HREC (Medical).

Who is organizing and funding the research?

The School of Anatomical Sciences of the University of the Witwatersrand has organised this study in partnership with the Department of Radiology of the Wits Donald Gordon Medical centre for the purposes of a master's degree. The researcher has received funding from the National Research Foundation (NRF) for tuition purposes study and this study will not be used for commercial gain.

Whom do I call if I have questions or problems?

If you would like more information, have any problems, concerns or questions about the study or experience research related injuries, please contact Mr. K. Meiring on 071 767 9566 alternatively my Supervisor Dr. N. Briers (Nanette.briers@wits.ac.za)

Human Research Ethics Committee (Medical), University of the Witwatersrand

HREC (Medical) contact details: Prof P Cleaton Jones, Tel: 011 717 2301, email: peter.cleaton-jones1@wits.ac.za

Ms. Z. Ndlovu/ Mr. Rhulani Mkansi/ Mr. Lebo Moeng (Administrative Officers) 011 717 2700/2656/1234/1252, zanele.ndlovu@wits.ac.za; Rhulani.mkansi@wits.ac.za; and Lebo.moeng@wits.ac.za

Appendix 4 - INFORMATION SHEET to act as an assessor for facial arrays

INFORMATION SHEET to act as an assessor for facial arrays

Title: Facial Soft Tissue Thickness and Facial Recognition in Black South African Adults

Introduction

Good day, my name is Keegan Meiring. I would like to invite you to take part in this study. I am required to conduct research as part of my studies in Masters of Science in Medicine Degree (Biological Anthropology) at the University of the Witwatersrand, Faculty of Health Sciences, and Department of Anatomical Sciences.

What is the purpose of this Research Study and why have I been chosen?

The purpose of the study is to assist in missing persons' cases when little to no information regarding the person is available. In South Africa very little has been done with regards to facial reconstruction and recognition, leaving a gap within an ever-growing database. You will be issued with 20 facial photographs, two photographs of facial reconstructions and a scoring sheet. You will be invited to state which of the 20 photographs is most like the facial reconstructions according to your opinion. You will also be requested to state which facial feature(s) between the reconstructions and photographs was/were influential in your decision. On a whole, this study will collect facial soft tissue thickness measurements from special x-rays (CT scans) for picturing the bones in head to compare with photographs of the face. We wish to invite individuals of African descent, both males and females, to participate in our study. We wish to invite individuals of black and white populations, males and females, and those who have and have no knowledge in anatomy of facial artistry to participate in this study.

What will happen to me if I take part in this study and what do I have to do?

If you wish to participate in this study you will issued with a scoring sheet and asked to match photographs to computed generated facial reconstructions. Secondly, you will be asked to indicate which feature(s) was/were influential in your decision and provided the opportunity to elaborate on your decision. The entire process will require 15 - 20 minutes. Once you have completed this task your participation in this study would be complete.

What are the side effects or risks of taking part?

There is no physical, emotional or financial harm if you wish to participate in this study. We will only require 15 – 20 minutes of your time.

What are the Benefits for participating in the study?

There are no direct benefits if you wish you to participate in this study. Participation, however, can help this and future research which can assist with missing persons' cases.

Is there reimbursement for study participation?

There will be no reimbursement.

Do I have to take part?

Your participation in this study is entirely voluntary. It is up to you to decide whether to take part. If you do decide to take part you will be given this Information Sheet to keep and will be asked to sign an informed consent form/ should you wish to discontinue participation, you may do so freely at any time and without giving a reason. If you choose not to participate or to withdraw from the study there will be no negative consequences

Will my taking part in this study be kept confidential?

All information obtained during this study will be kept confidential. Your scoring sheets will only be accessible to Dr. N. Briers & Mr. K. Meiring. Data may be reported in scientific journals and will not include any information that identifies you as a participant in this study. Data will be kept for five years if published or ten years if not published, after this period the will be destroyed. In the event of data sharing with other researchers for academic purposes written permission will be sought from HREC (Medical).

Who is organizing and funding the research?

The School of Anatomical Sciences of the University of the Witwatersrand has organised this study in partnership with the Department of Radiology of the Wits Donald Gordon Medical centre for the purposes of a master's degree. The researcher has received funding from the National Research Foundation (NRF) for tuition purposes and this study will not be used for commercial gain.

Whom do I call if I have questions or problems?

If you would like more information, have any problems, concerns or questions about the study or experience research related injuries, please contact Mr. K. Meiring on 071 767 9566 alternatively my Supervisor Dr. N. Briers (Nanette.briers@wits.ac.za) Alternatively, contact the Human Research Ethics Committee (Medical), University of the Witwatersrand, HREC (Medical) contact details: Prof P Cleaton Jones, Tel: 011 717 2301, email: peter.cleaton-jones1@wits.ac.za, Ms. Z. Ndlovu/ Mr. Rhulani Mkansi/ Mr. Lebo Moeng (Administrative Officers) 011 717 2700/2656/1234/1252, zanele.ndlovu@wits.ac.za; Rhulani.mkansi@wits.ac.za; and Lebo.moeng@wits.ac.za

Appendix 5 - INFORMATION SHEET for the use of CT scans

INFORMATION SHEET for the use of CT scans

Title: Facial Soft Tissue Thickness and Facial Recognition in Black South African Adults

Introduction

Good day, my name is Keegan Meiring. I would like to invite you to take part in this study. I am required to conduct research as part of my studies in Masters of Science in Medicine Degree (Biological Anthropology) at the University of the Witwatersrand, Faculty of Health Sciences, and Department of Anatomical Sciences.

What is the purpose of this Research Study and why have I been chosen?

The purpose of the study is to possibly assist in missing persons' cases when little to no information regarding the person is available In South Africa very little has been done with regards to facial reconstruction and recognition, leaving a gap within an ever-growing database. You will be asked to allow the researcher to use your special x-rays (CT scans) for the purposes of facial reconstruction. You will not be asked to tell the research any personal information such as: your name & surname, contact details or address. On a whole, this study will collect facial soft tissue thickness measurements from special x-rays (CT scans) for picturing the bones in head to compare with photographs of the face. We wish to invite individuals of African descent, both males and females, to participate in our study.

What will happen to me if I take part in this study and what do I have to do?

If you wish to participate in this study your CT scans will used for facial reconstruction. As you are already a patient of the Wits Donald Gordon Medical Centre, no extra CT scans will be needed. Nothing more is required of you except allowing the researcher to use the scans and no follow up is necessary for the purposes of this study.

What are the side effects or risks of taking part?

There is no physical, emotional or financial harm if you wish to participate in this study.

What are the Benefits for participating in the study?

There are no direct benefits if you wish you to participate in this study. Participation, however, can help this and future research which can assist with missing persons' cases.

Is there reimbursement for study participation?

There will be no reimbursement.

Do I have to take part?

Your participation in this study is entirely voluntary. It is up to you to decide whether to take part. If you do decide to take part you will be given this Information Sheet to keep and will be asked to sign an informed consent form. Should you wish to discontinue participation, you may do so freely at any time and without giving a reason. If you choose not to participate or to withdraw from the study, it will not affect the care and treatment that you receive from the clinic/doctor in any way.

Will my taking part in this study be kept confidential?

All information obtained during this study will be kept confidential. All your medical information will remain filed at the hospital and only authorized personnel (Dr. M. Haagensen, Dr. N. Briers & Mr. K. Meiring) will have access to this information. Data may be reported in scientific journals and will not include any information that identifies you as a participant in this study. Data will be kept for five years if published or ten years if not published, after this period the will be destroyed. In the event of data sharing with other researchers for academic purposes written permission will be sought from HREC (Medical).

Who is organizing and funding the research?

The School of Anatomical Sciences of the University of the Witwatersrand has organised this study in partnership with the Department of Radiology of the Wits Donald Gordon Medical centre for the purposes of a master's degree. The researcher has received funding from the National Research Foundation (NRF) for tuition purposes study and this study will not be used for commercial gain.

Whom do I call if I have questions or problems?

If you would like more information, have any problems, concerns or questions about the study or experience research related injuries, please contact Mr. K. Meiring on 071 767 9566 alternatively my Supervisor Dr. N. Briers (Nanette.briers@wits.ac.za)

Human Research Ethics Committee (Medical), University of the Witwatersrand

HREC (Medical) contact details: Prof P Cleaton Jones, Tel: 011 717 2301, email: peter.cleaton-jones1@wits.ac.za

Ms. Z. Ndlovu/ Mr. Rhulani Mkansi/ Mr. Lebo Moeng (Administrative Officers) 011 717 2700/2656/1234/1252, zanele.ndlovu@wits.ac.za; Rhulani.mkansi@wits.ac.za; and Lebo.moeng@wits.ac.za

Appendix 6 - INFORMED CONSENT FORM for the taking and using of facial photographs

INFORMED CONSENT FORM for the taking and using of facial photographs

Title of Study: Facial soft tissue thickness and facial recognition in black South African adults

I confirm that I have been informed by a Mr. K. Meiring about the nature of the study. I have also read/it was read to me and I understood the information sheet and have had the opportunity to ask questions.

I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, without my medical care or legal rights being affected.

I understand that photographs will be taken of my face for the purposes of compiling a facial array test and may be looked at by Mr. K. Meiring and Dr. N. Briers. I am aware that I my photographs will be displayed with several other photographs for the purposes of testing facial recognition. I understand that my photograph will be anonymously processed into a computerised system. Data will be kept for five years if published or ten years if not published, after this period the data will be destroyed.

Should you wish to contact us at any stage regarding consent, contact Mr. K. Meiring (547058@students.wits.ac.za) or Dr. N. Briers (Nanette.briers@wits.ac.za).

I agree to take part in the above mentioned study. I hereby give consent for my records to be used as per the above mentioned conditions and for the purposes of research.

Name and Surname Signature/Mark or Thumbprint Date:

Translator/Other Person Explaining Informed Consent (Designation)

Printed name Signature Date:

Witness (If applicable):

Printed name Signature Date:

Appendix 7 - INFORMED CONSENT FORM to act as an assessor for facial arrays

INFORMED CONSENT FORM to act as an assessor for facial arrays

Title of Study: Facial soft tissue thickness and facial recognition in black South African adults

I confirm that I have been informed by a Mr. K. Meiring about the nature of the study. I have also read/it was read to me and I understood the information sheet and have had the opportunity to ask questions.

I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, without my medical care or legal rights being affected.

I understand that that I am about to participate as an assessor for a facial array and that the results may be looked at by Mr. K. Meiring and Dr. N. Briers. I am aware that my assessment will be used for the purposes of testing facial recognition. Test results, including details such as: sex and ethnicity will be anonymously processed into a computerised system. Data will be kept for five years if published or ten years if not published, after this period the data will be destroyed.

Should you wish to contact us at any stage regarding consent, contact Mr. K. Meiring (547058@students.wits.ac.za) or Dr. N. Briers (Nanette.briers@wits.ac.za).

I agree to take part in the above mentioned study. I hereby give consent for my records to be used as per the above mentioned conditions and for the purposes of research.

Name and Surname Signature/Mark or Thumbprint Date:

Translator/Other Person Explaining Informed Consent(Designation)

Printed name Signature Date:

Witness (If applicable):

Printed name Signature Date:

Appendix 8 - INFORMED CONSENT FORM for the use of CT scans

INFORMED CONSENT FORM for the use of CT scans

Title of Study: Facial soft tissue thickness and facial recognition in black South African adults

I confirm that I have been informed by a Mr. K. Meiring about the nature of the study. I have also read/it was read to me and I understood the information sheet and have had the opportunity to ask questions.

I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, without my medical care or legal rights being affected.

I understand that sections of any of my medical records may be looked at by Mr. K. Meiring and authorized parties (Dr. M. Haagensen & Dr. N. Briers). I am aware that I my CT scans will be used to measure my facial soft tissue thickness. Test results, including details such as: sex, age and ethnicity will be anonymously processed into a computerised system. Data will be kept for five years if published or ten years if not published, after this period the data will be destroyed.

Should you wish to contact us at any stage regarding consent, contact Mr. K. Meiring (547058@students.wits.ac.za) or Dr. N. Briers (Nanette.briers@wits.ac.za).

I agree to take part in the above mentioned study. I hereby give consent for my records to be used as per the above mentioned conditions and for the purposes of research.

Name and Surname Signature/Mark or Thumbprint Date:

Translator/Other Person Explaining Informed Consent(Designation)

Printed name Signature Date:

Witness (If applicable):

Printed name Signature Date:

R14/49 Mr Keegan Oliver Meiring

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)

CLEARANCE CERTIFICATE NO. M170666

NAME: (Principal Investigator)	Mr Keegan Oliver Meiring
DEPARTMENT:	Anatomical Sciences University of the Witwatersrand Wits Donald Gordon Medical Centre
PROJECT TITLE:	Facial Soft Tissue Thickness and Facial Recognition Rates of Black South African Adults
DATE CONSIDERED:	30/06/2017
DECISION:	Approved unconditionally
CONDITIONS:	
SUPERVISOR:	Dr Nanette Briers
APPROVED BY:	CB Penny 1/9/2017 Professor CB. Penny Co-Chairperson, HREC (Medical)
DATE OF APPROVAL:	31/08/2017

This clearance certificate is valid for 5 years from date of approval. Extension may be applied for.

DECLARATION OF INVESTIGATORS

To be completed in duplicate and **ONE COPY** returned to the Research Office Secretary in Room 10004,10th floor, Senate House/3rd floor, Phillip Tobias Building, Parktown, University of the Witwatersrand. I/We fully understand the conditions under which I am/we are authorised to carry out the above-mentioned research and I/we undertake to ensure compliance with these conditions. Should any departure be contemplated, from the research protocol as approved, I/we undertake to resubmit to the Committee. **I agree to submit a yearly progress report**. The date for annual re-certification will be one year after the date of convened meeting where the study was initially reviewed. in this case, the study was initially review June and will therefore be due in the month of June each year. Unreported changes to the application may invalidate the clearance given by the HREC (Medical).

Principal Investigator Signature

Date 01 - 09 - 2017

PLEASE QUOTE THE PROTOCOL NUMBER IN ALL ENQUIRIES