THREE-DIMENSIONAL CRANIOMETRICS IDENTIFICATION MODEL AND CEPHALIC INDEX CLASSIFICATION OF MALAYSIAN SUB-ADULTS: A MULTI-SLICE COMPUTED TOMOGRAPHY STUDY

SHARIFAH NABILAH BT SYED MOHD HAMDAN

FACULTY OF DENTISTRY
UNIVERSITI MALAYA
KUALA LUMPUR

2024

THREE-DIMENSIONAL CRANIOMETRICS IDENTIFICATION MODEL AND CEPHALIC INDEX CLASSIFICATION OF MALAYSIAN SUB-ADULTS: A MULTI-SLICE COMPUTED TOMOGRAPHY STUDY

SHARIFAH NABILAH BT SYED MOHD HAMDAN

THESIS SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

FACULTY OF DENTISTRY
UNIVERSITI MALAYA
KUALA LUMPUR

2024

UNIVERSITI MALAYA

ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: **SHARIFAH NABILAH BT SYED MOHD HAMDAN**

Matric No: **S2007791**

Name of Degree: **DOCTOR OF PHILOSOPHY**

Title of Project Paper/Research Report/Dissertation/Thesis: **THREE-DIMENSIONAL CRANIOMETRICS IDENTIFICATION MODEL AND CEPHALIC INDEX CLASSIFICATION OF MALAYSIAN SUB-ADULTS: A MULTI-SLICE COMPUTED TOMOGRAPHY STUDY**

Field of Study: **HEALTH (DENTAL STUDIES)**

I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work;
(2) This Work is original;
(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
(4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
(5) I hereby assign all and every right in the copyright to this Work to the University of Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
(6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate's Signature Date: 6/7/2024

Subscribed and solemnly declared before,

Witness's Signature Date: 6/7/2024

Name:

Designation:

THREE-DIMENSIONAL CRANIOMETRICS IDENTIFICATION MODEL AND CEPHALIC INDEX CLASSIFICATION OF MALAYSIAN SUB-ADULTS: A MULTI-SLICE COMPUTED TOMOGRAPHY STUDY

ABSTRACT

Human identification is the main goal in anthropological and forensic investigations such as examination of ancient skeletons, investigations at criminal related scenes, or due to mass disasters. The primary focus is to determine the biological profile of unknown individuals by estimating their sex and ethnicity. Sex and ethnicity estimation methods utilised in adult are less effective in sub-adults due to varied cranium patterns during growth. Therefore, this study aimed to develop three-dimensional (3D) craniometric models in Malaysian sub-adults for sex and ethnicity estimation, and to establish a cephalic index (CI) classification for Malaysian sub-adults. A total of 521 cranial multi-slice computed tomography (MSCT) dataset of sub-adult Malaysians aged 0 to 20 with Malay, Chinese, and Indian ethnicities were obtained. MIMICS software version 21.0 (Materialise, Leuven, Belgium) was used to construct 3D models and plane-to-plane (PTP) protocol was utilised to measure 14 selected craniometric parameters. Discriminant function analysis (DFA), binary logistic regression (BLR), and several machine learning (ML) algorithms (random forest (RF), support vector machines (SVM), and linear discriminant analysis (LDA)) were used to statistically analyse the data. Additionally, CI was calculated according to the following equation: cephalic width/cephalic length×100. This present study demonstrated a minimal degree of sexual dimorphism in the cranium of individuals below the age of six, and the level was then increased with age. All the age groups, except for 0–2 years and 3–6 years, exhibited reliable sex estimation with a high accuracy percentage (≥75%) when tested using DFA and BLR. As for the ethnicity estimation models, a high similarity of craniometric measurements between Chinese and Malays (as compared to Indians and Malays, and Chinese and Indians) was demonstrated.

This resulted in the highest classification accuracy obtained by Indians, followed by Chinese and Malays in the age groups of 10–12 years and 16–20 years. Moreover, ML methods obtained slightly higher accuracy rates than classical methods for sex (RF: 73% vs BLR: 66.9% and DFA: 61.6%) and ethnicity estimation (LDA: 58% vs DFA: 57.5%) using sub-adults' crania. In addition, the modified CI of Malaysian sub-adults were found to be as follows: dolichocephalic, 78.8 or less; mesocephalic, 78.9–89.0; brachycephalic, 89.1–94.0; and hyperbrachycephalic, 94.1 or higher. Hence, the proposed CI index indicated that the dominating type of head for Malaysian sub-adults was mesocephalic (66.4%), followed by dolichocephalic (18.4%), brachycephalic (12.3%), and hyperbrachycephalic (2.9%). The present study has demonstrated that sex and ethnicity estimation of sub-adults can be effectively performed by assessing the cranium via 3D virtual anthropometry. To the best of our knowledge, this was the preliminary study that described craniometric variations of multi-ethnic groups in Malaysian sub-adult population using MSCT data. Ultimately, this present study has bridged the gap of population-specific cranial data in Malaysian sub-adults.

Keywords: Cranium, sub-adults, sex and ethnicity estimation, multi-slice computed tomography, cephalic index.

MODEL PENGENALAN KRANIOMETRIK TIGA DIMENSI DAN PENGELASAN INDEKS SEFALON SUB-DEWASA: KAJIAN TOMOGRAFI KOMPUTER BERBILANG HIRISAN

ABSTRAK

Pengenalpastian manusia ialah matlamat utama dalam penyiasatan antropologi dan forensik seperti pemeriksaan rangka purba, penyiasatan di tempat kejadian jenayah, atau akibat berlakunya bencana alam berskala besar. Fokus utama ialah untuk menentukan profil biologi individu-individu yang tidak dikenali dengan menganggarkan jantina dan kumpulan etnik. Corak kranium yang berbeza-beza semasa pertumbuhan menandakan bahawa pendekatan anggaran jantina dan kumpulan etnik yang lazim diguna pakai bagi orang dewasa adalah kurang berkesan bagi golongan sub-dewasa. Oleh itu, kajian ini bermatlamat untuk mengembangkan model kraniometrik tiga dimensi (3D) bagi golongan sub-dewasa Malaysia bagi tujuan anggaran jantina dan kumpulan etnik, serta membina pengelasan indeks sefalon (*cephalic index*, CI) bagi golongan sub-dewasa Malaysia. Sejumlah 521 dataset imbasan tomografi komputer berbilang hirisan (*multi-slice computed tomography*, MSCT) bagi sub-dewasa Malaysia berumur 0 hingga 20 tahun, daripada kumpulan etnik Melayu, Cina, dan India telah diperoleh. Perisian MIMICS versi 21.0 telah digunakan untuk membina model tiga dimensi sementara protokol satah-ke-satah telah digunakan untuk mengukur 14 parameter kraniometrik yang terpilih. Analisis fungsi diskriminan (*discriminant function analysis*, DFA), regresi logistik binari (*binary logistic regression*, BLR), dan beberapa algoritma pembelajaran mesin (*machine learning*, ML) iaitu *random forest* (RF), mesin sokongan vektor (*support vector machines*, SVM) dan analisis diskriminan linear (*linear discriminant analysis*, LDA) telah digunakan untuk menganalisis data secara statistik. Tambahan lagi, CI telah dihitung berdasarkan persamaan berikut: kelebaran sefalon/panjang sefalon x 100. Kajian ini mempamerkan takat minimum dimorfisme jantina dalam kranium individu bawah

umur 6 tahun, dan takat ini bertambah dengan peningkatan usia. Apabila sampel dibahagikan kepada kumpulan umur yang berbeza-beza, semua kumpulan umur, kecuali 0-2 tahun dan 3-6 tahun, mempamerkan anggaran jantina dengan peratusan ketepatan yang tinggi (≥75%) apabila diuji dengan menggunakan DFA dan BLR. Bagi model anggaran kumpulan etnik pula, terdapat persamaan ukuran kraniometrik yang tinggi telah dipamerkan antara Cina dan Melayu (berbanding antara India dan Melayu, dan antara Cina dan India). Ini menghasilkan ketepatan klasifikasi tertinggi yang diperoleh oleh orang India, diikuti dengan orang Cina dan orang Melayu dalam kumpulan umur 10-12 tahun dan 16-20 tahun. Tambahan pula, kaedah ML memperoleh kadar ketepatan yang lebih tinggi sedikit berbanding dengan kaedah klasik bagi jantina (RF: 73% vs BLR: 66.9% dan DFA: 61.6%) dan anggaran etnik (LDA: 58% vs DFA: 57.5%) menggunakan kranium sub-dewasa. Sementara itu, CI sub-dewasa Malaysia yang diubah suai adalah didapati seperti berikut: dolikosefalik, 78.8 atau kurang; mesosefalik, 78.9–89.0; brakhisefalik, 89.1–94.0; dan hiperbrakhisefalik, 94.1 atau lebih tinggi. Maka, index CI yang dicadangkan menandakan bahawa jenis kepala yang mendominasi bagi sub-dewasa orang Malaysia adalah mesosefalik (66.4%), diikuti dengan dolikosefalik (18.4%), brakhisefalik (12.3%), dan hiperbrakhisefalik (2.9%). Kajian ini telah menunjukkan bahawa anggaran jantina dan etnik sub-dewasa boleh dilakukan secara berkesan, dengan membuat penilaian terhadap kranium melalui antropometri secara maya. Setakat pengetahuan terbaik kami, ini kajian awal yang menerangkan variasi kraniometrik kumpulan multi-etnik dalam populasi sub-dewasa orang Malaysia menggunakan data MSCT. Pada dasarnya, kajian ini telah dapat merapatkan jurang data kranial khusus bagi populasi sub-dewasa di Malaysia.

Kata kunci: Kranium, sub-dewasa, anggaran jantina dan kumpulan etnik, tomografi komputer berbilang hirisan, indeks sefalon.

ACKNOWLEDGEMENTS

I am immensely grateful to the following individuals who have played pivotal roles in the completion of this project. Their unwavering support and encouragement have been truly invaluable. First and foremost, I extend my deepest appreciation to my exceptionally dedicated supervisors: Prof. Dr. Norliza Ibrahim, Dr. Rabiah Al-Adawiyah Rahmat, Prof. Dr. Fathilah Abdul Razak, Assoc. Prof. Dr. Khairul Azmi Abd Kadir, Dr. Erma Rahayu Mohd Faizal Abdullah, and Dr. Wan Izlina Wan Ibrahim. Your guidance, advice, and unwavering belief in the potential of this research project have been instrumental in shaping its outcome. I am truly grateful for the countless hours you have dedicated to perfecting my work, presentations, and manuscripts.

The conduct of this research would not have been possible without the help and support from the Radiology Department, University Malaya Medical Centre (UMMC). Thank you, Mr. Mohamad Zamri Mohamad Zin for providing your assistance in obtaining MSCT scans, and Mr. Ahmad Muammar Kasmin for the time that you have spent in acquiring and collecting the data during the data collection process.

Lastly, I would like to express my deep appreciation to my family and friends for their continuous support. Your staunch encouragement and support has been a constant source of strength and inspiration. To my beloved mother, Salasiah bt Abdul Murad, thank you for everything you have done and all the sacrifices you have made for me. To my husband, Mohd Hazwan bin Rusli, and my children, Harith Ayyash and Sofea Ayana, your presence, love, understanding, and encouragement have brought me joy and fuelled my determination to succeed. Thank you, everyone.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS AND ABBREVIATIONS

3D	:	Three-dimensional
AI	:	Artificial intelligence
ANN	:	Artificial neural network
BAW	:	Biasteronic width
BLR	:	Binary logistic regression
CBL	:	Cranial base length
C.C	:	Canonical correlation
CH	:	Cranial height
CI	:	Cephalic index
DFA	:	Discriminant function analysis
DICOM	:	Digital imaging and communications in medicine
DL	:	Deep learning
DNA	:	Deoxyribonucleic acid
DT	:	Decision tree
FC	:	Frontal chord
GUI	:	Graphic user interface
HR	:	Hit rates
FML	:	Foramen magnum length
FMW	:	Foramen magnum width
IPW	:	Interporion width
KNN	:	K-nearest neighbors
kV	:	Kilovolt
LCL	:	Lateral cranial length
LCL (R)	:	Lateral cranial length right

LCL (L)	:	Lateral cranial length left
LDA	:	Linear discriminant analysis
mAs	:	Milliampere-seconds
MCL	:	Maximum cranial length
MCW	:	Maximum cranial width
MIMICS	:	Materialise interactive medical image control system
ML	:	Machine learning
MLP	:	Multi-layer perceptron
MSCT	:	Multi-slice computed tomography
NOL	:	Nasio-occipital length
OC	:	Occipital chord
PC	:	Parietal chord
PCA	:	Principal component analysis
PLS	:	Partial least squares regression
PMCT	:	Post-mortem computed tomography
ProCoords	:	Procrustes coordinates
TEM	:	Technical error of the measurement
R	:	Coefficient of reliability
RBF	:	Radial basis function
RF	:	Random forest
rTEM	:	Relative technical error of the measurement
SD	:	Standard deviation
SDP	:	Sexual dimorphism percentage
SPSS	:	Statistical package for the social sciences
SVM	:	Support vector machines
U.C	:	Unstandardised coefficients

LIST OF APPENDICES

CHAPTER 1: INTRODUCTION

1.1 Background

Human identification is the main goal in anthropological and forensic investigations such as ancient skeletal examination, investigation at criminal-related scenes, or due to mass disasters (Gill et al., 1988; Klepinger, 2006; Ousley et al., 2009; Rhine, 1990). The primary focus of identification is to determine the biological profile of unknown individuals by estimating their age, sex, ethnicity, and stature (Kanchan et al., 2013). Various skeletal components have been used to identify these biological profiles in adults (Ekizoglu et al., 2017; Kranioti et al., 2018; Ramamoorthy et al., 2016; Singh & Pathak, 2013; Spradley 2021). However, methods to identify sub-adults are lacking. The complex biological profile of sub-adults may contribute to this observation. Furthermore, evidence of sex and ethnicity of a sub-adult skeleton is obscured by skeletal growth indicators, as the growth of sub-adults is an ongoing process (McDowell et al., 2012). This has resulted in the lack of established sex and ethnicity estimation methods for sub-adults.

The accuracy of the biological profile largely depends on the availability of the skeletal components and their state of preservation, such as the cranium and postcranial skeleton (Scheuer & Black, 2000). Cranium is universally recognised as the best bone structure for estimating ethnicity and the second-best for estimating sex (Byers, 2016; Krogman & Iscan, 1962; Pickering & Bachman, 2009). This is because the cranium is highly durable even in extreme environmental conditions, resulting in better-preserved morphological features (Gonzalez et al., 2011; Naikmasur et al., 2010; Sangvichien et al., 2007).

Traditionally, sex and ethnicity estimation using cranium is conducted using a nonmetric method through visual assessments and statistical analyses (Birkby et al., 2008; Hefner, 2009; Ousley & Jantz, 2005; Ousley et al., 2009; Rhine, 1990). However, this method is subjective and observer-dependent. To overcome these drawbacks, metric methods have been widely used since they are less dependant on the observer's judgement, rely on standardised measurements of the cranium, and the results can be statistically analysed (Gillet et al., 2020; Pretorius et al., 2006; Sierp & Henneberg, 2015; Tunis et al., 2017). Previous studies on sex and ethnicity estimation have been more geared towards craniometric data because of its methodical nature and ability to be performed under methods of statistical analysis such as discriminant function analysis (DFA) or binary logistic regression (BLR; Gonzalez, 2012; Hsiao et al., 2010; Noble et al., 2019; O'Donnell et al., 2017; Sierp & Henneberg, 2015; Teodoru-Raghina et al., 2017; Zaafrane et al., 2018). DFA and BLR provide relatively simple, objective, and accurate means to estimate sex and ethnicity. Variability in skeletal measurements enables mathematical analysis and statistical processing of the collected data (Santos et al., 2014). However, BLR is limited to sex estimation, while DFA can be utilised for estimating both sex and ethnicity (Nikita & Nikitas, 2020).

In recent times, artificial intelligence (AI) and machine learning (ML) techniques have been applied in diverse fields such as bioarcheology and forensic anthropology (Muzzall, 2021). In forensic anthropology, these techniques have been used to estimate the sex and ethnicity of adult individuals (Hefner et al., 2015; Imaizumi et al., 2020; Pozzi et al., 2020; Toneva et al., 2021). ML works by utilising computer-aided cranial measurements based on three-dimensional (3D) models and making predictions without being explicitly programmed to do so. Multiple ML techniques such as linear discriminant analysis (LDA), support vector machine (SVM), logistic regression (LR), artificial neural networks (ANN), decision trees (DT), and random forest (RF) have been used for estimating sex and ethnicity (Hefner et al., 2015; Imaizumi et al., 2020; Pozzi et al., 2020; Toneva et al., 2021; Toy et al., 2022). However, the comparison of classification accuracy between these ML techniques in sub-adults has been less explored (Ortega et al., 2021).

The cephalic index (CI) is an important index in assessing the head shape of individuals. The CI and head shape are greatly affected by factors of age, sex, and ethnicity (Khanduri et al., 2021). Furthermore, they are also reported to be influenced by geographical variation (Hanihara et al., 1982). There have been several reports measuring CI in children with normal development (Koizumi et al., 2010; Likus et al., 2014; Nam et al., 2021; Waitzman et al., 1992). However, studies have been conducted mostly in Caucasians (Likus et al., 2014; Waitzman et al., 1992). In the Asian population, CI studies have been conducted on Japanese and Korean children (Koizumi et al., 2010; Nam et al., 2021). With the understanding that CI varies between different populations, hence, applying CI of other populations to the Malaysian population would not be sensible given that Malaysia consists of three major ethnic groups, namely Malay, Chinese, and Indian. Therefore, it is necessary to develop the CI classification for the current Malaysian sub-adult population.

1.2 Problem statement

Determining sex and ethnicity is crucial for establishing skeletal profiles of unidentified individuals in forensic investigations (Pretorius et al., 2006). Sex and ethnic variances specific to the cranium have been widely studied in various populations around the world. However, the standard model for such estimations based on cranium has not been set for global populations. In addition, existing methods for sex determination using sub-adult cranium are less reliable and have yielded variable degrees of accuracy (Hsiao et al., 2010; Noble et al., 2019; O'Donnell et al., 2017; Teodoru-Raghina et al., 2017). Furthermore, skeletal samples of sub-adults used by anthropologists are often derived from antiquated and/or foreign skeletal collections that were amassed in the late 19th to early 20th century (Cardoso, 2008; Galdames et al., 2008). These collections predominantly comprise skeletons of European or African descent and primarily reflect the demographics of pre-1950s America, leaving Asians as a significantly understudied group (Hunt & Albanese, 2005; Komar & Buikstra, 2008; Komar & Grivas, 2008). Given the wide range of global secular change, human, temporal, and geographical variations that have occurred since, these samples may not be reliably applied to all human groups (Algee-Hewitt, 2016; Wescott & Jantz, 2005).

Temporal or secular changes are physical changes that may take place within a population as a result of abrupt shifts in lifestyle or exposure to a different environment. These changes may be due to an improvement or a decline in environmental conditions, mostly nutrition (Spradley, 2006). In addition, distinct genetic and physical characteristics between the modern and the old (>100 years) populations have been reported in various populations such as the Native American tribe (Sutphin et al., 2014), African and White (Moore-Jansen, 1989), Japanese (Hossain et al., 2005; Nagaoka et al., 2012), Mexican (Little et al., 2006), Croatian (Buretić-Tomljanović et al., 2006), and Indian (Saini, 2014). In a study by Nagaoka et al. (2012), significant variations were reported in several cranial measurements, including maximum cranial length, maximum cranial breadth, and upper facial height, when compared to data from populations of different time periods. Moreover, a shift toward brachy cephalisation or increased cranial breadth was reported in the contemporary Indian population after India achieved independence in 1947. This has resulted in varying accuracies of sex estimation models between two successive populations of the same geographical region (Saini, 2014). Therefore, these findings indicated the importance of updating local forensic anthropology database.

Sex and ethnicity estimation models are known to be population-specific (Franklin et al., 2013b). The accuracy rates of estimation models have been found to be low when applied to different populations (Kim, 2009; Toneva et al., 2020; Yang et al., 2020). In addition, low accuracy rates of sex and ethnicity estimation are influenced by inter-population disparities in sexual dimorphism expression, as some communities have robust characteristics while others have gracile characteristics of both sexes (Green & Curnoe, 2009). The need for population-specificity for sex estimation was demonstrated by Franklin et al. (2013a), where models developed for a South African population (İşcan & Steyn, 1999) were applied to a Western Australian sample. The findings reported a high sex bias in which misclassification of males occurred more than in females (>30%). Therefore, a population-specific model should be developed to accommodate the large variance of craniometric measurements across the globe.

Conventional radiographs have been widely used in sub-adult studies (Gonzalez, 2012; Hsiao et al., 2010; O'Donnell et al., 2017; Sprowl, 2013), but it comes with limitations such as superimposition of anatomical structures and inconsistent orientation. Additionally, plain film radiographs exhibit distortion due to parallax and magnification errors, which hinders accurate metric data collection (Bontrager & Lampignano, 2001; Mantini & Ripani, 2009; Schroeder et al., 1997). To address these limitations, computed tomography (CT) scans have emerged as a valuable alternative. CT scans allow researchers to access high-resolution skeletal data from diverse populations without the need for physical skeletal materials. Unlike plain film radiographs, CT scans provide unambiguous images of the anatomical region of interest without distortion, magnification errors, or superimposition of structures (Hildebolt et al., 1990). Moreover, CT scan data contain embedded technical information, such as slice thickness, voxel size, and user-selectable parameters, which can aid in the analysis. Retrospective data acquisition from clinical settings, where CT scans are routinely performed, offers an ideal opportunity for advancing forensic anthropological research, especially concerning the analysis of sub-adult skeletal remains (Lottering et al., 2014).

The variation across populations poses a significant challenge in establishing a standard CI for accurate anthropological assessment. The lack of a standardised CI can be associated with several factors such as genetic diversity, geographical location, historical migration patterns, and cultural practices. Therefore, developing a population-specific CI will facilitate and improve anthropological studies, medical practices, and a more inclusive understanding of human diversity. In view of the aforementioned gaps, the aim of this study is to fill a longstanding void in sub-adult sex and ethnicity estimation and CI classification by utilising a substantial sample of MSCT scans.

1.3 Research questions

To date, a significant research gap exists in the development of sex and ethnicity estimation models specifically for sub-adults using MSCT scans, with no emphasis on the Malaysian sub-adult population. Given the population-specific nature of sex and ethnicity estimation, it is imperative to establish more reliable and precise estimation methods for the growing sub-adult population. Therefore, the aim of this study was to address the following key research questions:

1) What are the craniometric models used for sex and ethnicity estimation in Malaysian sub-adults?
2) What are the differences between the accuracy of discriminant function analysis (DFA) and binary logistic regression (BLR) models for sex estimation in Malaysian sub-adults?
3) What is the accuracy of DFA models for ethnicity estimation in Malaysian sub-adults?
4) What are the differences between the validity of machine learning (ML) algorithms and classical statistical methods (DFA and BLR) for sex and ethnicity estimation models in Malaysian sub-adults?
5) What is the normal range of CI classification for Malaysian sub-adults?

1.4 Aims and objectives

The main aims of this multi-slice computed tomography (MSCT) study were two-fold:

1) To develop three-dimensional (3D) craniometric models for sex and ethnicity estimation in Malaysian sub-adults; and
2) To establish a cephalic index (CI) classification for Malaysian sub-adults.

This study was set out with the following objectives in view:

1) To compare the accuracy of discriminant function analysis (DFA) and binary logistic regression (BLR) models for sex estimation in Malaysian sub-adults.
2) To assess the accuracy of DFA models for ethnicity estimation in Malaysian sub-adults.
3) To compare the validity of sex and ethnicity estimation models between machine learning (ML) algorithms and classical statistical methods (DFA and BLR) in Malaysian sub-adults.
4) To determine the normative range of CI classification for Malaysian sub-adults.

1.5 Hypothesis

The following hypotheses were tested in this study to achieve the aims and objectives set out above:

1) Sexually dimorphic differences in the cranium are predictable for each age group when utilising statistical analyses.
2) Ethnic differences in the cranium are predictable across all age groups and between males and females when utilising DFA.

1.6 Significance of the study

The identification of an unknown individual is important in forensic investigations. Numerous methods have been introduced to improve the accuracy of sex and ethnicity estimation from an unknown cranium. However, these estimations lack population inclusivity, as most cranial studies are specific to international samples that do not reflect the biological variety and growth patterns of Malaysian sub-adults. Therefore, the significance of the present study is to develop reliable and population-inclusive sex and ethnicity estimation models for Malaysian sub-adults using craniometrics. These cranium estimation models will significantly assist in various medico-legal investigations, including identifying individuals with unknown identities, determining refugee status, investigating cases of human trafficking, and facilitating the identification of disaster victims.

This research will enhance and update the availability of normative reference data and growth changes in the cranium for the Malaysian sub-adult population. Such data should be of value for clinicians as they can use the 3D data and models to facilitate diagnosis, carry out treatment planning pertaining to different cranial abnormalities, and perform post-operative cranial surgery follow-ups. In addition, this age-specific normative data can provide important information during surgery and demonstrate morphologic variation that may be encountered by the clinicians. This will enrich their knowledge and skills in managing diverse cases. In addition, the data can be used to develop a protective headgear for children. Interestingly, this data can also contribute to the knowledge of human evolution, variation, and developmental biology.

The lack of recent documented sub-adult skeletal data has impeded new research in sex and ethnicity estimation. Therefore, the rapid evolution of technology such as MSCT scans has revolutionised forensic and medical investigations. Having emerged as one of the most widely used imaging methods today, MSCT has the capability to overcome the limitations associated with conventional radiographic databases, such as superimposition and image distortion. It utilises 3D multi-planar and volume-rendered reconstructions which act as virtual models of skeletal components. Most importantly, it should be recommended that the data obtained from this study be saved in a data bank so that it will be accessible for future use.

1.7 Definitions of postnatal life stages used in this study

Substantial differences in the terminologies used to define the stages of postnatal life is common between countries. Universally, the term 'sub-adult' refers to any stage of life that is not yet truly an adult, which overlaps with the achievement of final adult stature (females:16-18 years, males: 18-21 years). Additionally, this signifies the completion of the permanent molars' eruption and the closure of the spheno-occipital synchondrosis.

The term 'perinate' denotes the newborn stage, typically referring to the time immediately after birth. Infancy refers to the first three years of life which is characterised by rapid growth and development in various aspects such as motor skills, cognitive abilities, and language acquisition (Bogin, 2020; Scheuer & Black, 2000). The childhood stage generally refers to the period of human development between infancy and juvenile. It encompasses a wide range of ages, typically from three to seven years. During this stage, children experience significant physical, cognitive, and social development (Black & Maat, 2010).

The juvenile stage signifies the completion of brain growth and the eruption of permanent dentition. Females reach this stage between seven and 10 years of age, while males reach it at around 12 years of age (Black & Maat, 2010). Adolescence is the transitional period between childhood and adulthood. It is characterised by significant physical, psychological, and social changes, including the onset of puberty. Females experience puberty and the adolescent growth spurt between 10 and 16 years of age, while males go through these changes between 12 and 18 years of age (Sinclair, 1973).

1.8 Organisation of chapters

This dissertation has been organised into six chapters:

Chapter One of this thesis includes the background of study, research problem, aims and objectives, hypothesis, and significance of the study. **Chapter Two** provides a literature review on various topics related to the research study. It covers the methods used for sex and ethnicity estimation, applications of MSCT scanning, AI technology in forensic casework, methods for developing population-specific standards for different global populations, and CI classification. The purpose of this chapter is to provide a comprehensive overview of existing knowledge and research relevant to the present study.

Chapter Three outlines the materials studied and the methods used to analyse the data collected in this study. **Chapter Four** presents the results from the analysis performed on the data. It includes the findings and statistical analyses derived from the research study. The results will be presented in the form of tables and graphs to facilitate understanding and interpretation.

Chapter Five will then discuss the results presented in Chapter Four in further detail. It provides interpretation, analysis, and contextualisation of the results within the broader context of the research objectives and relevant literature. This chapter highlights patterns, trends, and significant findings, as well as discussing their implications and potential applications. It also addresses the study's limitations and recommends areas of improvement for future research. **Chapter Six** summarises the entire research project, including a recap of the research objectives, methodology, and findings.

CHAPTER 2: LITERATURE REVIEW

2.1 Growth and development of the cranium

Human growth is a revolutionised process of size and morphology during the development stage of an individual. The presence of variation in growth between males and females, as well as within population groups, has been documented (Bulygina et al., 2006). Thus, understanding the human skeletal anatomy is fundamental for forensic anthropologists when performing sex or ethnicity estimation. The following sections outline a description of the human skull to provide a better insight into its anatomy, growth, and developmental process.

2.1.1 Skull anatomy

The human skull is formed by two major components, namely the neurocranium and the viscerocranium. Neurocranium includes calvarium, which creates cranial vault, and basicranium, which develops the skull base (Soames, 1995). Viscerocranium is the structure that supports the facial skeleton, consisting of face, palate, pharyngeal, temporal, and auditory bones. However, the anatomy of viscerocranium will not be discussed as facial skeleton has not been included in this research.

2.1.2 Growth of neurocranium

Neurocranium, or the braincase, consists of calvarium and basicranium. Neurocranium constitutes most of the protective components of the brain (Scheuer, 2000).

2.1.2.1 Calvarium

Calvarium, or cranial vault, is primarily composed of flat bones: parietal bones and paired frontal, interparietal occipital, and squamous temporal bone segments. These bone plates are connected by fibrous connective tissue bands known as calvarial sutures. Sutures of the calvarium include the coronal, metopic, lambdoid, and sagittal sutures (Beederman et al., 2014).

Calvarial sutures inhibit early bone fusion and allow the expansion of the skull throughout the prenatal and newborn stages of development. Additionally, it enables the bones to move throughout the birth process, preventing the skull from distorting during delivery and functioning as shock absorbers to external pressures. Hence, suture patency is unquestionably necessary for appropriate craniofacial growth (Twigg & Wilkie, 2015). Sutures normally remain unossified well into adolescence. If any of the sutures close too early, growth in that area may be prevented, forcing the growth to occur in another direction. This will alter the shape of the skull and result in an abnormal head shape (Dias et al., 2020).

Fontanelles, or soft spots, are the gaps between bones that stay open in newborns. It is one of the most noticeable anatomical aspects of the infant's skull. These structures are the remnants of ectomeninx, the neural crest-derived tissue from which the calvarial bones emerge (Jiang et al., 2002). After birth, the apposition of bones along the borders of the fontanelles rapidly closes these open areas. The bones will remain separated for many years and finally merge in adulthood.

Calvarium grows fastest during the first two years after birth. When the temporal and parietal lobes of the brain expand, the calvarium elongates, and the cranial base descends (Cunningham et al., 2016). Calvarium is typically 25% of its adult size at birth, 50% at six months, 65% at one year, and 90% at seven years, with continuous growth until late childhood (Kamdar et al., 2009). The brain and calvarium reach adulthood at around the same time, which is at around 10 years (Lieberman, 2011).

2.1.2.2 Basicranium

Basicranium, or the cranial base, is the most complex substructure within the skull. It connects the vault above with the face below and serves as a major integration center of the skull (Lieberman et al., 2000). Basicranium begins as cartilage and changes into bone by replacing hyaline cartilage at their suture lines. Cartilage acts as a template for new bone to be entirely replaced. These cartilaginous sutures are called synchondroses. Spheno-occipital synchondrosis is the most important and active contributor to the growth of the cranial base. It remains active until mid-adolescence, with ossification completed at 20 years of age (Enlow, 1990).

Basicranium develops rapidly until five to seven years of age (Lieberman et al., 2000). Its development slows after seven years of age until it reaches around 95% of its adult size, which corresponds to the size of the brain at 12 years of age. Basicranium does not stop growing through adolescence and mid-pubescence; it continues to expand from the internal to exterior structures at a much slower rate (Bastir et al., 2006). It reaches its mature size and form at about 11 to 12 years of age when brain development ends (Neubauer et al., 2009).

Human growth variation is influenced by several factors, such as ethnic group, secular change, socioeconomic status, health, and diet. In addition, variation occurs between individuals, even within the same population. The present study focused on developing population-specific models, particularly for the cranium of Malaysian sub-adults.

2.2 Methods for sex and ethnicity estimation

In the past, forensic anthropologists relied on rudimentary traditional inter-landmark distances measured with calipers or subjective assessments, based on the expertise of evaluators to estimate the biological profile of human remains. However, the emergence of 3D technology and advanced image processing algorithms has revolutionised the field, enabling virtual measurements and comprehensive analysis of the entire skull. Despite these advancements, determining sex and ethnicity in sub-adult samples remains challenging due to the ongoing growth and development of the craniofacial complex. Consequently, methods that exhibit high accuracy in adults may not yield the same level of precision when applied to sub-adults. Three methods are commonly utilised for sex and ethnicity estimation, namely nonmetric analysis, metric analysis, and molecular biology. Each approach has its own benefits and drawbacks. A general comparison of these methods will be discussed in the next section.

2.2.1 Nonmetric analysis

Nonmetric analysis is the simplest and longest-practiced method for sex and ethnicity estimation. This method utilises a visual assessment to identify the shape or trait of bones. The assessment is based on visual images and a scoring system for dimorphic cranial traits. However, it tends to be subjective as it requires a highly experienced observer, resulting in various interpretations by different observers (Roberts & Connell, 2004). This method is advantageous since it is cost-efficient, simple, and can be performed quickly. However, it requires sufficient years of experience in observing skeletal variation, which could lead to poor inter-observer agreement and a lack of statistics to analyse the data.

2.2.2 Metric analysis

The metric analysis is objective since it involves measuring the distance between specific features and landmarks on a skull. The metric method employs special instruments such as sliding calipers, spreading calipers, radiometers, or imaging technologies such as magnetic resonance imaging (MRI) and computed tomography (CT) modalities. These tools are utilised to collect information on the morphological features of a skull.

Advanced technologies of 3D visualisations and mathematical works allow the establishment of statistical equations to estimate the sex and ethnicity of an individual. Examples of statistical approaches are DFA, LR, ANN, RF, SVM, and MLP (İşcan & Steyn, 2013). Accuracy depends on the reliability of the evaluator, precise skeletal measurements, and correct statistical analysis. However, this method is usually population-specific and commonly affected by several factors such as secular trends, geographic location, and nutrition (Kotěrová et al., 2017).

Metric analysis is essential to develop estimation programmes such as FORDISC, CRANID, AncesTrees, and 3D-ID. These programmes allow the researcher to enter their measurements into a database that will classify the sex or ethnicity of an unknown individual (DiGangi & Hefner, 2013). The programmes and database can facilitate the measurement of skeletons by inexperienced forensic anthropologists. These analytic programmes will be discussed further in the next section.

2.2.3 Molecular biology

Molecular biology is the most accurate approach for sex and ethnicity estimation. It can provide critical information to enable identification in several ways. It may be used to determine a person's sexual orientation and ethnic origin, probable association with cancer genes, and comparison of the DNA to those of family members in missing person cases (Lundy, 1998). Sex estimation is accomplished by analysing the X and Y sex chromosomes. The Y chromosome is often employed since it includes single male-specific genes (İşcan & Steyn, 2013).

However, the benefits are often obscured by the inevitable. These techniques need complex equipment that is both costly and time-consuming (Bašić et al., 2017; Mannucci et al., 1994; Tierney & Bird, 2015). Furthermore, there are certain limitations to obtaining DNA from a cadaver. Gonzalez (2012) noted the difficulty in obtaining a usable DNA sample because it degrades through decomposition and samples are easily contaminated (Gonzalez, 2012).

Even though genetic profiling seems effective in the identification of sex and ethnicity of sub-adult remains, the concerns mentioned above require the development of methodologies focusing on the skeletal profiling of sub-adults. If DNA samples are available, skeletal profiling techniques of sub-adult remains can validate the results. However, if DNA samples are unavailable, skeletal assessments serve as the only biological source of identification.

Sex and ethnicity can be estimated using the three anthropological methods mentioned above. However, forensic anthropologists favour metric and morphological analyses. These methods have been tested on dry bone, MSCT scans of living humans, and MSCT scans of dry bone. Several studies have compared craniometric measurements obtained on dry crania to those obtained on MSCT scans (Franklin et al., 2013a; Robinson & Terhune, 2017). The results reported that the observer errors are generally not significant. Therefore, craniometric measurements obtained from MSCT scans can be utilised for the purpose of identification of sub-adult remains.

2.2.4 Computed tomography (CT) scanning technology

The introduction of 3D computerised CT has revolutionised the field of craniofacial imaging, enabling comprehensive visualisation and precise analysis of the entire craniofacial complex. The principle of CT originated from the work of an Austrian mathematician named Radon. In 1917, Radon demonstrated that it was possible to reconstruct the image of a 3D object using an infinite number of two-dimensional projections of that object (Hendee, 1989). This fundamental insight laid the foundation for the development of CT as we know it today. The first CT scanning technology was developed by Hounsfield and Cormack in the early 1970s (Cormack, 1980; Hounsfield, 1980). CT is commonly utilised in medical practice and forensic cases including cancer, neurological disorders, heart disease, and autopsy. It has allowed imaging of the complete craniofacial skeleton. This reconstruction of CT slices using 3D technology allows the investigation to be performed using real-life information of the skull. Furthermore, digitally rendered images provide visual data of bone in a format that is easily transported and distributed. This can provide worldwide accessibility, especially when it is impossible to work hands-on with skeletons due to the fragility of the remains or when the skeletons are not present to reconstruct estimation models. Thus, clinicians and anthropologists can access CT databases to obtain appropriate and accurate data (Stull et al., 2014b).

2.2.5 MSCT validation studies

The use of MSCT scans has been proven to be an alternative method for traditional data collection of the skull, which is done mainly using dry bone (Decker et al., 2011; Richard et al., 2014; Ross, 2004; Stull et al., 2014b). Thus, MSCT scan has been used to construct identification methods without removing the soft tissue (Stull et al., 2014b). This advancement has allowed researchers to conduct investigations beyond in-person observation (Decker et al., 2011).

A recent study has shown that MSCT demonstrates better results as a non-invasive method that offers detailed anatomical description of bony structures (Franklin et al., 2013a). In a study by Lottering et al. (2014), virtual reconstructions of 10 sub-adult crania were utilised to assess observer errors using two cranial measurements: maximum cranial length and maximum cranial breadth. Low values of technical error of measurement (TEM) between 0.01 mm and 0.35 mm for intra-observer error and 0.01 mm and 1.14 mm for inter-observer error were observed. Additionally, the relative TEMs (rTEM) were below 0.4%, confirming the high reliability of MSCT.

Barbeito-Andrés et al. (2012) utilised 15 3D crania of individuals ranging in age from birth to 31 years. Two sets of wireframes traced were compared using 51 landmarks and 17 semi-landmarks. Low mean errors were found, and standard deviations (SD) of landmark coordinates were noted to be similar between observers (Barbeito-Andrés et al., 2012). Meanwhile, McIntosh et al. (2020) reported low intra- and inter-observer errors with TEM values below 0.80 mm and rTEM values below 1.80%. The study was conducted on 15 craniometric measurements obtained from 3D crania of 30 individuals from birth to 18 months old. Another recent study on 3D cranial reconstructions of 12 individuals ranging in age from birth to 20 years demonstrated low TEM value (0.05 mm) and rTEM percentage (0.073%) for all tested craniometric parameters (Corron et al., 2022). Thus, all these studies have confirmed that high reliability can be achieved when obtaining cranial measurements on 3D virtual renderings of sub-adult's cranium.

2.2.6 Machine learning (ML) algorithm

In recent years, there have been numerous studies exploring the application of ML in the field of forensic sciences (Bidmos et al., 2023; Imaizumi et al., 2020; Nikita & Nikitas, 2020; Toneva et al., 2020, 2021; Toy et al., 2022). These studies have presented new challenges and shed light on the potential advantages and disadvantages of utilising ML methodologies to address well-known forensic problems. One particular aspect where ML technology has shown promise is in overcoming the inherent subjective biases associated with traditional approaches in forensic anthropology, specifically in sex prediction and age estimation (Bidmos et al., 2023; du Jardin et al., 2009; Li et al., 2019; Navega et al., 2015b; Santos et al., 2014).

ML depends on the ability of a computer-controlled robot to perform tasks associated with human intelligence. Thus, ML performance mimics human cognitive capabilities that are trained by human brainpower (Leonardi et al., 2021). ML is driven by the idea of enabling machines to learn and improve from experience without being explicitly programmed (Valdes et al., 2017). This is accomplished through the utilisation of algorithms that can automatically analyse and interpret data, uncovering hidden patterns and making predictions based on the insights gained. The process begins with training data, namely a representative sample that serves as a foundation for the ML model to learn and develop mathematical representations of the underlying patterns. These representations form the basis for predicting outcomes when presented with new and unseen data (Zhou et al., 2017).

ML encompasses various techniques and algorithms, each suited to different types of problems and data. Numerous ML methods, both supervised and unsupervised, have been developed for forensic investigations. DT, RF, SVM, and ANN are examples of supervised methods for sex and ethnicity estimation studies using bone measurements (du Jardin et al., 2009; Klales et al., 2020; Mahfouz et al., 2007; Musilová et al., 2016; Santos et al., 2014; Yang et al., 2020). These algorithms employ statistical methods, optimisation techniques, and mathematical models to generalise patterns from the training data, enabling them to make accurate predictions or classifications when presented with new input. In contrast, unsupervised ML does not rely on labeled data for training. Instead, these algorithms aim to uncover patterns, relationships, and structures within the data without any predefined class labels. Clustering, anomaly detection, dimensionality reduction, and association rule mining are examples of unsupervised ML techniques that have found applications in the field (Sarker, 2021). One popular clustering algorithm is the fuzzy c-means clustering algorithm that has been used for the sex estimation of a skull (Gao et al., 2018). In conclusion, ML is a powerful subset of AI that enables machines to forecast outcomes by leveraging mathematical tools generated from training data. Its application extends across diverse fields to enhance decision-making processes.

2.2.7 Analytic programmes

In recent years, the development of computer programmes has significantly enhanced the accuracy and efficiency of sex and ethnicity estimation. Forensic anthropologists now have access to sophisticated programmes that utilise advanced algorithms and statistical models to generate customised measurements for sex and ethnicity estimation. These types of programmes will be discussed further in the next section.

2.2.7.1 FORDISC

FORDISC, developed by Stephen Ousley and Richard Jantz in 1993, is an interactive computer programme for identifying humans by sex and ethnicity using any combination of conventional cranial measurements. It classifies skulls of unknown origin into up to 13 ethnicities and sex-specific reference groups using discriminant or canonical variates analyses (Ousley & Jantz, 1993). FORDISC uses Howells' dataset with additional samples from Terry and Hamann-Todd Collection and American Forensic Data. Although it has been utilised worldwide, the application of FORDISC is heavily based on American data. Thus, it has been recommended for population-specific data to be developed for accurate estimations to be obtained (Elliott & Collard, 2009).

The accuracy of FORDISC application in forensic studies has produced different results in various populations. Guyomarc'h and Bruzek (2011) used FORDISC 3.0 to test the accuracy of sex and ethnicity estimation using French and Thai samples. Low classification accuracy was observed for the sex and ethnicity estimation of the Thai samples. Although the French samples had better accuracy rates than the Thai samples, neither were significantly accurate. Dudzik and Jantz (2016) found poor classification rates of Hispanic individuals as Japanese when using FORDISC. In addition, Hispanics were often misclassified as White or Native Americans. The common evolutionary histories between Hispanic and Japanese, White, and Native American populations were the reason for the poor classification rates. Another study by Elliott and Collard (2009) tested FORDISC classification for ethnicity estimation of unidentified human remains. The authors used 200 specimens of known ethnicities, with and without defining the sex. Only 1% of the samples were classified correctly. Hence, the FORDISC programme is evidently imprecise even though it is automatic and straightforward.

2.2.7.2 CRANID

Richard Wright developed CRANID in 1992 to determine the ethnicity of unknown adult human remains (Wright, 1992). The programme compares cranial measurements to a global craniometric database using DFA. It is based on Howell's 1973 cranial database and supplemented with samples from Italy, the United Kingdom, West Asia, Denmark, India, Patagonia, and Australia, totalling 3163 crania samples from 39 populations (Hughes et al., 2005). Given that the reference crania from Australia and Europe are more prominent, CRANID is more valid in those countries. Wright obtained an accuracy of 68.2% for the LDA classification of the 74 sex-differentiated reference samples in CRANID (Wright, 2010). In comparison, low percentages of accuracy were obtained (39% for local groups and less than 48% for regional groups) when samples from the University of Melbourne's Berry collection were used (Kallenberger & Pilbrow, 2012). In conclusion, if the cranium comes from mixed ethnicities or belongs to modern populations, CRANID may not be able to accurately estimate its geographic origin.

2.2.7.3 AncesTrees

AncesTrees was developed in 2015 by Navega and colleagues in Portugal to determine ethnicities using 23 cranial measurements (Navega et al., 2015a). This programme classifies human skulls using the ML algorithm technique known as RF. RF is a nonlinear, non-parametric, ensemble-based classification approach that employs hundreds to thousands of classification decision trees as base models. The programme has a database of around 3,000 adults from six major ethnic groups: Sub-Saharan African, Australo-Melanesian, East Asian, European, Native American, and Polynesian.

In a study by Skalic in 2018, the author reported a relatively low accuracy rate ranging from 37% to 40% when predicting the ethnicities of Portuguese and American samples that did not align with any of the predefined ethnic groupings in the database (Skalic, 2018). Similarly, in a validation study on Brazilian samples, it was found that the absence of Brazilian samples resulted in relatively low accuracy rates, ranging from 50% to 67.4% (Fernandes et al., 2021). However, in another validation study by Sieber and García-Donas (2023), the programme demonstrated high accuracy by correctly classifying the Spanish sample as Southwestern European or European, with accuracy rates ranging between 82.61% and 100%. Thus, the AncesTrees programme may not be suitable for estimating ethnicities in underrepresented groups in the database.

2.2.7.4 3D-ID

3D-ID is a programme developed by Dr. Dennis Slice and Ann Ross to estimate sex and ethnicity based on skulls of unknown individuals (Slice & Ross, 2009). It has a database from worldwide collections containing 2,300 modern sample crania of known sex and ethnicities. The programme assigns the unknown skull to the available populations in the database by supplying the Mahalanobis squared distance (D^2) and craniometric measurements with respect to each available reference group (Humphries et al., 2015; Slice & Ross, 2009). 3D-ID uses geometric morphometric techniques, generalised Procrustes analysis, and discrimination classification methods. Although 3D-ID relies on a smaller database compared to other programmes, according to King (2015), the use of geometric morphometrics provides a valuable alternative to traditional onedimensional measurement methods. Geometric morphometrics can account for curves and differences between points, which are limitations of traditional approaches (King, 2015). Hence, the value of the mathematics used in 3D-ID demonstrates its magnitude for forensic identification.

In a study conducted by Urbanová et al. (2014), the accuracy of 3D-ID in sex and ethnicity estimation was tested using Brazilian samples. The results reported that in terms of ethnicity estimation, 55% of the Brazilian crania were correctly assigned to their respective groups. The highest accuracy of 87% was achieved for the European group. However, a significant proportion of admixed individuals were assigned to reference groups of European descent (66%), although some were also assigned to groups of Mesoamerican and South American origin. For sex estimation, the accuracy ranged from 60.3% to 77.9% in correctly classifying individuals (Urbanová et al., 2014). These findings indicated that while 3D-ID reported some success in ethnicity estimation, there were challenges in accurately classifying individuals from mixed backgrounds. Additionally, the reliability of sex estimation varied across different programmes utilised in this study.

2.2.7.5 (hu)MANid

The Human Mandible Identification (hu)MANid programme was developed by Berg and Kenyhercz in 2017. The programme has a global sample of mandibular morphology and metric data that allow the end user to classify the sex and ethnicity of an individual. The sample consists of 1750 individuals from 15 main populations. The skeletal collections were obtained from various institutions such as the University of Tennessee Forensic collection, William W. Bass Donated (WBD) collection, Central Identification Laboratory, Pima Country Office of the Medical Examiner (PCOME), Guatemalan Foundation of Anthropology (FAFG), and the Hawaii (CILHI) collection.

The (hu)MANid programme successfully predicted the sex and ethnicity of an unknown adult's mandible (Berg & Kenyhercz, 2017; Sieber & García-Donas, 2023). In a study conducted by Sieber and García-Donas, the programme correctly estimated the ethnicity of the Spanish sample as being of white origin with accuracies ranging from 70.59% to 80%, while for sex estimation, accuracy ranged between 62.75% and 80% (Sieber & García-Donas, 2023). However, several studies had reported low accuracy rates when the programme was validated in a pediatric population (Farhi et al., 2023; Tomás, 2020). In a study conducted by Tomás (2020), it was reported that the accuracy rate for sex and ethnicity estimation of older children and adolescents was relatively low. Specifically, when predicting ethnicity, the study reported adequate accuracy when estimating individuals of White ethnicity. However, the accuracy significantly decreased when predicting the ethnicity for individuals of Hispanic and Black descent. In addition, the programme was inclined to predict White females in the sample (Tomás, 2020). In another study by Farhi et al. (2023), the author aimed to predict sex and ethnicity using samples from adults (aged 20-45) and adolescents (aged 15-17) from the United States of America. The results reported that the programme achieved a sex prediction accuracy of 75.52% in the adult sample. However, the accuracy for ethnicity estimation was less favourable, ranging from 19.3% to 50%. For the adolescent sample, the author reported an accuracy of 45% for sex estimation and 47.5% for ethnicity estimation (Farhi et al., 2023). These findings suggest that the programme may not be as useful in accurately identifying the sex and ethnicity of a younger individual.

In conclusion, numerous populations around the globe cannot be classified using the identification programmes mentioned above. These systems lack global demographic reference data, resulting in poor accuracy for individuals from geographical locations which are not completely represented in the programme databases (Cunha & Ubelaker, 2020; Kallenberger & Pilbrow, 2012). Malaysian communities with multi-ethnic diversity were not included in the database. Additionally, many programmes focus only on the adult population, leaving the sub-adult population underrepresented.

Furthermore, the skulls utilised to generate the database for these programmes were from a relatively distant era, as these museum collections include skeletal remains which had accumulated between 50 and 100 years ago. Numerous academics queried the usefulness of these old collections in a medicolegal context for determining the biological profile of the present population. Scheuer and Black (2000) mentioned two reports (in 1920 and 1985) that highlighted the observance of a large discrepancy in the recorded ages of individuals in the Terry and Hamann-Todd collection. Therefore, misinformation regarding age groups and the recorded demographic data can create inaccurate results in statistical analyses.

The utilisation of regularly collected radiographic images permits the development of new techniques, the validation of previous research, and the use of larger sample sizes. In contrast to previous studies on historical remains, the present study incorporates data from modern sub-adults that can be used in forensic analyses. Thus, Malaysia needs its own automated programme to be able to classify sex and ethnicity of sub-adult remains.

2.3 Sex estimation

Sex estimation is one of the first steps in establishing the biological profile of human remains (Spradley & Jantz, 2011). Sex estimation is an essential component because it serves as the foundation on which other identifications are developed. Furthermore, it helps to narrow the search on missing-person databases by approximately 50% (Spradley & Jantz, 2016).

2.3.1 Sexual dimorphism

Sex estimation is possible due to the existence of sexual dimorphism, which presents discernible morphological differences between sexes. A more specific definition of sexual dimorphism is the average difference in body size between male and female individuals (Charisi et al., 2011). These differences are due to several factors including hormonal fluctuation and production involved in puberty, which ultimately results in variation between males and females (Scheuer & Black, 2000).

2.3.2 Endocrine basis for sexual dimorphism

Hormones are crucial in controlling various aspects of the body's growth, development, reproduction, and metabolism. The gonads (ovaries in females and testes in males) are responsible for synthesising steroid sex hormones, namely estrogen and testosterone. Although these hormones are present in both sexes, they may act differently due to their varying levels (Lewis, 2018). These hormones are important for the development of both female and male reproductive organs, bone growth, and sexual identity.

Sex characteristics are classified into two phases, primary and secondary traits (Norris & Carr, 2020). The primary sex characteristics are the foundational traits that begin to develop during prenatal stages. These traits involve the formation of specific organs that define the male and female sexes. In females, the primary sex characteristics include the development of ovaries, which are responsible for producing eggs and releasing hormones such as estrogen. In males, the primary sex characteristics involve the development of testes, which produce sperm and secrete hormones, particularly testosterone. These primary sex characteristics form the fundamental components of the reproductive system and are crucial for fertility and sexual development.

There are notable differences in the hormonal patterns between males and females during early development. Females experience a longer duration of high estrogen production, which spans approximately 12 months. Estrogen plays a pivotal role in the development of female sexual characteristics, such as breast development, regulation of the menstrual cycle, and the maturation of reproductive organs. On the other hand, males have a shorter period of high testosterone production, lasting for around six months. Testosterone is primarily responsible for the development of male sexual characteristics, including the growth of facial hair, the deepening of the voice, and the maturation of the reproductive system. Following the initial hormonal surge during prenatal development, testosterone levels remain relatively low until around the age of six to eight years. It is during this time that testosterone levels begin to rise again, signaling the onset of puberty and the development of secondary sexual characteristics in males (Norris & Carr, 2020).

Puberty is defined by a skeletal growth spurt, the development of gonads, secondary sex characteristics, and changes in body composition (Garnett et al., 2004). Differences in sex hormone levels result in females achieving puberty earlier than males by about one to two years (Lewis & Flavel, 2006). However, males stay in puberty longer, namely for around five years compared to three and half years for females (Faulkner & Tanner, 1986). As a result, the development of males is slow but consistent until the start of puberty, and then it accelerates. Secondary sexual characteristics in females and males undergo distinct changes during puberty. Estrogen influences the development of secondary sexual characteristics in females, including breast growth, changes in body fat distribution, and the onset of the menstrual cycle. However, its bone turnover rate decreases, resulting in less growth in stature (Lieberman, 1982). Testosterone drives the development of secondary sexual characteristics in males, such as the growth of facial and body hair, the deepening of the voice, and increased muscle mass. Increased testosterone levels promote bone apposition and mass accumulation. Hence, increased size and muscle attachments in males result in a more prominent shape of glabella, supraorbital ridges, frontal sinus, and mastoid processes (Sprowl, 2013).

Understanding the development of sexual dimorphism can provide insights into the regulatory processes and the mechanisms determining it. One potential mechanism is that sexual dimorphism is a product of the differential actions of sex hormones on growth regulation between the sexes. Hence, these differences are usually assumed to emerge at puberty when differences in hormone levels are most pronounced. Nevertheless, several studies suggest that sexual dimorphism is evident in some skeletal components from an early age (Bulygina et al., 2006; Matthews et al., 2016; Sprowl, 2013). However, it is often not believed to reach levels sufficient for accurate sex estimation until after puberty (Cunningham et al., 2016). While there are challenges in determining the sex of sub-adult remains, several articles have investigated sexual dimorphism in skull features at different ages, both before and after puberty (Hsiao et al., 2010; Noble et al., 2019; O'Donnell et al., 2017; Sprowl, 2013; Teodoru-Raghina et al., 2017). These articles are further discussed in the next section.

2.3.3 Pre-pubertal sexual dimorphism

Sexual dimorphism in craniofacial shape has been reported to be present as early as when an individual is one year old (Baughan & Demirjian, 1978; Bulygina et al., 2006; Matthews et al., 2016). At birth, the growth rate is slightly slower in females; then, it equalises with males at seven months of age. The growth rate becomes faster in females until the age of four years. From this point, males and females grow at a similar rate until the stage of adolescent growth spurt (Rogol et al., 2000). By six to eight years, females are two to three months more advanced, and by adolescence, they can be as much as two years more advanced (Tanner, 1990). The slower growth rate in females at birth results in the greater neurocranial size of male infants and foetuses. Additionally, males have larger head circumference, head size, and greater body lengths compared to females (Bulygina et al., 2006; Joffe et al., 2005).

Several measurements in the craniofacial structure (neurocranial length and width, total facial height, facial depth and width, width of the nasal base, outer canthal width and palpebral fissure length, and philtrum length) were observed to be greater in males aged three to six years compared to females of the same age (Yusof, 2007). Similarly, males' cranial vault measurements were found to be significantly higher during infancy (Yusof, 2007). This may be the result of testosterone, which is higher in males during the first few months of the postnatal period and stabilises until puberty (Lieberman, 1982). Matthews et al. (2016) reported that measurements such as maximum cranial width and length, and cranial base width, showed evidence of sexual dimorphism as early as three years. In addition, a study conducted by Sprowl (2013) demonstrated that sexual dimorphism exists in sub-adults as early as six years. Furthermore, craniofacial sex differences were observed to increase with age, and significantly increase after puberty (Gaži-Čoklica et al., 1997; Kesterke et al., 2016; Sforza et al., 2010, 2011). Therefore, sexual dimorphism is present as early as the prenatal stage, reflecting brain growth until puberty (Gonzalez, 2012).

2.3.4 Post-pubertal sexual dimorphism

Several studies have reported that sexual dimorphism appears after puberty (Hsiao et al., 2010; Noble et al., 2019; Ramamoorthy et al., 2016). In Taiwanese children, sexual dimorphism becomes evident after puberty at between 12 and 17 years (Hsiao et al., 2010). Similarly, sexual dimorphism was observed after nine years in the Australian population (Noble et al., 2019). Furthermore, the mean levels of a few pre-pubertal sexual dimorphism parameters (maximum frontal breadth, anterior and posterior intraoccipital condylar distances, foramen magnum breadth) were found to be insignificant as they attain adult size after puberty (Humphrey, 1998). Additionally, due to differences in growth rate, sexual differences are not apparent until after the growth spurt occurring between six and eight years (Lejarraga, 2002).

2.3.5 Population-specific in sex estimation

Sexual dimorphism in the skull has been extensively studied in different populations and has been found to vary across populations (Franklin et al., 2013b; Gonzalez, 2012; Noble et al., 2019; O'Donnell et al., 2017; Ramamoorthy et al., 2016; Teodoru-Raghina et al., 2017). As a result, no standards have been set for global populations (Dillon, 2014). This could be due to sex estimation being population-specific. Therefore, standards developed in one population cannot be applied to another.

Genetics, nutrition, and other environmental factors contribute to the expression of sexual dimorphism (Charisi et al., 2011; Kimmerle et al., 2008). In the context of Malaysia's population, individuals of various ethnicities exist. Therefore, it is critical to take ethnicity into account when estimating sex. Sex estimation takes on a new dimension due to the country's various ethnic groupings: Malays, Chinese, and Indians in West Malaysia and other smaller ethnic groups in East Malaysia. This necessitates determining whether a universal norm applies to the population in Malaysia.

Worldwide variance highlights the critical need for local databases in forensic anthropology. Reduced accuracy has been observed when standards built for a specific population were applied to different populations (Kim, 2009; Toneva et al., 2020; Yang et al., 2020). Furthermore, inaccurate sex estimation can occur due to inter-population differences in sexual dimorphism expression. Some communities have robust traits of both sexes, while others have gracile traits of both sexes (Green & Curnoe, 2009). Therefore, this demonstrates the importance of creating specific standards for certain populations.

2.3.6 Binary logistic regression (BLR) vs. discriminant function analysis (DFA)

There are two main methods to develop sex estimation models, which are DFA and BLR. Both have been shown to be suitable for sex estimation using the skeleton (Ahmed et al., 2021; Ekizoglu et al., 2017; Macaluso Jr., 2010; Singh & Pathak, 2013). Previous research on sex estimation provides insights into the best types of analysis based on the constraints of the research project at hand. It is important to look back on this research and their methods when conducting new studies to make the best decision regarding which method to use for a project that utilises metric data for sex estimation.

When utilising DFA, each parameter must fulfil the assumption of normality and equal variance-covariance matrices. However, BLR is not restricted to a normal distribution of data nor does it assume equality of the variance-covariance matrices. Also, BLR is less sensitive to high correlations and outliers among the parameters, which is particularly important when conducting craniometric measurements. Hence, in cases where the dataset requires a less restrictive statistical analysis, BLR is preferred over DFA.

Several studies have been conducted to compare DFA and BLR in developing a sex estimation method (Ekizoglu et al., 2017; Macaluso Jr., 2010; Singh & Pathak, 2013). Macaluso Jr. (2010) developed metric sex estimation equations using sternal measurements for the South African population. The BLR models yielded slightly better accuracy compared to DFA in overall classification rates, with an increase of approximately 1.5% for most of the derived equations. However, these slightly enhanced classification rates were related to much higher sex bias rates, ranging between 4.7% and 24.3% (Macaluso Jr., 2010).

Singh and Pathak (2013) aimed to estimate sex in the Indian population using sternum measurements by employing two statistical methods (BLR and DFA). The study found that BLR produced better results compared to DFA in terms of overall accuracy rates. However, higher sex bias rates were observed in BLR compared to DFA. Therefore, the

DFA method has been reported to be a more reliable sex-discriminating method as it enables the development of sex estimation models that have both high classification accuracies and low sex bias rates (Singh & Pathak, 2013).

Another study proved that BLR is a viable option for sex estimation (Ekizoglu et al., 2017). Ekizoglu et al. (2017) developed a sex estimation method for the Turkish population based on MSCT scans of the calcaneus bone. The study utilised both statistical methods, namely DFA and BLR. When comparing the two statistical methods, BLR models yielded higher accuracies and the lowest sex bias in overall cases for the original sample and in three of the five equations for the validation sample (Ekizoglu et al., 2017). Based on the results of previous studies, it is recommended to use both statistical methods to develop sex estimation models.

2.3.7 Methods for sex estimation using sub-adult crania

Methods for sex estimation using sub-adult crania have been developed since the late 19th century (Bulygina et al., 2006; Hsiao et al., 2010; Matthews et al., 2016; Noble et al., 2019; Sprowl, 2013). Early studies had reported sex differences in the crania of young children (Bulygina et al., 2006; Matthews et al., 2016; Sprowl, 2013), while others emphasised that estimating sex for sub-adults is impossible until after puberty (Hsiao et al., 2010; Noble et al., 2019). This section summarises prior studies on sex estimation using sub-adult crania. Data were gathered using many approaches, including the measurement of dry bone specimens, two-dimensional radiography data, and virtual reconstruction measures. Despite the difficulties associated with attributing sex to sub-adult remains, several publications have examined sexual dimorphism in certain skeletal characteristics at various ages (Amin & Othman, 2014; Divakar et al., 2016; Gonzalez, 2012; Hsiao et al., 2010; Noble et al., 2019; O'Donnell et al., 2017; Sprowl, 2013;

Teodoru-Raghina et al., 2017; Veroni et al., 2010). A summary of previous studies in estimating the sex of sub-adults is tabulated in Table 2.1.

Table 2.1: Previous studies of sex estimation in sub-adults

Authors	Population	M	F	T	Age (years)	Method	Analysis	Accuracy (%)
Amin and Othman (2014)	Jordanian	47	99	146	13-27	Cephalo-metric	DFA	87
Divakar et al. (2016)	Indian	380	236	616	6.5-18	Cephalo-metric	DFA	100
Gonzalez (2012)	European	250	248	498	5-16	Cephalo-metric	DFA	71-90
Hsiao et al. (2010)	Taiwanese	50	50	100	12-18	Cephalo-metric	DFA	92-95
Noble et al. (2019)	Australian	83	69	152	0-19	CT	DFA	50-90
O'Donnell et al. (2017)	New Mexican	688	930	1618	<18	Cephalo-metric	DFA	57.8-100
Sprowl (2013)	Latino	118	185	303	6-18	Cephalo-metric	DFA	71.7-100
Teodoru-Raghina et al. (2017)	Romanian	334	266	600	5-18	CT	DFA	68-84
Veroni et al. (2010)	Portuguese	19	17	36	8-18	Dry skull	DFA	75.8

M=male, F=female, T=total; CT=computed tomography, DFA=discriminant function analysis.

2.3.7.1 Veroni et al. (2010)

Veroni et al. (2010) examined Holland's (1986) results using sub-adult skeletal samples of determined sex and age from Portugal's Bocage Museum. Individuals of European ethnicity born between 1805 and 1972 had been selected for the museum from Lisbon cemeteries. A total of 17 females and 19 males were included in the study, aged between eight and 18 years. The length and breadth of the foramen magnum, length and breadth of the occipital condyle, and breadth of the occipital bicondylar were determined. Males were found to have greater breadth and length of the foramen magnum, whereas females had the greatest dimorphism in the breadth of the left occipital condyle. The accuracy rate of sex estimation obtained from this study was 75.8%. However, the poor accuracy rate was due to demographic variability rather than age (Veroni et al., 2010).

2.3.7.2 Hsiao et al. (2010)

Hsiao et al. (2010) applied the same procedures used on adults to study 100 Taiwanese children and adolescents aged 12 to 17 years. Lateral cephalograms at Kaohsiung Medical University Orthodontic Department were obtained between January 2005 and June 2009. Throughout the radiograph, 22 cephalometric measurements and a cervical vertebral maturation stage were measured and analysed. The mean differences were statistically significant ($p<0.05$) for all measurements, demonstrating the presence of sexual dimorphism in the skull. Males had greater values than females for all linear measurements, whereas females had greater values than males for angular measurements. High accuracy rates of 92%-95% were achieved when four to seven parameters [Glabella-metopion to basion-nasion (GM–BaN), basi-bregmatic height (Ba-Br), mastoid height from cranial base (MaHt), foramen magnum length (Ba–O), glabella projection index (GPI), inion-opisthocranion to basion-nasion (IOP–BaN), and frontal sinus width on inner bregma to nasion line (FSWd)] were selected in a stepwise discriminant function. However, when two (GM– BaN and Ba–Br) or three parameters (GM–BaN, Ba–Br, and MaHt) were examined, the accuracy rates were between 84% and 90%, respectively. The glabella-metopion to basion-nasion alone estimated sex with 73% accuracy. Cross-validation resulted in nearly identical accuracy with a decrease from 1% to 4%. Although this study was quite effective in finding sexual dimorphism, the average age of the participants was determined after the commencement of puberty, thus it did not represent the pre-pubertal age range (Hsiao et al., 2010).

2.3.7.3 Gonzalez (2012)

Gonzalez (2012) retested Giles and Elliot’s 1963 methodology on sub-adults using lateral cephalometric radiographs at the University of Michigan School of Dentistry. Sexual dimorphism was assessed in 83 hand-traced lateral cephalograms of European sub-adult individuals aged five to 16 years. Eight landmarks were chosen: basion, sella, bregma, nasion, glabella, posterior nasal spine, prosthion, and opisthocranion. Canonical discriminant analysis was used to generate linear combinations of the parameters based on 20 measurements from these landmarks. Approximately three of the 20 functions generated were found to be statistically significant, accounting for 87.3% of the overall variance. The first canonical discriminant function, representing the facial growth changes, achieved 71.9% of the overall variance. The second canonical discriminant function, which accounts for the developmental trajectories between the sexes, achieved 10.6% of the overall variance. The third canonical discriminant function represents the presence of sex differences in the neurocranium, accounting for 4.8% of the overall variation. Apparent differences were noted between the sexes, with males having greater third canonical discriminant function dimensions than females. This resulted in males having longer and taller crania. Sexual dimorphism may occur at the age of six, prior to the completion of craniofacial development. These traits continue to develop as a result of craniofacial characteristics that are unique to each sex. The overall accuracy rate for sex estimation was 87.3%, implying that sexual dimorphism after puberty enhances the accuracy rate (Gonzalez, 2012).

2.3.7.4 Sprowl (2013)

The study analysed 303 lateral cephalograms of pre- and post-adolescent Latinos aged six to 18 years, divided into several age groups. Radiographs were obtained from the digital database at the University of Nevada, Las Vegas. A linear discriminant function and canonical correlation analyses were developed for the total samples and each age group. Sex estimation was achieved with a 74.6% accuracy rate, with 89% of females and 51.7% of males correctly classified using the stepwise function. For all the samples, all the parameters, i.e., distance between glabella and supraglabellare to nasion line (GSgN), distance between mastoidale and SN line (MaSN), glabella projection index (GPI), and ratio of total chin thickness to upper lip thickness (ULTc) contributed significantly ($p<0.05$) to sex estimation. The analyses revealed an accuracy of 100% for the age group of 6.5-8.5 years, 83.3% for the age group of 8.6-10.5 years, 71.7% for the age group of 10.6-12.5 years, 78.3% for the age group of 12.6-14.5 years, and 94.7% for the age group of 14.6-17.9 years. This study indicated that sexual dimorphism exists in sub-adults as early as six years (Sprowl, 2013).

2.3.7.5 Amin and Othman (2014)

This study randomly selected 146 digital lateral cephalometric radiographs of 47 males and 99 females aged 13 to 27 years. The DFA was performed, and an 85.6% cross-validated accuracy was obtained. The accuracy was increased to 87.0% when the stepwise technique was used. Males were found to have greater mean values of 3% to 41% than females for all metrics, except for frontal sinus width and mandibular body length. Mastoid height was the most accurate single parameter for sex estimation, with an accuracy rate of 82.2%. The stepwise procedure increased the classification accuracy to 87.7% by utilising four measurements: cranial base length, mastoid height, mastoid breadth, and glabella to supraglabellare-nasion distance (Amin & Othman, 2014).

2.3.7.6 Divakar et al. (2016)

This study in India utilised lateral cephalometric radiography and DFA to determine sex. A total of 616 lateral cephalograms of 380 males and 236 girls aged six years and six months to 18 years were included. Males had substantially greater mean angular and linear cephalometric measures than females ($p<0.05$) based on 24 cephalometric measurements. Additionally, significant differences ($p<0.05$) in all parameters were identified according to age group. The ratio of total chin thickness to upper lip thickness

(ULTc) parameter successfully predicted sex among the 24 parameters. DFA revealed 100% sex discrimination, with 100% females and 100% males accurately estimated. This study reported the existence of sexual dimorphism in the cranium as early as six years and six months of age, owing to the indigenous groups' distinct genetic and environmental composition. (Divakar et al., 2016).

2.3.7.7 O'Donnell et al. (2017)

This study estimated sex from samples of 688 males and 930 females in individuals ranging in age between six and 18 years. The samples were collected from various communities in New Mexico. When DFA was applied to all the samples, only 58.5% of males and 57.8% of females were accurately estimated. However, when the samples were divided by ethnicity, high accuracy rates were obtained in the African, Asian, and Native American groups. African American males achieved a higher rate of classification (85.71%) than females (66.7%). When the groups were divided by age, classification accuracies ranged from 50% to 80% were reported. The highest accuracy of 73% was obtained in the age range between 15 and 18 years. DFA demonstrated the greatest discriminating strength when age classes and ethnic groups were combined, with higher accuracy obtained in the age range between 15 and 18 years and in the Native Americans, Hispanics, and Asian Americans groups. This study indicated that dividing the samples by age and ethnicity enhances accuracy rates for older groups, and for Native Americans, and Asian Americans. Meanwhile, when age and ethnic groups were unknown, cephalometric measurements were ineffective in determining the sex of unidentified individuals (O'Donnell et al., 2017).

2.3.7.8 Teodoru-Raghina et al. (2017)

This study utilised 13 anthropometric metrics of known sub-adults (266 females and 334 males) ranging from five to 18 years. Population-specific DFA was conducted for sex estimation in Romanian sub-adults. Sex accuracies ranged from around 75% for ages eight-11 years to 84% for ages 16-18 years. The accuracy increased to 84% for the 16-18 age range based on two parameters (cranial breadth and cranial base length). However, the development pattern did not evolve linearly, as they achieved a significantly lower classification accuracy (between 67% and 70%) in the group with sub-adults ranging from ages 12 to 15 years. This could be due to the differences in the onset of puberty between males and females, which resulted in varying sexual dimorphism within that period. The cranial vault exhibited the highest degree of dimorphism. This is because the most important parameters in each discriminant function are located inside the cranial vault. The cranial width was shown to be the most dimorphic parameter exhibited in sub-adults between the ages of 12 and 18 years (Teodoru-Raghina et al., 2017).

2.3.7.9 Noble et al. (2019)

The study examined 152 sub-adult crania aged one to 19 years from Western Australia. A total of 52 3D landmarks were analysed using Procrustean geometric morphometrics. This study demonstrated that sexual dimorphism and age differences were apparent through geometric morphometric analysis of form, size, and shape. However, the analysis of size yielded better accuracy than shape. There was a substantial overlap in size analysis between the sexes, particularly in individuals below nine years. Resultantly, a low level of cranial sexual dimorphism before the commencement of puberty was reported. The accuracy of sex estimation increases with age, with average hit rates (HR) ranging from 50% to more than 90%. HRs increasing from less than 10% in young children to almost 70% in young adults indicated a high degree of confidence (posterior probability>0.8).

The average HRs were greater than the 95th percentile of random chance classification in sub-adults (12–17 years), showing that sex may be predicted significantly in that age group. While size and shape appear to be insignificant in predicting sex in the sub-adult human skull, basic size measurements and classic morphometric features may produce more accurate findings. Therefore, this study indicated that sexual dimorphism exists in the post-puberty cranial shape using a geometric morphometric analysis (Noble et al., 2019).

2.4 Ethnicity estimation

Ethnicity is known as the most challenging and highly debated biological profile. This is due to disagreements in forensic anthropology regarding the use of the term race or ethnicity (Klales & Kenyhercz, 2015). However, when the term ethnicity is used, the forensic anthropologist describes the ethnic group to which the individual would have belonged (Hefner, 2009). This is estimated by observing metric and nonmetric traits on the skeletal remains. The literature on ethnicity estimation in sub-adults is limited compared to adults. Assessing ethnicity in sub-adults is difficult because most ethnic-related differences develop fully only in adulthood. However, many scholars have worked in this direction (Christensen & Passalacqua, 2018) based on the assumption that cranial variety is present regardless of age (Byers, 2016). This section summarises previous studies on ethnicity estimation using measurements of dry bone, two-dimensional radiography data, and virtual reconstruction measures.

2.4.1 Human variation

Human variation is present in most populations due to factors such as culture, environment, migration, population divisions, and geography (Hunley et al., 2009). These factors intervene with the gene flow and result in more divided populations. Furthermore, genetic drift in a population occurs as genes of a smaller segment become excessively representative of the parent population. There is a complex interaction which takes place between phenotype, genotype, and environment. This interaction will never be fully divided. An individual is exposed to the environment (nutrition, exposure to pathogens, education, climate, and physical and psychosocial stress), which begins *in utero* and continues throughout life. The environment essentially influences the selection of traits that will be beneficial or harmful, and adaptation occurs as the environment changes (Stojanowski & Schillaci, 2006).

2.4.2 Malaysian population

Malaysia is an incredibly diverse country that consists of three major ethnic groups, namely Malays and indigenous populations (70.1%), Chinese (22.6%), and Indians (6.6%). The remaining 0.7% of the population consists of individuals classified as Eurasians and minor ethnic groups residing in East Malaysia (Current population estimates, Malaysia, 2024). Each ethnic group has its own population history within the country. The Malays and indigenous populations (also known as Orang Asli) are grouped into three main tribal groups: Negrito, Senoi, and Proto-Malay (Aghakhanian et al., 2015). The Negritos are the smallest group and the first inhabitants in Peninsular Malaysia about 25,000 years ago (Aghakhanian et al., 2015; Hill et al., 2006). The Senoi are the largest group that migrated from the mainland of Southeast Asia around 8,000 years ago (Hill et al., 2006). Meanwhile, Proto-Malays or aboriginal Malays have a more diverse origin since prehistoric times, dating back to as early as the 1800s. Thereafter, the historic influx and multiples admixture with Indian, Siamese, Sumatran, Javanese, Arab, Thai, and Chinese traders (Comas et al., 1998) introduced Deutero-Malays. Deutero-Malays are descendants of Proto-Malays that constitute the majority of people residing in Peninsular Malaysia (Simon, 2012). Nonetheless, according to Fix, the original Deutero-Malays migrated from Southern China over 1,500 years ago and their intermarriages with the Proto-Malays and traders have resulted in the diverse Deutero-Malay, namely the contemporary Malay population today (Fix, 1995).

Multiple historical events have led to further genetic variations within Peninsular Malaysia. In the early centuries, Peninsular Malaysia was home to various distinct dynasties that brought in the Indian and Buddhist cultures. Between the 6th and 7th century, the Straits of Malacca became a significant maritime commerce route connecting the Indian Ocean to the South China Sea and the Pacific Ocean (Baker, 2020). This contributed to the economic development of the region as a major entrepôt and business centre for international traders from China, India, and Southeast Asia (Belle, 2014). Although intermarriages had occurred between the locals and the traders, they did not create significant migration. This period was followed by European colonisation by the Portuguese, Dutch, and British starting from the early 16th century (Saw, 2007). In the 19th century, during the reign of the British Empire, there was a substantial influx of Chinese from Southern China and Indians from South Indian who came to work in tin mines and rubber plantation industries (Tan, 2001). Consequently, the inflow of genes between Malays and Indians, Arabs, and Chinese traders in addition to the European colonists during the last 500-600 years is likely to have had a substantial impact on their gene pools (Deng et al., 2015; Hatin et al., 2011). Hence, the modern Malays today showed some admixture of genetic components from Arabia, India, China, Java, Sumatra, and Thailand populations (Norhalifah et al., 2016). The multi-ethnicity history of Malaysia is summarised in Figure 2.1.

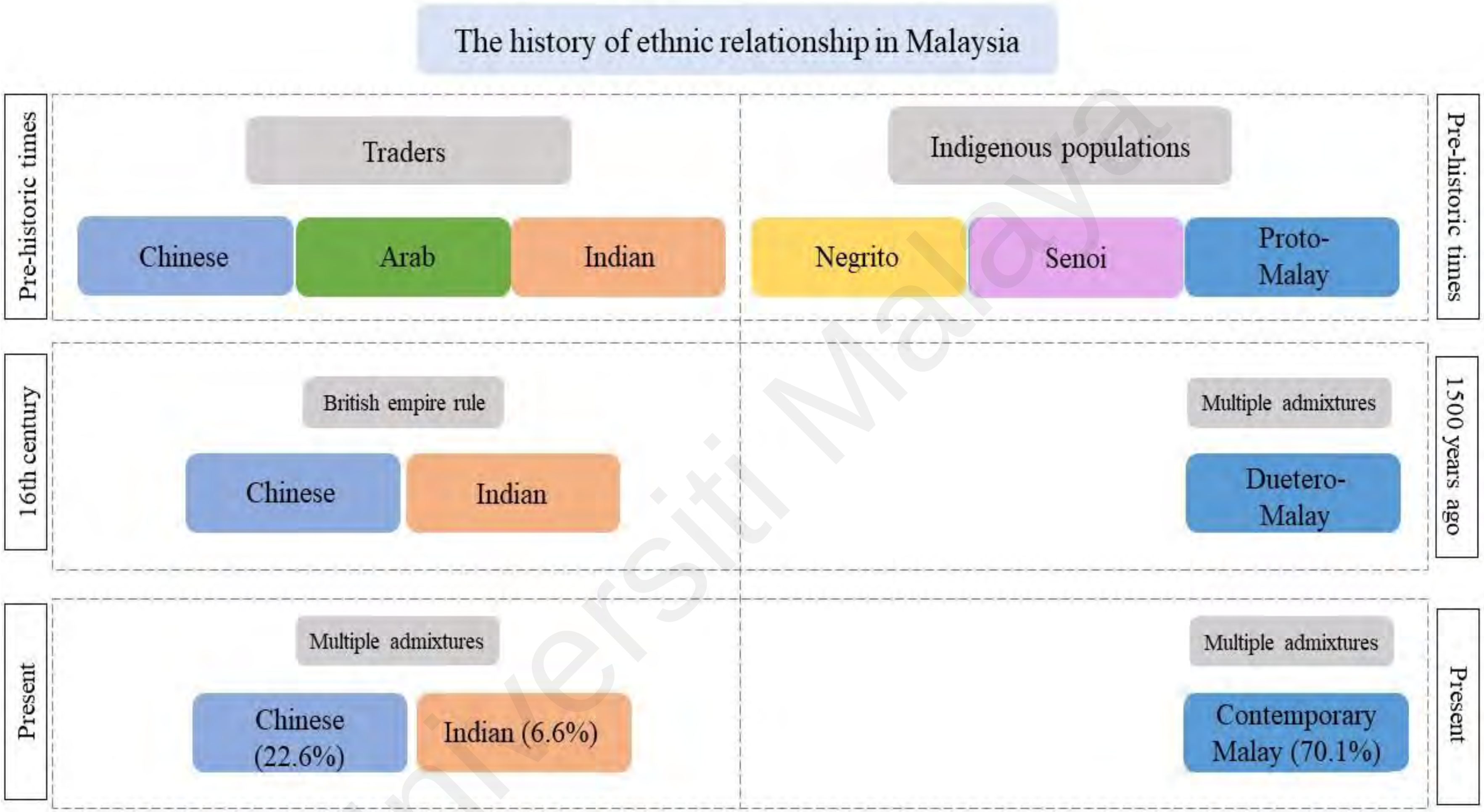


Figure 2.1: The multi-ethnicity history of Malaysia

2.4.3 Ethnic differences in the skull

The presence of ethnic differences in sub-adults is noted from an early age. One of the earliest examples of research relating to possible ethnic differences in sub-adults was the research conducted by Steyn and Henneberg in 1997. The skull of Black South Africans aged above five years is narrower and longer than the Europeans (Steyn & Henneberg, 1997). Similarly, interpopulation differences in the shape and size of the skull were found in children above five years in the South American population (Gonzalez et al., 2010). The biasterionic breadths of Asian and African crania are smaller compared to that of the Europeans. However, Asian and European crania exhibit similar nuchal and occipital angles, plane proportions, and lower inion positions than Africans (Zhang & Schepartz, 2021).

Variation in the growth of nasal floor shape was noted in Euro-Americans, African Americans, Africans, Europeans, Australians, and Asians. Variation in the nasal floor depression appears at around the age of three (Nicholas & Franciscus, 2014). However, variations of the nasal spine, nasal border, and palate shape were noted from the fetal period in black and white perinates (Limson, 1932; Weinberg et al., 2005). This is because variability in nasal bone contour is present in the prenatal period (Limson, 1932; Wood, 2015). In conclusion, variation in craniofacial traits is present at birth and becomes more apparent during postnatal development (Vioarsdóttir et al., 2002; Vioarsdóttir & Cobb, 2004).

2.4.4 Methods of ethnicity estimation in adults

2.4.4.1 Stull et al. (2014a)

This study consists of 377 crania of South African black, white, and coloured individuals. Data were analysed by generalised Procrustes analysis (PC) and Procrustes coordinates (ProCoords), a system that produces size-free shape parameters.

Craniometric measurements, ProCoords, and ProCoords PCs were subjected to LDA. LDA of the ProCoords PCs obtained an accuracy of 79% using 18 stepwise selected parameters. Meanwhile, LDA of craniometric analysis (D^2 matrix) demonstrated similar sizes between South African coloured and South African whites. However, both geometric morphometric analyses demonstrated more similarity in shape between the coloured and the blacks. Despite complex genetic admixture found to be present in the South African populations, high cross-validated accuracy (89%) was obtained from LDA using craniometrics and geometric morphometrics. Geometric morphometrics outperformed traditional craniometrics and principal component analysis (PCA) of shape variables in estimating the ethnic group of unknown individuals in South Africa (Stull et al., 2014a).

2.4.4.2 Kranioti et al. (2018)

This study analysed 297 crania of Greek and Cyprus origin and 380 CT scans of Turkish individuals. A total of 12 craniometric measurements were taken in both dry and virtual skulls. Data were subjected to PCA and DFA. Classification accuracy was obtained between 74.1% and 97.9%, with the highest accuracy obtained by both sexes in the Turks sample. In the PCA and DFA analyses, Cretans and Cypriots clustered more closely than the Turks sample. Thus, combining the Cretan and Cypriot samples resulted in 96%–98% increased accuracy (Kranioti et al., 2018). The findings of this study agree with genetic data suggesting a higher genetic similarity between Cypriots and Greeks compared to between Cypriots and Turks, as well as between Greeks and Turks (Hellenthal et al., 2014).

2.4.4.3 Herrera and Tallman (2019)

This study analysed 190 cranial CT scans of the Dominican Republic and Haiti ethnicities. A total of 28 craniometric measurements were taken on the skull. Classification accuracy of between 53.6% and 71.4% for both sexes was obtained. Additionally, when 12 canonical DFA were developed, cross-validated accuracies increased to 72.0%–77.8% for both sexes. Males achieved higher classification rates (71.8%–87.5%) than females (73.7%–78.6%). Haitians were more correctly classified compared to the Dominicans. The percentage of accuracy reflects the genetic heterogeneity between Haitians and Dominicans. Hence, in certain forensic contexts, Latin Americans should not be grouped under the broad umbrella term "Hispanic" (Herrera & Tallman, 2019).

2.4.4.4 Spradley (2021)

This study utilised craniometric measurements to estimate the ethnicity of migrant remains at the United States/Mexico border. This research consisted of 690 positively identified and unidentified migrants from Arizona and Texas. The craniometric data were compared to skeletal data representing Guatemalan and Mexican Mayans by utilising biological distances and DFA. The biological distances between the groups demonstrated that Guatemalan and Mexican migrants were close to each other but varied from Guatemalan Mayans and Mexican Mayans. Guatemalan Mayans obtained the highest overall accuracy rate at 90%, followed by Mexican Mayans at 77%, Mexican migrants at 57%, and Guatemalan migrants with the lowest accuracy rate of 40%. Therefore, it is possible to estimate an unknown individual from Mexican Mayans and Guatemalan Mayans groups. However, estimating the origin between Mexican migrants and Guatemalan migrants is currently impossible. This is because the migrants share a population history that incorporates varied levels of European and African admixture, creating similar population structures that cross geopolitical borders (Spradley, 2021).

2.4.4.5 Kongkasuriyachai et al. (2022)

This study aimed to estimate ethnicity in Thai and Japanese individuals using craniometric measurements. A total of 440 modern Thai and modern Japanese skulls were measured to develop discriminating formulae. Stepwise DFA was utilised to develop four formulae consisting of a complete skull, cranium, male skull, and female skull. The accuracy obtained ranged between 84.3% and 92.0% with the highest accuracy obtained when using a complete skull. The difference in skull characteristics between Thai and Japanese reflected a complex Asian population with regional differences between the northern and southern continents. This is demonstrated by the outcome of this study where Japanese or northern skulls were found to be longer and narrower than Thai or southern skulls. Hence, geographical differences do not only influence genetics but also affect the adaptation process due to different climates (Kongkasuriyachai et al., 2022).

2.4.5 Methods of ethnicity estimation in sub-adults

2.4.5.1 Vioarsdóttir et al. (2002)

Vioarsdóttir et al. (2002) conducted research to examine variations in the facial skeleton of 334 sub-adult individuals aged one to 19. The samples were from 10 discrete populations: Polynesians, Arikara Plains, Australians, Papua New Guineans, Alaskan Inupiaq Eskimo, Egyptian, Aleutians, African Americans, West African Ashanti, and French/British Caucasians. Facial landmarks and geometric morphometric methods were utilised. On average, 71% were correctly assigned using DFA, with two populations (Caucasians and Polynesians) being the most significant in the analysis. This demonstrated the presence of facial shape differences between the populations regardless of age. In conclusion, variation in facial shape is present at birth and continues to be more pronounced throughout growth and development (Vioarsdóttir et al., 2002).

2.4.5.2 Buck and Vidarsdottir (2004)

Buck and Vidarsdottir (2004) used geometric morphometric methods to examine ethnic differences in sub-adults' mandibles. The sample comprised 174 individuals from the ethnic groups African Americans, Native Americans, Caucasians, Inuit, and Pacific Islanders. Statistically significant differences were obtained with an accuracy of 70.1% of individuals correctly assigned to the group. Additionally, they found that comparisons between groups provided greater accuracy in the correct ethnic group assignment when the number of groups being compared was reduced to three populations (African American, Native American, and Caucasian). The African American sample had the lowest accuracy rate of correctly assigned individuals (84.62%) while the Caucasian sample had the highest accuracy (90.48%). However, cross-validation accuracy decreased to 56% and 79% when partial mandibular remains were used, with an average of 73% of correctly identified individuals for the mandibular ramus and 67.2% for the mandibular corpus (Table 2.2). Although accuracy decreased when using a partial mandible, the accuracy rate remains high. This suggests that ethnicity can be determined using either complete or partial mandible (Buck & Vidarsdottir, 2004).

Table 2.2: Cross-validation accuracy (%) of the mandibular ramus and mandibular corpus between five populations and three populations

		African American	Native American	Caucasian	Inuit	Pacific Islander
Complete mandible	**Five populations**	69.23	73.17	76.20	63.64	68.48
	Three populations	84.62	87.80	90.48	-	-
Partial mandible	**Ramus**	79.49	68.29	71.43	-	-
	Corpus	56.41	65.00	79.07	-	-

2.4.5.3 Weinberg et al. (2005)

Weinberg et al. (2005) analysed 13 nonmetric craniofacial traits in white and black Americans with a mean age of nine months. White infants were found to exhibit narrow occipital squamae, pronounced vomers and anterior nasal spines, "deep" sub-nasal margins, and semi-circular temporal squamae compared to black infants. Based on these five traits, an accuracy of 79.1% was obtained by using stepwise logistic regression. However, the accuracy decreased to 67.5% when an independent sample was utilised for cross-validation. The low cross-validation accuracy could be due to black perinates obtaining only 53.8% compared to white perinates with 100% accuracy, resulting from an unequal distribution of black perinates in the validation sample. In addition, the methods used were biased toward identifying white perinates. Unfortunately, only a limited number of nonmetric traits can be utilised due to disarticulated skull bones. Moreover, some bones were too small and not fully developed in perinates, resulting in some difficulty locating the landmarks. In addition, some perinates had been diagnosed with diseases such as tuberculosis, rickets, and syphilis, thus affecting skeletal morphology. Furthermore, the samples do not fully represent early 21st-century perinates, as they were derived from a skeletal collection in the first quarter of the 20th century. However, this study may be useful for anthropologists encountering unidentified cranial material from this age range (Weinberg et al., 2005).

2.4.5.4 Smith et al. (2013)

Smith et al. (2013) examined the temporal bone morphology of 133 sub-adult individuals. The samples were collected from modern sub-adults of New World populations (Alaskan, Mexican, Peruvian, and European Americans) as well as Old-World populations (Austrian, Egyptian, and Polynesian). DFAs were developed, and the correct classification of individuals was obtained at rates of between 7.7% and 65.8%.

The highest misclassification rates occurred within the New World populations, with Utah and Mexico obtaining the lowest correct classification rates of 7.7% and 28.6%, respectively. The results reported that groups of closer geographic origin could be greatly misclassified compared to groups in geographically distant populations. Additionally, there were strong correlations between sub-adult morphology and the molecular distance matrix. This indicated that sub-adult temporal bone shape reflects genetic affinity, especially in the age group before the first molar eruption. Although the samples used were small, variation in the temporal bone structure was exhibited as early as at five years of age (Smith et al., 2013).

2.4.5.5 Szen (2018)

This study examined ethnicity in sub-adult skeletons using metric analysis and 3D digitiser. The sample included 169 individuals ranging in age from birth to 21 years obtained from the Hamann-Todd, Terry and American Museum of Natural History Osteological Collections. The analysis yielded no statistically significant results. However, this does not necessarily mean that the ethnicity of the individuals could not be determined. It is possible that the ethnic-related craniofacial differences might not be metrically visible until after the completion of puberty. Alternatively, the ethnicity-related variations may be too minute for quantification with metric methods such as the 3D digitisation used in the study (Szen, 2018).

2.5 Sex and ethnicity estimation using machine learning (ML) algorithm

In recent years, the method of sex and ethnicity estimation has been aided by several ML classification algorithms. ML has the capacity to make predictions without being explicitly programmed to do so. Mathematical models generated from training data can be used to provide a versatile and robust solution for sex and ethnicity estimation.

Researchers have investigated several ML strategies for estimating sex and ethnicity in adults with high accuracies (D'Oliveira Coelho & Curate, 2019; Imaizumi et al., 2020; Nikita & Nikitas, 2020; Yang et al., 2019b). However, limited research has been conducted using ML to estimate sex and ethnicity in sub-adult individuals (Ortega et al., 2021). In the following section, a selection of literature that focuses on different classification methods employed for sex and ethnicity estimation in adult and sub-adult individuals worldwide will be reviewed. A summary of previous studies on sex and ethnicity estimation in adults and sub-adults is presented in Table 2.3.

Table 2.3: Previous studies of sex and ethnicity estimation using machine learning methods in adults and sub-adults

Authors	Population	M	F	T	Age (years)	Analysis	Accuracy (%)
Bertsatos et al. (2020)	European	181	143	324	19-99	SVM	71.7-96.7
Bewes et al. (2019)	Australia, Italy, German, China, India, Vietnam, Ireland, and Scotland	500	500	1000	18-60	ANN	95
Hefner et al. (2015)	White, Black, Hispanic	373	244	718	16-99	ANN, SVM, OSSA, RF, DT, KNN, QDFA, LR, KPD, LDFA, NBC	76.5-87.8
Imaizumi et al. (2020)	Japan	50	50	100	23-65	SVM	90-100
Ortega et al. (2021)	Mediterranean	83	52	135	Five months -six years	ML (SVM, RF, AdaBoost) and DL (VGG16, ResNet50)	49-61
Pozzi et al. (2020)	Italian	265	265	530	18-70	DT, KNN, LDA	40-85.5
Toneva et al. (2021)	Bulgarian	169	224	393	18-94	SVM, ANN, LR	95
Yang et al. (2019b)	China	114	153	267	18-88	MLP	94

M=male, F=female, T=total, ANN=artificial neural network, DL=deep leaning, DT= decision tree, KNN=k-nearest neighbors, KPD=kernel probability density, LDFA=linear discriminant function analysis, LR=logistic regression, MLP=multi-layer perceptron, NBC=naive bayes classifier, OSSA=optimized summed scored attributes, QDFA=quadratic discriminant function analysis, RF=random forest, SVM=support vector machines, VGG16=visual geometry group 16.

2.5.1 Hefner et al. (2015)

This study aimed to estimate the ethnicity of unknown cranium of White, Black, and Hispanic groups using morphoscopic traits. Samples from a total of 373 males and 244 females aged 16 to 99 years were obtained. Out of the 11 ML methods used (ANN, SVM, OSSA, RF, DT, KNN, quadratic discriminant function analysis (QDFA), LR, kernel probability density (KPD), linear discriminant function analysis (LDFA), and naive bayesian (NBC)), ANN provided the highest overall classification accuracy (87.8%), followed by SVM (86.4%) and RF (85.5%). LR recorded the lowest overall classification accuracy of 76.5% (Hefner et al., 2015).

2.5.2 Yang et al. (2019a)

This study aimed to estimate the sex from the Uighur ethnic group in northern China. A total of 267 complete skull CT scans (consisting of 153 females and 114 males) ranging in age from 18 to 88 years were utilised in the study. An ensemble of shallow multilayer perceptron (MLP) was performed to conduct sex estimation based on six cranial measurements (cranial sagittal arc and chord, apical sagittal arc and chord, occipital sagittal arc and chord). All cases achieved higher than 94% accuracy (Yang et al., 2019a).

2.5.3 Bewes et al. (2019)

This study randomly selected 500 males and 500 females aged 18 to 60 years for estimation of sex using the skull. CT head scans from the Royal Adelaide Hospital were obtained which comprised individuals from Australia, Italy, German, China, India, Vietnam, Ireland, and Scotland. The ANN was trained using images of 900 skulls to virtually reconstruct CT scans obtained from the hospital. When evaluated using previously unknown skull images, the ANN achieved a sex classification accuracy of 95%. Males have a precision of 96% and females have a precision of 94%, with an overall accuracy of 95%. This demonstrated that AI systems based on ANN are well-suited for sex estimation of skeletal remains (Bewes et al., 2019).

2.5.4 Imaizumi et al. (2020)

This study aimed to estimate sex based on the skull, using ML on 3D skull shapes. A total of 100 skulls (50 males and 50 females) ranging in age from 23 to 65 years were obtained from the post-mortem computed tomography (PMCT) data in Tsukuba Medical Centre, Japan. The ML approach was adopted using an SVM with a radial basis function kernel. Partial least squares regression (PLS) and principal component analysis (PCA) were used to decrease the dimensionality of the data set, which increased the accuracy rate. The accuracy rates achieved for sex estimation in PCA+SVM and PLS+SVM were between 90% and 100% (Imaizumi et al., 2020).

2.5.5 Pozzi et al. (2020)

This research was conducted to determine the feasibility of estimating sex on two Italian populations (Sardinia and Bologna) using ML and craniometric data. The skeletal material aged 18 to 70 years were obtained from the Anthropological Museum of the University of Bologna, Italy. ML algorithms such as DT and KNN were compared with LDA. KNN obtained the highest accuracy of 84.5% and 85.5% for Sardinia and Bologna populations, respectively. DT obtained an accuracy of 75% and 81%, and LDA obtained the lowest accuracy of 40% and 83.3% for Sardinia and Bologna populations, respectively. Although Italian crania were sexually dimorphic, the characteristics that contribute to this dimorphism vary amongst groups. Indeed, whether ML or LDA was utilised, each method was accurate only with individuals of the same population. While ML and LDA frequently achieve comparable accuracy, it is noted that ML models provide greater consistency (Pozzi et al., 2020).

2.5.6 Bertsatos et al. (2020)

This research aimed to present an automated approach for sex estimation using cranial sex diagnostic characteristics. The proposed technique was developed and validated using two European population samples: a Czech sample of 170 crania constructed from anonymised CT scans, and a Greek sample of 156 crania from the Athens Collection. The automated technique involved extracting and assessing specific morphometric features from 3D models of glabella, mastoid process supraorbital ridge, and occipital protuberance. Then, further processing was carried out using computer vision and ML algorithm such as SVM. The classification accuracy was determined using a two-way cross-validation technique involving population-specific and population generic classification. Population-specific categorisation accuracy varied between 78.5% and 96.7%, but population generic classification accuracy varied between 71.7% and 90.8%. Overall, classification accuracy higher than 91% was obtained when all sex diagnostic parameters were combined. However, a 100% accuracy rate was achieved when about 75% of the sample was used with posterior probability sex estimations (Bertsatos et al., 2020).

2.5.7 Toneva et al. (2021)

This study selected 393 individuals (169 males and 224 females) aged 18 to 94 years from Bulgaria. Classification models for estimating sex were developed using 64 cranial measurements and 22 cranial indices. Two ML algorithms (SVM and ANN) and a traditional statistical analysis method (LR) were utilised. All three techniques recorded more than 95% accuracy rates, with SVM achieving the highest accuracy of 96.1% (Toneva et al., 2021).

2.5.8 Ortega et al. (2021)

This research aimed to estimate the sex of infants using macroscopic examinations of the ilium. Photographs of the ilium from a collection of identified infant skeletons at the University of Granada were obtained, comprising 135 individuals (83 males and 52 females) ranging in age from five months to six years. The efficiency of classic MLs (SVM, RF, and AdaBoost) and DL (VGG16 and ResNet50) were developed and compared to an expert's eye evaluation. DL approaches achieved a precision of 59%, greater than that of ML methods (49%). Therefore, the accuracy of DL was almost similar to the expert's observation under the same conditions (61%). This demonstrated the possibility of using AI algorithms to enhance methods for estimating sex. This development has received massive attention from AI researchers as various reported algorithms have primarily focused on sex estimation in adult individuals (Ortega et al., 2021).

2.5.9 Summary

This section has provided summaries of previous studies on population-specific sex and ethnicity estimation standards for several populations worldwide. The applicability of medical technology such as CT to formulate population-specific standards has been demonstrated. This is important as such technology can be used as a substitution for actual bone specimens, where there is lack of documented skeletal collections available for study. Traditional statistical analysis (such as DFA and BLR) is useful as it accurately estimates sex and ethnicity in sub-adults. However, it is noted that the ML algorithm could potentially outperform the traditional statistical analysis. Hence, choosing the right statistical tool is vital to maximise accuracy rate. Furthermore, the sex and ethnicity estimation models generated were population-specific and can only be applied with the highest accuracy to populations similar to the sample populations (Spradley, 2021;

Walker, 2008). This demonstrated the necessity for population-specific models for estimating sex and ethnicity.

2.6 Cephalic index (CI)

The CI is an important parameter in assessing sexual and ethnic differences. CI was invented by Professor Anders Rezitus and was first used to classify human remains in Europe (Rakosi et al., 1993). The concept of CI can be defined as the ratio of the maximum width of the skull to the maximum length, multiplied by 100 (van Lindert et al., 2013). The width is the distance between the most projecting points at the sides of the head, and the length is the distance from the glabella to the most projecting point at the back of the head (Farkas et al., 1992; Standring & Gray, 2008; van Lindert et al., 2013). The original technique to measure CI was by using a caliper; however, various imaging technologies, such as CT scan and MRI, are now available instead of taking manual measurements (van Lindert et al., 2013).

Head shapes are classified based on the CI value and divided into four categories: dolichocephalic, mesocephalic, brachycephalic, and hyperbrachycephalic (Soames, 1995). A dolichocephalic head, or “long head,” describes an individual with a narrow cranial breadth. It usually presents a long, narrow shape and a high mandibular plane angle. A mesocephalic head, or “medium head,” describes an individual with an average cranial breadth. A brachycephalic head, or “short broad head,” describes an individual with a longer cranial breadth. It is usually associated with a broad, square head shape and low mandibular plane angle. A hyperbrachycephalic head, or “very short broad head,” describes a larger cranial breadth. It also presents a broader square head shape than a brachycephalic head. Correlation between CI as well as head shapes and age, sex, and ethnicity is important. It is valuable in plastic and reconstructive surgeries associated with craniofacial deformities and the prognosis of orthodontic treatment (Soames, 1995).

2.6.1 CI classifications

Various CI classifications have been introduced by different researchers regarding the head shape. Standring and Gray (2008) classified dolichocephalic as having CI of less than 74.9, mesocephalic (75.0–79.9), brachycephalic (80.0–84.9), and hyperbrachycephalic as having CI of more than 85.6. Meanwhile, Cohen and MacLean (2000) defined head shape categories as follows: dolichocephalic (less than 75.9), mesocephalic (76.8– 80.9), brachycephalic (81.0–85.4), and hyperbrachycephalic (more than 85.5). Koizumi et al. (2010) introduced a classification according to which CI under 76 signifies dolichocephalic, CI of 76–80.9 signifies mesocephalic, while CI exceeding 81.0 signifies brachycephalic. Nam et al. (2021) classified the CI of Korean children as follows: dolichocephalic (80.1 or less), mesocephalic (80.2–93.4), brachycephalic (93.5–100), and hyperbrachycephalic (100.1 or higher). Therefore, these classifications should be interpreted within the context of the specific population or sample being studied. There may be variations between head shape categories among ethnic groups, geographic regions, or age range. Thus, researchers should consider population-specific norms. Furthermore, considering reports that the standards of the CI may change over time, it is not sensible to apply any previously reported normal range to the current clinical decision-making in the Malaysian population (Koizumi et al., 2010).

2.6.1.1 CI in Indian population

Several studies have been conducted to measure the CI of the Indian population in different geographic regions (Ghosh, 2018; Khanduri et al., 2021; Yagain et al., 2012). A study by Ghosh (2018) reported that the dominant head shape in East India was brachycephalic. The overall mean CI was 81.09±3.42, with males obtaining higher CI (81.2±5.23) than females (80.76±5.88). In Central India, males were mesocephalic, while females were brachycephalic, with mean CI values of 77.92 and 80.85, respectively, and an overall mean CI of 77.92 ± 5.20 (Yagain et al., 2012).

These studies suggested that there is a correlation between CI and climatic variation within the Indian population. The head shape appears to be influenced by environmental conditions, with brachycephalic head shapes being more prevalent in temperate zones as compared to dolichocephalic head shapes being more prevalent in tropical zones (Bharati et al., 2001; Khanduri et al., 2021). It is important to note that India encompasses both temperate and tropical regions, which contributes to the diversity of head shapes observed within the population.

2.6.1.2 CI in Nigerian population

Several studies have been conducted to identify the CI of the Nigerian population (Akinbami, 2014; Eroje et al., 2010; Musa & Danfulani, 2015). Eroje et al. (2010) studied school children from the Ogbia tribe in Nigeria. A total of 440 sub-adults (219 males and 221 females) were randomly selected with their ages ranging from two to 18 years. The cranial length and width were measured with a spreading caliper. The overall mean CI was 72.96±6.12 with males obtaining higher CI (73.68±6.53) than females (72.24±5.60). Based on the CI, most of the head shapes were dolichocephalic (66.82%), followed by mesocephalic (21.59%), brachycephalic (10.23%), and hyperbrachycephalic (1.36%; Eroje et al., 2010).

A similar study was carried out by Akinbami (2014) from the Ogbia tribe in Nigeria. A total of 700 sub-adults aged 11 to 20 years were selected. The overall mean CI was 76.56, with males having a higher CI (77.21) than females (76.50). Based on this study, the majority of individuals were mesocephalic (78.68%), followed by dolichocephalic (11.4%), brachycephalic (9.0%), and hyperbrachycephalic (0.43%; Akinbami, 2014).

Another study was conducted by Musa and Danfulani (2015) on 76 children (42 males and 34 females) aged four to seven years. The authors reported that the mean CI value for the entire sample was 79.12±3.37 with higher CI values in females (80.08±3.34) than in males (78.35±3.23). By applying Strandring's head shape classification, the most frequently occurring type of head shape was mesocephalic (55.26%), followed by brachycephalic (28.95%), dolichocephalic (7.89%), and hyperbrachycephalic (7.89%; Musa & Danfulani, 2015). These findings differed from those of previous studies conducted on Nigerian children by Eroje et al. (2010) and Akinbami (2014), which reported lower CI values. This suggested that age range differences in the study sample may contribute to the observed variation in CI values. Furthermore, it is important to consider that the results obtained from this age group may not be directly comparable to those obtained by studies with broader age ranges. This is because younger age groups tend to have higher CI values than older age groups (Pereira et al., 2008).

2.6.1.3 CI in Polish Population

Likus et al. (2014) aimed to develop a CI classification for Polish children with normal brain development from birth to three years of age. The authors found that the mean value of CI in children was 81.45±7.06 with the CI of males (82.22±6.87) being higher than females (80.54±7.20). The study also examined the variation in CI values across different age groups. The CI value for children under three months old was recorded as 80.19, which increased to 81.45 between the ages of four and six months. In the age range of seven to 12 months, the CI value further increased to 83.15. However, for children aged two years, the CI value decreased to 81.05; for children aged three years, it further decreased to 79.76. No significant difference in CI values was observed between males and females. Nevertheless, within the age group of seven to 12 months, males had higher

CI values than females. By applying Cohen's skull shape classification, mesocephalic (34%) was the dominant head shape of Polish children (Likus et al., 2014).

2.6.1.4 CI in Siberian population

The study was conducted to compare data from two groups of students aged seven to 15 years from Southeastern Serbia, measured in two different years, namely 1983 and 2010. Data for the first group were obtained in 1983 and involved 968 students, while data from 1,037 students were collected in 2010. The students were divided into three age groups: seven to nine years, 10 to 12 years, and 13 to 15 years. In 1989, the overall mean value of CI was 82.45±4.66, with higher values observed in males (82.62±4.84) compared to females (82.27±4.47). The specific CI values reported for each age group were 81.34 for the age group of 7-9 years, 83.21 for the age group of 10-12 years, and 82.80 for the age group of 13-15 years. In 2010, the overall CI value was lower at 80.68±4.9, with higher CI values observed in females (80.91±4.7) compared to males (80.45±5.1). The specific CI values reported for each age group in 2010 were as follows: 80.37 for the age group of 7-9 years, 81.51 for the age group of 10-12 years, and 80.17 for the age group of 13-15 years. Notably, higher CI values were noted in children aged 10 to 12 across both the 1989 and 2010 studies. The dominant head shape in 1983 was brachycephalic, whereas in 2010, it was mesocephalic. The increase in the width and length of the head had caused the debrachycephalisation process to occur (Cvetković et al., 2014). Therefore, this supports the reported study of debrachycephalisation in Europe in recent decades (Buretić-Tomljanović et al., 2004; Grbeša et al., 2007; Paulová et al., 2000).

2.6.1.5 CI in Iranian population

Several studies have been conducted on the Iranian population to determine their CI (Golalipour, 2006a, 2006b; Golalipour et al., 2003, 2005). Golalipour et al. (2003) sought to determine the CI value in male newborns from Gorgan-North of Iran. This study was conducted on 420 male newborns of Fars and Turkman ethnic groups within 12-24 hours after delivery. The CI value in the Turkman group was 77.00±5.91, and the CI value in the Fars group was 77.97±5.34 with no significant difference between the two groups. The dominant type of head shape in both the native Fars and Turkman groups was mesocephalic (36.5% and 38.2%, respectively; Golalipour et al., 2003).

The authors repeated the same study for female newborns in 2005. The study was conducted on 423 female newborns from Fars and Turkman ethnic groups in Iran. The CI value in the Turkman group was 77.8±58.7 and the CI value in the Fars group was 78.63±4.7. The CI of female newborns was found to be higher than male newborns. In this case, the dominant type of head shape in both the native Fars and Turkman group was mesocephalic (41.98% and 38.86%), similar to the male newborns (Golalipour et al., 2005).

Golalipour (2006b) conducted research in Gorgan, North of Iran, focusing on young Fars and Turkman males and females aged 17 to 20 years. In the female Turkman group, the mean CI was found to be 82.8±3.6, while in the female native Fars group, it was 85±4.5. The dominant head type in the Fars group was hyperbrachycephalic, accounting for 53.6% of the individuals, while the Turkman group had a dominant brachycephalic head type at 58.1% (Golalipour, 2006a). For the male Turkman and Fars groups, the mean CI values were slightly lower compared to females, with 80.4±4 for Turkman males and 84.8±6.9 for Fars males. As with the female groups, hyperbrachycephalic was the dominant head type in both the male Turkman and Fars groups (Golalipour, 2006b).

2.6.1.6 CI in Brazilian population

This study was conducted to evaluate the CI of Brazilian sub-adults ranging in age from birth to 18 years. Overall CI values reported no significant differences between male (85.45±4.08) and female (85.45±4.6) individuals. In addition, a marked decrease in the mean value of CI was observed in the first year of life, followed by a milder decrease between the ages of four and five years. Thereafter, the CI values stabilised up to the age of 18 years (Pereira et al., 2008). This indicated that CI is a good indicator for skull growth in the early years due to the increased rate of skull growth during that time. When the skull growth has stabilised thereafter, it can become a good indicator for follow-up.

2.6.1.7 CI in Asian population

Several studies conducted on Asian populations (such Japanese, Korean, and Malaysian populations) have been found to exhibit higher CI values compared to other populations, such as Caucasians (Koizumi et al., 2010; Nam et al., 2021; Swamy et al., 2013; Yusof, 2007). In a study on Japanese children ranging from birth to four years, Koizumi et al. (2010) reported a relatively high CI value of 86.5±7.3. Among the participants, males reported slightly higher CI values (87±7.5) compared to females (86.3±6.5; Koizumi et al., 2010). The significant difference in CI between Asian and Caucasian populations rendered Cohen's classification unsuitable for Japanese children, as it would categorise them as hyperbrachycephalic. To address this discrepancy, Koizumi et al. (2010) developed a new CI classification specifically for Japanese children using the obtained data, where the mean value of CI±1 SD was defined as mesocephalic. Therefore, the most prevalent head shape among Japanese children was found to be brachycephalic. This has provided a more accurate categorisation based on their population-specific characteristics (Koizumi et al., 2010).

A study on Korean children revealed a high average CI value of 86.82±6.66. In this population, males exhibited slightly higher CI values (86.80±6.32) compared to females (86.84±7.13). Notably, Korean children showed even higher CI values than the Japanese population (Koizumi et al., 2010). Therefore, following the methodology established for Japanese children, a modified CI classification was developed by Nam et al. (2021)

specifically for Korean children. Based on the new CI classification, the most prevalent head shape among Korean children was found to be mesocephalic (Nam et al., 2021).

Several studies have been conducted to measure CI in Malaysian children, specifically within the Malay ethnic group (Swamy et al., 2013; Yusof, 2007). Yusof (2007) conducted a study to measure CI values in the Malaysian population aged from birth to 25 years. The author reported a relatively high CI among Malaysian children, with CI values of 84.8±5 for males and 85.2±4 for females (Yusof, 2007). On the other hand, Swamy et al. (2013) conducted a study on school children aged six to 16 years and reported slightly lower CI values, averaging 81.48±10.55 for males and 79.51±10.04 for females. The variation in CI value between the two studies could potentially be attributed to differences in sample size, age range, and measurement techniques employed. However, these studies solely focused on the Malay ethnic group and did not include individuals from Indian and Chinese ethnic backgrounds. Further research incorporating individuals from the Indian and Chinese ethnic backgrounds would enable a more comprehensive assessment of CI and its potential variations across the diverse ethnic groups in Malaysia. This would provide a deeper understanding of the cranial characteristics within the Malaysian population as a whole and contribute to the field of craniofacial anthropology and forensic sciences. A summary of CI in different populations is presented in Table 2.4.

Table 2.4: Cephalic index in different populations

Authors	Populations	Age (years)	Mean CI		Overall mean CI	Head shapes
			Male	Female		
Akinbami (2014) Eroje et al. (2010) Musa and Danfulani (2015)	Nigerian	11-20 2-18 4-7	77.21 73.68±6.53 78.35±3.23	76.50 72.24±5.60 80.08 ±3.34	76.56 72.96±6.12 79.12±3.37	Mesocephalic Dolichocephalic Mesocephalic
Cvetković et al. (2014)	Siberian	7-15	82.62±4.84 (1989) 80.45±5.1 (2010)	82.27±4.47 (1989) 80.91±4.7 (2010)	82.45±4.66 80.68±4.9	Brachycephalic Mesocephalic
Ghosh (2018)	East India	17-20	81.2±5.23	80.76±5.88	81.09±3.42	Brachycephalic
Golalipour et al. (2003) Golalipour et al. (2005)	Iranian	Newborn	Turkman group: 77.00±5.91 Fars group: 77.97±5.34	Turkman group: 77.8 5±8.7 Fars group: 78.63±4.7	Turkman group: 77.43± 7.31 Fars group: 78.3±5.02	Mesocephalic Mesocephalic
(Golalipour, 2006a) (Golalipour, 2006b)		17-20	Turkman group: 80.4±4 Fars group: 84.8±6.9	Turkman group: 82.8±3.6 Fars group: 85±4.5	Turkman group: 82.8±3.6 Fars group: 85±4.5	F=Brachycephalic (Turkman group), Hyperbrachycephalic (Fars group) M=Hyperbrachy-cephalic
Koizumi et al. (2010)	Japanese	0-4	87±7.5	86.3±6.5	86.5±7.3	Brachycephalic
Likus et al. (2014)	Polish	0-3	82.22±6.87	80.54±7.20	81.45±7.06	Mesocephalic
Nam et al. (2021)	Korean	0-7	86.80±6.32	86.84±7.13	86.82±6.66	Mesocephalic
Pereira et al. (2008)	Brazilian	0-18	85.45±4.08	85.45±4.6	85.45±4.34	-
Swamy et al. (2013) Yusof (2007)	Malaysian	6-16 0-25	81.48±10.55 84.8±5	79.51±10.04 85.2±4	80.5±10.3 85.2±5.01	- -
Yagain et al. (2012)	Central India	18-22	77.92	80.85	77.92±5.20	M=Mesocephalic F=Brachycephalic

CHAPTER 3: MATERIALS AND METHODS

This chapter outlines the materials studied in the current project, the methods applied for data acquisition, and the statistical approaches required to achieve the project aims. The first aim was to develop sex and ethnicity estimation models specifically for Malaysian sub-adults. This was achieved by classical statistical analyses such as DFA and BLR, by identifying the most meaningful combination of measurements that can be used to develop accurate classification models. Recently, there has been growing interest in exploring the application of ML algorithms to improve the accuracy of sex and ethnicity estimation models. Therefore, another aspect of the study involved comparing the validity and accuracy of the sex and ethnicity estimation models developed using ML algorithms with those developed using classical statistical methods (DFA and BLR) specifically for Malaysian sub-adults.

The second aim of the study was to propose a new classification for CI in the Malaysian population. It has been well-established that there are apparent sex and ethnicity differences in cranial dimensions and CI among different populations. Furthermore, the standards of CI may change over time due to various factors. Hence, it is impractical to rely on previously reported normal ranges for CI when studying the modern Malaysian population. Therefore, the study aimed to gather updated CI data specific to the Malaysian population, considering sex and ethnicity variations, in order to propose a new classification system for CI which will be applicable to the present-day population.

3.1 Ethics

The Medical Research Ethics Committee, University Malaya Medical Centre (UMMC) has approved the human research ethics aspect of this study (MREC ID NO: 202147-10039).

3.2 Materials

3.2.1 Study design

This retrospective study has been carried out based on CT records of male and female individuals of known age and ethnicity at the Universiti Malaya Medical Centre (UMMC). MSCT records used for this study were from July 2010 to March 2021.

3.2.2 Study sample

The sample utilised in this study included sub-adults from birth to 20 years of age. Initially, 682 MSCT scans were collected for this study. However, due to the radiation risks involved in acquiring MSCT scans, many cranial scans in the Picture Archiving Computerised System (PACS) database did not have all the landmarks necessary for the study and were therefore excluded. Thus, only 521 out of 682 MSCT scans were used in this study, which comprised 279 males and 242 females. The patient details, including those of the father and mother, were obtained from the hospital record-keeping system. Patients and their parents were ensured to have a Malaysian identification number (IC) in the registration information. All MSCT scans received were anonymised beforehand to maintain the patients' privacy. Therefore, birth date, scan date, sex, and ethnicity were the only available information that can be obtained from each MSCT scan. Each sample was categorised as belonging to one of six age groups as described by Klales and Burns (2017) with slight modifications: 0-2 years, 3-6 years, 7-9 years, 10-12 years, 13-15 years, and 16-20 years. The age groups were defined as follows: 0-2 years: day 1 of life to 2 years, 11 months, and 29 days, 3-6 years: 3 years to 6 years, 11 months, and 29 days, 7-9 years: 7 years to 9 years, 11 months, and 29 days, 10-12 years: 10 years to 12 years, 11 months, and 29 days,13-15 years: 13 years to 15 years, 11 months, and 29 days, 16-20 years: 16 years to 20 years, 11 months, and 29 days. Figure 3.1 shows the distribution of age, sex, and ethnicities in the study samples.

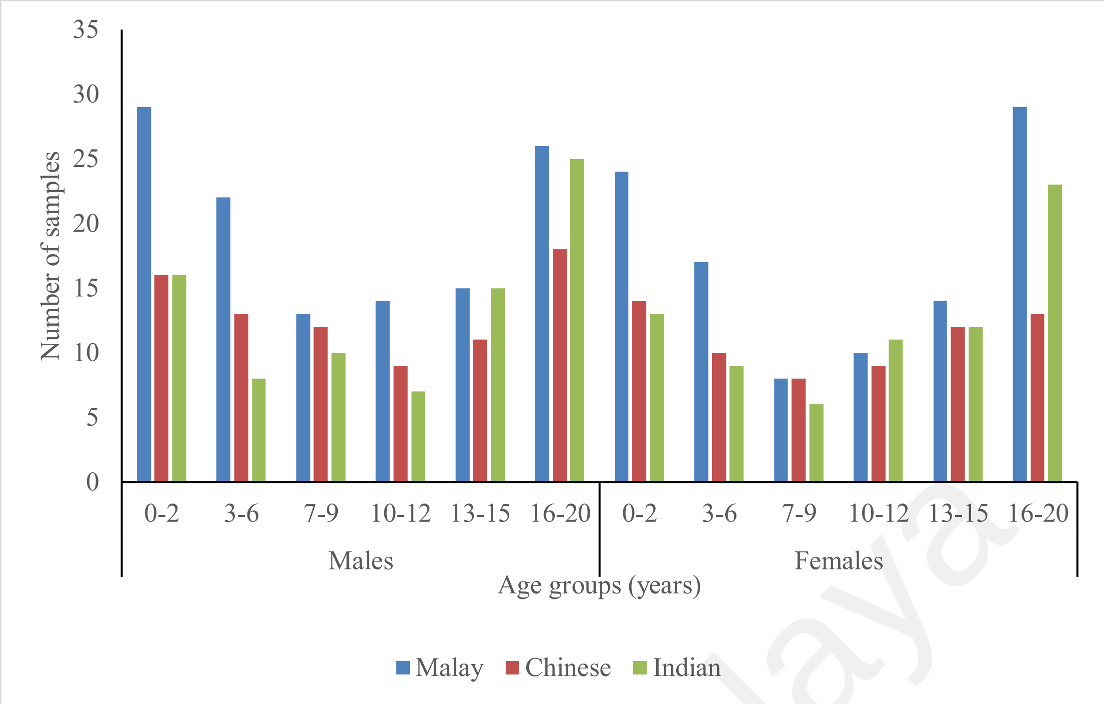


Figure 3.1: Distribution of age, sex and ethnicities in the study samples

3.2.3 Inclusion and exclusion criteria

A strict criterion was imposed when selecting MSCT images from the UMMC PACS database. MSCT images of individuals ranging in age from birth to 20 years with Malay, Chinese, and Indian ethnicities were included. In addition, good quality MSCT images with slice thickness of 1 mm were used. Parental ethnicity was confirmed to match that of the sample population; individuals whose parents did not share the same ethnicity were excluded from the study. Meanwhile, individuals above the age of 20 years and from other ethnicities were excluded from this study. MSCT images presenting a history of trauma, surgical procedures, deformities, or any pathological condition were excluded as these might have influenced normal cranial morphology. In addition, MSCT images with artifacts preventing landmark recognition and low resolution were also excluded.

3.2.4 3D imaging technique and software

The MSCT images were further enhanced with the Materialise Interactive Medical Image Control System (MIMICS) software to assess 3D cranial models.

3.2.4.1 Multi-slice computed tomography (MSCT) images

The MSCT images were retrieved from the PACS database, which is a repository of medical scan images for Malaysians that is monitored by the Radiology Department of UMMC. MSCT images were selected based on several parameters, such as tube voltage of 120 kV, effective 110-450 mAs, slice thickness of 1 mm, and exposure time of 0.4s. The present study included the acquisition of craniometric measurements in 3D volume of Malaysian sub-adult individuals.

3.2.4.2 Materialise interactive medical image control system (MIMICS) software

Materialise interactive medical image control system (MIMICS) version 21.0 (Materialise, Leuven, Belgium) software is a commercially accessible third-party software. MIMICS was used in the present study to obtain primary reconstructed images on multiplanar reconstruction views (axial, coronal, and sagittal). In addition, 3D reconstructions of cranial MSCT images were obtained for landmark identification, location, and measurements. The advanced and automated segmentation tools allowed for faster and easier reconstruction of 3D cranial models.

3.3 Methods: Craniometric measurements

3.3.1 Landmarks definition

Landmarks are defined as geometric locations of osteological points that are biologically homologous (Bookstein, 1991). Landmarks can be identified and measured using several methods: with calipers on a physical specimen, collected digitally from a physical specimen using a digitiser, or collected from a CT scan using computer programmes that allow image visualisation and manipulation. Bookstein (1991) defined three different types of landmarks: Types 1, 2, and 3. All three types are discussed below.

3.3.1.1 Type 1 landmarks

Type 1 landmarks have the strongest biological homology in tissues, such as the intersection of two sutures. They are the most desirable in biological studies because they are more likely to be homologous and therefore more precise (Ross & Williams, 2008). The most accurate and precise statistical quantification can be achieved when using Type 1 landmarks due to the ability to identify landmarks accurately and repeatedly at the junction of two or more anatomical structures. An example of a Type 1 landmark in the present study is bregma, which is the intersection between coronal and sagittal sutures (O'Higgins, 2000).

3.3.1.2 Type 2 landmarks

Type 2 landmarks are the maximum curvature points along tissue boundaries or with biomechanical forces. An example of a Type 2 landmark is prosthion, which is defined as the point on the maxillary alveolar process that projects most anteriorly in the midline (O'Higgins, 2000).

3.3.1.3 Type 3 landmarks

Type 3 landmarks are the least desirable for biological studies due to the lack of homology between specimens. They are landmarks with at least one deficient coordinate, such as the bottom of a concavity or the tip of a rounded bump. An example of a Type 3 landmark used in the present study is euryon, which is defined as the greatest transverse diameter of the head (O'Higgins, 2000).

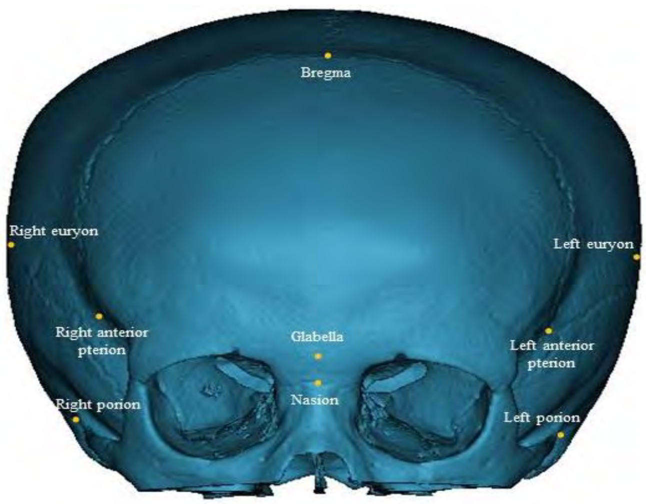
3.3.2 Landmark acquisitions

Materialise interactive medical image control system (MIMICS) software was used to identify 18 homologous landmarks in each cranial scan. The selection of the homologous landmarks and the measurements calculated thereafter were based on previous ontogenetic cranial growth and skeletal development research (Bastir et al., 2006; Baughan & Demirjian, 1978; Braga & Treil, 2007). Landmark definitions were accordingly adapted from several sources (Bass, 1987; Haas et al., 1994; Corner & Richtsmeier, 1991; Franklin et al., 2012; Howells, 1973). Locating some cranial landmarks in younger subjects could be challenging due to the presence of open spaces between the bones or open fontanelles, such as bregma and asterion. Hence, the landmarks were approximated by extending the present sutures to a meeting point (Noble, 2015). The definitions of these landmarks are listed in Table 3.1 and illustrated in Figure 3.2.

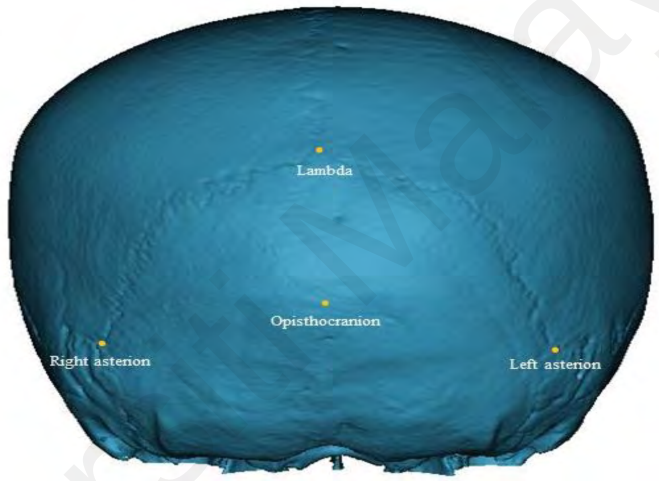
Table 3.1: Definitions of anatomical landmarks used for anthropometric analysis

Landmarks	Code	Definitions
Midline landmarks		
Basion	ba	Midpoint of the anterior margin of the foramen magnum[1].
Bregma	br	Intersection of the coronal suture and the sagittal suture[1].
Glabella	g	Most anterior point in the median sagittal plane above the naso-frontal suture; can be instrumentally determined[1].
Lambda	l	Point where the lambdoidal sutures and the sagittal sutures meet[1].
Nasion	n	Point where the naso-frontal suture meets the midsagittal plane; junction of the naso-frontal suture and the intranasal suture[1].
Opisthicranion	op	The most posterior point on the skull excluding the external occipital protuberance; instrumentally determined[1].
Opisthion	o	Midpoint of the posterior margin of the foramen magnum[1].
Sella	s	The midpoint of sella turcica[1].
Bilateral landmarks		
Anterior pterion	pt.r/pt.l	Intersection of the fronto-parietal and fronto-spehnoid sutures[3].
Asterion	ast.l/ast.r	Junction of the occipital, temporal and parietal bones[5].
Euryon	eu.l/eu.r	The two points on the opposite sides of the cranium that form maximum cranial breadth[2].
Foramen magnum lateralis	fml.l /fml.r	The most lateral point on the side of the foramen magnum[4].
Porion	po.l/po.r	The uppermost lateral point on the superior margin of the external auditory meatus[2].

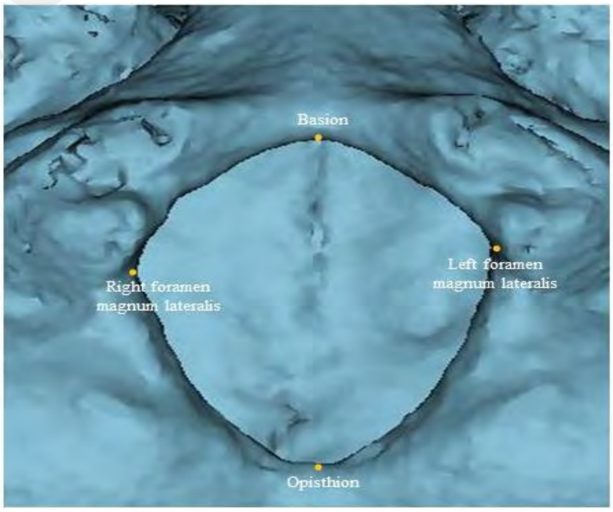
(Bass, 1987) [1], (Haas et al., 1994) [2], (Corner & Richtsmeier, 1991) [3], (Franklin et al., 2012) [4], (Howells, 1973) [5].



(A) Anterior view



(B) Posterior view



(C) Base view

Figure 3.2: Cranium anatomical landmarks used for anthropometric analysis from (A) anterior, (B) posterior, and (C) base view

3.3.3 3D cranial model generation

The raw images were acquired in Digital Imaging and Communications in Medicine (DICOM) format and uploaded into the MIMICS software. MIMICS was used to reconstruct primary images on multiplanar reconstruction views for purposes of landmark identification and measurements in 3D images. A gantry tilt correction (±30 degrees) was applied to images to normalise the slice angle. This was due to the position of the paediatric head, which may have slanted while the scans were performed. Frankfurt horizontal (FH; Li et al., 2015) and midsagittal (MSP; Zhang et al., 2018) planes were used to ensure that the craniums were aligned consistently. MSP was constructed by using three reference landmarks, nasion, sella turcica, and basion (Ono et al., 1992), and making them perpendicular to the FH (Li et al., 2015).

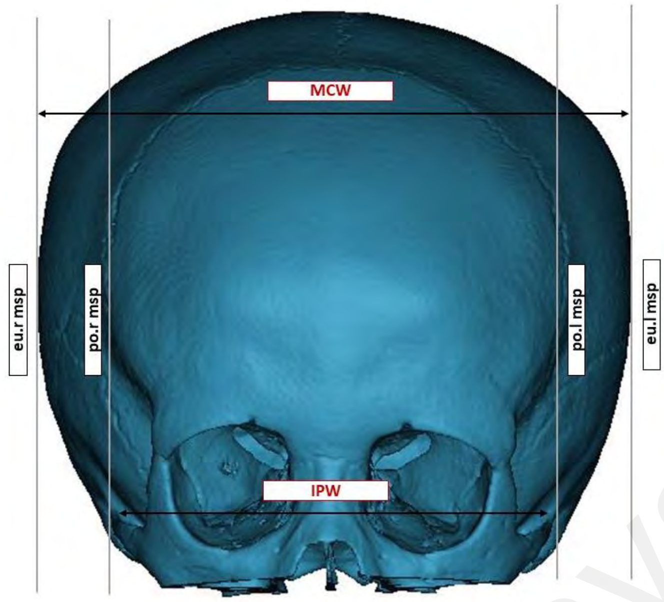
3.3.4 Plane-to-plane concept (PTP)

Plane-to-plane (PTP) protocol is dependent on the curve extraction and the placing of semi-automated planes between the curves' planes in offset or extreme positions. The protocol uses a reference plane system to define the extremities of surface-rendered models. In the present study, the reference system was generated using anatomical planes such as MSP, FH, and coronal planes. Then, offset planes (planes parallel to each anatomical plane) and extreme planes were created on a 3D cranial model or curve. Inter-plane linear measurements were taken between the offset planes for 14 craniometric measurements, as listed in Table 3.2. Examples of craniometric measurements are illustrated in Figure 3.3 to Figure 3.6. Pythagoras' theorem was used to calculate linear measurement of CH, CBL, OC, PC, and FC.

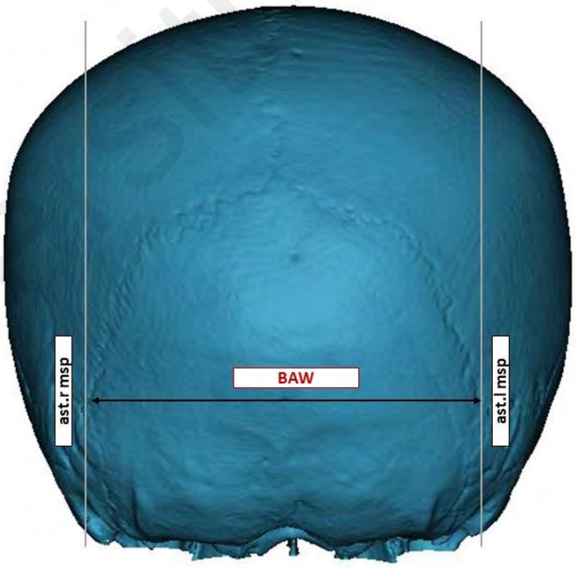
Table 3.2: Definitions of the 14 craniometric measurements

Parameters	Abbreviations	Landmarks	Definitions
Biasteronic width	BAW	ast.r-ast.l	The straight-line distance from left to right asterion[1].
Cranial base length	CBL	ba-n	Distance between the basion and nasion[1].
Cranial height	CH	br-ba	Linear distance from the bregma to the basion[1].
Foramen magnum length	FML	ba-o	Distance between the basion and the opisthion[1].
Foramen magnum width	FMW	fml.l-fml.r	The distance from left to right foramen magnum lateralis[1].
Frontal cord	FC	n-b	Distance from the nasion to the bregma measure in the MSP[1].
Interporion width	IPW	po.r-po.l	The distance from the right to the left porion[1].
Lateral cranial length (right & left)	LCL R/L	pt.r/pt.l-ast.r/ast.l	Linear distance from the anterior pterion to the asterion[1].
Maximum cranial length	MCL	ant frontal-op	Linear distance from the anterior point of the frontal to the opisthocranion in the MSP measured in a straight line[1].
Maximum cranial width	MCW	eu-eu	Maximum width of the cranium in the coronal plane[1].
Nasio–occipital length	NOL	n-op	Maximum length measured in the mid-sagittal plane from nasion to opisthocranion[1].
Occipital cord	OC	l-o	The distance from lambda to opisthion taken in the MSP[1].
Parietal cord	PC	br-l	The distance from bregma to lambda taken in the MSP[1].

(Langley et al., 2016)[1].

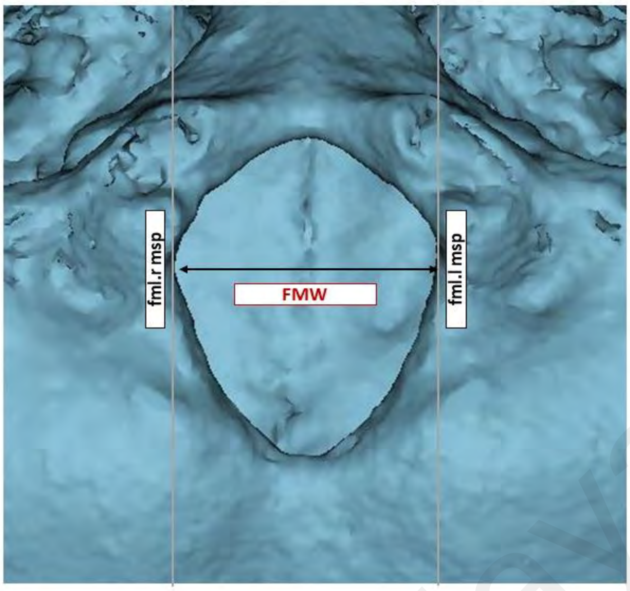


(A) MCW and IPW

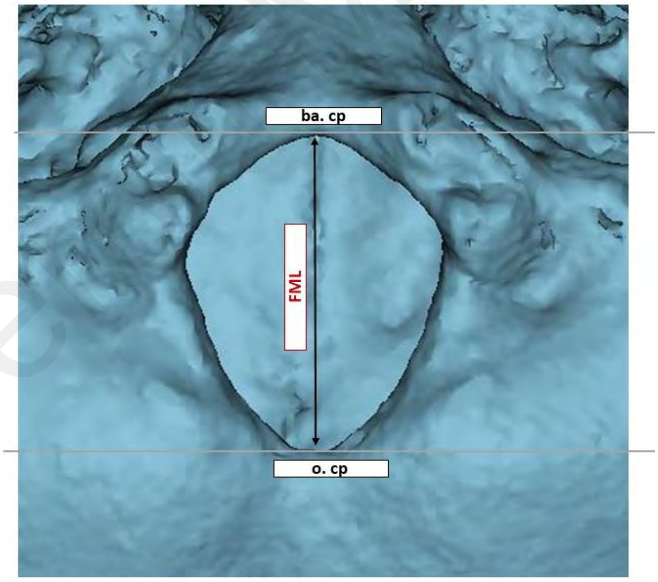


(B) BAW

Figure 3.3: Craniometric measurements from anterior and posterior view of (A) MCW and IPW and (B) BAW

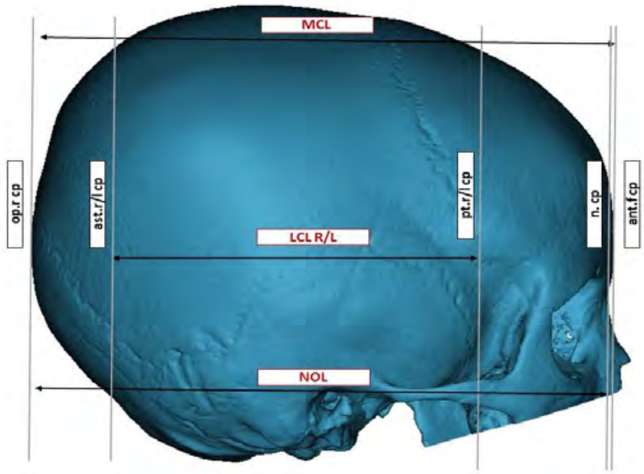


(A) FMW

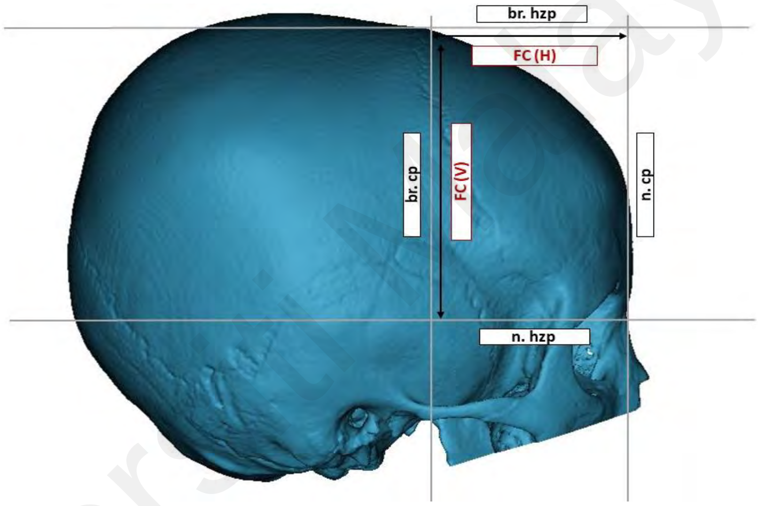


(B) FML

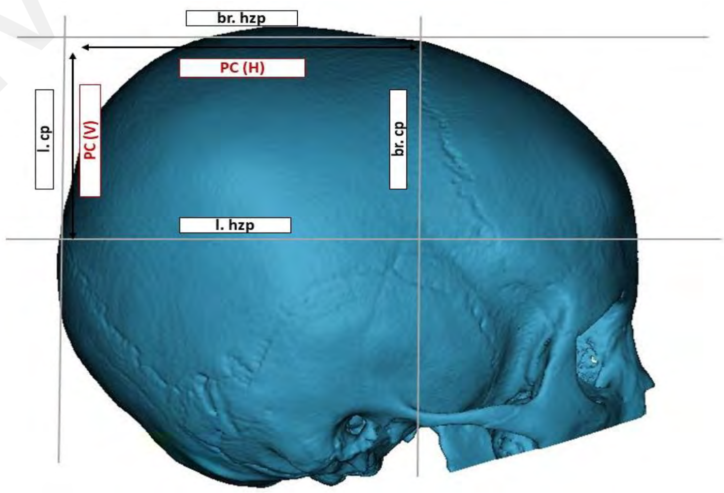
Figure 3.4: Craniometric measurements from base view of (A) FMW and (B) FML



(A) MCL, LCL (R/L), NOL

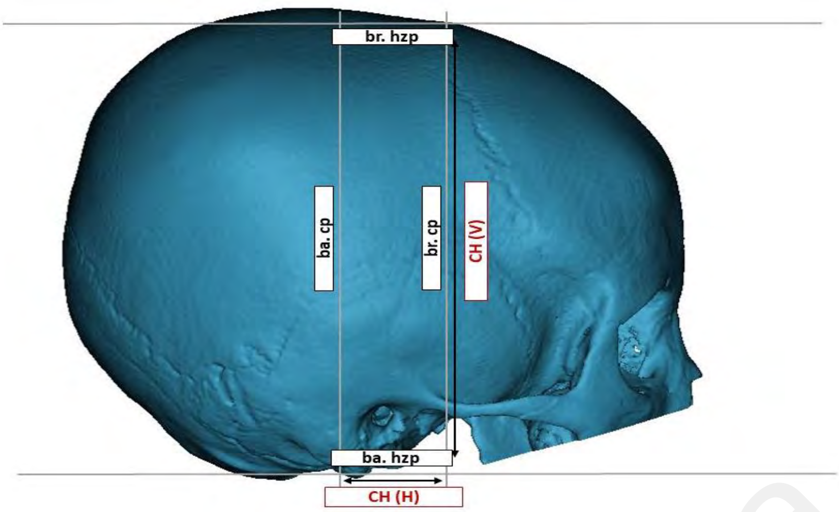


(B) FC

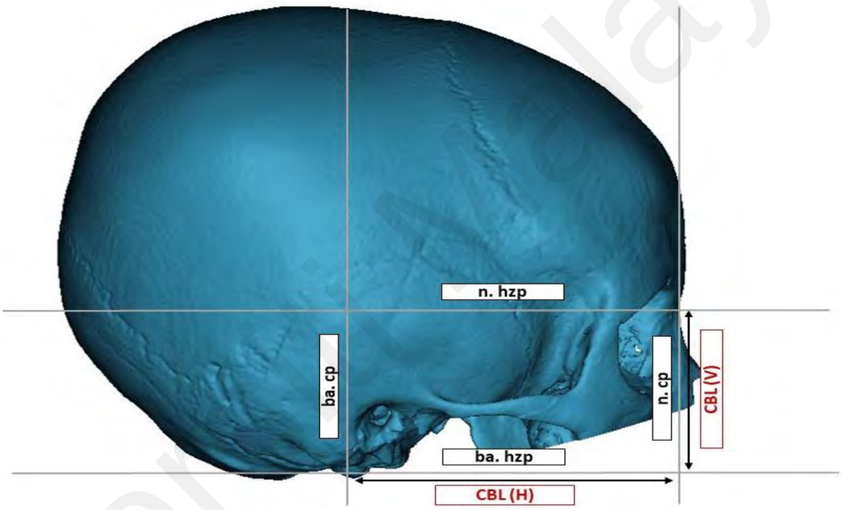


(C) PC

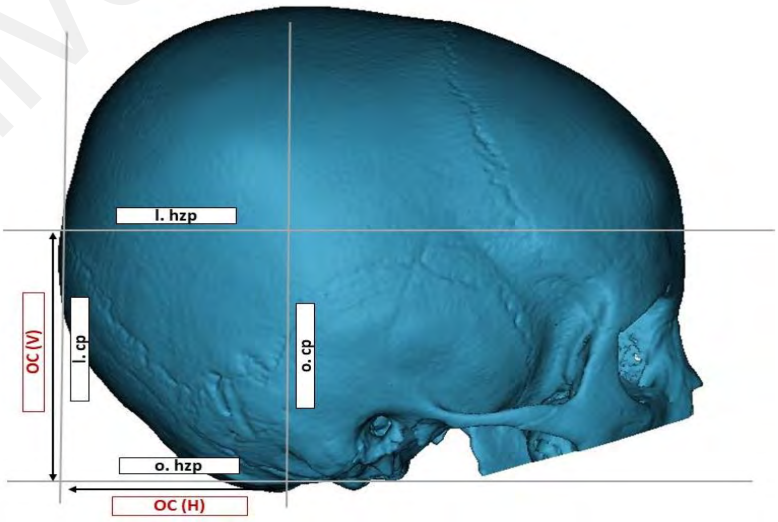
Figure 3.5: Craniometric measurements from lateral view of (A) MCL, LCL(R/L), NOL, (B) FC, and (C) PC



(A) CH



(B) CBL



(C) OC

Figure 3.6: Craniometric measurements from lateral view of (A) CH, (B) CBL, and (C) OC

3.4 Statistical analysis

The statistical package for the social sciences (SPSS; version 26.0; IBM, Armonk, NY) was used to conduct the statistical analyses.

3.4.1 Methods for sex estimation

The objective of this section is to present methods for sex estimation in Malaysian sub-adults using cranium. Two discriminant methods (DFA and BLR) were tested and compared. The accuracy was evaluated in estimating the sex of sub-adults.

3.4.1.1 Correlation test

A Pearson correlation test was conducted between all pairs of parameters. One parameter will be removed or combined in each pair of highly correlated parameters. This will decrease high correlations in the dataset and retain other parameters to generate classification models.

3.4.1.2 Descriptive statistics

All craniometric measurements were tested for normality by skewness and kurtosis calculations, along with histograms and normal distribution curves. Means±SD values were calculated for all parameters by age groups and sex.

3.4.1.3 Independent *t*-test

An independent sample *t*-test compares the means of two independent groups to determine the differences in a population. Significant values ($p<0.05$) were calculated using the mean±SD, minimum value, and maximum value.

3.4.1.4 Multivariate analysis of variance (MANOVA)

Multivariate analysis of variance (MANOVA) was conducted to test the hypotheses in the present study. It evaluates the significance of sex, age groups, and the relationship between sex and age groups on the craniometric measurements. In addition, it determined whether differences were statistically meaningful according to age and sex.

3.4.1.5 Sexual dimorphism percentage (SDP)

The SDP was calculated to identify how much males and females differ from each other throughout their development. This was calculated using the formula below as proposed by Ricklan and Tobias (1986), with a ratio between the mean values of males and females.

$$\frac{(\text{Male mean} - \text{Female mean})}{(\text{Male mean})} \times 100$$

3.4.1.6 Discriminant function analysis (DFA)

Discriminant function analysis (DFA) is a procedure that statistically assigns an unknown individual to the most accurate group. DFA uses independent parameters to estimate which group an individual belongs to (Pietrusewsky, 2000). To conduct DFA, assumptions of multivariate analysis must be fulfilled to ensure the accuracy of the results. Two main assumptions made are multivariate normality and homogeneity of variances/covariances across groups (Pietrusewsky, 2000). Multivariate normality refers to the distributional metric parameters and their combinations. The skewness and kurtosis test and histograms with normal distribution curves will determine the normal distribution. The second assumption is the homogeneity of covariance matrix. If the level of heterogeneity within groups is high, then exploring other statistical procedures, such as logistic regression, is highly encouraged (Ousley & Jantz, 2005). The log determinants and Box's M procedure will test the equality of the group covariance matrices.

Discriminant function analysis (DFA), by way of canonical correlations, was conducted for all age groups. Canonical discriminant analyses allowed the study of between-group variation by analysing multiple parameters simultaneously (Klecka et al., 1980). The accuracy of the discriminant functions was calculated based on standardised coefficients, unstandardised coefficients, and structural matrix. The unstandardised coefficients in each of the canonical discriminant analyses provided the parameters in the equation. The sectioning points between males and females were indicated using the average function centroids in all the discriminant formulae. Individuals were categorised as males when the sectioning point was less than the score value and categorised as females in the opposite case (Patriquin et al., 2005).

Wilks' lambda is a multivariate test that provides information and assumptions regarding differences between and among groups. Wilks' lambda analysis will classify the input of each parameter and its significance in discriminating each group and the overall sample size. Wilks' lambda displays each parameter's contribution based on a scale from 0.000 to 1.000, with 1.000 being the least and 0.000 representing a highly influential parameter contributing to discrimination. The significance of lambda is also based on a scale from 0.000 to 1.000, with 0.000 demonstrating the highest significance of data. A canonical correlation discriminant function (CAN1) displays the strength of correlation between the discriminant score and the set of 14 independent parameters with a minimum acceptable level of 0.05. Additionally, the F statistic compares the variation among the sample means to the variation within the samples. The variation among the samples is large if the value of the F statistic is large (SPSS Inc., 1999).

Stepwise DFA was used to identify the most suitable parameters to classify the group (Wilks' lambda, F to enter 3.84, F to remove 2.71). In this analysis, a parameter was eliminated if it made the least contribution to identifying the sex. Therefore, the discriminant function equations were formulated using certain chosen parameters (Pietrusewsky, 2000). The final stage involved performing cross-validation on all the classification models using the analysis procedure's leave-one-out function.

3.4.1.7 Binary logistic regression (BLR)

Binary logistic regression (BLR) is a technique that is utilised to predict the relationship between independent and dependent parameters. BLR is a suitable approach when the data is non-parametric and not bound by data distribution assumptions. In the present study, BLR was conducted on each age group and the overall sample to determine the probability of accurately estimated sex. A stepwise procedure was used to choose the best predictor parameters. A significance value of $p \leq 0.05$ was required for entry into the model and a significance value of $p > 0.10$ was required for removing a parameter. Through this process, a group of parameters would be narrowed to only the most significant for classification. The corresponding coefficients (β) and constant from the overall sample and each separate age group were used to derive the regression function equations. Individuals were categorised as males when the scores were greater than the 0.5 sectioning point and categorised as females in the opposite case. A second sample consisting of 75% from the original sample was used for cross-validation.

3.4.2 Methods of ethnicity estimation in Malaysian sub-adults

The objective of this section is to present methods for ethnicity estimation (Malay, Chinese and Indian) of the Malaysian sub-adult population.

3.4.2.1 Correlation test

This method was carried out as previously described in Section 3.4.1.1.

3.4.2.2 Descriptive statistics

All craniometric measurements were assessed for normality using skewness and kurtosis calculations, along with histograms and normal distribution curves. The calculations of means±SD value were performed separately for each age group and ethnic group.

3.4.2.3 Analysis of variance (ANOVA)

Analysis of variance (ANOVA) was used to identify which craniometric parameters were significant with ethnicities. Furthermore, ANOVA can also be used to examine the means of the independent parameters across the dependent parameters. The results of this analysis helped to determine whether ethnic differences could be metrically estimated across age groups.

3.4.2.4 Multivariate analysis of variance (MANOVA)

Multivariate analysis of variance (MANOVA) was conducted to test the hypothesis in the present study. It evaluates the significance of ethnicity, sex, age groups, and the relationship between ethnicity and sex as well as between ethnicity and age groups on the craniometric measurements.

3.4.2.5 Discriminant function analysis (DFA)

In conducting DFA, two important statistical assumptions must be fulfilled. The assumptions are multivariate normality and homogeneity of variances/covariances. The log determinants and Box's M test were used to compare multivariate sample variation and determine if two or more covariance matrices were equal. Therefore, the log determinants should be relatively equal and the Box's M test should be non-significant (p>0.001), suggesting that the covariance matrices were equal (Hahs-Vaughn, 2016).

Canonical discriminant analyses were performed using craniometric measurements to compare differences among the Malay, Chinese, and Indian groups. A forward stepwise was conducted in each DFA model to select the best parameter for group separation. The canonical discriminant analysis produces the canonical correlation, canonical structure, and canonical coefficients. Pairwise comparisons were calculated on the craniometric measurements in each age group to indicate levels of similarity and dissimilarity among the ethnicities. Discriminant score plots were produced to demonstrate multivariate separation among the three ethnic groups. Then, DFAs were validated with the leave-one-out cross-validation function.

3.4.3 Machine learning (ML) model development

Machine learning (ML) techniques can be applied to gain insights and knowledge from craniometric data. These techniques are expressed through model building used to interpret patterns and relationships in the data (Williams, 2011). ML techniques, which may be used for classification problems or regression, consist of three steps: 1) initial exploration, 2) model building or pattern identification with validation/verification, and 3) deployment. The first step involves data preparation, such as data cleaning and any exploratory research. Second, the appropriate model for assessing the data is selected depending on which model has the best predictive power. Third, the chosen model is applied to new data to assess how well it predicts outcomes.

In the present study, ML techniques for classification were applied to the craniometric data for the entire sample. These analyses included RF, SVM, and LDA with GridSearchCV. A random 70% of the individuals (N=364) in the total sample were used as the training data, while the remaining 30% (N=157) were used to test the model. All models were compared, and the models which produced the highest accuracy for sex and ethnicity estimation were chosen to develop the interface programme.

Support vector machines (SVM) is one of the most popular methods of analysis (Imaizumi et al., 2020; Toneva et al., 2021; Ortega et al., 2021) followed by RF (Hefner et al., 2015; Ortega et al., 2021), and LDA (Gao et al., 2018; Nikita & Nikitas, 2020; Toy et al., 2022). SVM is a powerful and effective tool for sex and ethnicity estimation in ML, especially when dealing with high-dimensional data and non-linear relationships (Toneva et al., 2021). Similarly, RF is ideal for problems where the relationship between features and target variables is complex and non-linear. Meanwhile, LDA is widely used due to its simplicity and efficiency, and its ability to handle multicollinearity. Furthermore, these methods appear to produce overall accuracies that are high compared to alternative machine classifiers such as single DTs and k-NN (Maxwell et al., 2018). Therefore, these ML methods are well-suited for the task of sex and ethnicity estimation.

GridSearchCV from scikit-learn was used in training the datasets to determine the optimal hyperparameters for ML models. Hyperparameters are settings or configurations that are not learnt from the data, but they are instead specified by the user before training a ML model. These hyperparameters are crucial for optimising the model's performance and improving its accuracy. GridSearchCV automates the process of systematically searching through a predefined set of hyperparameters and selecting the combination that yields the best performance for the given evaluation metric. By specifying the number of cross-validation iterations, the reliability of the evaluation can be controlled (Kurachka & Sukhoi, 2017; Zöller & Huber, 2021).

The most important hyperparameters for SVM are kernel, penalty coefficient (C), and gamma. The radial basis function (RBF) was used in the present study where the bias term is not required in RBF kernel (Ghezelbash et al., 2021). C is a hyperparameter that decides the level of penalty over the outliers. It is directly inverse to regularisation parameter. When C is large, outliers will be given high penalty, and a hard margin is formed. When C is small, the outliers are ignored, and the margin will be wide. Hyperparameters for RF are n_estimators and max_depth where n_estimators is the number of trees in the forest. The higher the number of trees, the better it is to learn the data. While max_depth is the maximum number of features allowed in an individual tree. If the number of splits is too low, the model underfits the data and if it is too high the model overfits (Daviran et al., 2021). On the other hand, LDA uses shrinkage estimator to regularise the covariance matrices of the classes.

3.4.3.1 Performance metrics

To evaluate the model with standard procedure, the true positive (TP), true negative (TN), false positive (FP), and false negative (FN) values were determined. Based on these values, accuracy, precision, recall, and F1-score were calculated (Kurachka & Sukhoi, 2017). Accuracy is the most intuitive performance measure and it is simply a ratio of correctly predicted observations to the total observations. The accuracy of the model on the basis of TP, TN, FP, and FN was calculated based on the ratio:

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+FP+TN+FN}}$$

Recall is the ratio of correctly predicted positive observations to all observations in the actual class.

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}}$$

Precision is the ratio of correctly predicted positive observations to the total predicted positive observations.

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}}$$

F1-Score is the weighted average of Precision and Recall.

$$\text{F1-Score} = 2\text{x} \left[\frac{\text{Precision X Recall}}{\text{Precision+Recall}}\right]$$

3.4.3.2 Interface programme development

The interface programme is a web-based graphical user interface (GUI) developed using the Streamlit application. Streamlit is an open-source app framework for building ML and data science web apps in Python that allows developers to create interactive and visually appealing programmes. It can calculate made-to-order craniometric measurements to estimate sex and ethnicity. Furthermore, it is user-friendly as the end user only needs to input the values into the programme and press the "estimate" button. The output includes the predicted sex and ethnicity from an unknown cranium. A user acceptance test was performed to evaluate the system. The GUI acts as a bridge to transfer knowledge gained in research to practical application in the field and it will be made freely available to other researchers. The workflow for model and GUI development is shown in Figure 3.7.

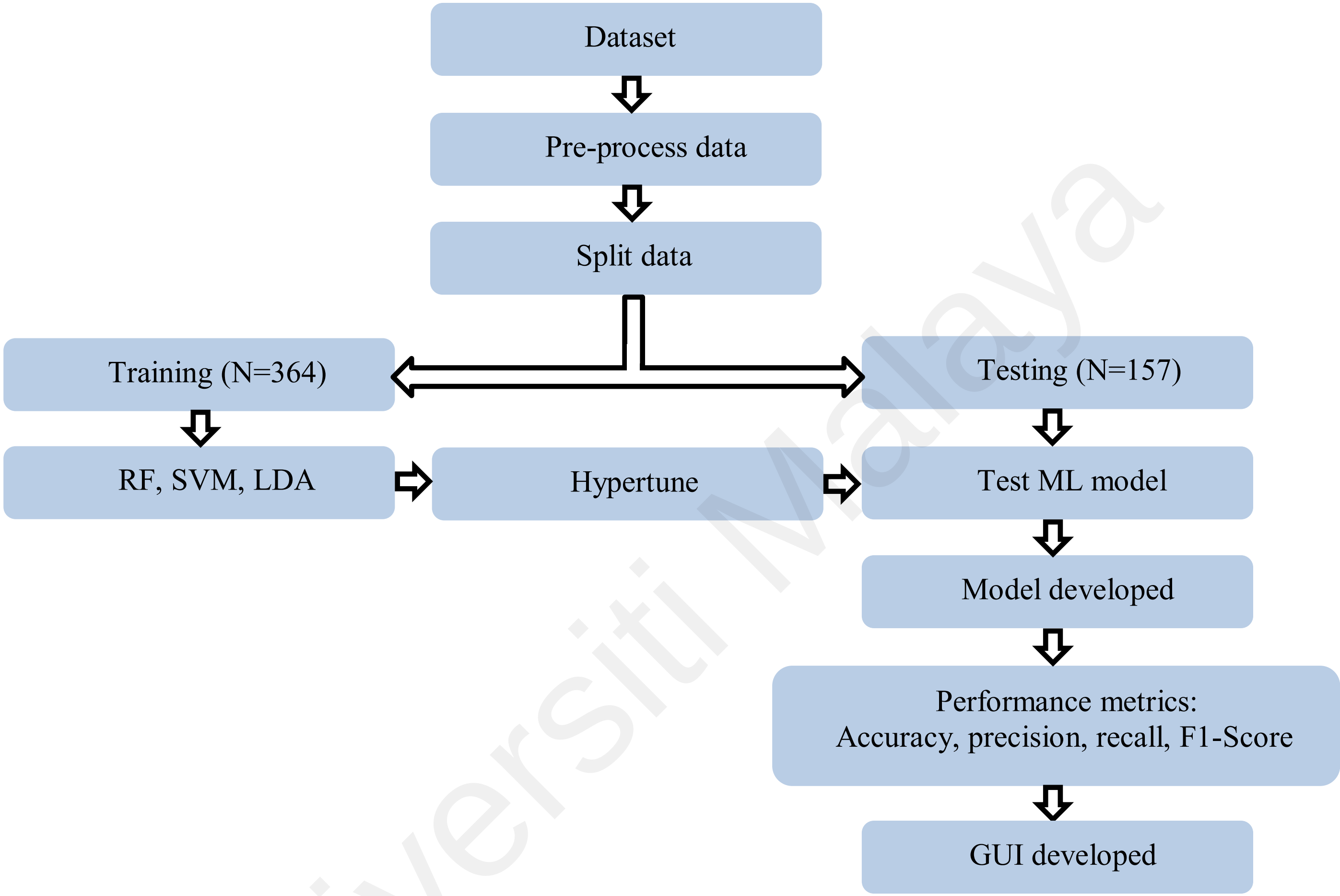


Figure 3.7: Workflow for model and graphic user interface (GUI) development

3.4.4 Cephalic index (CI)

The maximum width and length of the cranium were measured to obtain CI. CI was calculated based on the following equation:

$$\frac{\text{Cephalic width}}{\text{Cephalic length}} \times 100$$

All data were demonstrated as mean±SD. Then, using the obtained data, CI was classified and grouped into dolichocephalic (mean CI ≤ -1 SD), mesocephalic (mean CI ±1 SD), brachycephalic (mean CI +1 to 2 SD), and hyperbrachycephalic (mean CI ≥ 2 SD; Koizumi et al., 2010).

3.4.5 Observer error assessment

To ensure that results are reliable and repeatable, it is essential to quantify the error rate (Braga & Treil, 2007). Intra-observer reliability was assessed using 25 cranial MSCT images of different age groups, sex, and ethnicities. The 14 modified cranial measurements were assessed as listed in Table 3.2, with a two-week interval between re-measurement. Similarly, precision testing of the PTP protocol was conducted on 25 randomly selected cranial MSCT images by an oral and maxillofacial radiologist having more than five years of experience. The following statistics were calculated: technical error of measurement (TEM), relative TEM (rTEM), and the coefficient of reliability (R) with acceptable levels for measurement at rTEM <1.5% and R>0.95 (Ulijaszek & Kerr, 1999).

3.4.5.1 Technical error of measurement (TEM)

The TEM represents the most frequent measure of error. It is the SD between repeated measurements (Harris & Smith, 2009; Weinberg et al., 2005). It is quantified by taking multiple measurements of the same object and then calculating the variance between the repeat values (Harris & Smith, 2009). The following equation was used to calculate TEM:

$$\frac{\sum D^2}{2\mathrm{N}} \times 100$$

D is the difference between the first and second measurements and N is the total number of individuals. TEM was recorded in the same unit as the measurement.

3.4.5.2 Relative technical error of measurement (rTEM)

To obtain the error demonstrated as a percentage, TEM was transformed into rTEM, which corresponded to the total average of the parameter to be analysed. Therefore, obtaining the average value of the parameter (X) is necessary. The equation used to calculate rTEM is as follows:

$$\frac{\mathrm{TEM}}{\mathrm{X}} \times 100$$

3.4.5.3 Coefficient of reliability (R)

The coefficient of reliability (R) measures the amount of between-subject variation that is not due to observer error. The values range from 0 to 1. Higher R-values thus indicate higher precision. The following equation was used to calculate the coefficient of reliability:

$$1 - \left(\frac{(\mathrm{TEM})^2}{\mathrm{SD}^2}\right)$$

CHAPTER 4: RESULTS

This chapter sets out the results obtained from statistical analyses to achieve the aims and objectives of this research. The first aim was to develop 3D craniometric models in Malaysian sub-adults for sex and ethnicity estimation, and the second was to establish a CI classification for Malaysian sub-adults. The objectives of this study were to compare the accuracy of DFA and BLR models for sex estimation in Malaysian sub-adults, to assess the accuracy of DFA models for ethnicity estimation in Malaysian sub-adults, to compare the validity of sex and ethnicity estimation models between ML algorithm and classical statistical methods in Malaysian sub-adults, and to determine the normative range of CI classification for Malaysian sub-adults.

4.1 Demographic profile

Multi-slice computed tomography data (MSCT) scans of the cranium were obtained from 521 individuals and grouped into six age groups. The demographic characteristics of the individuals based on sex, ethnicity, and age groups are presented in Table 4.1.

Table 4.1: Sociodemographic summary based on age groups, sex and ethnicity

Variables		Percentage % (N)
Age groups (years)	0-2	21.5 (112)
	3-6	15.2 (79)
	7-9	10.9 (57)
	10-12	11.5 (60)
	13-15	15.2 (79)
	16-20	25.7 (134)
Sex	Male	53.6 (279)
	Female	46.4 (242)
Ethnicity	Malay	42.4 (221)
	Chinese	27.8 (145)
	Indian	29.8 (155)

N=number of samples.

4.2 Intra-observer and inter-observer errors

Intra-observer and inter-observer errors, calculated by determining the means of TEM, rTEM, and R values, are provided in Table 4.2 and Table 4.3, respectively. Table 4.2 presents the results for intra-observer errors of the 14 craniometric measurements utilised for this study. All measurements showed extremely low errors, indicating consistency in the data collection and the quality of being easily repeatable. All measurements had rTEM and R values of <1% and >0.99, respectively. The lowest TEM value was recorded as 0.085mm (FMW) and the highest was 0.315mm (CH). Meanwhile, the lowest rTEM value was calculated as 0.097% (MCW) and the highest rTEM value was 0.551% (FML). Finally, the lowest R value was 0.994 (FMW and CBL) and the highest R value was 0.99 (MCL and NOL).

Table 4.2: Intra-observer errors (TEM; rTEM; R) of the 14 craniometric measurements

Intra-observer errors	TEM (mm)	rTEM (%)	R
MCL	0.230	0.131	0.999
LCL (R)	0.182	0.192	0.996
LCL (L)	0.225	0.236	0.996
NOL	0.270	0.157	0.999
MCW	0.140	0.097	0.997
BAW	0.219	0.201	0.997
IPW	0.183	0.154	0.998
PC	0.256	0.232	0.996
OC	0.284	0.290	0.998
FC	0.257	0.229	0.998
CH	0.315	0.229	0.995
CBL	0.264	0.261	0.994
FML	0.189	0.551	0.998
FMW	0.085	0.289	0.994
Mean	**0.221**	**0.232**	**0.997**

TEM=technical error of measurement, rTEM=relative technical error of measurement, R=coefficient of reliability.

Table 4.3 presents the results for inter-observer errors of the 14 craniometric measurements utilised for this study. Similar to the intra-observer errors, all measurements showed extremely low inter-observer errors, indicating consistency in the data collection. All measurements had rTEM and R values of <1.5% and >0.99, respectively. The lowest TEM value was 0.142mm (LCL (L)) and the highest was 0.448mm (FMW). For the rTEM values, the lowest was recorded at 0.111% (MCL) and the highest was 1.3% (FML). The lowest R value was 0.990 (IPW and PC) and the highest was 0.995 (LCL (L)).

Table 4.3: Inter-observer errors (TEM; rTEM; R) of the 14 craniometric measurements

Inter-observer errors	TEM (mm)	rTEM (%)	R
MCL	0.192	0.111	0.992
LCL (R)	0.207	0.223	0.993
LCL (L)	0.142	0.152	0.995
NOL	0.210	0.124	0.992
MCW	0.218	0.153	0.992
BAW	0.286	0.264	0.991
IPW	0.253	0.215	0.990
PC	0.292	0.268	0.990
OC	0.295	0.308	0.992
FC	0.327	0.293	0.992
CH	0.246	0.179	0.991
CBL	0.228	0.226	0.994
FML	0.448	1.300	0.992
FMW	0.272	0.917	0.992
Mean	**0.258**	**0.338**	**0.992**

TEM=technical error of measurement, rTEM=relative technical error of measurement, R=coefficient of reliability.

4.3 Sex estimation in Malaysian sub-adults

4.3.1 Correlation analysis

Pearson product-moment correlation coefficient (Pearson correlation) revealed that most parameters have strong positive correlations with two independent parameters (Table 4.4). Since LCL (R) and LCL (L) were the most highly correlated parameters (r=0.982), these parameters were combined as LCL.

Table 4.4: Pearson correlation coefficient of the 14 craniometric measurements for sex and ethnicity estimation

	MCL	LCL(R)	LCL(L)	NOL	MCW	BAW	IPW	PC	OC	FC	CH	CBL	FML	FMW
MCL	1	0.850	0.852	0.896	0.721	0.782	0.839	0.789	0.706	0.884	0.880	0.879	0.327	0.656
LCL(R)		1	0.982	0.834	0.791	0.748	0.881	0.609	0.689	0.819	0.850	0.865	0.307	0.671
LCL(L)			1	0.838	0.786	0.745	0.882	0.608	0.694	0.818	0.849	0.868	0.303	0.666
NOL				1	0.698	0.766	0.846	0.755	0.695	0.883	0.866	0.891	0.309	0.645
MCW					1	0.845	0.836	0.580	0.672	0.781	0.786	0.681	0.242	0.648
BAW						1	0.821	0.661	0.667	0.780	0.804	0.722	0.283	0.679
IPW							1	0.602	0.700	0.854	0.891	0.890	0.304	0.736
PC								1	0.425	0.691	0.723	0.606	0.255	0.476
OC									1	0.660	0.724	0.643	0.225	0.573
FC										1	0.897	0.840	0.306	0.654
CH											1	0.896	0.312	0.711
CBL												1	0.286	0.695
FML													1	0.309
FMW														1

4.3.2 Descriptive statistics

Descriptive statistics of mean values±SD for the male and female cranial measurements are presented in Table 4.5.

4.3.3 Independent sample *t*-test

The means of the 14 craniometric measurements were subjected to independent sample *t*-test analysis (Table 4.5). For the entire sample, apart from the OC parameter, significant variations ($p<0.05$) were observed in all analysed parameters. Males exhibited considerably greater mean values than females. The measurements expressing the most significant dimorphism as indicated by *t*-values and *p*-values were LCL ($t=4.62$; $p<0.001$), followed by MCW and MCL ($t=4.56$; $p<0.001$ and $t=4.55$; $p<0.001$). For descriptive analysis by age groups, significant variations were found in five parameters (MCW, FC, CH, CBL, FML) and six parameters (LCL, MCW, IPW, FC, CBL, FMW) for the age groups of 0-2 years and 3-6 years, respectively. Moreover, eight parameters (age group of 7-9 years: MCL, LCL, MCW, BAW, IPW, FC, FML, FMW), 10 parameters (age group of 10-12 years: MCL, LCL, NOL, BAW, IPW, PC, FC, CH, FML, FMW), and 12 parameters (age groups of 13-15 years and 16-20 years: MCL, LCL, NOL, MCW, BAW, IPW, PC, FC, CH, CBL, FML, FMW) were found to differ significantly ($p<0.05$).

Table 4.5: Descriptive analysis and mean variation between sexes of the entire sample (N=521) and by age groups

Parameters	N	Entire sample (M=279 F=242)	0-2 years (M=61 F=51)	3-6 years (M=43 F=36)	7-9 years (M=35 F=22)	10-12 years (M=30 F=30)	13-15 years (M=41 F=38)	16-20 years (M=69 F=65)
MCL	M	168.45±14.07	149.25±11.99	163.90±6.17	170.09±5.90	172.11±5.84	176.88±8.14	180.84±6.19
	F	162.99±13.16	145.92±16.19	162.48±7.11	166.09±5.88	165.57±6.93	170.18±6.60	170.22±6.22
	t	**4.549**	1.245	0.954	**2.493**	**3.946**	**3.979**	**9.888**
	P	**<0.001**	0.216	0.343	**0.016**	**<0.001**	**<0.001**	**<0.001**
LCL	M	90.56±10.25	76.33±8.19	87.32±4.53	91.78±4.53	95.61±4.83	96.39±5.14	98.68±5.89
	F	86.57±9.34	73.48±8.09	83.41±4.51	88.64±6.41	92.47±5.44	91.47±4.21	91.90±4.67
	t	**4.620**	1.838	**3.824**	**2.164**	**2.356**	**4.629**	**7.351**
	P	**<0.001**	0.069	**<0.001**	**0.035**	**0.022**	**<0.001**	**<0.001**
NOL	M	163.55±15.58	142.65±11.34	159.03±6.42	164.12±14.62	168.70±6.43	172.85±8.72	176.78±6.08
	F	158.62±14.20	138.68±15.79	156.80±6.75	162.35±4.80	162.48±6.84	166.70±6.28	167.48±6.22
	t	**3.756**	1.541	1.498	0.558	**3.627**	**3.567**	**8.740**
	P	**<0.001**	0.126	0.138	0.579	**0.001**	**0.001**	**<0.001**
MCW	M	140.73±10.00	128.81±9.42	140.72±6.71	143.58±7.49	143.97±5.72	144.61±6.80	146.10±7.65
	F	136.77±9.71	125.00±10.72	136.08±6.61	139.02±6.37	140.84±6.76	140.52±6.34	141.57±5.68
	t	**4.562**	**2.003**	**3.081**	**2.364**	1.935	**2.758**	**3.871**
	P	**<0.001**	**0.048**	**0.003**	**0.022**	0.058	**0.007**	**<0.001**
BAW	M	106.54±8.94	95.42±10.10	106.58±4.90	110.41±5.03	109.60±4.92	110.69±6.06	110.57±5.27
	F	103.82±9.03	92.80±11.80	104.48±4.55	106.98±5.37	107.03±4.37	107.44±5.18	107.44±5.23
	t	**3.436**	1.261	1.966	**2.439**	**2.136**	**2.552**	**3.445**
	P	**0.001**	0.210	0.053	**0.018**	**0.037**	**0.013**	**0.001**

Table 4.5, continued

Parameters	N	Entire sample (M=279 F=242)	0-2 years (M=61 F=51)	3-6 years (M=43 F=36)	7-9 years (M=35 F=22)	10-12 years (M=30 F=30)	13-15 years (M=41 F=38)	16-20 years (M=69 F=65)
IPW	M	108.77±16.44	83.99±10.32	102.85±6.09	111.16±5.31	115.00±5.44	120.11±6.47	123.71±6.35
	F	104.09±15.01	81.13±11.00	97.84±6.30	107.21±5.83	111.83±6.56	114.27±5.20	114.97±5.54
	t	**3.375**	1.417	**3.569**	**2.629**	**2.034**	**4.394**	**8.458**
	P	**0.001**	0.159	**0.001**	**0.011**	**0.047**	**<0.001**	**<0.001**
PC	M	108.13±7.98	99.42±8.35	108.66±4.94	111.05±5.98	109.67±5.95	110.59±6.48	111.88±5.89
	F	105.09±8.20	96.84±9.97	108.40±6.82	108.38±5.99	105.07±5.69	107.29±5.49	107.33±5.84
	t	**4.283**	1.489	0.200	1.638	**3.056**	**2.435**	**4.490**
	P	**<0.001**	0.139	0.842	0.107	**0.003**	**0.017**	**<0.001**
OC	M	95.66±8.12	86.73±7.50	93.87±4.86	95.83±4.37	99.60±6.38	100.98±6.79	99.70±6.12
	F	94.40±8.21	85.89±8.61	93.19±5.02	95.44±5.39	97.49±6.15	98.11±6.62	97.80±6.85
	t	1.755	0.557	0.607	0.301	1.302	1.897	1.695
	P	0.080	0.579	0.545	0.765	0.198	0.062	0.092
FC	M	106.08±10.05	92.13±8.09	103.98±4.50	107.26±4.58	109.04±4.01	111.96±6.36	114.33±5.41
	F	102.37±9.80	88.61±10.06	101.24±4.89	104.38±3.73	105.61±5.31	107.37±4.53	108.70±5.22
	t	**4.248**	**2.057**	**2.590**	**2.473**	**2.823**	**3.665**	**6.099**
	P	**<0.001**	**0.042**	**0.011**	**0.017**	**0.007**	**<0.001**	**<0.001**
CH	M	129.56±12.17	111.60±9.04	126.91±4.95	131.58±4.99	133.47±4.30	137.15±5.83	139.87±5.75
	F	125.68±11.93	107.35±10.44	124.91±5.24	129.14±5.95	130.61±5.74	132.11±5.45	133.28±5.03
	t	**3.665**	**2.306**	1.738	1.670	**2.183**	**3.951**	**7.035**
	P	**<0.001**	**0.023**	0.086	0.101	**0.033**	**<0.001**	**<0.001**

Table 4.5, continued

Parameters	N	Entire sample (M=279 F=242)	0-2 years (M=61 F=51)	3-6 years (M=43 F=36)	7-9 years (M=35 F=22)	10-12 years (M=30 F=30)	13-15 years (M=41 F=38)	16-20 years (M=69 F=65)
CBL	M	92.27±11.76	75.21±6.05	86.53±4.15	92.93±4.55	96.24±4.73	100.79±4.77	103.79±4.68
	F	88.70±10.62	72.45±7.44	84.29±4.77	90.97±4.13	94.21±4.49	95.57±4.33	96.58±3.85
	t	**3.606**	**2.163**	**2.233**	1.641	1.704	**5.075**	**9.701**
	P	**<0.001**	**0.033**	**0.028**	0.107	0.094	**<0.001**	**<0.001**
FML	M	35.12±8.84	31.99±3.17	33.79±2.32	35.91±2.36	35.88±2.28	36.10±2.40	37.40±16.82
	F	32.74±3.07	29.71±3.37	33.02±2.47	33.68±2.33	33.51±1.91	33.97±2.46	33.57±2.63
	t	**3.980**	**3.680**	1.417	**3.477**	**4.348**	**3.883**	**2.034**
	P	**<0.001**	**<0.001**	0.161	**0.001**	**<0.001**	**<0.001**	**0.047**
FMW	M	28.82±2.97	25.20±2.47	28.56±1.68	29.62±2.22	29.83±1.76	29.97±2.40	30.66±2.23
	F	27.59±2.73	24.34±2.49	27.50±1.51	27.64±1.55	28.51±2.16	28.73±2.08	29.06±2.18
	t	**2.226**	**1.821**	**2.931**	**3.652**	2.589	**2.448**	**4.175**
	P	**<0.001**	**0.07**	**0.004**	**0.001**	0.012	**0.017**	**<0.001**

Independent sample *t*-test values are presented as mean±SD, unit of measurement: millimetres (mm), N=number of samples expressed as M=male and F=female, bold indicates statistical significance ($p<0.05$).

4.3.4 Multivariate analysis of variance (MANOVA)

The results of the MANOVA analysis are presented in Table 4.6. The MANOVA analysis tested the hypothesis of 'sexually dimorphic differences in the cranium are predictable for each age group when utilising statistical analyses'. MANOVA analysis accepted the hypotheses (p<0.05) demonstrating statistically significant differences were obtained in all measurements, except for PC, across all age groups. This suggests that the measured parameters varied significantly across the different age groups. Additionally, there were significant differences between males and females for two measurements, namely MCW and FML. This indicated that these two measurements showed sex-related differences. Furthermore, there was significant interaction between sex and age groups for the measurement of FML.

Table 4.6: Interactions between age groups, sex, and age groups*sex

Parameters	Age groups		Sex		Age groups*sex	
	***F*-value**	**Sig.**	***F*-value**	**Sig.**	***F*-value**	**Sig.**
MCL	12.090	**0.001**	0.541	0.462	0.024	0.877
LCL	19.301	**0.000**	3.767	0.053	0.258	0.612
NOL	11.798	**0.001**	0.031	0.860	0.004	0.951
MCW	6.452	**0.011**	4.495	**0.034**	0.000	0.998
BAW	4.203	**0.041**	2.014	0.156	0.001	0.980
IPW	35.412	**0.000**	1.911	0.167	0.442	0.507
PC	3.420	0.065	0.090	0.764	0.025	0.875
OC	6.770	**0.010**	0.000	0.998	0.147	0.701
FC	11.757	**0.001**	1.996	0.158	0.003	0.955
CH	17.124	**0.000**	1.101	0.295	0.185	0.667
CBL	37.044	**0.000**	0.507	0.477	0.724	0.395
FML	8.969	**0.003**	5.633	**0.018**	4.013	**0.046**
FMW	8.938	**0.003**	1.646	0.200	0.455	0.500

MANOVA test, bold indicates statistical significance (p<0.05).

4.3.5 Sexual dimorphism percentage (SDP)

Table 4.7 presents the results of SDP divided by age groups. The SDP was used as an indicator to demonstrate the differences between males and females. The pattern of SDP varied according to age group and region of the cranium. Most of the parameters showed that males were more prominent than females. The majority of parameters exhibited a pattern of constant increase in sexual dimorphism, with males getting progressively larger than females. The age groups of 0-2 years and 3-6 years represented the stages at which sexually dimorphic differences were least expressed, contrary to the age group of 16-20 years, where sexually dimorphic differences were highly expressed.

Table 4.7: Sexual dimorphism percentage (%) of the 14 craniometric measurements for all age groups

Parameters	0-2 years	3-6 years	7-9 years	10-12 years	13-15 years	16-20 years
MCL	2.225	0.870	2.353	3.799	3.771	5.871
LCL (R)	3.759	4.498	3.147	3.394	4.654	6.882
LCL (L)	3.686	4.452	3.691	3.158	5.557	6.887
NOL	2.777	1.400	1.098	3.686	3.554	5.258
MCW	2.962	3.296	3.176	-2.175	2.830	3.103
BAW	2.736	1.979	3.105	2.345	2.937	2.831
IPW	3.404	4.871	3.552	2.755	4.860	7.065
PC	2.594	0.244	2.403	4.193	2.987	4.073
OC	0.976	0.722	0.410	-2.116	2.841	1.908
FC	2.831	2.636	2.685	3.147	4.102	4.931
CH	2.804	1.575	1.859	2.143	3.670	4.711
CBL	2.669	2.592	2.113	2.109	5.177	6.948
FML	4.138	2.264	6.204	6.596	5.896	10.252
FMW	3.404	3.735	6.680	4.427	4.150	5.205

4.3.6 Discriminant function analysis (DFA)

To conduct DFA, two assumptions must be fulfilled: normality and homogeneity of variance/covariance matrices.

4.3.6.1 Normality of the data

Upon analysing the skewness and kurtosis values presented in Appendix A, most parameters' skewness and kurtosis ratios fell well within the normal range (skewness: -2 to 2 and kurtosis: -7 to 7), suggesting that the data were normally distributed (Byrne, 2010). The histogram with normal distribution curves further demonstrated the normal distribution of the sample.

4.3.6.2 Homogeneity of covariance matrices

The results of the log determinants and Box's M test are presented in Table 4.8 and Table 4.9, respectively. The log determinants demonstrated a degree of similarity between males and females, while the Box's M test did not show statistical significance (p>0.001), thus suggesting that the covariance matrices were equal.

Table 4.8: Log determinants of covariance matrices for sex estimation

Group	Rank	Log determinants
Male	1	-2.498
Female	1	-2.712

Table 4.9: Box's M test of equality of covariance matrices for sex estimation

Box's M		
F	Approx	2.910
	df1	1
	df2	781847.017
	Significant	0.088

4.3.6.3 Stepwise discriminant function analysis (DFA)

The 14 cranial measurements from the Malaysian sub-adult population were subjected to stepwise DFA. In Table 4.10, the discriminant coefficients, eigenvalues, F statistics, Wilk's lambda, and canonical correlation are presented.

The Wilks' lambda analysis was conducted to assess the discriminatory power of the independent parameters in different age groups for sex estimation. In the age group of 0-12 years, the Wilks' lambda values ranged from 0.946 to 0.754, with a significance of 0.000. This indicated a low but statistically significant contribution of the independent parameters to the discrimination of sex in this age group. Furthermore, the canonical correlation coefficients were reported to be between 0.232 and 0.582. These coefficients represented the correlation between the discriminant function and the independent parameters. A moderate correlation between the discriminant function and the independent parameters suggested that the measured parameters have a moderate influence on accurately discriminating between sexes in this age group.

In the age groups of 13-15 years and 16-20 years, high Wilks' lambda values of 0.647 and 0.412, respectively, were obtained, both with a significance of 0.000. This indicated a high and statistically significant contribution of the independent parameters to the discrimination of sex in these age groups. Moreover, the canonical correlation coefficients for the age groups of 13-15 years and 16-20 years were reported to be 0.594 and 0.767, respectively. These high correlation coefficients indicated a strong relationship between the discriminant function and the independent parameters, suggesting that the measured parameters have a strong influence on accurately discriminating between sexes in these age groups. Overall, these findings indicated that the independent parameters have varying levels of contribution to the discrimination of sex in different age groups, with a low but significant contribution in the age group of 0-12 years, and a high and significant contribution in the age groups of 13-15 years and 16-20 years.

Table 4.10: Discriminant coefficients, eigenvalues, F statistics, Wilks's lambda and canonical correlation for the entire sample and all age groups for sex estimation

Age groups (years)	Para-meters	U.C	Eigen-values	F statistics	Wilk's lambda	*p*-value	C.C
Entire sample	PC FMW Constant	0.059 0.241 -13.077	0.057	24.198	0.946	0.000	0.232
0-2	FML Constant	0.306 -9.466	0.123	13.539	0.890	0.000	0.331
3-6	LCL Constant	0.221 -18.910	0.190	14.622	0.840	0.000	0.399
7-9	MCL FMW Constant	0.092 0.411 -27.290	0.342	13.340	0.805	0.000	0.505
10-12	MCL FML Constant	0.095 0.331 -27.551	0.511	18.907	0.754	0.000	0.582
13-15	IPW CBL FML Constant	0.074 0.130 0.163 -27.195	0.545	25.755	0.647	0.000	0.594
16-20	MCL NOL IPW CBL Constant	0.357 -0.301 0.051 0.129 -29.917	1.424	45.936	0.412	0.000	0.767

U.C=unstandardised coefficients, C.C=canonical correlation.

Table 4.11 shows the group centroids, sectioning point values, and the percentage accuracy for both original and cross-validated data across the entire sample and all age groups. The stepwise DFA was calculated for the entire sample, resulting in an accuracy of 61.6% after cross-validation with a sex bias of 14.8%. The accuracy percentage for each age group ranged between 62.5% and 88.1%, with improved accuracy as the age increased. The age group of 16-20 years recorded the highest classification accuracy of 88.1%, with a low sex bias of 0.7%, followed by the age group of 13-15 years with 79.5% accuracy and 1.1% sex bias. In contrast, the age group of 0-2 years showed the lowest classification accuracy of 62.5%, with a sex bias of 3.1%, followed by the age group of 3-6 years with 67.1% accuracy and 9.4% sex bias. The age group of 7-9 years recorded an accuracy of 75.4%, with sex bias of 1.4%, and the age group of 10-12 years obtained 75% accuracy and -13.6% sex bias. Males yielded higher accuracy rates than females in all age groups, except the age group of 10-12 years.

Table 4.11: Group centroids, sectioning point values, and percentage accuracy for original and cross-validation samples for the entire sample and age groups for sex estimation

Age groups (years)	Centroids	Sectioning point	Classification accuracy of original samples (%)	Classification accuracy of cross-validation samples (%)
Entire sample	M: 0.222 F: -0.255	-0.017	M: 68.6 F: 54.5 Sex bias: 14.3 Total: 62.2	M: 68.5 F: 53.7 Sex bias: 14.8 Total: 61.6
0-2	M: 0.318 F: -0.380	-0.031	M: 67.5 F: 59.3 Sex bias: 8.2 Total: 63.4	M: 63.9 F: 60.8 Sex bias: 3.1 Total: 62.5
3-6	M: 0.394 F: -0.470	-0.038	M: 72.2 F: 62.8 Sex bias: 9.4 Total: 67.1	M: 72.2 F: 62.8 Sex bias: 9.4 Total: 67.1
7-9	M: 0.455 F: -0.725	-0.135	M: 77.3 F: 77.1 Sex bias: 0.2 Total: 77.2	M: 76.1 F: 74.7 Sex bias: 1.4 Total: 75.4
10-12	M: 0.703 F: -0.703	0	M: 66.7 F: 80.3 Sex bias: -13.6 Total: 75.0	M: 66.7 F: 80.3 Sex bias: -13.6 Total: 75.0
13-15	M: 0.701 F: -0.757	-0.028	M: 80.7 F: 78.9 Sex bias: 1.8 Total: 79.7	M: 80.0 F: 78.9 Sex bias: 1.1 Total: 79.5
16-20	M: 1.150 F: -1.220	-0.035	M: 89.9 F: 87.7 Sex bias: 2.2 Total: 88.8	M: 88.4 F: 87.7 Sex bias: 0.7 Total: 88.1

M=male, F=female.

4.3.6.4 Discriminant scores

The discriminant scores were calculated using the discriminant function equations derived from the unstandardised coefficients and the constant. Seven equations were derived from stepwise DFA for the entire sample and all age groups are summarised in Table 4.12. The sex estimation model was developed using two parameters (PC and FMW) for the entire sample. In the age groups of 0-2 years and 3-6 years, FML and LCL were selected for the model development, respectively. Moreover, two parameters (age group of 7-9 years: MCL and FMW, age group of 10-12 years: MCL and FML), three parameters (age group of 13-15 years: IPW, CBL, FML), and four parameters (age group of 16-20 years: MCL, NOL, IPW, CBL) were selected for the development of the model.

These equations allow for the prediction of sex based on the measured parameters for each age group. If an individual's discriminant score is less than the sectioning point, the unknown is the closest to the female sample. On the other hand, if the value is greater than the sectioning point, this indicates that the unknown is the closest to the male sample. By applying the predictive equations and comparing the discriminant scores to the sectioning points, the sex of an unknown individual can be estimated based on their proximity to the male or female sample.

Table 4.12: Discriminant model equations for sex estimation

Age groups (years)	Discriminant model equations
Entire sample	-13.077+0.059(PC)+0.241(FMW)
0-2	-9.466+0.306(FML)
3-6	-18.910+0.221(LCL)
7-9	-27.290+0.092(MCL)+0.411(FMW)
10-12	-27.551+0.095(MCL)+0.331(FML)
13-15	-27.195+0.074(IPW)+0.130(CBL)+0.163(FML)
16-20	-29.917+0.357(MCL)-0.301(NOL)+0.051(IPW)+0.129(CBL)

The histograms in Figure 4.1 illustrate the discriminant function scores for all age groups in sex estimation. The horizontal axis represents each individual in the analysis, while the vertical axis represents the discriminant score assigned to each individual. In the histogram, a positive score indicated a prediction of male (represented by the black column), while a negative score indicated a prediction of female (represented by the orange column). However, misclassification had resulted in the opposite prediction, where a positive score was assigned to females and a negative score was assigned to males.

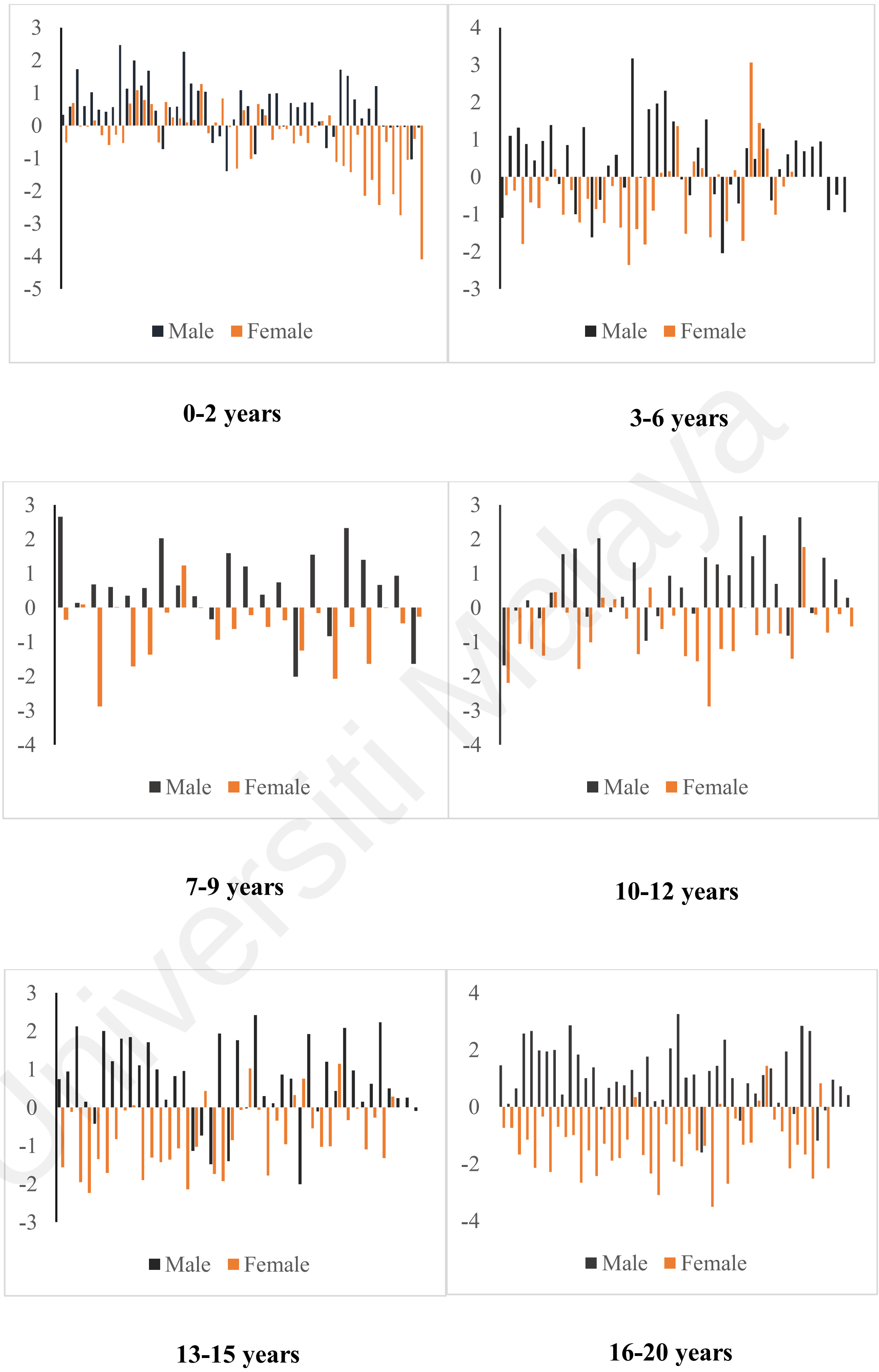


Figure 4.1: Histograms of discriminant function scores for all age groups for sex estimation

4.3.7 Binary logistic regression (BLR)

Table 4.13 presents the results of the stepwise analysis conducted for BLR. For the entire sample, MCL, NOL, MCW, OC, and FML were selected for the development of sex estimation model. Cross-validation accuracy was recorded at 66.9%, with males obtaining higher accuracy (72.9%) than females (60.1%). The lowest accuracy rate was obtained by the age group of 0-2 years, with an accuracy of 67.9% and a sex bias of 2.8% by utilising FML. Higher accuracy was obtained for males (69.5%) compared to females (66.7%).

The highest classification accuracy was obtained by the older age groups (age group of 16-20 years followed by age group of 13-15 years). Three parameters (MCL, NOL, FML) were selected for the development of the model in the age group of 16-20 years to produce an accuracy of 90.3% with a sex bias of 0.3%. Males obtained higher accuracy (90%) than females (89.7%). In the age group of 13-15 years, an accuracy of 82.3% was obtained by utilising two parameters (MCL, FML). Males obtained higher accuracy (82.5%) than females (82.0%) with a sex bias of 0.5%.

Table 4.13: Binary logistic regression analysis for the entire sample and according to age groups for sex estimation

Age groups (years)	Para-meters	*β* coefficients	Standard error	Sectioning point	Classification accuracy of original samples (%)	Classification accuracy of cross-validation samples (%)
Entire sample	MCL NOL MCW OC FML (Constant)	-4.986 4.193 -0.979 1.516 -2.119 1.049	0.048 0.051 0.020 0.029 0.045 1.899	0.5	M: 73.1 F: 61.6 Sex bias: 11.5 Total: 67.4	M: 72.9 F: 60.1 Sex bias: 12.8 Total: 66.9
0-2	FML (Constant)	-3.127 0.589	0.118 5.330	0.5	M: 72.1 F: 66.7 Sex bias: 5.4 Total: 69.6	M: 69.5 F: 66.7 Sex bias: 2.8 Total: 67.9
3-6	MCW (Constant)	-3.314 1.376	0.052 8.559	0.5	M: 75.9 F: 68.5 Sex bias: 7.4 Total: 72.2	M: 75.7 F: 68.7 Sex bias: 7.0 Total: 72.2
7-9	FML (Constant)	-4.123 2.021	0.212 15.049	0.5	M: 82.0 F: 80.4 Sex bias: 1.6 Total: 81.2	M: 81.0 F: 80.4 Sex bias: 1.4 Total: 80.7
10-12	MCL FML (Constant)	-5.165 -4.800 5.505	0.63 0.175 12.094	0.5	M: 77.8 F: 88.0 Sex bias: -10.2 Total: 82.9	M: 77.8 F: 86.2 Sex bias: -8.4 Total: 82.0
13-15	MCL FML (Constant)	-3.508 -2.884 4.025	0.147 0.134 9.721	0.5	M: 83.7 F: 82.1 Sex bias: 1.6 Total: 82.9	M: 82.5 F: 82.0 Sex bias: 0.5 Total: 82.3
16-20	MCL NOL FML (Constant)	-27.853 18.932 -1.979 7.313	0.287 0.292 0.164 17.043	0.5	M: 92.1 F: 90.8 Sex bias: 1.3 Total: 91.4	M: 90.0 F: 89.7 Sex bias: 0.3 Total: 90.3

M=male, F=female.

Table 4.14 presents the overall accuracy of DFA and BLR for original and cross-validation samples. In general, the BLR yielded slightly higher overall classification accuracy rates with lower sex bias rates in all cases for the validation sample compared to the DFA. Overall, for the validation sample, the formula with the highest accuracy rate was obtained by BLR (90.3%), followed by DFA (88.1%) in the age group of 16-20 years.

Table 4.14: Classification accuracy of original and cross-validation samples between DFA and BLR for sex estimation

Age groups (years)		Classification accuracy of original samples (%)				Classification accuracy of cross-validation samples (%)			
		Males	Females	Sex bias	Total	Male	Female	Sex bias	Total
Entire sample	DFA	68.6	54.5	14.3	62.2	68.5	53.7	14.8	61.6
	BLR	73.1	61.6	11.5	67.4	72.9	60.1	12.8	66.9
0–2	DFA	67.5	59.3	8.2	63.4	63.9	60.8	3.1	62.5
	BLR	72.1	66.7	5.4	69.6	69.5	66.7	2.8	67.9
3–6	DFA	72.2	62.8	9.4	67.1	72.2	62.8	9.4	67.1
	BLR	75.9	68.5	7.4	72.2	75.7	68.7	7.0	72.2
7–9	DFA	77.3	77.1	0.2	77.2	76.1	74.7	1.4	75.4
	BLR	82.0	80.4	1.6	81.2	81.0	80.4	1.4	80.7
10–12	DFA	66.7	80.3	-13.6	75.0	66.7	80.3	-13.6	75.0
	BLR	77.8	88.0	-10.2	82.9	77.8	86.2	-8.4	82.0
13–15	DFA	80.7	78.9	1.8	79.7	80.0	78.9	1.1	79.5
	BLR	83.7	82.1	1.6	82.9	82.5	82.0	0.5	82.3
16–20	DFA	89.9	87.7	2.2	88.8	88.4	87.7	0.7	88.1
	BLR	92.1	90.8	1.3	91.4	90.0	89.7	0.3	90.3

4.4 Ethnicity estimation in the Malaysian sub-adult population

4.4.1 Correlation analysis

The results obtained were similar to that of Section 4.3.1 (Table 4.4).

4.4.2 Descriptive statistics

Descriptive statistics of mean values±SD for Malay, Chinese, and Indian cranial measurements are presented in Table 4.15.

4.4.3 Analysis of Variance (ANOVA)

ANOVA was conducted to test whether the mean size of each ethnicity was significantly different when comparing the inter-population variation for the entire sample and all age groups. As presented in Table 4.15, all parameters were found to vary significantly (p<0.05), except for MCL, PC, FC, CH, and FML, for the entire sample. The measurements expressing the most significant ethnic differences as indicated by F-values and p-values were MCW (F=20.532; p<0.001), followed by BAW (F=9.439; p<0.001) and OC (F=8.502; p<0.001). For descriptive analysis by age group, no measurement was found to be statistically significant in the age group of 0-2 years. Significant variations were found in four parameters (BAW, MCW, LCL, IPW) for the age group of 3-6 years and four parameters for the age group of 7-9 years (CBL, MCW, BAW, IPW). Moreover, nine parameters (age group of 10-12 years: BAW, MCW, LCL, IPW, PC, OC, CH, CBL, FMW), five parameters (age group of 13-15 years: BAW, MCW, LCL, IPW, OC), and seven parameters (age group of 16-20 years: BAW, MCW, LCL, IPW, OC, CH, FMW) were found to differ significantly between Malay, Chinese, and Indian (p<0.05).

Table 4.15: Descriptive analysis and mean variation of ethnic groups for the entire sample (N=521) and age groups

Para-meters	**N**	**Entire sample** (M=221 C=145 I=155)	**0-2 years** (M=53 C=30 I=29)	**3-6 years** (M=39 C=23 I=17)	**7-9 years** (M=21 C=20 I=16)	**10-12 years** (M=24 C=18 I=18)	**13-15 years** (M=29 C=27 I=23)	**16-20 years** (M=55 C=31 I=48)
MCL	M	164.35±13.56	147.07±12.32	162.12±6.30	169.56±6.95	168.19±6.59	171.20±7.85	175.29±6.77
	C	166.70±15.07	146.99±17.46	165.00±7.18	167.71±5.37	171.0839±6.48	176.47±8.23	176.58±9.45
	I	167.42±13.11	149.70±13.57	163.48±6.35	168.25±6.20	167.47±8.33	173.87±7.76	175.58±8.84
	F	2.547	0.378	1.400	0.478	1.319	2.843	0.249
	P	0.079	0.686	0.253	0.622	0.275	0.064	0.780
LCL	M	87.69±9.91	74.84±7.53	85.11±4.65	90.00±7.15	92.12±4.68	93.61±5.11	95.67±4.95
	C	90.97±10.85	76.02±9.97	87.72±5.23	91.7513±4.78	97.42±5.14	96.52±5.09	99.09±6.46
	I	88.01± 9.06	74.36±7.67	83.57±4.08	89.84±3.59	93.21±5.01	92.34±5.06	92.69±6.41
	F	**5.288**	0.321	**4.095**	0.701	**6.312**	**4.353**	**11.280**
	P	**0.005**	0.726	**0.020**	0.501	**0.003**	**0.016**	**<0.001**
NOL	M	159.24±15.35	139.80±11.84	157.28±6.70	161.95±18.09	164.758±6.8	167.24±7.59	171.69±6.37
	C	162.13±15.63	140.82±16.94	159.31±7.36	163.56±5.40	168.10±6.30	172.29±8.78	172.90±8.74
	I	163.31±14.05	142.76±13.14	157.96±5.40	165.18±6.30	164.20±8.48	170.70±7.80	172.50±8.47
	F	**3.674**	0.436	0.675	0.335	1.590	2.743	0.279
	P	**0.026**	0.647	0.512	0.717	0.213	0.071	0.757
MCW	M	139.03±9.30	128.32±8.74	138.16±5.94	142.12±7.95	142.79±4.83	142.65±5.89	145.25±5.73
	C	142.49±11.05	128.04±12.47	142.89±6.98	145.33±4.93	146.41±5.83	146.79±7.22	148.89±6.89
	I	135.32± 8.85	123.80±9.59	133.85±6.18	137.04±6.85	137.90±6.19	139.09±5.59	139.14±5.81
	F	**20.532**	2.080	**10.230**	**6.814**	**10.573**	**9.540**	**26.793**
	P	**<0.001**	0.130	**<0.001**	**0.002**	**<0.001**	**<0.001**	**<0.001**

Table 4.15, continued

Para-meters	N	Entire sample (M=221 C=145 I=155)	0-2 years (M=53 C=30 I=29)	3-6 years (M=39 C=23 I=17)	7-9 years (M=21 C=20 I=16)	10-12 years (M=24 C=18 I=18)	13-15 years (M=29 C=27 I=23)	16-20 years (M=55 C=31 I=48)
BAW	M	105.02±8.44	94.92±8.85	105.06±4.58	110.01±5.25	109.14±5.41	108.24±6.03	109.32±4.09
	C	107.76±10.38	94.86±14.19	108.73±4.36	111.61±4.45	109.84±3.80	112.66±6.09	112.22±6.52
	I	103.32± 8.11	92.31±10.71	102.73±3.78	104.71±4.03	105.71±3.88	107.06±3.97	106.70±5.06
	F	**9.439**	0.595	**9.889**	**10.396**	**4.360**	**7.168**	**11.214**
	P	**<0.001**	0.554	**<0.001**	**<0.001**	**0.017**	**0.001**	**<0.001**
IPW	M	105.28±15.95	82.72±9.58	100.36±5.79	108.58±6.36	112.96±4.47	116.49±6.78	119.98±6.54
	C	109.84±17.21	83.27±12.17	102.98±7.64	113.24±4.20	117.74±5.97	121.78±5.72	124.99±6.89
	I	105.44±14.32	82.01±11.30	97.77±6.31	106.52±4.49	109.69±5.95	114.35±4.90	115.33±6.10
	F	**4.205**	0.101	**3.182**	**8.174**	**10.081**	**10.317**	**21.230**
	P	**0.015**	0.904	**0.047**	**0.001**	**<0.001**	**<0.001**	**<0.001**
PC	M	106.16±8.13	97.52±8.18	108.68±5.25	110.92±7.31	107.62±5.45	107.22±5.20	109.69±6.19
	C	107.33± 8.65	97.66±10.96	107.81±6.22	110.24±5.56	110.00±5.41	110.73±6.05	110.38±5.44
	I	106.93± 7.92	100.17±8.93	109.22±6.79	108.57±4.88	104.39±6.93	109.45±7.02	109.19±6.92
	F	0.961	0.859	0.302	0.689	**4.070**	2.215	0.334
	P	0.383	0.426	0.740	0.507	**0.022**	0.116	0.716
OC	M	95.13±7.86	87.05±7.61	93.20±5.10	96.29±3.68	99.52±5.54	100.76±5.52	98.97±5.97
	C	97.01± 8.57	87.84±9.68	95.2157±4.18	95.99±3.52	100.88±6.31	102.82±6.96	101.34±6.47
	I	93.17± 7.87	83.52±6.09	92.17±5.07	94.48±6.94	94.89±5.94	95.60±6.21	96.91±6.73
	F	**8.502**	2.604	2.144	0.718	**5.181**	**9.209**	**4.594**
	P	**<0.001**	0.079	0.124	0.493	**0.009**	**<0.001**	**0.012**
CBL	M	88.92±11.18	73.29±5.73	85.17±4.10	91.20±4.38	93.40±4.29	97.48±5.73	99.30±4.85
	C	90.47± 11.77	73.23±8.28	85.49±4.81	91.33±4.03	95.65±4.19	98.37±5.04	101.44±6.14
	I	93.15± 10.87	75.92±6.91	86.33±5.33	94.50±4.48	97.23±4.95	99.05±4.90	100.69±5.96
	F	**6.460**	1.631	0.382	**3.264**	**3.890**	0.630	1.636
	P	**0.002**	0.201	0.684	**0.046**	**0.026**	0.535	0.199

Table 4.15, continued

Para-meters	N	Entire sample (M=221 C=145 I=155)	0-2 years (M=53 C=30 I=29)	3-6 years (M=39 C=23 I=17)	7-9 years (M=21 C=20 I=16)	10-12 years (M=24 C=18 I=18)	13-15 years (M=29 C=27 I=23)	16-20 years (M=55 C=31 I=48)
FC	M	103.64±9.88	90.72±7.60	102.00±4.97	107.11±5.88	107.04±4.41	108.19±5.16	112.05±5.56
	C	104.49±10.91	89.39±11.69	103.86±5.08	105.59±3.38	108.10±3.63	110.50±6.43	112.31±6.93
	I	105.25± 9.60	91.35±9.08	102.87±4.15	105.58±3.44	106.94±6.75	110.80±6.29	110.63±5.92
	F	1.176	0.355	1.071	0.765	0.306	1.597	0.988
	P	0.309	0.702	0.348	0.470	0.737	0.209	0.375
CH	M	126.72±11.74	110.02±7.82	125.96±5.41	130.85±6.19	131.28±4.07	133.93±6.13	135.97±6.06
	C	129.10±13.56	108.87±12.81	126.28±5.32	131.37±4.71	135.01±4.87	136.82±6.35	140.16±6.14
	I	127.99±11.39	109.85±10.15	125.68±4.55	129.46±5.47	130.09±5.89	133.79±5.81	135.22±6.01
	F	1.716	0.134	0.065	0.558	**4.999**	1.927	**6.862**
	P	0.181	0.875	0.937	0.576	**0.010**	0.153	**0.001**
FML	M	34.09±9.89	30.85±2.54	33.60±2.28	35.00±2.89	34.22±2.24	34.50±2.35	36.93±19.02
	C	34.07± 3.34	31.50±4.38	33.30±2.70	34.74±2.24	35.39±2.77	35.38±3.16	34.95±2.08
	I	33.86± 3.23	30.57±3.83	33.26±2.39	35.50±2.63	34.65±2.18	35.44±2.44	34.33±2.58
	F	0.055	0.574	0.160	0.384	1.235	1.090	0.612
	P	0.946	0.565	0.852	0.683	0.298	0.341	0.544
FMW	M	27.92±2.89	24.63±2.15	27.92±1.73	29.03±2.32	29.18±1.88	29.18±2.55	29.44±2.37
	C	28.84± 3.05	25.10±3.06	28.37±1.63	29.17±2.27	30.16±1.89	29.56±2.22	31.27±1.77
	I	28.17±2.77	24.84±2.54	28.05±1.68	28.24±1.94	28.18±2.09	29.41±2.22	29.50±2.32
	F	**4.453**	0.340	0.504	0.886	**4.670**	0.173	**7.759**
	P	**0.012**	0.712	0.606	0.418	**0.013**	0.841	**0.001**

ANOVA test, values are presented as mean±SD, Unit of measurement: millimetres (mm), N=number of samples expressed as M=Malay, C=Chinese, I=Indian, bold indicates statistical significance (p<0.05).

4.4.4 Multivariate analysis of variance (MANOVA)

The results of the MANOVA analysis are presented in Table 4.16. The MANOVA analysis tested the hypothesis of 'ethnic differences in the cranium are predictable across all age groups and between males and females when utilising DFA'. MANOVA analysis rejected the null hypotheses (p>0.05) demonstrating no statistically significant differences in all measurements among the Malay, Chinese, and Indian ethnicities except for MCW and CBL. In addition, no significant differences in all measurements between males and females except for MCW and FML. Furthermore, no significant interaction was observed between ethnicity and sex, or between ethnicity and age groups, for all parameters.

Table 4.16: Interactions between age groups, sex, ethnicity, sex*ethnicity, and age groups*ethnicity

Para-meters	Age groups		Sex		Ethnicity		Sex*ethnicity		Age groups* ethnicity	
	***F*-value**	**Sig.**	***F*-value**	**Sig.**	***F*-value**	**Sig.**	***F*-value**	**Sig.**	***F*-value**	**Sig.**
MCL	12.090	**0.001**	0.541	0.462	3.690	0.055	1.577	0.210	1.158	0.282
LCL	19.301	**0.000**	3.767	0.053	0.041	0.840	0.001	0.982	0.002	0.962
NOL	11.798	**0.001**	0.031	0.860	5.369	0.212	2.306	0.129	0.774	0.379
MCW	6.452	**0.011**	4.495	**0.034**	1.560	**0.021**	0.078	0.780	0.054	0.816
BAW	4.203	**0.041**	2.014	0.156	0.219	0.640	0.000	0.982	0.018	0.893
IPW	35.412	**0.000**	1.911	0.167	0.017	0.897	0.000	0.998	0.031	0.860
PC	3.420	0.065	0.090	0.764	3.442	0.064	2.560	0.110	0.023	0.879
OC	6.770	**0.010**	0.000	0.998	0.012	0.914	0.601	0.438	0.008	0.927
FC	11.757	**0.001**	1.996	0.158	0.748	0.388	0.134	0.714	0.190	0.663
CH	17.124	**0.000**	1.101	0.295	0.734	0.392	0.249	0.618	0.017	0.895
CBL	37.044	**0.000**	0.507	0.477	4.006	**0.046**	0.766	0.382	0.074	0.785
FML	8.969	**0.003**	5.633	**0.018**	0.738	0.391	0.672	0.413	3.687	0.055
FMW	8.938	**0.003**	1.646	0.200	1.303	0.254	0.650	0.421	0.147	0.702

MANOVA test**,** bold indicates statistical significance (p<0.05).

4.4.5 Discriminant function analysis (DFA)

Two assumptions must be fulfilled before conducting DFA, which are normality of the data and homogeneity of variance/covariance matrices.

4.4.5.1 Normality of the data

Upon analysing the skewness and kurtosis values presented in Appendix B, most parameters' skewness and kurtosis ratios fell well within the normal range (skewness: -2 to 2 and kurtosis: -7 to 7), suggesting that the data were normally distributed (Byrne, 2010). The histogram with normal distribution curves further demonstrated the normal distribution of the sample.

4.4.5.2 Homogeneity of covariance matrices

The results of the log determinants and the Box's M test are presented in Table 4.17 and Table 4.18, respectively. Log determinants were found to be relatively equal between Malay, Chinese, and Indian. In addition, the Box's M test was not statistically significant (p>0.001), indicating that the covariance matrices were equal.

Table 4.17: Log determinants of covariance matrices for ethnicity estimation

Group	Rank	Log determinants
Malay	2	-5.565
Chinese	2	-5.297
Indian	2	-5.990

Table 4.18: Box's M test of equality of covariance matrices for ethnicity estimation

Box's M		
F	Approx	2.772
	df1	6
	df2	3651521.573
	Significant	0.011

4.4.5.3 Pairwise comparison among the three ethnic groups

Pairwise comparison was calculated on the craniometric measurements in each sample to indicate the levels of similarity and dissimilarity among the three ethnic groups. The results of the pairwise comparison are presented in Table 4.19, showing that the samples were significantly different from each other (p<0.001). Overall, Malay and Chinese samples were the most similar to each other (5.729), followed by Malay and Indian samples (10.986), and finally, Chinese and Indian samples (14.637). Similar results were obtained when the samples were divided by age groups, where the distance between Malay and Chinese pair was the smallest, followed by Malay and Indian pair, and finally, Chinese and Indian pair.

Table 4.19: Pairwise comparison among the three ethnic groups for ethnicity estimation

Age groups (years)	Ethnicity	Malay	Chinese	Indian
Entire sample	Malay	0		
	Chinese	5.729	0	
	Indian	10.986	14.637	0
0-2	Malay	0		
	Chinese	1.260	0	
	Indian	7.616	7.698	0
3-6	Malay	0		
	Chinese	4.031	0	
	Indian	5.315	12.293	0
7-9	Malay	0		
	Chinese	3.022	0	
	Indian	8.830	12.790	0
10-12	Malay	0		
	Chinese	3.416	0	
	Indian	7.300	10.367	0
13-15	Malay	0		
	Chinese	2.964	0	
	Indian	7.170	10.113	0
16-20	Malay	0		
	Chinese	4.264	0	
	Indian	7.858	10.884	0

4.4.5.4 Canonical discriminant function analysis for the entire sample

Using stepwise DFA, canonical discriminant functions were developed for the entire sample and for each age group. The present study relied on canonical discriminant function analysis to explore the pattern of possible ethnic differences in a sample of sub-adult craniometric measurements for the age group of 0-20 years. For the entire sample, six parameters (LCL, CBL, IPW, CH, FC, MCW) were selected for the development of ethnicity estimation model. Canonical correlations for the first and second canonical variate were obtained, the values being 0.578 and 0.150, respectively. Accounting for 95.6% of the total variance, the first canonical discriminant function displayed high loadings for MCW. The second canonical function accounted for 4.4% of the total variance and displayed high loadings for LCL, CBL, IPW, CH, and FC. This within-group variation is illustrated in the canonical variates plot in Figure 4.2, demonstrating the general relationship between the three ethnic groups. A DFA structure matrix was used to interpret the results of the canonical variates plot, suggesting that the Chinese and Malay samples have longer cranial width (MCW) than Indian sample (represented on CAN 1). Additionally, Chinese and Indian have longer LCL, CBL, IPW, CH, and FC, than Malay (represented on CAN 2). The results of the structure matrix are shown in Table 4.20, and associated information for each axis is presented in Table 4.21 (Section 4.4.5.5), including unstandardised coefficients, functions at group centroids, canonical correlations, eigenvalues, and percentage of variance (%).

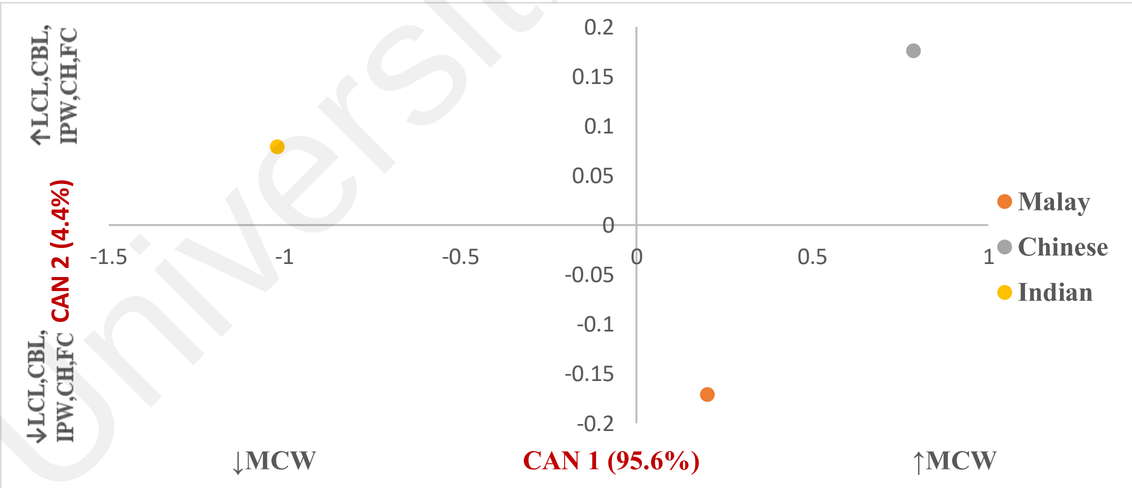


Figure 4.2: CAN 1 and CAN 2 representing 100% of the variation among the study samples

Table 4.20: Discriminant function analysis structure matrix for ethnicity estimation

Parameters	Function 1	Function 2
MCW	**0.389**	0.370
LCL	0.123	**0.741**
CBL	-0.168	**0.683**
IPW	0.120	**0.627**
CH	0.022	**0.527**
FC	-0.060	**0.344**
OC	0.250	**0.250**

Bold indicates the largest absolute correlation between each parameter and any discriminant functions.

4.4.5.5 Canonical discriminant function analysis for all age groups

The canonical discriminant function results for all age groups are presented in Table 4.21. Canonical discriminant function for the age group of 0-2 years selected five parameters (LCL, MCW, OC, FC, CBL) for the development of ethnicity estimation model. Canonical correlations for the first and second canonical variate were obtained, with the recorded values being 0.557 and 0.227, respectively. The first canonical discriminant function accounts for 89.2% of the total variance, and the second canonical discriminant function accounts for 10.8% of the total variance.

For the age group of 3-6 years, two parameters (MCW and FC) were selected for the development of the model. The first canonical variate had a canonical correlation of 0.497, explaining 93.5% of the total variance. The second canonical variate had a canonical correlation of 0.150, explaining 6.5% of the total variance. Three parameters (BAW, IPW, CBL) were selected for the age group of 7-9 years. The results of the canonical discriminant function provided canonical correlation for the first and second canonical variate, with the recorded values being 0.662 and 0.351, respectively. The first canonical variate represented 84.8% of the variation, while the second canonical variate represented 15.2% of the variation, forming a total of 100% of the variation.

For the age group of 10-12 years, four parameters (LCL, BAW, IPW, CBL) were selected for the development of the model. Canonical correlation for the first (80%) and second (20%) canonical variate represented 100% of the variation in the craniometric parameters used for the analysis. For the age group of 13-15 years, three parameters (MCL, MCW OC) were selected. Canonical correlations for the first and second canonical variate were obtained, with the recorded values being 0.555 and 0.299, respectively. The first canonical discriminant function accounts for 81.9% of the total variance and the second canonical discriminant function accounts for 18.1% of the total variance. Finally, for the age group of 16-20 years, six parameters (MCL, LCL, BAW, IPW, CBL, FMW) were selected for the development of the model. Canonical correlations for the first (91.9%) and second (8.1%) canonical variate represent 100% of the variation in the craniometric parameters used for the analysis.

Table 4.21: Canonical discriminant function including unstandardised coefficients, functions at group centroids, canonical correlations, eigenvalues, and percentage of variance (%)

Age groups (years)	0-2		3-6		7-9		10-12	
U.C	**F1**	**F2**	**F1**	**F2**	**F1**	**F2**	**F1**	**F2**
LCL	0.053	-0.257					0.061	0.141
MCW	0.118	0.073	0.188	-0.025				
BAW					0.182	-0.177	0.108	-0.081
IPW					0.043	0.239	0.143	0.024
OC	0.075	0.013						
FC	-0.070	0.182	-0.108	0.222				
CBL	-0.213	-0.048			-0.177	0.045	-0.211	0.093
Constant	-3.309	-4.062	-14.967	-19.397	-8.213	-11.032	-13.503	-16.189
Functions at group centroids								
Malay	0.318	0.216	-0.006	-0.151	0.295	-0.461	0.292	-0.518
Chinese	0.510	-0.336	0.685	0.145	0.763	0.376	0.897	0.500
Indian	-1.110	-0.047	-0.912	0.149	-1.341	0.134	-1.287	0.191
C.C	0.557	0.227	0.497	0.150	0.662	0.351	0.670	0.411
Eigenvalue	0.449	0.054	0.328	0.023	0.782	0.140	0.813	0.204
Percentage of variance (%)	89.2	10.8	93.5	6.5	84.8	15.2	80.0	20.0

Table 4.21, continued

Age groups (years)	13-15		16-20		Entire sample	
U.C	**F1**	**F2**	**F1**	**F2**	**F1**	**F2**
MCL	-0.053	0.122	-0.072	-0.018		
LCL			0.129	-0.021	0.063	0.126
MCW	0.103	0.061			0.090	-0.038
BAW			0.085	0.020		
IPW			0.110	-0.054	0.053	0.030
OC	0.125	-0.074			0.038	-0.056
FC					-0.116	-0.116
CH					0.050	0.045
CBL			-0.122	0.127	-0.170	0.019
FMW			0.052	0.357		
Constant	-18.037	-22.407	-11.432	-13.903	-6.117	0.944
Functions at group centroids						
Malay	0.277	-0.383	0.241	-0.305	0.200	-0.171
Chinese	0.684	0.357	1.223	0.309	0.786	0.176
Indian	-0.880	0.107	-1.065	0.150	-1.021	0.079
C.C	0.555	0.299	0.665	0.256	0.578	0.150
Eigenvalue	0.446	0.098	0.794	0.070	0.502	0.023
Percentage of variance (%)	81.9	18.1	91.9	8.1	95.6	4.4

U.C=unstandardised coefficients, C.C=canonical correlation.

4.4.5.6 Discriminant scores

Discriminant scores were calculated for the entire sample and each age group, and the scores were then plotted (Figure 4.3). Discriminant score equations are presented in Table 4.22. For the entire sample, Malay clustered between Chinese and Indian. Interestingly, Malay clustered closer to Chinese compared to Indian, and Chinese and Indian clustered furthest from each other. Similar patterns were also observed in the age groups of 0-2 years, 7-9 years, 10-12 years, and 13-15 years. However, in the age groups of 3-6 years and 16-20 years, Malay clustered relatively evenly between Chinese and Indian. Additionally, the centroids between Malay and Chinese were plotted more closely than the centroids between Chinese and Indian.

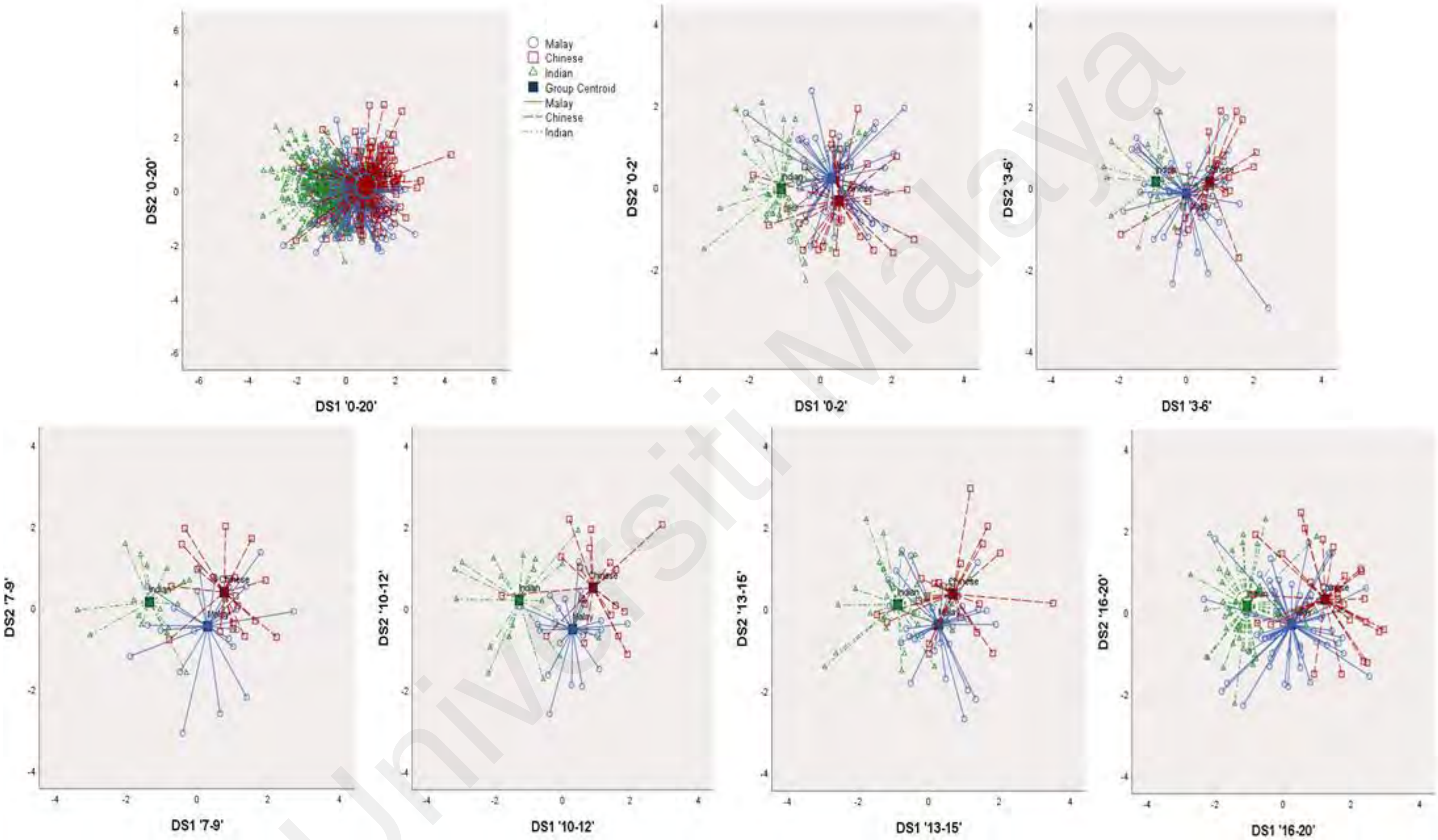


Figure 4.3: Discriminant scores for functions 1 and 2 for the entire sample and all age groups for ethnicity estimation

Table 4.22: Discriminant model equations for the entire sample and all age groups for ethnicity estimation

Age groups (years)	Discriminant score equation 1	Discriminant score equation 2
0-2	-3.309+(0.053*LCL)+(0.118*MCW)+(0.075*OC)+(-0.070*FC)+(-0.213*CBL)	-4.062+(-0.257*LCL)+(0.073*MCW)+(0.013*OC)+(0.182*FC)+(-0.048*CBL)
3-6	-14.967+(0.188*MCW) + (-0.108*FC)	-19.397+(-0.025*MCW)+(0.222*FC)
7-9	-8.213+(0.182*BAW)+(0.043*IPW)+(-0.177*CBL)	-11.032 +(-0.177*BAW)+(0.239*IPW)+(0.045*CBL)
10-12	-13.503+(0.061*LCL)+(0.108*BAW)+(0.143*IPW)+(-0.211*CBL)	-16.189+(0.141*LCL)+(-0.081*BAW)+(0.024*IPW)+(0.093*CBL)
13-15	-18.037+(-0.053*MCL)+(0.103*MCW)+(0.125*OC)	-22.407+(0.122*MCL)+(0.061*MCW)+(-0.074*OC)
16-20	-11.432+(-0.072*MCL)+(0.129LCL)+(0.085*BAW +(0.110*IPW)+(-0.122CBL)+(0.052*FMW)	-13.903+(-0.018*MCL)+(-0.021*LCL)+(0.020*BAW)+(-0.054*IPW)+(0.127*CBL)+(0.357*FMW)
Entire sample	-6.117+(0.063*LCL)+(0.090*MCW)+(0.053*IPW)+(0.038*OC)+(-0.116*FC)+(0.050*CH)+(-0.170*CBL)	0.944+(0.126*LCL)+(-0.038*MCW)+(0.030*IPW)+(-0.056*OC)+(-0.116*FC)+(0.045*CH) +(0.019*CBL)

4.4.5.7 Classification accuracy in ethnicity estimation for pooled-sex

The results of the classification accuracy for both original and cross-validated data are presented in Table 4.23. Classification accuracy for cross-validated data ranged from 56.0% to 67.4%. For the entire sample, ethnicity estimation yielded an accuracy of 57.5% for cross-validated data. Indian obtained the highest accuracy (74.2%), followed by Chinese (62.1%) and then Malay (36.2%). When the samples were divided by age groups, the best classification accuracy was obtained by the age group of 10-12 years with 67.4% accuracy. Indian obtained the highest accuracy (72.2%), followed by Chinese (71.8%) and Malay (58.3%). The lowest classification accuracy was obtained by the age group of 3-6 years with 56% accuracy. Indian achieved the highest accuracy (69.7%) followed by Chinese (65.0%), while Malay achieved the lowest accuracy (33.3%).

The misclassification rate varied among the ethnic groups, with Malay obtaining the highest misclassification rate compared to Chinese and Indian. For the entire sample, Malays were highly misclassified as Chinese (38%) and Indians (25.8%). In comparison, Chinese were misclassified as Malays (24.1%) and Indians (13.8%). Indians achieved the lowest misclassification rate as they were misclassified as Malays by 17.4% and Chinese by 8.4%. Similar patterns were obtained in all age groups except for the age groups of 3-6 years and 13-15 years, where Malays were misclassified relatively evenly with Chinese and Indians. This indicated a less distinct differentiation between the ethnic groups within these specific age ranges.

Table 4.23: Classification accuracy of original and cross-validation samples for the entire sample and all age groups for ethnicity estimation

Age groups (years)		Ethnicity	Classification accuracy of original and cross-validation samples (%)			
			Malay	Chinese	Indian	Total
Entire sample	Original	Malay	36.7	38.0	25.3	58.6
		Chinese	21.4	64.8	13.8	
		Indian	17.4	8.4	74.2	
	Cross-validation	Malay	36.2	38.0	25.8	57.5
		Chinese	24.1	62.1	13.8	
		Indian	17.4	8.4	74.2	
0-2	Original	Malay	50.9	28.3	20.8	61.2
		Chinese	33.3	53.3	13.3	
		Indian	10.3	10.3	79.3	
	Cross-validation	Malay	50.9	28.3	20.8	60.7
		Chinese	33.3	53.3	13.3	
		Indian	11.8	10.3	77.9	
3-6	Original	Malay	33.3	33.3	33.3	56.4
		Chinese	26.1	65.2	8.7	
		Indian	17.6	11.8	70.6	
	Cross-validation	Malay	33.3	33.3	33.3	56.0
		Chinese	26.3	65.0	8.7	
		Indian	18.5	11.8	69.7	
7-9	Original	Malay	38.1	38.1	23.8	63.5
		Chinese	25.0	65.0	10.0	
		Indian	6.3	6.3	87.5	
	Cross-validation	Malay	38.1	38.1	23.8	63.0
		Chinese	25.0	65.0	10.0	
		Indian	7.8	6.3	85.9	

Table 4.23, continued

Age groups (years)		Ethnicity	Classification accuracy of original and cross-validation samples (%)			
			Malay	Chinese	Indian	Total
10-12	Original	Malay	58.3	25.0	16.7	67.6
		Chinese	22.2	72.2	5.6	
		Indian	16.7	11.1	72.2	
	Cross-validation	Malay	58.3	25.0	16.7	67.4
		Chinese	22.6	71.8	5.6	
		Indian	19.7	11.1	72.2	
13-15	Original	Malay	44.8	27.6	27.6	57.5
		Chinese	21.7	60.9	17.4	
		Indian	25.9	7.4	66.7	
	Cross-validation	Malay	44.8	27.6	27.6	57.0
		Chinese	22.5	60.1	17.4	
		Indian	26.5	7.4	66.1	
16-20	Original	Malay	41.8	32.7	25.5	65.4
		Chinese	16.1	71.0	12.9	
		Indian	12.5	4.2	83.3	
	Cross-validation	Malay	41.8	32.7	25.5	65.1
		Chinese	16.1	71.0	12.9	
		Indian	13.3	4.2	82.5	

4.4.5.8 Classification accuracy in ethnicity estimation for males and females

F1m, F2m, F1f, F2f, and group centroids for each ethnicity are presented in Table 4.24. When each sex was treated separately, five parameters (LCL, MCW, IPW, FC, CBL) and two parameters (MCW, CBL) were obtained for males and females, respectively.

Table 4.24: F1m, F2m, F1f, F2f and group centroids for Malay, Chinese, and Indian

Sex	Parameters	F1	F2
Males	LCL	0.075	0.082
	MCW	0.094	0.039
	IPW	0.071	-0.047
	FC	-0.088	-0.111
	CBL	-0.156	0.124
	(Constant)	-3.901	-7.318
Function at group centroid			
Malay		0.229	-0.195
Chinese		0.799	0.207
Indian		-1.116	0.085
Females	MCW	0.156	0.040
	CBL	-0.130	0.064
	(Constant)	-9.845	-11.181
Function at group centroid			
Malay		0.188	-0.060
Chinese		0.532	0.068
Indian		-0.734	0.023

The results of the classification accuracy for both original and cross-validated data are presented in Table 4.25. Cross-validated accuracy of 54.5% and 55.4% were obtained for males and females, respectively.

Table 4.25: Classification accuracy for original and cross-validation samples for males and females for ethnicity estimation

	Classification accuracy of original and cross-validation samples (%)			
	Malay (N=119/102)	**Chinese (N=79/66)**	**Indian (N=81/74)**	**Total**
Males				
Original	38.7	63.3	77.8	57
Cross-validation	37.0	60.8	74.1	54.5
Females				
Original	40.2	65.2	70.3	56.2
Cross-validation	39.2	63.6	70.3	55.4

N=number of males/number of females.

4.5 Machine learning methods

4.5.1 ML methods with GridSearchCV results

The results of each ML method with GridSearchCV are presented in Table 4.26. For sex estimation, RF achieved the optimal hyperparameters: max_depth=6, max_samples=40, and n_estimators=150. SVM obtained the best hyperparameters of learning rate (C)=10, gamma=0.01, and kernel=rbf, while LDA yielded a shrinkage value of 0.02. For ethnicity estimation, RF obtained the optimal hyperparameters with max_depth=6 and n_estimators=45, whereas SVM achieved the best hyperparameters with a learning rate (C) of 1000, gamma of 0.001, and a kernel of rbf. Additionally, LDA resulted in a shrinkage value of 0.01.

Table 4.26: Machine learning methods with GridSearchCV for sex and ethnicity estimation

	RF		SVM		LDA	
	Hyper-parameters	Value	Hyper-parameters	Value	Hyper-parameters	Value
Sex estimation	Max_depth N_estimators	6 150	C Gamma Kernel	10 0.01 RBF	Shrin-kage	0.02
Ethnicity estimation	Max_depth N_estimators	6 45	C Gamma Kernel	1000 0.001 RBF	Shrin-kage	0.01

RBF=radial basis function

4.5.2 Performance metrics of different machine learning (ML) methods for sex and ethnicity estimation

Performance metrics such as accuracy, precision, recall, and F1-score of three ML methods for sex and ethnicity estimation models are summarised in Table 4.27 and Table 4.28, respectively. For sex estimation, RF had the highest accuracy ratio of 0.73, while LDA had the lowest accuracy ratio of 0.65. In contrast, for ethnicity estimation, LDA had the highest accuracy ratio of 0.58, while RF had the lowest accuracy ratio of 0.52. In accordance with recent efforts to make statistical programmes freely available for practitioners (Berg & Kenyhercz, 2017), a GUI was developed utilising the RF and LDA models for sex and ethnicity estimation, respectively. This application, SNA estimator version 1.0, was built using streamlit application and is freely available at https://snaestimator.streamlit.app/ (Figure 4.4).

Table 4.27: Performance metrics of different machine learning methods for sex estimation

ML methods	Sex	Performance metrics			
		Accuracy	Precision	Recall	F1-Score
RF	Male	**0.73**	0.71	0.80	0.75
	Female		0.75	0.64	0.70
SVM	Male	0.67	0.66	0.75	0.70
	Female		0.69	0.58	0.63
LDA	Male	0.65	0.64	0.72	0.68
	Female		0.66	0.58	0.62

Bold signifies the highest percentage of classification accuracy, RF=random forest, SVM=support vector machines, LDA=linear discriminant analysis.

Table 4.28: Performance metrics of different machine learning methods for ethnicity estimation

ML methods	Ethnicity	Performance metrics			
		Accuracy	Precision	Recall	F1-Score
RF	Malay	0.52	0.49	0.68	0.57
	Chinese		0.41	0.31	0.35
	Indian		0.69	0.51	0.59
SVM	Malay	0.57	0.49	0.71	0.58
	Chinese		0.60	0.36	0.45
	Indian		0.72	0.58	0.65
LDA	Malay	**0.58**	0.51	0.69	0.59
	Chinese		0.67	0.43	0.52
	Indian		0.69	0.58	0.63

Bold signifies the highest percentage of classification accuracy, RF=random forest, SVM=support vector machines, LDA=linear discriminant analysis.

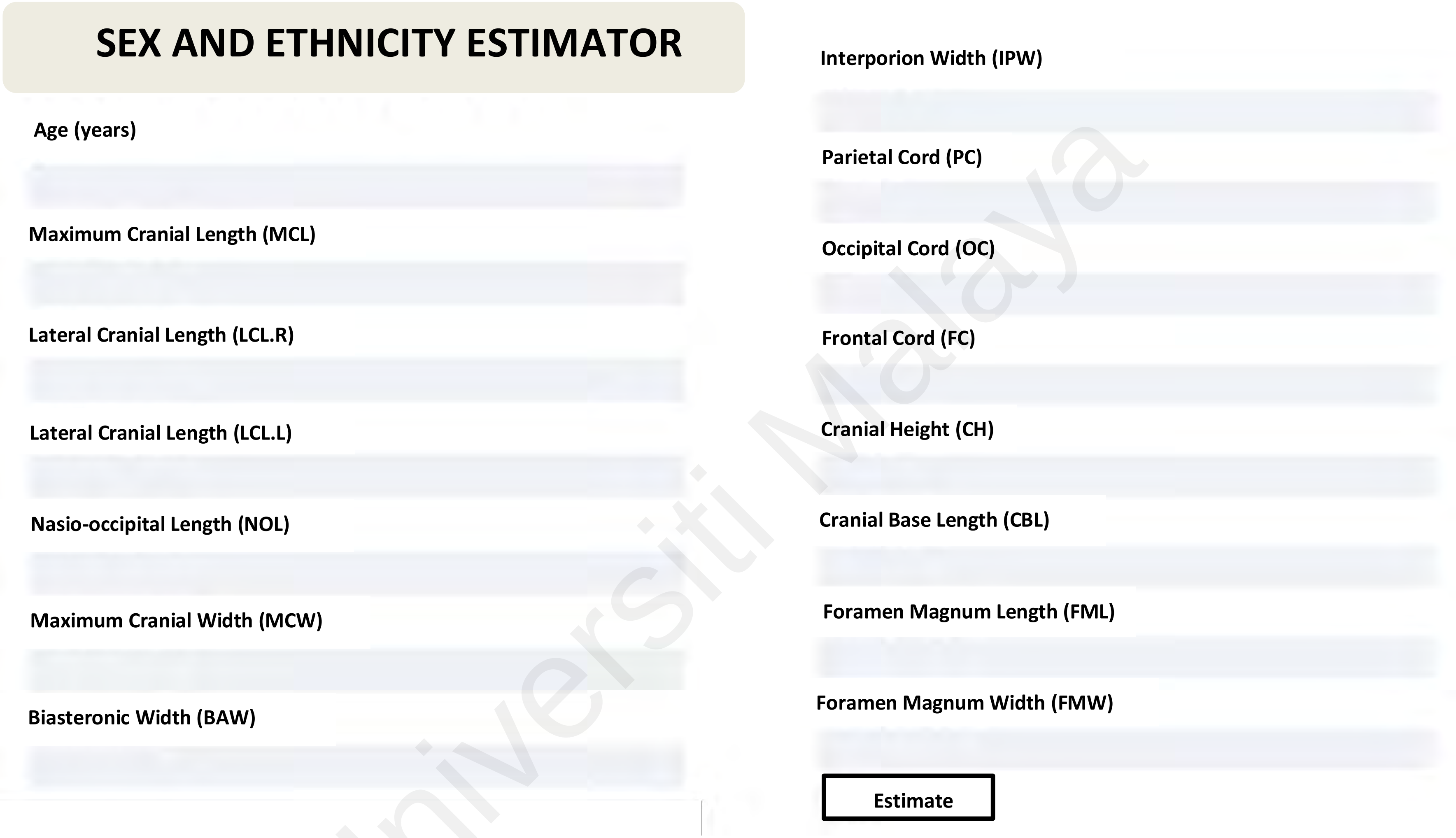


Figure 4.4: GUI design for sex and ethnicity estimation, available for free at https://snaestimator.streamlit.app

4.5.3 Validity of sex and ethnicity estimation models between machine learning (ML) and classical statistical methods (DFA and BLR)

Overall classification accuracies between ML and classical methods for the entire sample are summarised in Table 4.29. For sex estimation models, ML method, specifically RF, achieved the highest classification accuracy of 73%. On the other hand, the classical method of DFA obtained the lowest accuracy of 61.6%. For ethnicity estimation models, ML method (LDA) obtained a slightly higher accuracy of 58% compared to the classical method of DFA, which yielded an accuracy of 57.5%. These results suggested that ML methods, particularly RF for sex estimation and LDA for ethnicity estimation, outperformed the classical methods in terms of overall classification accuracy.

Table 4.29: Overall classification accuracies between machine learning and classical methods for sex and ethnicity estimation

Statistical methods	Accuracy (%)
Sex estimation	
DFA	61.6
BLR	66.9
RF	**73.0**
SVM	67.0
LDA	65.0
Ethnicity estimation	
DFA	57.5
RF	52.0
SVM	57.0
LDA	**58.0**

Bold signifies the highest percentage of classification accuracy, DFA=discriminant function analysis, BLR=binary logistic regression, RF=random forest, SVM=support vector machines, LDA=linear discriminant analysis.

4.6 Cephalic index (CI)

4.6.1 Descriptive statistics of CI

The mean±SD of the CI for each age group are shown in Table 4.30. The overall mean CI was 84.04±4.71. The mean CI for males was 83.91±4.63, while the mean CI for females was 84.13±4.72. No significant differences were found between the male CI and female CI for all age groups. Based on the values observed in the present study, a marked decrease in the mean value for CI for the first few years of life was noted, followed by a milder decrease for the age range between three and six years, with stabilisation achieved after that.

Table 4.30: Descriptive statistics and mean variation of cephalic index between males and females for all age groups

Age groups (years)	CI			*p*-value (<0.05)
	Total (N=521)	Male (N=279)	Female (N=242)	
0-2	86.31±5.37 (75.2-101.7)	86.52±5.22 (75.9-96.9)	86.06±5.59 (75.2-101.7)	0.835
3-6	84.99±4.71 (72.3-96.8)	85.93±4.48 (75.7-93.6)	83.87±4.79 (72.3-96.8)	0.763
7-9	84.22±4.76 (72.9-92.6)	84.50±4.86 (72.9-92.6)	83.78±4.67 (75.0-92.1)	0.984
10-12	84.44±4.18 (75.6-93.3)	83.72±3.92 (75.6-93.3)	85.15±4.37 (76.2-92.2)	0.389
13-15	82.27±4.70 (72.0-92.3)	81.89±4.64 (72.0-92.3)	82.67±4.79 (74.9-91.2)	0.962
16-20	82.03±4.54 (71.0-93.8)	80.87±4.66 (71.0-91.5)	83.25±4.10 (74.5-93.8)	0.962
Total	**84.04 ± 4.71**	**83.91 ± 4.63**	**84.13 ± 4.72**	**0.815**

Values are presented as mean±SD (min-max), N=number of samples.

4.6.2 Proposed CI of the Malaysian population

Based on the data obtained in these results, the modified CI range for the current Malaysian population is categorised as shown in Table 4.31. The modified ranges of the current CI of the Malaysian sub-adult population were as follows: dolichocephalic, 78.8 or less; mesocephalic, 78.9–89.0; brachycephalic, 89.1–94.0; and hyperbrachycephalic, 94.1 or higher.

Table 4.31: The proposed Malaysian classification of cephalic index

CI index	Skull morphology
$X \leq 78.8$	Dolichocephalic
78.9 – 89.0	Mesocephalic
89.1 – 94.0	Brachycephalic
$94.1 \geq X$	Hyperbrachycephalic

Figure 4.5 illustrates the percentages of cranial shape in Malaysian sub-adult population. Analysis of the cranial shape in the Malaysian sub-adult population indicated that the dominating type was mesocephalic (66.4%), followed by dolichocephalic (18.4%), and brachycephalic (12.3%). Hyperbrachycephalic shape was found to be the least frequently observed category (2.9%).

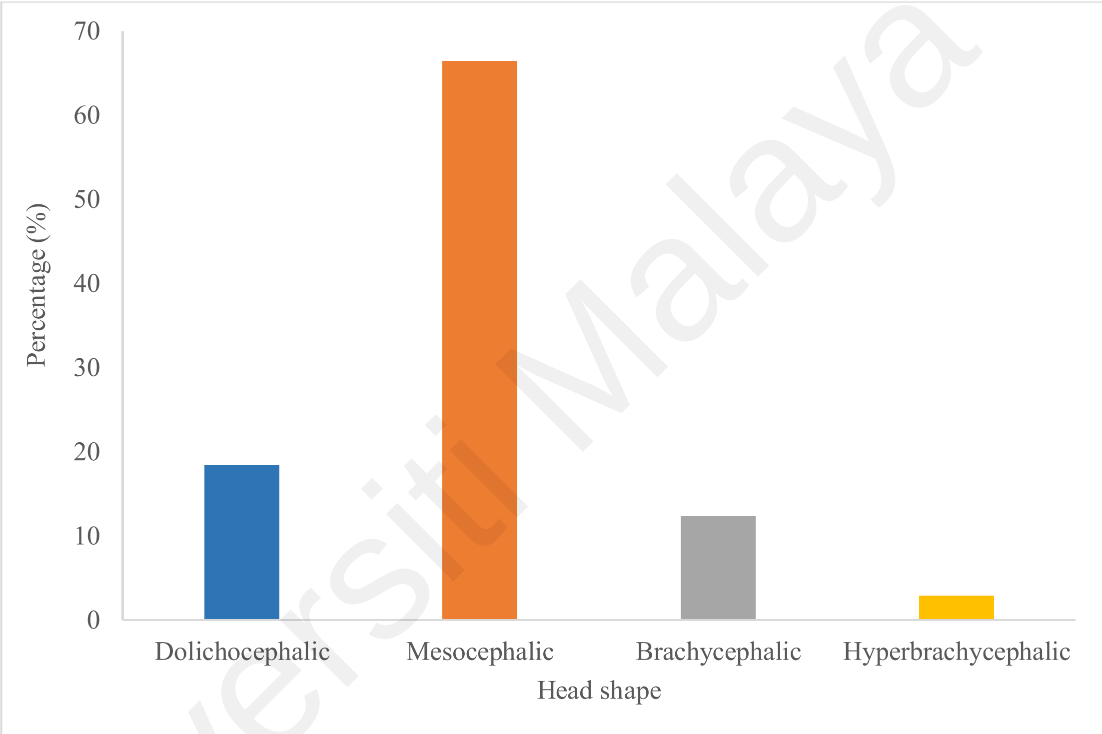


Figure 4.5: The cranial shapes of Malaysian sub-adults

Figure 4.6 presents the distribution of cranial shape percentages within the different age groups. It provides a breakdown of the cranial shape distribution among sub-adults of varying ages. Among the younger age groups (0-2 years, 3-6 years, 7-9 years, and 10-12 years), the prevalent head shape was mesocephalic, followed by brachycephalic and dolichocephalic forms. Meanwhile, among the older age groups (13-15 years and 16-20 years), mesocephalic remained the prevalent head shape, followed by dolichocephalic and brachycephalic shapes.

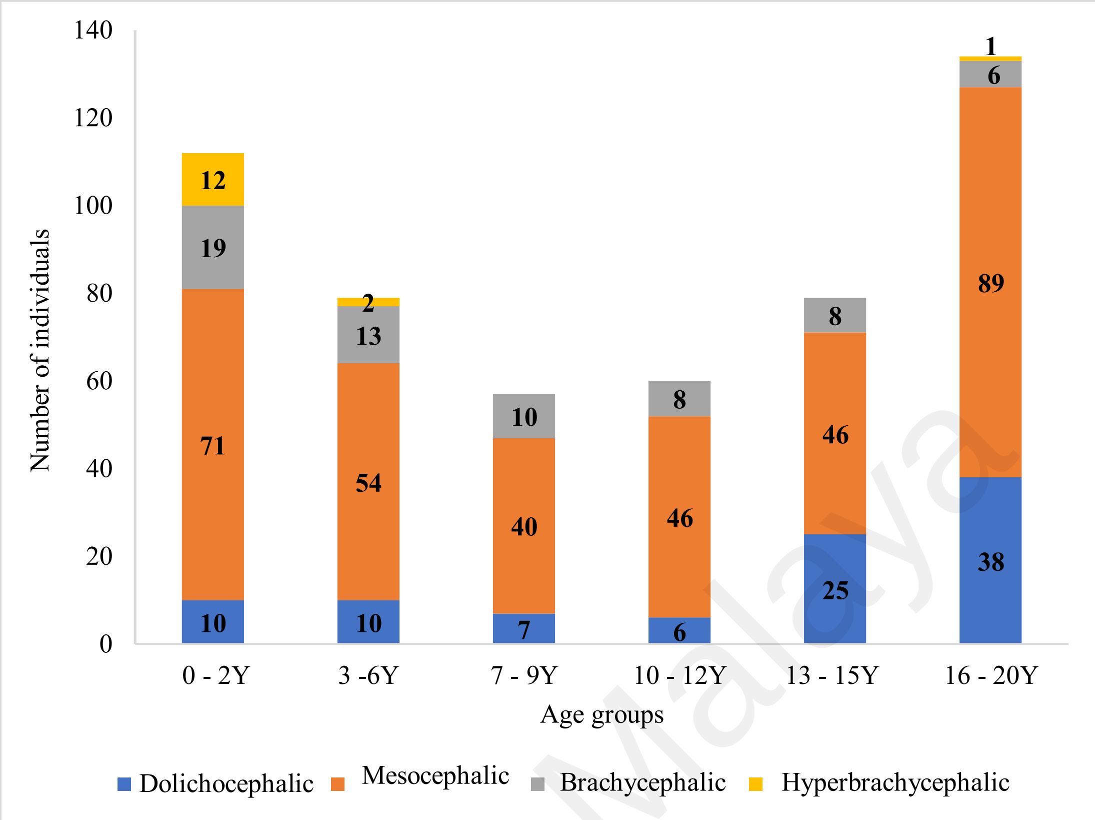


Figure 4.6: The cranial shapes of Malaysian sub-adults by age groups

4.6.3 A comparison of the CI between the Malaysian population and other populations

The average CI values for the Malaysian population and other populations worldwide are presented in Table 4.32.

Table 4.32: Comparison of cephalic index values between the Malaysian population and other populations

Age groups (years)	0-2	3-6	7-9	10-12	13-15	16-20
Malaysian	86.31	84.99	84.22	84.44	82.27	82.03
Korean	89.01	85.79	85.84			
Japanese	86.5					
Brazilian	83	80.48	82.07	83.17	82.17	82.2
European	82.7	80.47	81.9	82.77	82.3	82.45
Siberian			80.37	81.51	80.17	
Polish	81.46					
Caucasian	77.87	75	82.5	74.4	75	76.5

4.6.4 Cephalic index (CI) classifications

A comparison between the proposed CI classification for the Malaysian population and other existing CI classifications are presented in Table 4.33. The proposed Malaysian CI classification ranges for all cranial shapes were found to be higher than Cohen's and Standring's CI classifications, but lower than Nam's and Koizumi's classifications.

Table 4.33: Comparison between the proposed cephalic index classification and other existing cephalic index classifications

Cranial shape	Malaysia	Cohen	Standring	Nam	Koizumi
Dolicho-cephalic	X≤ 78.8	X≤ 75.9	*X≤ 74.9*	*X≤ 80.1*	X≤ 79.1
Meso-cephalic	78.9-89.0	76.0-80.9	75.0-79.9	80.2-93.4	79.2-93.8
Brachy-cephalic	89.1-94.0	81.0-85.4	80.0-84.9	93.5-100.0	93.9-101.1
Hyper-brachy-cephalic	94.1 ≥X	85.5 ≥X	85.0 ≥X	100.1 ≥X	101.2 ≥X

CHAPTER 5: DISCUSSION

This chapter will discuss the results obtained from the statistical analyses in four main sections. The first two sections of this chapter will discuss the development of sex and ethnicity estimation models for Malaysian sub-adults. The estimation models were developed using classical statistical analyses, such as DFA and BLR, based on the most meaningful combinations of measurements. However, it was proposed that ML algorithm may increase the accuracy percentage of sex and ethnicity estimation (Nikita & Nikitas, 2020). Thus, the third section will compare the validity of the sex and ethnicity estimation models developed using ML algorithms (RF, SVM, and LDA) and classical statistical methods (DFA and BLR). The fourth section will discuss the new CI classification for the Malaysian sub-adult population. This new classification system aimed to provide a more accurate and refined approach for interpreting CI measurements of the Malaysian sub-adult population considering the variations observed within this specific group. By achieving these goals, this study aimed to contribute to the field of forensic anthropology and population studies by providing more accurate and updated sex and ethnicity estimation methods, as well as a refined CI classification system for the Malaysian sub-adult population. Additionally, this chapter will address the limitations of this study and provide recommendations for future research.

5.1 Sex estimation models

5.1.1 Sexual dimorphism

Sexual dimorphism is the differences in size, shape, or other physical characteristics between males and females. These differences are frequently influenced by the secretion of hormones during growth and development (İşcan, 2005; Klales & Burns, 2017; Scheuer & Black, 2000). Prior to birth, females would receive more estrogen than males would receive testosterone during the same period of development. Differences in the secretion of sex hormones between males and females provide the initial requirements for the development of sexual dimorphism early in life. During puberty, there is an increase in the levels of hormonal activities. For example, estradiol is responsible for pubertal growth spurts and increases in body fat in females. Meanwhile, testosterone is responsible for growth spurts, decreased body fat, and increased muscle mass in males. These sex hormones are also responsible for the development of secondary sexual characteristics and the acceleration of skeletal maturation (Roche & Sun, 2003). Therefore, the traits between females and males are very distinctive due to the development of the female secondary sexual characteristics that starts to occur approximately two years earlier than males. This difference would result in an early growth spurt in females, whereas males have a delayed growth spurt (Humphrey, 1998). However, the levels of these hormones can fluctuate throughout development and can be influenced by numerous genetic and environmental factors. Therefore, these factors must be considered when studying the development and expression of sexual dimorphism in different populations.

The neurocranium, which includes the cranial base and cranial vault, is the most sexually dimorphic region of the sub-adult craniofacial skeleton. This is because the early growth trajectory of the neurocranium is mainly related to neurological development and brain growth. During the first two years of life, the brain undergoes substantial change, leading to an increase in head circumference from 65% to 90% of the adult size. The growth of the cranial vault is completed by around five years of age, while the increase of the cranial base happens gradually and reaches adult size at around 13 years of age. Because of these growth patterns, the neurocranium develops faster and reaches adult size earlier than the facial complex, which includes the maxilla, mandible, and other structures related to mastication and speech. This difference in growth can contribute to the development of sexual dimorphism in measurements associated with the neurocranium, as the rates of neurological development and brain growth may differ between males and females (Farkas et al., 1992; Waitzman et al., 1992). However, the specific patterns of sexual dimorphism in the neurocranium can vary depending on the specific traits being studied and the population being analysed, as these patterns can differ among ethnic groups and geographic regions.

This study found that the growth of neurocranium is most significant within the first six years of age and slowly continues until 20 years of age. Specifically, six measurements were found to be statistically significant ($p<0.05$) in the youngest age group of 0–2 years (MCW, FC, CH, CBL, FML, and FMW) and 3-6 years (LCL, MCW, IPW, FC, CBL, and FMW), and the number of significant measurements increased with age (Table 4.5). A parallel growth pattern for males and females prior to six years had resulted in low levels of sexual dimorphism. Conversely, differences in the onset of puberty between males and females aged 10 to 12 had increased the level of sexual dimorphism in the Malaysian sub-adults (Table 4.7). Sexual dimorphism became more evident in the older age group (>16 years), as they had reached the end stage of craniofacial development before skeletal maturity (Gonzalez, 2012). Hence, the findings for the age group of 16-20 years were comparable with the results of other studies on adult populations (Franklin et al., 2013b; Gillet et al., 2020; Ramamoorthy et al., 2016). The existence of sex differences could also be due to age-related growth changes, as observed when the accuracy of sex estimation improved with age from 60% (62.5% to 67.9%) to higher than 85% (88.1% to 90.3%; Table 4.14).

The present study's findings were consistent with two previous studies conducted by Gonzalez (2012) and Noble et al. (2019). These studies reported that the level of sexual dimorphism was minimal in the age range of 3–6 years, which then increased with age. The results showed that the differences between males and females, in terms of various measurements used for sex estimation, would become more pronounced as individuals grow older. Baughan and Demirjian (1978) contributed to the understanding of sexual dimorphism by demonstrating that it was present as early as at six years of age, with a 5% difference between males and females. The authors observed that the level of sexual dimorphism diminished during puberty (around 4% difference) before becoming more pronounced post-puberty at the age of 18 (8% difference). These findings indicated that the level of sexual dimorphism would fluctuate throughout different stages of development and maturation. The magnitude of sexual dimorphism can also vary between and within populations. This variation can be influenced by several factors, including differences in body composition, biochemical factors, environmental conditions, nutrition, and genetics (Kimmerle et al., 2008).

The effects of population differences were evident in this study, particularly in cranial length and height. The cranium of the Malaysian sub-adult population was found to be 4%–17% smaller than the standard Romanian (Teodoru-Raghina et al., 2017) and Taiwanese (Hsiao et al., 2010) populations. These discrepancies suggested that foreign indices are unsuitable for the Malaysian population. However, data for the Malaysian sub-adult population are currently unavailable. Thus, sex estimation should be conducted based on data derived from the respective population.

5.1.2 Discriminant function analysis (DFA) vs binary logistic regression (BLR)

The accuracy of sex estimation in the present study depended on the selection of suitable statistical methods. DFA and BLR are commonly employed to analyse craniometric data. DFA is widely utilised in sex estimation and is known to be population-specific. However, the accuracy of this method can vary based on the specific population under study (İşcan & Steyn, 2013). Previous researchers have used DFA to estimate the sex of sub-adults' crania and achieved different ranges of classification accuracies (Gonzalez, 2012; Hsiao et al., 2010; Noble et al., 2019; O'Donnell et al., 2017; Teodoru-Raghina et al., 2017; Veroni et al., 2010). On the other hand, BLR was reported to have outperformed DFA in assessing cranial sexual dimorphism due to its increased flexibility and robustness (Toneva et al., 2018). In the present study, DFA and BLR showed similar classification accuracies between the original and validation samples. However, BLR consistently achieved a higher classification accuracy of approximately 4.9% and lower sex bias rates compared to DFA (Table 4.14). In addition, BLR has also successfully determined sex using fewer parameters (Table 4.13). This observation is aligned with previous works that produced marginally better rates of classification when BLR was used (Ekizoglu et al., 2017; Macaluso Jr., 2010; Santos et al., 2014; Singh & Pathak, 2013).

A minimum standard of accuracy for adult sex estimation methods was set at 80% to 85% (Klales & Burns, 2017). However, it has been recommended that sub-adults' accuracy be ≥75% (Stull et al., 2020). In the present study, DFA and BLR recorded accuracies of ≥75% in all age groups, except for 0–2 years and 3–6 years (Table 4.14). The age group of 0-2 years showed the lowest accuracy rate for DFA and BLR due to the low level of sexual dimorphism in foetal and infants' crania (Klales & Burns, 2017). In the age group of 5-6 years, males and females experienced similar changes that included the end of neurocranial growth and the decreasing rate of growth in the transition from being a child to a juvenile (Scheuer & Black, 2000). These factors could have contributed to a reduced ability to accurately estimate sex using DFA and BLR in these age groups. The findings of the present study are aligned with the results of another study that found sex estimation as being unreliable in this age range (Noble et al., 2019). However, the present study contradicted their findings by demonstrating that DFA and BLR can be confidently employed for sex estimation of sub-adults from seven to 20 years, which is earlier than previously reported (Noble et al., 2019). Overall, this study has highlighted the limitations of DFA and BLR for sex estimation in the early years of life. The findings supported the notion that these methods could become more reliable and accurate for sub-adults above the age of six years.

Previous methods have demonstrated higher sex biases toward male classification (Gonzalez, 2012; Teodoru-Raghina et al., 2017). Similarly, the current findings also demonstrated higher classification accuracies in males than females. This could be due to a slightly larger male sample size than the female sample size, which influenced the predictive power of sex estimation methods (Hanifah et al., 2015). However, the accuracy rate for females in this study was higher than for males aged 10 to 12. In this age range, the traits between females and males would be very distinctive due to the development of the female secondary sexual characteristics, which starts to occur approximately two years earlier than males (Humphrey, 1998).

The classification accuracy of sub-adults' crania in this study was comparable to the accuracy of previous studies (Gonzalez, 2012; Hsiao et al., 2010; Noble et al., 2019). Discriminant equations were developed and more than 80% classification accuracy was achieved by several studies (Gonzalez, 2012; Hsiao et al., 2010). The accuracy of sex estimation can be influenced by two main factors: age and population differences. Studies which focused on older age groups, particularly after puberty, obtained higher accuracy rates (Hsiao et al., 2010) compared to studies involving younger age groups (Noble et al., 2019). This observation can be attributed to the lower level of sexual dimorphism present in younger individuals prior to the onset of puberty (Noble et al., 2019). Population homogeneity can also impact the classification accuracy of sex estimation. A study on the homogeneous population of Taiwan reported a higher accuracy rate of up to 95% (Hsiao et al., 2010). However, a study on the New Mexican population with diverse ethnicities reported an accuracy rate of lower than 60% (O'Donnell et al., 2017). Moreover, previous research has shown that ethnic variations can influence the classification accuracy of sex estimation of the neurocranium (Holland, 1986; Wescott & Moore-Jansen, 2001).

Considering Malaysia's diverse and multi-ethnic composition, it is important to use data derived from the Malaysian population instead of referring to population data from other countries. Therefore, the lower overall classification accuracy observed in the present study could be attributed to the age and ethnic variations present within the Malaysian population. These factors highlighted the importance of using population-specific data, as well as to account for age and ethnic variations, when developing and applying sex estimation models based on cranial measurements in Malaysia.

5.2 Ethnicity estimation models

Metric and nonmetric anthropological analyses would often be deployed when assessing ethnicity from skeletal remains (İşcan, 2005). As a nonmetric method, the assessment of morphoscopic traits in sub-adults can be challenging in terms of reproducibility and accuracy (Weinberg et al., 2005). This is because most sub-adult bones are small and not fully developed, causing difficulties in locating the anthropological landmarks (Weinberg et al., 2005). Therefore, this present study has focused on the metric method of using MSCT data to determine the differences in cranial measurements in the Malaysian multi-ethnic sub-adult population. This approach allows for more reliable, objective, accurate, and reproducible measurements, which can facilitate the assessment of cranial differences among various ethnic groups within the sub-adult population in Malaysia.

Cranial features were found to differ significantly in the different ethnic groups in Malaysia (Kranioti et al., 2018). The present study observed that Malays and Chinese cranium shapes were shorter and wider than that of Indians. Additionally, several parameters, such as MCW, BAW, and OC, showed significant differences between Malays, Chinese, and Indians (Table 4.15). Previous studies have also reported that these parameters indicated significant differences between ethnicities in different populations, such as Japanese and Thai (Kongkasuriyachai et al., 2022) and Turks, Cypriots, and Cretans (Kranioti et al., 2018). Hence, these parameters could be essential in devising robust classification models for cranium-based ethnicity estimations for sub-adults.

Cranial growth during the sub-adult age is a dynamic process influenced by various factors, including differences in growth patterns and the onset of puberty between males and females (Norris & Carr, 2020). These factors contribute to cranial variations between younger and older sub-adults. The present study observed that ethnic variations were not particularly significant in the younger age groups (below six years; Table 4.15). This level of variation could be attributed to a parallel growth pattern observed in males and females of Malay, Chinese, and Indian ethnicities during the early developmental period. The similarities in cranial growth trajectories among the different ethnic groups may lead to the lack of ethnic variations in cranial measurements in the younger age groups (0-2 years and 3-6 years).

The overall classification accuracy for ethnicity estimation in the sub-adult population was 57.5%. However, there were notable differences in the accuracy between the different ethnic groups, with Malays exhibiting a substantially lower accuracy rate (36.2%) compared to Chinese (62.1%) and Indians (74.2%). These differences can be attributed to a high similarity of craniometric measurements between Chinese and Malays compared to the measurements between Indians and Malays, and Chinese and Indians, as is evident in the discriminant score plots where the Malay cluster was more inclined towards the Chinese cluster (Figure 4.3). These pairwise comparisons also showed that several differences existed between Malays and Chinese (Table 4.19). Genetic data studies have demonstrated higher genetic similarities between Malays and Chinese than between Malays and Indians. The presence of more than 20% of Chinese DNA markers in Malay DNA was apparently adequate to manifest resemblances in cranial morphology (Deng et al., 2014). In contrast, only 4.3% and 1.3% of Malay and Chinese DNA markers were found in Indians, respectively, which resulted in a great partition between the two groups (Malays/Chinese vs Indians; Wong et al., 2014). Similarly, a study of the Singaporean population found that Malays and Chinese have a higher anthropological association with each other than with Indians (Yong et al., 2006).

The present study obtained low accuracy rates in the age groups of 0-2 years and 3-6 years (Table 4.23). However, as individuals reached the age group of 10-12 years, differences in the onset of puberty between males and females became more apparent. Differences in the timing and progression of puberty can contribute to increased levels of ethnic variations in the Malaysian sub-adults' crania. The divergent effects of puberty on cranial growth and development between males and females within different ethnic groups may lead to greater disparities in cranial measurements, as individuals progress through the sub-adult age range. This can be observed when the highest classification accuracy of 67.4% was obtained by the age group of 10-12 years, which corresponded to the emergence of secondary sex characteristics and implied the presence of ethnic differences in cranial morphology during this stage of development (Table 4.23). Therefore, the observed variations in cranial measurements and the influence of ethnic factors would become more pronounced with age, particularly during the period close to the onset of puberty. These findings highlighted the importance of considering age-specific cranial variation and development patterns when performing ethnicity estimation in sub-adult populations.

The low overall accuracy in ethnicity classification could be attributed to the unequal distribution of samples across different age groups, as DFA tends to be more accurate when applied to small sample sizes (Buck & Vidarsdottir, 2004). Imbalanced sample sizes across ethnicities or age groups could influence the overall accuracy rates. In addition, a combination of factors, including genetics, inter-racial marriages, and migrations, can also influence ethnic variations in cranial measurements. In inter-racial marriages, the estimation of an individual's ethnicity may be incorrect if their biological ethnicity differs from their self-identified or cultural ethnicity. The current findings have highlighted the importance of considering both age and ethnicity factors when studying cranial variations and developing sex estimation models for Malaysian sub-adults.

The sex of a cranium can affect the accuracy of ethnicity estimation due to the differences in size between male and female craniums. However, the present study found that the accuracy percentages between pooled-sex were higher than the accuracy between males and females (Table 4.23 and Table 4.25). This could be due to the low level of sexual dimorphism in the Malaysian sample. This reduced variability in crania within the Asian population was also observed in Japanese, Thai, and Filipino samples (Green & Curnoe, 2009; Kongkasuriyachai et al., 2022; Tallman, 2019). These findings contrast with the results of studies on the Caribbean (Herrera & Tallman, 2019) and Mediterranean (Kranioti et al., 2018) populations, where higher sex-specific cross-validated classification accuracies were observed in males than in females. These findings emphasised the importance of developing population-specific methods for ethnicity estimation in Asian groups. Hence, it would not be sensible to generalise these findings as a standard for other population groups.

The current classification accuracies obtained in this study were comparable to those obtained by Smith et al. (2013) and Buck and Vidarsdottir (2004), which ranged between 7.7% and 76.2%. Smith et al. (2013) studied the ethnicity estimation of sub-adults using temporal bone traits. The lower classification accuracies observed in Utahn (7.7%) and Mexican groups (28.6%) compared to Austrian (65.8%), Egyptian (56.7%), and Polynesian (53.3%) groups were due to their close geographic proximity and genetic similarities. Meanwhile, Buck and Vidarsdottir (2004) obtained over 70% accuracy when estimating the ethnicity of sub-adults of five distinct population groups, namely African Americans, Native Americans, Caucasians, Inuit, and Pacific Islanders, using mandibular morphology and DFA. Their findings supported the current findings that ethnic groups within a close geographical area could be more prone to being wrongly classified compared to populations that are located far from each other.

Previous research on sub-adults mostly utilised the geometric morphometrics method (GMM) as the primary method, as it is a robust statistical approach. It is often used to analyse variations in skeleton shapes among different populations. However, this approach requires an enormous sample size to ensure an accurate estimation (Noble et al., 2019). Using a large sample of children and adolescents can be challenging; therefore, acquiring measurements that are linear and easy to process would be useful. Furthermore, the accuracy obtained in the present study was not much different from that obtained using geometric methods in previous studies (Buck & Vidarsdottir, 2004; Smith et al., 2013). Consequently, it was recommended that forensic estimation studies focus on measurements of size and traditional inter-landmark distances, as they only need a modest number of individuals to produce accurate results (Noble et al., 2019).

5.3 Machine learning (ML) algorithm vs classical statistical methods (DFA and BLR)

Statistical methods have been employed more frequently in recent years, as they offer objective and robust approaches for sex and ethnicity estimation. Classical statistical methods, such as DFA and BLR, are widely used for sex and ethnicity classifications in forensic anthropology (Hsiao et al., 2010; Noble et al., 2019; O'Donnell et al., 2017; Smith et al., 2013; Sprowl, 2013; Teodoru-Raghina et al., 2017). However, BLR is restricted to sex classification, whereas DFA can be used for both sex and ethnicity classifications. ML algorithms recently emerged as a popular modeling approach that offers an alternative class of models with more computational flexibility (Steyerberg et al., 2014). ML is a subset of AI with the capacity to make predictions without being explicitly programmed to do so, using mathematical models generated from training data (Williams, 2011). Over the last few years, ML algorithms achieved significant success across a broad range of fields due to their superiority, such as their ability to model nonlinear relations and the accuracy of their overall predictions (Nikita & Nikitas, 2020). Nevertheless, most of the existing sex and ethnicity estimation models based on craniometrics in sub-adults are founded on classical statistical methods (Hsiao et al., 2010; Noble et al., 2019; O'Donnell et al., 2017; Sprowl, 2013; Teodoru-Raghina et al., 2017). Therefore, this study has compared the validity of sex and ethnicity estimation models that utilised ML algorithms (RF, SVM, and LDA) and classical statistical methods (DFA and BLR), with the aim of their applicability to the Malaysian sub-adults.

The key difference between ML and classical statistical methods lies in their objectives: the former emphasizes maximizing prediction accuracy, whereas the latter are geared towards deducing relationships between variables (Azzolina et al., 2019). ML methods have been reported to obtain a higher accuracy compared to classical methods, especially for complex data with a high-dimensional feature space (Pozzi et al., 2020; Bidmos et al., 2023). In the present study, ML methods obtained higher accuracy rates than classical methods for sex and ethnicity estimation using sub-adults' crania (Table 4.29). The high accuracies obtained by ML methods can be attributed to the ability of the model to capture complex class signatures, accept a variety of input predictor data, and operate without making assumptions about the data distribution. In a study by Bidmos et al. (2023), the authors evaluated their models' performance using classification accuracy and reported that ML methods achieved a significantly higher accuracy rate (90.77%) than DFA (81.9-84.2%). Similarly, Pozzi et al. (2020) reported that ML methods achieved higher accuracy (74-85.5%) than LDA (40%) in the Sardinia dataset. This demonstrates that the application of the ML algorithms enables a more accurate classification of sex (Pozzi et al., 2020).

This present study has demonstrated that RF achieved a higher accuracy in sex estimation of sub-adult's crania, followed by SVM and LDA (Table 4.27). However, previous studies in adult sex estimation reported contradictory results (Gao et al., 2018; Nikita & Nikitas, 2020; Toy et al., 2022). A study that explored sex differences using cranial and pelvic traits proposed that LDA has a distinct, albeit slight, edge over other ML methods (LR and ANN; Nikita & Nikitas, 2020). LDA has also shown superior performance when used with a skull measurement dataset compared with most ML methods (SVM, DT, and BP neural networks; Gao et al., 2018). In addition, LDA has reportedly obtained higher accuracy rates than DT and RF when estimating sex using cranium measurements (Toy et al., 2022). The different classification accuracies obtained by these studies can be attributed to several factors, such as population differences, sexual dimorphism expression, methodological methods used, and differences in statistical analyses.

This study provided the first methodological approach to estimate the sex and ethnicity of sub-adults' crania using different ML methods. In addition, this represents the first instance of a comparison between ML and classical methods for estimating sex and ethnicity in sub-adult crania. This present study has employed several performance metrics, including precision, recall, F1 score, and accuracy, have been employed to confirm the algorithms' reliability and results (Table 4.27 and Table 4.28). Furthermore, ML methods used in the present study were designed as 70% training and 30% test set. By allowing ML algorithms to learn from the training dataset, and assessing the models' performance using the testing dataset, this has increased the prediction reliability and value of the study in comparison with classical methods.

5.3.1 Advantages of computed tomography (CT) images

Advances in imaging techniques, such as CT, offer an anatomically precise characterisation of skeletal architecture that surpasses conventional morphometric practices (Franklin et al., 2013a). The MSCT scan used in this study has proven to be a suitable approach for collecting data, and a viable alternative for dry bone analysis. Despite the ambiguity of virtual model measurements that are prone to errors when compared with dry bone measurements, recent investigations have validated that the measurement error of virtual bones of sub-adult cranium was ≤ 2 mm, which showed high reliability (Corron et al., 2022; McIntosh et al., 2020). The overall differences between virtual cranial measurements and dry bone measurements were generally small and negligible at an average of 1.5% (Franklin et al., 2013a). In addition, the conversion from the DICOM files to 3D models and their subsequent postprocessing was found to be appropriate for craniometric purposes since it has negligible effect on the overall geometry of the cranium (Bertsatos et al., 2020). Hence, CT scans' skull measurements were relatively comparable to direct dry skull measurements. Data derived from contemporary individuals should be used to support the identification process in present-day forensic cases. In this regard, medical imaging allows samples of contemporary populations to be studied and helps to keep forensic standards up to date.

5.4 Cephalic index (CI)

CI is a valuable measurement for determining the shape of a skull. It plays a crucial role when comparing individuals of different ages, sexes, and ethnic backgrounds in forensic and clinical contexts. CI allows for a quantitative assessment of cranial characteristics and can provide insights into population-specific variations. Nonetheless, a single standard for CI may not be applicable or accurate for populations worldwide. By considering population-specific CI, forensic and clinical practitioners can enhance the accuracy of sex and ethnicity estimation and improve the precision of craniofacial comparisons across diverse populations. This helps to account for the inherent variations in cranial morphology observed among different ethnic groups and aids in making more reliable and appropriate assessments.

The present study demonstrated that skull growth undergoes accelerated development during the first two years of life, as shown by the high CI value of 86.31±5.37 (Table 4.30). However, it was found that the rate of change in CI gradually slows down after the initial three years of life, leading to a decrease in CI value of 84.99±4.71. Subsequently, growth velocity further decreases and eventually reaches a plateau from six years of age into adulthood. These findings indicated that the most significant changes in cranial proportions occur during early infancy, providing valuable insights into cranial proportions and growth patterns. Moreover, as growth decelerates from early childhood to adulthood, CI has proven to be a reliable indicator for monitoring skull development in both females and males.

In the present study, females are found to have a higher CI value (83.91 ± 4.63) than males (84.13 ± 4.72; Table 4.30). This showed that female skulls were narrower and shorter than male skulls, possibly resulting from the effects and interactions between growth and sex hormones. This observation indicated that the CI could be greater in either sex, depending on the characteristics of the population under study. Differences in skull morphology across the sexes underscored the significance of customising assessments of anatomical variations to specific individuals within a group. Since the development of an individual's skeleton is influenced by hormones, dietary status, cultural variances, and environmental conditions, males and females will acquire skeletal maturity at different ages (Norris & Carr, 2020). Therefore, it is an essential feature to consider when comparing the sexes.

The findings of the present study demonstrated that the CI values of Malaysian sub-adults' exhibit variations across diverse populations. Therefore, according to Cohen and MacLean (2000) and Standring and Gray (2008), a higher frequency of hyperbrachycephalic head shape in the Malaysian sub-adult population was observed when applying the skull shape classification. This high frequency was due to the percentages used for cranial classification derived from European populations, where the Caucasian race predominates and exhibits physical characteristics distinct from other populations (Halazonetis, 2007). This observation suggested that the current classification system might not accurately represent the diversity of cranial shapes observed in non-European populations (Halazonetis, 2007). The CI classifications proposed by Koizumi et al. (2010) and Nam et al. (2021) may not be suitable for application in the Malaysian context due to the lack of inclusion in the country's multi-ethnic diversity. Therefore, the proposed Malaysian CI classification in the present study is more suitable for capturing the unique cranial proportions and growth patterns presented in Malaysia's multi-ethnic population (Table 4.31).

Differences in CI values among populations have been reported in various studies (Haas, 1952; Koizumi et al., 2010; Likus et al., 2014; Nam et al., 2021; Pereira et al., 2008; Waitzman et al., 1992; Table 4.32). The findings of the present study implied that Malaysians have relatively higher CI values compared to Caucasians (Waitzman et al., 1992), Brazilians (Haas, 1952), Europeans (Pereira et al., 2008), and Polish (Likus et al., 2014) populations. While differences in CI values were also observed in Asian populations, the extent of these differences was generally found to be smaller than in Caucasian populations. Specifically, the CI values of the Japanese (Koizumi et al., 2010) were relatively similar to those of Malaysians, whereas Koreans (Nam et al., 2021) had higher CI values than Malaysians. These differences proved that cultural differences could have a more significant impact on the shape of Asian skulls.

The differences in CI value were also reported within the Malaysian population. The present study observed higher CI values (males: 83.91, females: 84.13) than a previous study conducted on Malaysians in the age group of 7-17 years (males: 81.55, females: 79; Swamy et al., 2013). In contrast, higher CI was reported by a previous study of Malaysian sub-adults in the age group of 0-25 years (males: 84.8, females: 85.2; Yusof, 2007). These differences could be due to variations in CI values across studies that can be influenced by several factors such as sample size, demographics, measurement techniques, and the composition of the population sample. Previous studies had exclusively focused on Malay individuals and thus may not have captured the full diversity and potential variations in CI value that may occur across different ethnic groups (Swamy et al., 2013; Yusof, 2007). In contrast, the present study had advantages such as larger sample size, greater ethnic diversity, and improved statistical analyses, allowing for a more comprehensive investigation of CI values in the Malaysian population.

Using the proposed Malaysian CI developed in this study, a greater proportion of Malaysians presented as mesocephalic (66.4%), followed by dolichocephalic (18.4%; Figure 4.5). Similar patterns of more prevalent mesocephalic cranial types have been observed in other populations, including the inhabitants of Southern Iran (Golalipour et al., 2005), Poland (Likus et al., 2014), Europe (Haas, 1952), Korea (Nam et al., 2021), and Siberia (Cvetković et al., 2014). Meanwhile, brachycephalic skulls are more common in Iranians (Golalipour, 2006a) and Japanese (Koizumi et al., 2010), and dolichocephalic heads are more prevalent in Africa (Eroje et al., 2010) and Colombia (Torres-Restrepo et al., 2014). In India, different regions presented with varying shapes of head, with dolichocephalic skulls being more common in North India (Khanduri et al., 2021), while brachycephalic skulls are found to be typical in East India (Ghosh, 2018). However, in Central India, male head shapes are more often mesocephalic, while female head shapes are more commonly brachycephalic (Yagain et al., 2012; Table 2.4). The different head shapes in India suggested that CI values vary substantially between locations. The heads of tropical inhabitants are longer and narrower than their temperate counterparts. In addition, the dietary patterns of a population could also influence the craniofacial anatomy of its people (Kasai et al., 1993).

The variation in head shape observed between younger and older sub-adults within the Malaysian population is likely influenced by a combination of genetic, developmental, environmental, and hormonal factors (Figure 4.6). Throughout childhood and early adolescence, there is substantial growth and development of the brain, which impacts the shape of the skull to accommodate the expanding brain (Humphrey, 1998). At puberty, hormonal changes trigger the release of growth hormones and sex hormones, which are pivotal not only for overall growth and development but also for the development of secondary sexual characteristics and the acceleration of skeletal maturation (İşcan, 2005; Klales & Burns, 2017; Scheuer & Black, 2000; Roche & Sun, 2003). These factors interact dynamically to shape cranial morphology throughout the stages of younger and older sub-adults.

Cranial measurements obtained in this study revealed that some CI values were accurate markers for skull development and can be used to monitor growth throughout time. These values could enable a thorough evaluation of abnormally tiny or large heads. In addition, the CI is a coded measure of cranial capacity that is assumed to be able to measure brain volume and predict cognitive ability (Maina et al., 2011). A higher CI value reflects a faster skull growth rate in the first year of life, indicating that this phase of rapid development is essential for appropriate brain maturation. After the first three years of life, the rate of change in cranial diameter will stabilise. As the pace of growth decreases and stabilises past this age, while the statistical significance of measurements diminishes, cranial size or tables with skull diameters could become reliable indicators of skull development for both sexes.

5.5 Limitations

Several limitations have been identified in the present study. First, this study lacked an equally distributed sample across age, sex, and ethnicity. This limitation is inherent in any study using MSCT scans, especially in studies involving sub-adults. Given the radiation risk to patients, it was challenging to obtain MSCT scans with an appropriate level of resolution and the correct landmarks. In terms of ethnicity, the Malaysian population comprises a mix of cultures and inter-racial marriages. Thus, the exact ethnic origin of subjects in inter-racial marriages was challenging to trace. Furthermore, individuals often identify their ethnicity as a cultural element rather than attributing it to population affinity.

Second, the socioeconomic status of individuals included in this study was unknown. Thus, it was assumed that all individuals have access to healthcare. However, socioeconomic status, nutrition, and stress are factors that should be considered, as they can affect skeletal growth and development (Schmeling et al., 2005). Third, the sample for this study came from a very specific part of the Malaysian population, which might not be fully representative of the entire population and could limit the generalisability of the findings to other regions or ethnic groups within Malaysia. Thus, it is vital to acknowledge the regional specificity of the sample and to interpret the results with caution when applying them to broader populations.

Fourth, the use of a cross-sectional sample in this study could be viewed as a limitation. Longitudinal studies are preferred because they can reveal accurate growth patterns of individuals, as this type of study can collect repeated measurements of the same individuals over time. However, longitudinal data can take many years to manage, which makes longitudinal studies not cost-effective. It is also impossible to conduct a longitudinal study on a sample large enough to represent a population accurately.

5.6 Future research

The present study has presented valuable findings that lay the groundwork for future research in the development of sex and ethnicity identification methods. Expanding the research to include measurements of the pelvis, long bones, and mandibles in addition to the crania, can indeed provide a more comprehensive understanding of sexual dimorphism and ethnicity identification. The pelvis, long bones, and mandibles have been widely recognised as reliable indicators of sex, often exhibiting greater sexual dimorphism than the cranium alone. Including these skeletal elements in the analysis could potentially improve the success rates of sex and ethnicity estimation. Furthermore, comparing the accuracy rates between cranial and postcranial measurements can yield insights into which skeletal regions are more reliable for sex and ethnicity identification. This information would be valuable for forensic practitioners and anthropologists working with incomplete or fragmentary remains, allowing them to prioritise certain skeletal elements in their analysis.

To assess the generalisability of the proposed technique, future studies should increase the number of individuals from various geographic regions in Malaysia. This can help researchers to identify population-specific patterns and develop more accurate models for sex and ethnicity identification within each region. In addition, it helps to reduce demographic composition bias and ensures that the models are reliable and valid across diverse population groups. Future studies should also include the adult population to ensure that the research captures craniofacial sexual dimorphism and ethnic characteristics that are fully developed and stable. This can contribute towards providing more accurate reference standards and predictive models specific to adult individuals.

Finally, the integration of deep learning methods in sex and ethnicity estimation for sub-adult skeletal remains holds great promise in advancing the field of forensic anthropology. With the development of large and diverse datasets, deep learning models have the potential to provide more accurate, efficient, and objective estimations.

CHAPTER 6: CONCLUSION

The availability of modified standards that offer reliable sex and ethnicity estimation, with a high and quantified degree of accuracy, is essential to biological and forensic anthropology. However, research pertaining to sex and ethnicity estimation using cranium in the Malaysian sub-adult population is lacking. The most sensitive method that can be adopted to address this issue is by acquiring cranial measurements from the MSCT dataset. Data obtained from contemporary individuals should be used to support identification processes in present-day forensic cases. In this regard, medical imaging methods, such as CT, allow samples of contemporary populations to be studied and forensic standards to be up to date. Metric estimations of sex and ethnicity are more robust when multivariate techniques are employed, such as a classification function derived from classical methods like DFA or BLR. Nevertheless, the focus has shifted towards utilising ML classification algorithms. Based on the methods tested in this study, ML models were able to provide greater classification accuracies compared to classical models for estimating both sex and ethnicity. Furthermore, the SNA estimator version 1.0 allows forensic anthropologists to utilise the developed models and it also offers measures of a model's success, which are essential for quantifying the obtained results.

CI values tend to show sex and ethnicity differences. In the present study, the mean CI values were found to be higher for females than males. Also, the CI of the Malaysian population was higher than most CI values of other populations worldwide, such as Europeans, Nigerians, Caucasians, and Iranians, but lower than most Asian populations, such as Japanese and Koreans. Hence, the present CI classification was proposed for the Malaysian population ranging in age from birth to 20 years. Based on the proposed classification, the dominating head type for the Malaysian population was mesocephalic (66.4%), followed by dolichocephalic (18.4%), and brachycephalic (12.3%). Hyperbrachycephalic was the least frequently observed category (2.9%). The CI of the

Malaysian sub-adults reported in this study will provide a valuable reference to diagnose and plan for surgery of cranial deformities.

REFERENCES

Aghakhanian, F., Yunus, Y., Naidu, R., Jinam, T., Manica, A., Hoh, B. P., & Phipps, M. E. (2015). Unravelling the genetic history of Negritos and indigenous populations of Southeast Asia. *Genome Biology and Evolution*, *7*(5), 1206–1215. https://doi.org/10.1093/gbe/evv065

Ahmed, A. A., Koko, A. O., & Elsayed, M. M. (2021). Multi-detector computed tomography for the estimation of sex from sternal medullary cavity measurements in Sudanese individuals. *Forensic Science International: Reports, 4*, 100244. https://doi.org/10.1016/j.fsir.2021.100244

Akinbami, B. O. (2014). Measurement of cephalic indices in older children and adolescents of a Nigerian population. *BioMed Research International*, *2014*, 527473. https://doi.org/10.1155/2014/527473

Algee-Hewitt, B. F. (2016). Population inference from contemporary American craniometrics. *American Journal of Physical Anthropology*, *160*(4), 604–624. https://doi.org/10.1002/ajpa.22959

Amin, W., & Othman, D. (2014). Gender dimorphism of crania in young Jordanian adults: Discriminant function analysis approach for gender prediction. *Journal of Current Surgery, 4*(3), 76–85. https://doi.org/10.14740/jcs241w

Azzolina, D., Baldi, I., Barbati, G., Berchialla, P., Bottigliengo, D., Bucci, A., Calza, S., Dolce, P., Edefonti, V., Faragalli, A., Fiorito, G., Gandin, I., Giudici, F., Gregori, D., Gregorio, C., Ieva, F., Lanera, C., Lorenzoni, G., Marchioni, M., … Vezzoli, M. (2019). Machine learning in clinical and epidemiological research: Isn't it time for biostatisticians to work on it? *Epidemiology, Biostatistics, and Public Health*, *16*(4), e13245-1–e13245-3. https://doi.org/10.2427/13245.

Baker, J. (2020). *Crossroads: A popular history of Malaysia and Singapore*. Marshall Cavendish International Asia Pte Ltd.

Barbeito-Andrés, J., Anzelmo, M., Ventrice, F., & Sardi, M. L. (2012). Measurement error of 3D cranial landmarks of an ontogenetic sample using Computed Tomography. *Journal of Oral Biology and Craniofacial Research*, *2*(2), 77–82. https://doi.org/10.1016/j.jobcr.2012.05.005

Bašić, Ž., Kružić, I., Jerković, I., Anđelinović, D., & Anđelinović, Š. (2017). Sex estimation standards for medieval and contemporary Croats. *Croatian Medical Journal*, *58*(3), 222–230. https://doi.org/10.3325/cmj.2017.58.222

Bass, W. M. (1987). *Human osteology: A laboratory and field manual* (Vol. 2). Missouri Archaeological Society.

Bastir, M., Rosas, A., & O'Higgins, P. (2006). Craniofacial levels and the morphological maturation of the human skull. *Journal of Anatomy*, *209*(5), 637–654. https://doi.org/10.1111/j.1469-7580.2006.00644.x

Baughan, B., & Demirjian, A. (1978). Sexual dimorphism in the growth of the cranium. *American Journal of Physical Anthropology,* *49*(3), 383–390. https://doi.org/10.1002/ajpa.1330490311

Beederman, M., Farina, E. M., & Reid, R. R. (2014). Molecular basis of cranial suture biology and disease: Osteoblastic and osteoclastic perspectives. *Genes & Diseases*, *1*(1), 120–125. https://doi.org/10.1016/j.gendis.2014.07.004

Belle, C. V. (2014). *Tragic orphans: Indians in Malaysia* (Vol. 496). Institute of Southeast Asian Studies.

Berg, G. E., & Kenyhercz, M. W. (2017). Introducing human mandible identification [(hu) MANid]: A free, web-based GUI to classify human mandibles. *Journal of Forensic Sciences*, *62*(6), 1592–1598. https://doi.org/10.1111/1556-4029.13479

Bertsatos, A., Chovalopoulou, M. E., Brůžek, J., & Bejdová, Š. (2020). Advanced procedures for skull sex estimation using sexually dimorphic morphometric features. *International Journal of Legal Medicine*, *134*(5), 1927–1937. https://doi.org/10.1007/s00414-020-02334-9

Bewes, J., Low, A., Morphett, A., Pate, F. D., & Henneberg, M. (2019). Artificial intelligence for sex determination of skeletal remains: Application of a deep learning artificial neural network to human skulls. *Journal of Forensic and Legal Medicine*, *62*, 40–43. https://doi.org/10.1016/j.jflm.2019.01.004

Bharati, S. P., Som, S., Bharati, P., & Vasulu, T. S. (2001). Climate and head form in India. *American Journal of Human Biology*, *13*(5), 626–634. https://doi.org/10.1002/ajhb.1101

Bidmos, M. A., Olateju, O. I., Latiff, S., Rahman, T., & Chowdhury, M. E. H. (2023). Machine learning and discriminant function analysis in the formulation of generic models for sex prediction using patella measurements. *International Journal of Legal Medicine*, *137*(2), 471–485. https://doi.org/10.1007/s00414-022-02899-7

Birkby, W. H., Fenton, T. W., & Anderson, B. E. (2008). Identifying Southwest Hispanics using nonmetric traits and the cultural profile. *Journal of Forensic Sciences*, *53*(1), 29–33. https://doi.org/10.1111/j.1556-4029.2007.00611.x

Black, S., & Maat, G. (2010). Principles of physical age estimation. In S. Black, A. Aggrawal & J. Payne-James (Eds.), *Age estimation in the living: The practioner's guide* (pp. 77–94). John Wiley & Sons Inc.

Bogin, B. (2020). *Patterns of human growth* (3rd ed.). Cambridge University Press. https://doi.org/10.1017/978110837997

Bontrager, K. L., & Lampignano, J. (2001). *Radiographic positioning and related anatomy* (5th ed., Vol. 1). Mosby Incorporated.

Bookstein, F. L. (1991). *Morphometric tools for landmark data: Geometry and biology*. Cambridge University Press.

Braga, J., & Treil, J. (2007). Estimation of pediatric skeletal age using geometric morphometrics and three-dimensional cranial size changes. *International Journal of Legal Medicine*, *121*(6), 439–443. https://doi.org/10.1007/s00414-007-0170-x

Buck, T. J., & Vidarsdottir, U. S. (2004). A proposed method for the identification of race in sub-adult skeletons: a geometric morphometric analysis of mandibular morphology. *Journal of Forensic Sciences*, *49*(6), 1159–1164. https://doi.org/10.1520/JFS2004074

Bulygina, E., Mitteroecker, P., & Aiello, L. (2006). Ontogeny of facial dimorphism and patterns of individual development within one human population. *American Journal of Physical Anthropology*, *131*(3), 432–443. https://doi.org/10.1002/ajpa.20317

Buretić-Tomljanović, A., Ostojić, S., & Kapović, M. (2006). Secular change of craniofacial measures in Croatian younger adults. *American Journal of Human Biology*, *18*(5), 668–675. https://doi.org/10.1002/ajhb.20536

Buretić-Tomljanović, A., Ristić, S., Brajenović-Milić, B., Ostojić, S., Gombac, E., & Kapović, M. (2004). Secular change in body height and cephalic index of Croatian medical students (University of Rijeka). *American Journal of Physical Anthropology*, *123*(1), 91–96. https://doi.org/10.1002/ajpa.10306

Byers, S. N. (2016). *Introduction to forensic anthropology* (5th ed.). Routledge.

Byrne, B. M. (2010). *Structural equation modeling with AMOS: Basic concepts, applications, and programming* (2nd ed.). Routledge. https://doi.org/10.4324/9780203805534

Cardoso, H. F. (2008). Sample-specific (universal) metric approaches for determining the sex of immature human skeletal remains using permanent tooth dimensions. *Journal of Archaeological Science*, *35*(1), 158–168. https://doi.org/10.1016/j.jas.2007.02.013

Charisi, D., Eliopoulos, C., Vanna, V., Koilias, C. G., & Manolis, S. K. (2011). Sexual dimorphism of the arm bones in a modern greek population. *Journal of Forensic Sciences*, *56*(1), 10–18. https://doi.org/10.1111/j.1556-4029.2010.01538.x

Christensen, A., & Passalacqua, N. V. (2018). *A laboratory manual for forensic anthropology* (1st ed.). Academic Press. https://www.sciencedirect.com/book/9780128122013/a-laboratory-manual-for-forensic-anthropology

Cohen, M. M., & MacLean, R. E. (2000). *Craniosynostosis: Diagnosis, evaluation, and management* (2nd ed.). Oxford University Press.

Comas, D., Calafell, F., Mateu, E., Pérez-Lezaun, A., Bosch, E., Martínez-Arias, R., Clarimon, J., Facchini, F., Fiori, G., Luiselli, D., Pettener, D., & Bertranpetit, J. (1998). Trading genes along the silk road: mtDNA sequences and the origin of central Asian populations. *American Journal of Human Genetics*, *63*(6), 1824–1838. https://doi.org/10.1086/302133

Cormack, A. M. (1980). Early two-dimensional reconstruction and recent topics stemming from it. *Science*, *209*(4464), 1482–1486. https://doi.org/10.1126/science.6254142

Corner, B. D., & Richtsmeier, J. T. (1991). Morphometric analysis of craniofacial growth in *Cebus apella*. *American Journal of Physical Anthropology*, *84*(3), 323–342. https://doi.org/10.1002/ajpa.1330840308

Corron, L. K., Broehl, K. A., Chu, E. Y., Vlemincq-Mendieta, T., Wolfe, C. A., Pilloud, M. A., Scott, G. R., Spradley, M. K., & Stull, K. E. (2022). Agreement and error rates associated with standardized data collection protocols for skeletal and dental data on 3D virtual subadult crania. *Forensic Science International*, *334*, 111272. https://doi.org/10.1016/j.forsciint.2022.111272

Cunha, E., & Ubelaker, D. H. (2019). Evaluation of ancestry from human skeletal remains: a concise review. *Forensic Sciences Research*, *5*(2), 89–97. https://doi.org/10.1080/20961790.2019.1697060

Cunningham, C., Scheuer, L., & Black, S. (2016). *Developmental juvenile osteology* (2nd ed.). Academic Press.

Cvetković, M., Najman, S., & Nikolić, M. (2014). Secular changes in cephalic index-a study of Serbian school children. *Genetika*, *46*(2), 561–568. https://doi.org/10.2298/GENSR1402561C

D'Oliveira Coelho, J., & Curate, F. (2019). CADOES: An interactive machine-learning approach for sex estimation with the pelvis. *Forensic Science International*, *302*, 109873. https://doi.org/10.1016/j.forsciint.2019.109873

Daviran, M., Maghsoudi, A., Ghezelbash, R., & Pradhan, B. (2021). A new strategy for spatial predictive mapping of mineral prospectivity: Automated hyperparameter tuning of random forest approach. *Computers & Geosciences*, *148*, 104688. https://doi.org/10.1016/j.cageo.2021.104688

Decker, S. J., Davy-Jow, S. L., Ford, J. M., & Hilbelink, D. R. (2011). Virtual determination of sex: Metric and nonmetric traits of the adult pelvis from 3D computed tomography models. *Journal of Forensic Sciences*, *56*(5), 1107–1114. https://doi.org/10.1111/j.1556-4029.2011.01803.x

Deng, L., Hoh, B. P., Lu, D., Fu, R., Phipps, M. E., Li, S., Nur-Shafawati, A. R., Hatin, W. I., Ismail, E., Mokhtar, S. S., Jin, L., Zilfalil, B. A., Marshall, C. R., Scherer, S. W., Al-Mulla, F., & Xu, S. (2014). The population genomic landscape of human genetic structure, admixture history and local adaptation in Peninsular Malaysia. *Human Genetics*, *133*(9), 1169–1185. https://doi.org/10.1007/s00439-014-1459-8

Deng, L., Hoh, B. P., Lu, D., Saw, W. Y., Twee-Hee Ong, R., Kasturiratne, A., de Silva, H. J., Zilfalil, B. A., Kato, N., Wickremasinghe, A. R., Teo, Y. Y., & Xu, S. (2015). Dissecting the genetic structure and admixture of four geographical Malay populations. *Scientific Reports*, *5*, 14375. https://doi.org/10.1038/srep14375

Dias, M. S., Samson, T., Rizk, E. B., Governale, L. S., & Richtsmeier, J. T. (2020). Identifying the misshapen head: Craniosynostosis and related disorders. *Pediatrics*, *146*(3). https://doi.org/10.1542/peds.2020-015511

DiGangi, E. A., & Hefner, J. T. (2013). Chapter 5 – Ancestry estimation. *Research Methods in Human Skeletal Biology*, 117–149. https://doi.org/10.1016/B978-0-12-385189-5.00005-4

Dillon, A. (2014). *Cranial sexual dimorphism and the population specificity of anthropological standards* [Master's thesis, University of Western Australia]. https://api.research-repository.uwa.edu.au/ws/portalfiles/portal/3367364/Dillon_Alexandra_2014.pdf

Divakar, D. D., John, J., Al Kheraif, A. A., Mavinapalla, S., Ramakrishnaiah, R., Vellappally, S., Hashem, M. I., Dalati, M. H., Durgesh, B. H., Safadi, R. A., & Anil, S. (2016). Sex determination using discriminant function analysis in Indigenous (Kurubas) children and adolescents of Coorg, Karnataka, India: A lateral cephalometric study. *Saudi Journal of Biological Sciences*, *23*(6), 782–788. https://doi.org/10.1016/j.sjbs.2016.05.008

du Jardin, P., Ponsaillé, J., Alunni-Perret, V., & Quatrehomme, G. (2009). A comparison between neural network and other metric methods to determine sex from the upper femur in a modern French population. *Forensic Science International*, *192*(1–3), 127.e1–127.e1276. https://doi.org/10.1016/j.forsciint.2009.07.014

Dudzik, B., & Jantz, R. L. (2016). Misclassifications of Hispanics using Fordisc 3.1: Comparing cranial morphology in Asian and Hispanic populations. *Journal of Forensic Sciences*, *61*(5), 1311–1318. https://doi.org/10.1111/1556-4029.13123

Ekizoglu, O., Inci, E., Palabiyik, F. B., Can, I. O., Er, A., Bozdag, M., Kacmaz, I. E., & Kranioti, E. F. (2017). Sex estimation in a contemporary Turkish population based on CT scans of the calcaneus. *Forensic Science International*, *279*, 310.e1–310.e6. https://doi.org/10.1016/j.forsciint.2017.07.038

Elliott, M., & Collard, M. (2009). FORDISC and the determination of ancestry from cranial measurements. *Biology Letters*, *5*(6), 849–852. https://doi.org/10.1098/rsbl.2009.0462

Enlow, D. H. (1990). *Facial growth* (3rd ed.). WB Saunders.

Eroje, M. A., Fawehinmi, H. B., Jaja, B. N., & Yaakor, L. (2010). Cephalic index of Ogbia tribe of Bayesla state. *International Journal of Morphology*, *28*(2), 389–392. https://www.scielo.cl/pdf/ijmorphol/v28n2/art09.pdf

Farhi, M., Tomas, N., Marion, I., Avenetti, D., Atsawasuwan, P., da Fonseca, M., Al Atabbe, M., & Nicholas, C. L. (2023). Evaluation of the (hu)MANid program for sex and ancestry estimation in a diverse, contemporary CT scan-based sample. *Journal of Forensic Sciences*, *68*(1), 242–251. https://doi.org/10.1111/1556-4029.15176

Farkas, L. G., Posnick, J. C., & Hreczko, T. M. (1992). Anthropometric growth study of the head. *The Cleft Palate-craniofacial Journal*, *29*(4), 303–308. https://doi.org/10.1597/1545-1569_1992_029_0303_agsoth_2.3.co_2

Faulkner, F. T., & Tanner, J. M. (1986). *Human growth: A comprehensive treatise* (Vol. 2 – Postnatal growth, neurobiology). Plenum Press.

Fernandes, L., Bento, M. I. C., Rabello, P. M., Soriano, E. P., Navega, D., Daruge Jr., E., & Cunha, E. (2021). Analysis of the accuracy of ancestrees software in ancestry estimation in Brazilian identified sample. *Advances in Anthropology*, *11*(02), 163–178. https://doi.org/10.4236/aa.2021.112011.

Fix, A. G. (1995). Malayan paleosociology: Implications for patterns of genetic variation among the Orang Asli. *American Anthropologist*, *97*(2), 313–323. https://doi.org/10.1525/aa.1995.97.2.02a00090

Franklin, D., Cardini, A., Flavel, A., & Kuliukas, A. (2012). The application of traditional and geometric morphometric analyses for forensic quantification of sexual dimorphism: Preliminary investigations in a Western Australian population. *International Journal of Legal Medicine*, *126*(4), 549–558. https://doi.org/10.1007/s00414-012-0684-8

Franklin, D., Cardini, A., Flavel, A., & Kuliukas, A. (2013a). Estimation of sex from cranial measurements in a Western Australian population. *Forensic Science International*, *229*(1–3), 158. e151–158. e158. https://doi.org/10.1016/j.forsciint.2013.03.005

Franklin, D., Cardini, A., Flavel, A., Kuliukas, A., Marks, M. K., Hart, R., Oxnard, C., & O'Higgins, P. (2013b). Concordance of traditional osteometric and volume-rendered MSCT interlandmark cranial measurements. *International Journal of Legal Medicine*, *127*(2), 505–520. https://doi.org/10.1007/s00414-012-0772-9

Galdames, I. C. S., Matamala, D. A. Z., & Smith, R. L. (2008). Blind test of mandibular morphology with sex indicator in subadult mandibles. *International Journal of Morphology*, *26*(4), 845–848. http://dx.doi.org/10.4067/S0717-95022008000400011

Gao, H., Geng, G., & Yang, W. (2018). Sex determination of 3D skull based on a novel unsupervised learning method. *Computational and Mathematical Methods in Medicine*, *2018*, Article ID 4567267. https://doi.org/10.1155/2018/4567267

Garnett, S. P., Högler, W., Blades, B., Baur, L. A., Peat, J., Lee, J., & Cowell, C. T. (2004). Relation between hormones and body composition, including bone, in prepubertal children. *The American Journal of Clinical Nutrition*, *80*(4), 966–972. https://doi.org/10.1093/ajcn/80.4.966

Gaži-Čoklica, V., Muretić, Z., Brcić, R., Kern, J., & Milicić, A. (1997). Craniofacial parameters during growth from the deciduous to permanent dentition – A longitudinal study. *European Journal of Orthodontics*, *19*(6), 681–689. https://doi.org/10.1093/ejo/19.6.681

Ghezelbash, R., Maghsoudi, A., Bigdeli, A., & Carranza, E. J. M. (2021). Regional-scale mineral prospectivity mapping: Support vector machines and an improved data-driven multi-criteria decision-making technique. *Natural Resources Research*, *30,* 1977–2005. https://doi.org/10.1007/s11053-021-09842-4

Ghosh, R. (2018). A study of Cephalic index among the young age group of West Bengal in relation to sex and geographical factors. *Indian Journal of Basic and Applied Medical Research*, *7*(3), 239–245. https://www.ijbamr.com/assets/images/issues/pdf/June%202018%20239%20-%20245.pdf.pdf

Gill, G. W., Hughes, S. S., Bennett, S. M., & Gilbert, B. M. (1988). Racial identification from the midfacial skeleton with special reference to American Indians and whites. *Journal of Forensic Sciences*, *33*(1), 92–99.

Gillet, C., Costa-Mendes, L., Rérolle, C., Telmon, N., Maret, D., & Savall, F. (2020). Sex estimation in the cranium and mandible: A multislice computed tomography (MSCT) study using anthropometric and geometric morphometry methods. *International Journal of Legal Medicine*, *134*(2), 823–832. https://doi.org/10.1007/s00414-019-02203-0

Golalipour, M. J. (2006a). The effect of ethnic factor on cephalic index in 17-20 years old females of North of Iran. *International Journal of Morphology*, *24*(3), 319–322. https://pdfs.semanticscholar.org/88ab/67209fafe093cd206eb9b9f531797b335be4.pdf

Golalipour, M. J. (2006b). The variation of head shapes in 17-20 years old native Fars male in Gorgan-North of Iran. *International Journal of Morphology*, *24*(2), 187–190. https://www.scielo.cl/pdf/ijmorphol/v24n2/art11.pdf

Golalipour, M., Haidari, K., Jahanshahi, M., & Farahani, R. (2003). The shapes of head and face in normal male newborns in South-East of Caspian Sea (Iran-Gorgan). *Journal of Anatomical Society of India*, *52*(1), 28–31.

Golalipour, M., Jahanshahi, M., & Haidari, K. (2005). The variation of head and face shapes in female newborns in the South-East of the Caspian Sea (Iran-Gorgan). *European Journal of Anatomy*, *9*(2), 95–98. https://core.ac.uk/download/pdf/52205147.pdf

Gonzalez, P. N., Bernal, V., & Perez, S. I. (2011). Analysis of sexual dimorphism of craniofacial traits using geometric morphometric techniques. *International Journal of Osteoarchaeology*, *21*(1), 82–91. https://doi.org/10.1002/oa.1109

Gonzalez, P. N., Perez, S. I., & Bernal, V. (2010). Ontogeny of robusticity of craniofacial traits in modern humans: A study of South American populations. *American Journal of Physical Anthropology*, *142*(3), 367–379. https://doi.org/10.1002/ajpa.21231

Gonzalez, R. A. (2012). Determination of sex from juvenile crania by means of discriminant function analysis. *Journal of Forensic Sciences*, *57*(1), 24–34. https://doi.org/10.1111/j.1556-4029.2011.01920.x

Grbeša, Đ., Pezerović-Panijan, R., Nadim Kalaya, M., Goršić, I., Čavčić, A., Žura, N., & Berberović, B. (2007). Craniofacial characteristics of Croatian and Syrian populations. *Collegium Antropologicum*, *31*(4), 1121–1125. https://hrcak.srce.hr/26980

Green, H., & Curnoe, D. (2009). Sexual dimorphism in Southeast Asian crania: A geometric morphometric approach. *Homo: Journal of Comparative Human Biology*, *60*(6), 517–534. https://doi.org/10.1016/j.jchb.2009.09.001

Guyomarc'h, P., & Bruzek, J. (2011). Accuracy and reliability in sex determination from skulls: a comparison of Fordisc® 3.0 and the discriminant function analysis. *Forensic Science International*, *208*(1–3), 180.e1–180.e1806. https://doi.org/10.1016/j.forsciint.2011.03.011

Haas, J., Buikstra, J. E., & Ubelaker, D. H., & Aftandilian, D. (1994). *Standards for data collection from human skeletal remains: Proceedings of a seminar at the Field Museum of Natural History* (Series 44). Arkansas Archaeological Survey Research.

Haas, L. L. (1952). Roentgenological skull measurements and their diagnostic applications. *The American Journal of Roentgenology, Radium Therapy, And Nuclear Medicine*, *67*(2), 197–209.

Hahs-Vaughn, D. L. (2016). *Applied multivariate statistical concepts* (1st ed.). Routledge. https://doi.org/10.4324/9781315816685

Halazonetis, D. J. (2007). Morphometric correlation between facial soft-tissue profile shape and skeletal pattern in children and adolescents. *American Journal of Orthodontics and Dentofacial Orthopedics*, *132*(4), 450–457. https://doi.org/10.1016/j.ajodo.2005.10.033

Hanifah, F.S., Wijayanto, H., & Kurnia, A. (2015). SMOTE Bagging Algorithm for Imbalanced Dataset in Logistic Regression Analysis (Case: Credit of Bank X). *Applied Mathematical Sciences*, *9*, 6857–6865. http://dx.doi.org/10.12988/ams.2015.58562

Hanihara, K., Kouchi, M., & Koizumi, K. (1982). A preliminary analysis of geographical variations in the male Japanese crania. *Journal of the Anthropological Society of Nippon*, *90*(Supplement), 139–152. https://doi.org/10.1537/ase1911.90.Supplement_139

Harris, E. F., & Smith, R. N. (2009). Accounting for measurement error: a critical but often overlooked process. *Archives of Oral Biology*, *54*(Suppl. 1), S107–S117. https://doi.org/10.1016/j.archoralbio.2008.04.010

Hatin, W. I., Nur-Shafawati, A. R., Zahri, M. K., Xu, S., Jin, L., Tan, S. G., Rizman-Idid, M., Zilfalil, B. A., & HUGO Pan-Asian SNP Consortium. (2011). Population genetic structure of peninsular Malaysia Malay sub-ethnic groups. *PloS One*, *6*(4), e18312. https://doi.org/10.1371/journal.pone.0018312

Hefner, J. T. (2009). Cranial nonmetric variation and estimating ancestry. *Journal of Forensic Sciences*, *54*(5), 985–995. https://doi.org/10.1111/j.1556-4029.2009.01118.x

Hefner, J. T., Pilloud, M. A., Black, C. J., & Anderson, B. E. (2015). Morphoscopic trait expression in "Hispanic" populations. *Journal of Forensic Sciences*, *60*(5), 1135–1139. https://doi.org/10.1111/1556-4029.12826

Hellenthal, G., Busby, G. B. J., Band, G., Wilson, J. F., Capelli, C., Falush, D., & Myers, S. (2014). A genetic atlas of human admixture history. *Science*, *343*(6172), 747–751. https://doi.org/10.1126/science.1243518

Hendee, W. R. (1989). Cross sectional medical imaging: a history. *Radiographics*, *9*(6), 1155–1180. https://doi.org/10.1148/radiographics.9.6.2685939

Herrera, M. D., & Tallman, S. D. (2019). Craniometric variation and ancestry estimation in two contemporary Caribbean populations. *Forensic Science International*, *305*, 110013. https://doi.org/10.1016/j.forsciint.2019.110013

Hildebolt, C. F., Vannier, M. W., & Knapp, R. H. (1990). Validation study of skull three-dimensional computerized tomography measurements. *American Journal of Physical Anthropology*, *82*(3), 283–294. https://doi.org/10.1002/ajpa.1330820307

Hill, C., Soares, P., Mormina, M., Macaulay, V., Meehan, W., Blackburn, J., Clarke, D., Raja, J. M., Ismail, P., Bulbeck, D., Oppenheimer, S., & Richards, M. (2006). Phylogeography and ethnogenesis of aboriginal Southeast Asians. *Molecular Biology and Evolution*, *23*(12), 2480–2491. https://doi.org/10.1093/molbev/msl124

Holland, T. D. (1986). Race determination of fragmentary crania by analysis of the cranial base. *Journal of Forensic Science*, *31*(2), 719–725.

Hossain, M. D., Lestrel, P. E., & Ohtsuki, F. (2005). Secular changes in head dimensions of Japanese adult male students over eight decades. *Homo: Internationale Zeitschrift fur die Vergleichende Forschung am Menschen*, *55*(3), 239–250. https://doi.org/10.1016/j.jchb.2003.04.001

Hounsfield, G. N. (1980). Computed medical imaging. *Science*, *210*(4465), 22–28. https://doi.org/10.1126/science.6997993

Howells, W. W. (1973). Cranial variation in man: A study by multivariate analysis of patterns of difference among recent human populations. *Papers of the Peabody Museum of American Archaeology and Ethnology* (Vol. 67). Peabody Museum of Archeology and Ethnology.

Hsiao, T. H., Tsai, S. M., Chou, S. T., Pan, J. Y., Tseng, Y. C., Chang, H. P., & Chen, H. S. (2010). Sex determination using discriminant function analysis in children and adolescents: A lateral cephalometric study. *International Journal of Legal Medicine*, *124*(2), 155–160. https://doi.org/10.1007/s00414-009-0412-1

Hughes, S., Wright, R. & Barry, M. (2005). Virtual reconstruction and morphological analysis of the cranium of an ancient Egyptian mummy. *Australasian Physics & Engineering Sciences in Medicine*, *28*, 122–127. https://doi.org/10.1007/BF03178703

Humphrey, L. T. (1998). Growth patterns in the modern human skeleton. *American Journal of Physical Anthropology*, *105*(1), 57–72. https://doi.org/10.1002/(SICI)1096-8644(199801)105:1<57::AID-AJPA6>3.0.CO;2-A

Humphries, A. L., Maxwell, A. B., Ross, A. H., & Ubelaker, D. H. (2015). A geometric morphometric study of regional craniofacial variation in Mexico. *International Journal of Osteoarchaeology*, *25*(6), 795-–04. https://doi.org/10.1002/oa.2345

Hunley, K. L., Healy, M. E., & Long, J. C. (2009). The global pattern of gene identity variation reveals a history of long-range migrations, bottlenecks, and local mate exchange: implications for biological race. *American Journal of Physical Anthropology*, *139*(1), 35–46. https://doi.org/10.1002/ajpa.20932

Hunt, D. R., & Albanese, J. (2005). History and demographic composition of the Robert J. Terry anatomical collection. *American Journal of Physical Anthropology*, *127*(4), 406–417. https://doi.org/10.1002/ajpa.20135

Imaizumi, K., Bermejo, E., Taniguchi, K., Ogawa, Y., Nagata, T., Kaga, K., Hayakawa, H., & Shiotani, S. (2020). Development of a sex estimation method for skulls using machine learning on three-dimensional shapes of skulls and skull parts. *Forensic Imaging*, *22*, 200393. https://doi.org/10.1016/j.fri.2020.200393

İşcan, M. Y. (2005). Forensic anthropology of sex and body size. *Forensic Science International*, *147*(2–3), 107–112. https://doi.org/10.1016/j.forsciint.2004.09.069

İşcan, M. Y., & Steyn, M. (1999). Craniometric determination of population affinity in South Africans. *International Journal of Legal Medicine*, *112*(2), 91–97. https://doi.org/10.1007/s004140050208

İşcan, M. Y., & Steyn, M. (2013). *The human skeleton in forensic medicine* (3rd ed.). Charles C. Thomas.

Jiang, X., Iseki, S., Maxson, R. E., Sucov, H. M., & Morriss-Kay, G. M. (2002). Tissue origins and interactions in the mammalian skull vault. *Developmental Biology*, *241*(1), 106–116. https://doi.org/10.1006/dbio.2001.0487

Joffe, T. H., Tarantal, A. F., Rice, K., Leland, M., Oerke, A. K., Rodeck, C., Geary, M., Hindmarsh, P., Wells, J. C., & Aiello, L. C. (2005). Fetal and infant head circumference sexual dimorphism in primates. *American Journal of Physical Anthropology*, *126*(1), 97–110. https://doi.org/10.1002/ajpa.20035

Kallenberger, L., & Pilbrow, V. (2012). Using CRANID to test the population affinity of known crania. *Journal of Anatomy*, *221*(5), 459–464. https://doi.org/10.1111/j.1469-7580.2012.01558.x

Kamdar, M. R., Gomez, R. A., & Ascherman, J. A. (2009). Intracranial volumes in a large series of healthy children. *Plastic and Reconstructive Surgery*, *124*(6), 2072–2075. https://doi.org/10.1097/PRS.0b013e3181bcefc4

Kanchan, T., Gupta, A., & Krishan, K. (2013). Estimation of sex from mastoid triangle - a craniometric analysis. *Journal of Forensic and Legal Medicine*, *20*(7), 855–860. https://doi.org/10.1016/j.jflm.2013.06.016

Kasai, K., Richards, L. C., & Brown, T. (1993). Comparative study of craniofacial morphology in Japanese and Australian aboriginal populations. *Human Biology*, *65*(5), 821–834.

Kesterke, M. J., Raffensperger, Z. D., Heike, C. L., Cunningham, M. L., Hecht, J. T., Kau, C. H., Nidey, N. L., Moreno, L. M., Wehby, G. L., Marazita, M. L., & Weinberg, S. M. (2016). Using the 3D Facial Norms Database to investigate craniofacial sexual dimorphism in healthy children, adolescents, and adults. *Biology of Sex Differences*, *7*, 23. https://doi.org/10.1186/s13293-016-0076-8

Khanduri, S., Malik, S., Khan, N., Patel, Y. D., Khan, A., Chawla, H., Singh, V., Gupta, A., Shaikh, J., & Siddiqui, S. (2021). Establishment of cephalic index using cranial parameters by computed tomography in a sampled North Indian population. *Cureus*, *13*(6), e15421. https://doi.org/10.7759/cureus.15421

Kim, J.-H. (2009). Estimating classification error rate: Repeated cross-validation, repeated hold-out and bootstrap. *Computational Statistics & Data Analysis*, *53*(11), 3735–3745. https://doi.org/10.1016/j.csda.2009.04.009

Kimmerle, E. H., Ross, A., & Slice, D. (2008). Sexual dimorphism in America: geometric morphometric analysis of the craniofacial region. *Journal of Forensic Sciences*, *53*(1), 54–57. https://doi.org/10.1111/j.1556-4029.2007.00627.x

King, R. E. (2015). *Estimating ancestry in South Africa: A comparison of geometric morphometrics and traditional craniometrics* [Master's thesis, Boston University]. https://open.bu.edu/handle/2144/13951

Klales, A. R., Ousley, S. D., & Passalacqua, N. V. (2020). Statistical approaches to sex estimation. In A. R. Klales (Ed.), *Sex estimation of the human skeleton* (pp. 203–217). Academic Press. https://doi.org/10.1016/B978-0-12-815767-1.00013-4

Klales, A. R., & Burns, T. L. (2017). Adapting and applying the Phenice (1969) Adult Morphological Sex Estimation Technique to subadults. *Journal of Forensic Sciences*, *62*(3), 747–752. https://doi.org/10.1111/1556-4029.13332

Klales, A. R., & Kenyhercz, M. W. (2015). Morphological assessment of ancestry using cranial macromorphoscopics. *Journal of Forensic Sciences*, *60*(1), 13–20. https://doi.org/10.1111/1556-4029.12563

Klecka, W. R., Iversen, G. R., & Klecka, W. R. (1980). *Discriminant analysis* (Vol. 19). Sage.

Klepinger, L. L. (2006). *Fundamentals of forensic anthropology*. John Wiley & Sons. https://doi.org/10.1002/0470007729

Koizumi, T., Komuro, Y., Hashizume, K., & Yanai, A. (2010). Cephalic index of Japanese children with normal brain development. *The Journal of Craniofacial Surgery*, *21*(5), 1434–1437. https://doi.org/10.1097/SCS.0b013e3181ecc2f3

Komar, D. A., & Buikstra, J. E. (2008). *Forensic anthropology: Contemporary theory and practice*. Oxford University Press.

Komar, D. A., & Grivas, C. (2008). Manufactured populations: what do contemporary reference skeletal collections represent? A comparative study using the Maxwell Museum documented collection. *American Journal of Physical Anthropology*, *137*(2), 224–233. https://doi.org/10.1002/ajpa.20858

Kotěrová, A., Velemínská, J., Dupej, J., Brzobohatá, H., Pilný, A., & Brůžek, J. (2017). Disregarding population specificity: its influence on the sex assessment methods from the tibia. *International Journal of Legal Medicine*, *131*(1), 251–261. https://doi.org/10.1007/s00414-016-1413-5

Kranioti, E. F., García-Donas, J. G., Can, I. O., & Ekizoglu, O. (2018). Ancestry estimation of three Mediterranean populations based on cranial metrics. *Forensic Science International*, *286*, 265.e1–265.e8. https://doi.org/10.1016/j.forsciint.2018.02.014

Krogman, W., & Iscan, M. (1962). *The human skeleton in forensic medicine*. Charles C. Thomas.

Kurachka, K. S., & Sukhoi, T. I. M. P. (2017, November). *Vertebrae detection in X-ray images based on deep convolutional neural networks* (pp. 194–196). 2017 IEEE

14th International Scientific Conference on Informatics. https://doi.org/10.1109/INFORMATICS.2017.8327245.

Langley, N. R., Jantz, L. M., Ousley, S. D., Jantz, R. L., & Milner, G. (2016). *Data collection procedures for forensic skeletal material 2.0*. Forensic Anthropology Center, University of Tennessee. https://fac.utk.edu/wp-content/uploads/2016/03/DCP20_webversion.pdf

Lejarraga, H. (2002). Growth in infancy and childhood: A pediatric approach. In N. Cameron (Ed.), *Human growth and development* (pp. 21–44), Academic Press.

Leonardi, R., Giudice, A. L., Isola, G., & Spampinato, C. (2021). Deep learning and computer vision: Two promising pillars, powering the future in orthodontics. *Seminars in Orthodontics*, *27*(2), 62–68. https://doi.org/10.1053/j.sodo.2021.05.002

Lewis, M. E. (2018). Children in bioarchaeology: Methods and interpretations. In M. A. Katzenberg & A. L. Grauer (Eds.), *Biological anthropology of the human skeleton* (3rd ed., pp. 117–144). https://doi.org/10.1002/9781119151647.ch4

Lewis, M. E., & Flavel, A. (2006). Age assessment of child skeletal remains in forensic contexts. In A. Schmitt, E. Cunha & J. Pinheiro (Eds.), *Forensic anthropology and medicine* (pp. 243–257). Humana Press. https://doi.org/10.1007/978-1-59745-099-7_10

Li, Y., Huang, Z., Dong, X., Liang, W., Xue, H., Zhang, L., Zhang, Y., & Deng, Z. (2019). Forensic age estimation for pelvic X-ray images using deep learning. *European Radiology*, *29*(5), 2322–2329. https://doi.org/10.1007/s00330-018-5791-6

Li, Z., Park, B. K., Liu, W., Zhang, J., Reed, M. P., Rupp, J. D., Hoff, C. N., & Hu, J. (2015). A statistical skull geometry model for children 0-3 years old. *PloS One*, *10*(5), e0127322. https://doi.org/10.1371/journal.pone.0127322

Lieberman, D. (2011). *The evolution of the human head*. Harvard University Press.

Lieberman, D. E., Pearson, O. M., & Mowbray, K. M. (2000). Basicranial influence on overall cranial shape. *Journal of Human Evolution*, *38*(2), 291–315. https://doi.org/10.1006/jhev.1999.0335

Lieberman, L. S. (1982). Normal and abnormal sexual dimorphic patterns of growth and development. In R. L. Hall (Ed.), *Sexual dimorphism in Homo sapiens. A question of size* (pp. 281–290). Praeger.

Likus, W., Bajor, G., Gruszczyńska, K., Baron, J., Markowski, J., Machnikowska-Sokołowska, M., Milka, D., & Lepich, T. (2014). Cephalic index in the first three years of life: study of children with normal brain development based on computed tomography. *The Scientific World Journal*, *2014*, 502836. https://doi.org/10.1155/2014/502836

Limson, M. (1932). Observations on the bones of the skull in white and negro fetuses and infants. *Contributions to embryology* (Vol. 23, pp. 205–222). Carnegie Institution of Washington.

Little, B. B., Buschang, P. H., Peña Reyes, M. E., Tan, S. K., & Malina, R. M. (2006). Craniofacial dimensions in children in rural Oaxaca, southern Mexico: Secular change, 1968-2000. *American Journal of Physical Anthropology*, *131*(1), 127–136. https://doi.org/10.1002/ajpa.20406

Lottering, N., MacGregor, D. M., Barry, M. D., Reynolds, M. S., & Gregory, L. S. (2014). Introducing standardized protocols for anthropological measurement of virtual subadult crania using computed tomography. *Journal of Forensic Radiology and Imaging*, *2*(1), 34–38. https://doi.org/10.1016/j.jofri.2013.11.005

Lundy, J. K. (1998). Forensic anthropology: What bones can tell us. *Laboratory Medicine*, *29*(7), 423–427, https://doi.org/10.1093/labmed/29.7.423

Macaluso, P. J., Jr. (2010). The efficacy of sternal measurements for sex estimation in South African blacks. *Forensic Science International*, *202*(1–3), 111.e1–111.e1117. https://doi.org/10.1016/j.forsciint.2010.07.019

Mahfouz, M., Badawi, A., Merkl, B., Fatah, E. E. A., Pritchard, E., Kesler, K., Moore, M., Jantz, R., & Jantz, L. (2007). Patella sex determination by 3D statistical shape models and nonlinear classifiers. *Forensic Science International*, *173*(2–3), 161-170. https://doi.org/10.1016/j.forsciint.2007.02.024

Maina, M., Shapu, Y., Garba, S., Muhammad, M., Garba, A., Yaro, A., & Omoniyi, O. (2011). Assessments of cranial capacities in a North-Eastern adult Nigerian population. *Journal of Applied Sciences*, *11*(14), 2662–2665. https://doi.org/10.3923/jas.2011.2662.2665

Mannucci, A., Sullivan, K. M., Ivanov, P. L., & Gill, P. (1994). Forensic application of a rapid and quantitative DNA sex test by amplification of the X-Y homologous gene amelogenin. *International Journal of Legal Medicine*, *106*(4), 190–193. https://doi.org/10.1007/BF01371335

Mantini, S., & Ripani, M. (2009). Modern morphometry: New perspectives in physical anthropology. *New Biotechnology*, *25*(5), 325–330. https://doi.org/10.1016/j.nbt.2009.03.009

Matthews, H., Penington, T., Saey, I., Halliday, J., Muggli, E., & Claes, P. (2016). Spatially dense morphometrics of craniofacial sexual dimorphism in 1-year-olds. *Journal of Anatomy*, *229*(4), 549–559. https://doi.org/10.1111/joa.12507

Maxwell, A. E., Warner, T. A., & Fang, F. (2018). Implementation of machine-learning classification in remote sensing: An applied review. *International Journal of Remote Sensing*, *39*(9), 2784–2817. https://doi.org/10.1080/01431161.2018.1433343

McDowell, J. L., L'Abbé, E. N., & Kenyhercz, M. W. (2012). Nasal aperture shape evaluation between black and white South Africans. *Forensic Science International*, *222*(1–3), 397.e1–397.e3976. https://doi.org/10.1016/j.forsciint.2012.06.007

McIntosh, K., Williams, N., Anderson, P., & Lottering, N. (2020). A semi-automated protocol for craniometric analysis on the subadult cranial isosurface models. *Forensic Imaging*, *23*, 200410. https://doi.org/10.1016/j.fri.2020.200410

Moore-Jansen, P. H. (1989). *A multivariate craniometric analysis of secular change and variation among recent North American populations* [Doctoral dissertation, University of Tennessee]. https://trace.tennessee.edu/utk_graddiss/4042/

Musa, M., & Danfulani, M. (2015). Cranial index determination using computed tomography (CT) scans of children with normal brain development: A retrospective study. *Cukurova Medical Journal*, *40*(2), 239–244. https://doi.org/10.17826/cutf.98563

Musilová, B., Dupej, J., Velemínská, J., Chaumoitre, K., & Bruzek, J. (2016). Exocranial surfaces for sex assessment of the human cranium. *Forensic Science International*, *269*, 70–77. https://doi.org/10.1016/j.forsciint.2016.11.006

Muzzall, E. (2021). A novel ensemble machine learning approach for bioarchaeological sex prediction. *Technologies*, *9*(2), 23. https://doi.org/10.3390/technologies9020023

Nagaoka, T., Kawakubo, Y., & Hirata, K. (2012). Evidence for temporal and social differences in cranial dimensions in Edo-period Japanese. *International Journal of Osteoarchaeology*, *22*(4), 409–422. https://doi.org/10.1002/oa.1215

Naikmasur, V. G., Shrivastava, R., & Mutalik, S. (2010). Determination of sex in South Indians and immigrant Tibetans from cephalometric analysis and discriminant functions. *Forensic Science International*, *197*(1–3), 122.e1–122.e1226. https://doi.org/10.1016/j.forsciint.2009.12.052

Nam, H., Han, N., Eom, M. J., Kook, M., & Kim, J. (2021). Cephalic index of Korean children with normal brain development during the first 7 years of life based on computed tomography. *Annals of Rehabilitation Medicine*, *45*(2), 141–149. https://doi.org/10.5535/arm.20235

Navega, D., Coelho, C., Vicente, R., Ferreira, M. T., Wasterlain, S., & Cunha, E. (2015a). AncesTrees: ancestry estimation with randomized decision trees. *International Journal of Legal Medicine*, *129*(5), 1145–1153. https://doi.org/10.1007/s00414-014-1050-9

Navega, D., Vicente, R., Vieira, D. N., Ross, A. H., & Cunha, E. (2015b). Sex estimation from the tarsal bones in a Portuguese sample: A machine learning approach. *International Journal of Legal Medicine*, *129*(3), 651–659. https://doi.org/10.1007/s00414-014-1070-5

Neubauer, S., Gunz, P., & Hublin, J. J. (2009). The pattern of endocranial ontogenetic shape changes in humans. *Journal of Anatomy*, *215*(3), 240–255. https://doi.org/10.1111/j.1469-7580.2009.01106.x

Nicholas, C. L., & Franciscus, R. G. (2014). The ontogeny of nasal floor shape variation in extant humans. *American Journal of Physical Anthropology*, *155*(3), 369–378. https://doi.org/10.1002/ajpa.22570

Nikita, E., & Nikitas, P. (2020). On the use of machine learning algorithms in forensic anthropology. *Legal Medicine*, *47*, 101771. https://doi.org/10.1016/j.legalmed.2020.101771

Noble, J. (2015). *Morphometrics of the juvenile skull: Analysis of ontogenetic growth patterns* [Master's thesis, University of Western Australia]. https://api.research-repository.uwa.edu.au/ws/portalfiles/portal/9181787/Noble_Jacqueline_2015.pdf

Noble, J., Cardini, A., Flavel, A., & Franklin, D. (2019). Geometric morphometrics on juvenile crania: Exploring age and sex variation in an Australian population. *Forensic Science International*, *294*, 57–68. https://doi.org/10.1016/j.forsciint.2018.10.022

Norhalifah, H. K., Syaza, F. H., Chambers, G. K., & Edinur, H. A. (2016). The genetic history of Peninsular Malaysia. *Gene*, *586*(1), 129–135. https://doi.org/10.1016/j.gene.2016.04.008

Norris, D. O., & Carr, J. A. (2020). *Vertebrate endocrinology* (6th ed.). Academic Press

O'Donnell, A., Daneshvari Berry, S., & Edgar, H. J. H. (2017). Can Cephalometrics Discriminate Between the Sexes in a Diverse Juvenile Sample? *Journal of Forensic Sciences*, *62*(3), 735–740. https://doi.org/10.1111/1556-4029.13300

O'Higgins, P. (2000). The study of morphological variation in the hominid fossil record: biology, landmarks and geometry. *Journal of Anatomy*, *197* (Pt 1), 103–120. https://doi.org/10.1046/j.1469-7580.2000.19710103.x

Ono, I., Ohura, T., Narumi, E., Kawashima, K., Matsuno, I., Nakamura, S., Ohhata, N., Uchiyama, Y., Watanabe, Y., & Tanaka, F. (1992). Three-dimensional analysis of craniofacial bones using three-dimensional computer tomography. *Journal of Cranio-maxillo-facial Surgery*, *20*(2), 49–60. https://doi.org/10.1016/s1010-5182(05)80468-8

Ortega, R. F., Irurita, J., Campo, E. J. E., & Mesejo, P. (2021). Analysis of the performance of machine learning and deep learning methods for sex estimation of infant individuals from the analysis of 2D images of the ilium. *International Journal of Legal Medicine*, *135*(6), 2659–2666. https://doi.org/10.1007/s00414-021-02660-6

Ousley, S. D., & Jantz, R. L. (1993). *FORDISC 1.0–3.0: Personal computerized forensic discriminant functions*. Forensic Anthropology Center, University of Tennessee.

Ousley, S. D., & Jantz, R. L. (2005). *FORDISC 3.0: Personal computer forensic discriminant functions*. Forensic Anthropology Center, University of Tennessee.

Ousley, S., Jantz, R., & Freid, D. (2009). Understanding race and human variation: why forensic anthropologists are good at identifying race. *American Journal of Physical Anthropology*, *139*(1), 68–76. https://doi.org/10.1002/ajpa.21006

Patriquin, M. L., Steyn, M., & Loth, S. R. (2005). Metric analysis of sex differences in South African black and white pelves. *Forensic Science International*, *147*(2–3), 119–127. https://doi.org/10.1016/j.forsciint.2004.09.074

Paulová, M., Bláha, P., Vignerová, J., & Riedlová, J. (2000). Influence of positioning of infants on long-term changes of cephalic dimensions. *Central European Journal of Public Health*, *8*(2), 83–87.

Pereira, I. M. R., Barros Filho, A. d. A., Alvares, B. R., Palomari, E. T., & Nanni, L. (2008). Radiological determination of cranial size and index by measurement of skull diameters in a population of children in Brazil. *Radiologia Brasileira*, *41*, 229–234. https://doi.org/10.1590/S0100-39842008000400006

Pickering, R. B., & Bachman, D. (2009). *The use of forensic anthropology* (2nd ed.). CRC Press.

Pietrusewsky, M. (2000). Metric analysis of skeletal remains: Methods and applications. In M. A. Katzenberg & S. R. Saunders (Eds.), *Biological anthropology of the human skeleton* (2nd ed., pp. 375–415). Wiley-Liss.

Pozzi, A., Raffone, C., Belcastro, M., & Camilleri-Carter, T. (2020). Sex estimation in cranial remains: A comparison of machine learning and discriminant analysis in Italian populations. *bioRxiv* [preprint]. https://doi.org/10.1101/2020.04.30.071597

Pretorius, E., Steyn, M., & Scholtz, Y. (2006). Investigation into the usability of geometric morphometric analysis in assessment of sexual dimorphism. *American Journal of Physical Anthropology*, *129*(1), 64–70. https://doi.org/10.1002/ajpa.20251

Kongkasuriyachai, N. P., Prasitwattanaseree, S., D. Troy Case, D. T., & Mahakkanukrauh, P. (2022). Craniometric estimation of ancestry in Thai and Japanese individuals. *Australian Journal of Forensic Sciences*, *54*(3), 294–310, https://doi.org/10.1080/00450618.2020.1789219

Rakosi, T., Jonas, I., & Graber, T. M. (1993). *Orthodontic diagnosis* (Vol. 8). Thieme Verlag.

Ramamoorthy, B., Pai, M. M., Prabhu, L. V., Muralimanju, B. V., & Rai, R. (2016). Assessment of craniometric traits in South Indian dry skulls for sex determination. *Journal of Forensic and Legal Medicine*, *37*, 8–14. https://doi.org/10.1016/j.jflm.2015.10.001

Rhine, S. (1990). Non-metric skull racing. In G. Gill & S. Rhine (Eds.), *Skeletal attribution of race: Methods for forensic anthropology – Anthropological Papers of the Maxwell Museum of Anthropology* (no. 4, pp. 9–20). University of New Mexico.

Richard, A. H., Parks, C. L., & Monson, K. L. (2014). Accuracy of standard craniometric measurements using multiple data formats. *Forensic Science International*, *242*, 177–185. https://doi.org/10.1016/j.forsciint.2014.06.015

Ricklan, D. E., & Tobias, P. V. (1986). Unusually low sexual dimorphism of endocranial capacity in a Zulu cranial series. *American Journal of Physical Anthropology*, *71*(3), 285–293. https://doi.org/10.1002/ajpa.1330710304

Roberts, C. A., & Connell, B. (2004). Guidance on recording palaeopathology. In M. Brickley & J. I. McKinley (Eds.), *Guidelines to the standards for recording human remains* (pp. 34–39). British Association for Biological Anthropology and Osteoarchaeology (BABAO) and the Institute of Field Archaeologists. https://dro.dur.ac.uk/6160/1/6160.pdf?DDD6+drk0car+d67a9y

Robinson, C., & Terhune, C. E. (2017). Error in geometric morphometric data collection: Combining data from multiple sources. *American Journal of Physical Anthropology*, *164*(1), 62–75. https://doi.org/10.1002/ajpa.23257

Roche, A. F., & Sun, S. S. (2003). *Human growth: Assessment and interpretation*. Cambridge University Press Cambridge. https://doi.org/10.1017/cbo9780511525681

Rogol, A. D., Clark, P. A., & Roemmich, J. N. (2000). Growth and pubertal development in children and adolescents: Effects of diet and physical activity. *The American Journal of Clinical Nutrition*, *72*(2 Suppl), 521S–8S. https://doi.org/10.1093/ajcn/72.2.521S

Ross, A. H. (2004). Cranial evidence of pre-contact multiple population expansions in the Caribbean. *Caribbean Journal of Science*, *40*(3), 291–298.

Ross, A. H., & Williams, S. (2008). Testing repeatability and error of coordinate landmark data acquired from crania. *Journal of Forensic Sciences*, *53*(4), 782–785. https://doi.org/10.1111/j.1556-4029.2008.00751.x

Saini, V. (2014). Significance of temporal changes on sexual dimorphism of cranial measurements of Indian population. *Forensic Science International*, *242*, 300.e1–300.e8. https://doi.org/10.1016/j.forsciint.2014.06.017

Sangvichien, S., Boonkaew, K., Chuncharunee, A., Komoltri, C., Piyawinitwong, S., Wongsawut, A., & Namwongsa, S. (2007). Sex determination in Thai skulls by using craniometry: Multiple logistic regression analysis. *Siriraj Medical Journal*, *59*(5), 216–221. https://he02.tci-thaijo.org/index.php/sirirajmedj/article/view/246135

Santos, F., Guyomarc'h, P., & Bruzek, J. (2014). Statistical sex determination from craniometrics: Comparison of linear discriminant analysis, logistic regression, and support vector machines. *Forensic Science International*, *245*, 204.e1–204.e2048. https://doi.org/10.1016/j.forsciint.2014.10.010

Sarker, I. H. (2021). Machine learning: Algorithms, real-world applications and research directions. *SN Computer Science*, *2*(3), 160. https://doi.org/10.1007/s42979-021-00592-x

Saw, S.-H. (2007). *The population of Peninsular Malaysia*. Institute of Southeast Asian Studies.

Scheuer, L. (2002). Application of osteology to forensic medicine. *Clinical Anatomy*, *15*(4), 297–312. https://doi.org/10.1002/ca.10028

Scheuer, L., & Black, S. (2000). *Developmental juvenile osteology*. Academic Press.

Schmeling, A., Olze, A., Reisinger, W., & Geserick, G. (2005). Forensic age estimation and ethnicity. *Legal Medicine*, *7*(2), 134–137. https://doi.org/10.1016/j.legalmed.2004.07.004

Schroeder, C. F., Schmidtke, S. Z., & Bidez, M. W. (1997). Measuring the human pelvis: a comparison of direct and radiographic techniques using a modern United States--based sample. *American Journal of Physical Anthropology*, *103*(4), 471–479. https://doi.org/10.1002/(SICI)1096-8644(199708)103:4<471:AID-AJPA5>3.0.CO;2-Q

Sforza, C., Grandi, G., Binelli, M., Dolci, C., De Menezes, M., & Ferrario, V. F. (2010). Age- and sex-related changes in three-dimensional lip morphology. *Forensic Science International*, *200*(1–3), 182.e1–182.e1827. https://doi.org/10.1016/j.forsciint.2010.04.050

Sforza, C., Grandi, G., De Menezes, M., Tartaglia, G. M., & Ferrario, V. F. (2011). Age- and sex-related changes in the normal human external nose. *Forensic Science International*, *204*(1–3), 205.e1–205.e2059. https://doi.org/10.1016/j.forsciint.2010.07.027

Sieber, K. S., & García-Donas, J. G. (2023). Population affinity estimation on a Spanish sample: Testing the validity and accuracy of cranium and mandible online software methods. *Legal Medicine*, *60*, 102180. https://doi.org/10.1016/j.legalmed.2022.102180

Sierp, I., & Henneberg, M. (2015). The difficulty of sexing skeletons from unknown populations. *Journal of Anthropology*, *2015*, Article ID 908535. https://doi.org/10.1155/2015/908535

Simon, T. W. (2012). *Ethnic identity and minority protection: Designation, discrimination, and brutalization*. Lexington Books.

Sinclair, D. (1973). *Human growth after birth* (2nd ed.). Oxford University Press.

Singh, J., & Pathak, R. K. (2013). Morphometric sexual dimorphism of human sternum in a north Indian autopsy sample: sexing efficacy of different statistical techniques and a comparison with other sexing methods. *Forensic Science International*, *228*(1–3),. https://doi.org/10.1016/j.forsciint.2013.03.020

Skalic, C. B. (2018). *A critical approach to ancestry in forensic anthropology: An assessment of Fordisc 3.1 and Ancestrees* [Master's thesis, University of Windsor]. https://scholar.uwindsor.ca/etd/7571/

Slice, D. E., & Ross, A. (2009). *3D-ID: Geometric morphometric classification of crania for forensic scientists*. National Institute of Justice. https://nij.ojp.gov/library/publications/3d-id-geometric-morphometric-classification-crania-forensic-scientists

Smith, H. F., Ritzman, T., Otárola-Castillo, E., & Terhune, C. E. (2013). A 3-D geometric morphometric study of intraspecific variation in the ontogeny of the temporal bone in modern Homo sapiens. *Journal of Human Evolution*, *65*(5), 479–489. https://doi.org/10.1016/j.jhevol.2013.01.017

Soames, R. W. (1995). Skeletal system. In P. L. Williams, L. H. Bannister, M. M. Berry, P. Collins, M. Dyson, J. E. Dussek & M. W. Ferguson (Eds.), *Gray's anatomy* (38th ed., pp 271–275). Churchill Livingston.

Spradley, M. K. (2006). *Biological anthropological aspects of the African diaspora: Geographic origins, secular trends, and plastic versus genetic influences utilizing craniometric data* [Doctoral dissertation, University of Tennessee]. https://trace.tennessee.edu/utk_graddiss/1864/

Spradley, M. K. (2021). Use of craniometric data to facilitate migrant identifications at the United States/Mexico border. *American Journal of Physical Anthropology*, *175*(2), 486–496. https://doi.org/10.1002/ajpa.24241

Spradley, M. K., & Jantz, R. L. (2011). Sex estimation in forensic anthropology: Skull versus postcranial elements. *Journal of Forensic Sciences*, *56*(2), 289–296. https://doi.org/10.1111/j.1556-4029.2010.01635.x

Spradley, M. K., & Jantz, R. L. (2016). Ancestry estimation in forensic anthropology: Geometric morphometric versus standard and nonstandard interlandmark distances. *Journal of Forensic Sciences*, *61*(4), 892–897. https://doi.org/10.1111/1556-4029.13081

Sprowl, A. E. (2013). *Sex determination using discriminant function analysis in Hispanic children and adolescents: A lateral cephalometric study* [Master's thesis, University of Nevada]. http://dx.doi.org/10.34917/5363948

SPSS Inc. (1999). *SPSS Base 9.0 applications guide* (1st ed.). SPSS Inc.

Standring, S., & Gray, H. (2008). *Gray's anatomy: The anatomical basis of clinical practice* (40th ed. anniversary). Churchill Livingstone/Elsevier.

Steyerberg, E. W., van der Ploeg, T., & Van Calster, B. (2014). Risk prediction with machine learning and regression methods. *Biometrical Journal (Biometrische Zeitschrift)*, *56*(4), 601–606. https://doi.org/10.1002/bimj.201300297

Steyn, M., & Henneberg, M. (1997). *Cranial growth in the prehistoric sample from K2 at Mapungubwe (South Africa) is population specific*. https://oa.mg/work/2613385800

Stojanowski, C. M., & Schillaci, M. A. (2006). Phenotypic approaches for understanding patterns of intracemetery biological variation. *American Journal of Physical Anthropology*, *Suppl 43*, 49–88. https://doi.org/10.1002/ajpa.20517

Stull, K. E., Cirillo, L. E., Cole, S. J., & Hulse, C. N. (2020). Subadult sex estimation and KidStats. In A. R. Klales (Ed.), *Sex estimation of the human skeleton* (pp. 219-242). Elsevier. https://doi.org/10.1016/B978-0-12-815767-1.00014-6

Stull, K. E., Kenyhercz, M. W., & L'Abbé, E. N. (2014a). Ancestry estimation in South Africa using craniometrics and geometric morphometrics. *Forensic Science International*, *245*, 206.e1–206.e2067. https://doi.org/10.1016/j.forsciint.2014.10.021

Stull, K. E., Tise, M. L., Ali, Z., & Fowler, D. R. (2014b). Accuracy and reliability of measurements obtained from computed tomography 3D volume rendered images. *Forensic Science International*, *238*, 133–140. https://doi.org/10.1016/j.forsciint.2014.03.005

Sutphin, R., Ross, A. H., & Jantz, R. L. (2014). Secular trends in Cherokee cranial morphology: Eastern vs Western bands. *Annals of Human Biology*, *41*(6), 511–517. https://doi.org/10.3109/03014460.2014.902991

Swamy, K. B., Zubaidi, A. L., Suwaibah, A., Azmi, H., Norizahar, K., Husbani, M. A. R., & Rohayah, H. (2013). The craniofacial indices correlate with age, gender and environmental influences – A study in Malaysian school children. *World Applied Sciences Journal*, *27*(2), 250–256. https://doi.org/10.5829/idosi.wasj.2013.27.02.81129

Szen, A. L. (2018). *Estimation of ancestry in non-adults* [Master's thesis, Binghamton University].

Tallman, S. D. (2019). Cranial nonmetric sexual dimorphism and sex estimation in East and Southeast Asian individuals. *Forensic Anthropology*, *2*(4), 204–221. https://doi.org/10.5744/fa.2019.1010

Tan, S. G. (2001). Genetic relationships among sixteen ethnic groups from Malaysia and Southeast Asia. In L. Jin, M. Seielstad & C. Xiao (Eds.), *Genetic, Linguistic and Archaeological Perspectives on Human Diversity in Southeast Asia* (pp. 83–92). World Scientific. https://doi.org/10.1142/9789812810847_0007

Tanner, J. M. (1990). *Foetus into man: Physical growth from conception to maturity* (revised ed.). Harvard University Press.

Teodoru-Raghina, D., Marinescu, M., Diaconeasa, A., Olteanu, B., Perlea, P., & Dragu, M. (2017). Crania sexual dimorphism of an European subpopulation: CT scan discriminant function analysis in a Romanian subadult casuistry. *Romanian Journal of Legal Medicine*, *25*(4), 373–378. https://doi.org/10.4323/rjlm.2017.373

Tierney, S. N., & Bird, J. M. (2015). Molecular sex identification of juvenile skeletal remains from an Irish medieval population using ancient DNA analysis. *Journal of Archaeological Science*, *62*, 27–38. https://doi.org/10.1016/j.jas.2015.06.016

Tomás, N. R. (2020). *Is the (hu)Manid program valid in a pediatric population?* [Master's thesis, University of Illinois at Chicago]. https://doi.org/10.25417/uic.12380453.v1

Toneva, D., Nikolova, S., Agre, G., Zlatareva, D., Hadjidekov, V., & Lazarov, N. (2021). Machine learning approaches for sex estimation using cranial measurements. *International Journal of Legal Medicine*, *135*(3), 951–966. https://doi.org/10.1007/s00414-020-02460-4

Toneva, D., Nikolova, S., Harizanov, S., Georgiev, I., Zlatareva, D., Hadjidekov, V., Dandov, A., & Lazarov, N. (2018). Sex estimation by size and shape of foramen magnum based on CT imaging. *Legal Medicine*, *35*, 50–60. https://doi.org/10.1016/j.legalmed.2018.09.009

Toneva, D. H., Nikolova, S. Y., Agre, G. P., Zlatareva, D. K., Hadjidekov, V. G., & Lazarov, N. E. (2020). Data mining for sex estimation based on cranial measurements. *Forensic Science International*, *315*, 110441. https://doi.org/10.1016/j.forsciint.2020.110441

Torres-Restrepo, A. M., Quintero-Monsalve, A. M., Giraldo-Mira, J. F., Rueda, Z. V., Vélez-Trujillo, N., & Botero-Mariaca, P. (2014). Agreement between cranial and facial classification through clinical observation and anthropometric measurement among Envigado school children. *BMC Oral Health*, *14*, 50. https://doi.org/10.1186/1472-6831-14-50

Toy, S., Secgin, Y., Oner, Z., Turan, M. K., Oner, S., & Senol, D. (2022). A study on sex estimation by using machine learning algorithms with parameters obtained from computerized tomography images of the cranium. *Scientific Reports*, *12*(1), 4278. https://doi.org/10.1038/s41598-022-07415-w

Tunis, T. S., Sarig, R., Cohen, H., Medlej, B., Peled, N., & May, H. (2017). Sex estimation using computed tomography of the mandible. *International Journal of Legal Medicine*, *131*(6), 1691–1700. https://doi.org/10.1007/s00414-017-1554-1

Twigg, S. R., & Wilkie, A. O. (2015). A genetic-pathophysiological framework for craniosynostosis. *American Journal of Human Genetics*, *97*(3), 359–377. https://doi.org/10.1016/j.ajhg.2015.07.006

Ulijaszek, S. J., & Kerr, D. A. (1999). Anthropometric measurement error and the assessment of nutritional status. *The British Journal of Nutrition*, *82*(3), 165–177. https://doi.org/10.1017/s0007114599001348

Urbanová, P., Ross, A. H., Jurda, M., & Nogueira, M. I. (2014). Testing the reliability of software tools in sex and ancestry estimation in a multi-ancestral Brazilian sample. *Legal Medicine*, *16*(5), 264–273. https://doi.org/10.1016/j.legalmed.2014.06.002

Valdes, G., Chan, M. F., Lim, S. B., Scheuermann, R., Deasy, J. O., & Solberg, T. D. (2017). IMRT QA using machine learning: A multi-institutional validation. *Journal of Applied Clinical Medical Physics*, *18*(5), 279–284. https://doi.org/10.1002/acm2.12161

van Lindert, E. J., Siepel, F. J., Delye, H., Ettema, A. M., Bergé, S. J., Maal, T. J., & Borstlap, W. A. (2013). Validation of cephalic index measurements in scaphocephaly. *Child's Nervous System*, *29*(6), 1007–1014. https://doi.org/10.1007/s00381-013-2059-y

Veroni, A., Nikitovic, D., & Schillaci, M. A. (2010). Brief communication: Sexual dimorphism of the juvenile basicranium. *American Journal of Physical Anthropology*, *141*(1), 147–151. https://doi.org/10.1002/ajpa.21156

Vioarsdóttir, U. S., O'Higgins, P., & Stringer, C. (2002). A geometric morphometric study of regional differences in the ontogeny of the modern human facial skeleton. *Journal of Anatomy*, *201*(3), 211–229. https://doi.org/10.1046/j.1469-7580.2002.00092.x

Vioarsdóttir, U. S., & Cobb, S. (2004). Inter- and intra-specific variation in the ontogeny of the hominoid facial skeleton: Testing assumptions of ontogenetic variability. *Annals of Anatomy*, *186*(5–6), 423–428. https://doi.org/10.1016/s0940-9602(04)80076-1

Waitzman, A. A., Posnick, J. C., Armstrong, D. C., & Pron, G. E. (1992). Craniofacial skeletal measurements based on computed tomography: Part II. Normal values and growth trends. *The Cleft Palate-craniofacial Journal*, *29*(2), 118–128. https://doi.org/10.1597/1545-1569_1992_029_0118_csmboc_2.3.co_2

Walker, P. L. (2008). Sexing skulls using discriminant function analysis of visually assessed traits. *American Journal of Physical Anthropology*, *136*(1), 39–50. https://doi.org/10.1002/ajpa.20776

Weinberg, S. M., Putz, D. A., Mooney, M. P., & Siegel, M. I. (2005). Evaluation of non-metric variation in the crania of black and white perinates. *Forensic Science International*, *151*(2–3), 177-185. https://doi.org/10.1016/j.forsciint.2005.02.009

Wescott, D. J., & Jantz, R. L. (2005). Assessing craniofacial secular change in American blacks and whites using geometric morphometry. In D. E. Slice (Ed.), *Modern morphometrics in physical anthropology – Developments in primatology: Progress and prospects*. Springer. https://doi.org/10.1007/0-387-27614-9_10

Wescott, D. J., & Moore-Jansen, P. H. (2001). Metric variation in the human occipital bone: forensic anthropological applications. *Journal of Forensic Sciences*, *46*(5), 1159–1163.

Williams, G. (2011). *Data mining with Rattle and R: The art of excavating data for knowledge discovery*. Springer.

Wong, L. P., Lai, J. K., Saw, W. Y., Ong, R. T., Cheng, A. Y., Pillai, N. E., Liu, X., Xu, W., Chen, P., Foo, J. N., Tan, L. W., Koo, S. H., Soong, R., Wenk, M. R., Lim, W. Y., Khor, C. C., Little, P., Chia, K. S., & Teo, Y. Y. (2014). Insights into the genetic structure and diversity of 38 South Asian Indians from deep whole-genome sequencing. *PLoS Genetics*, *10*(5), e1004377. https://doi.org/10.1371/journal.pgen.1004377

Wood, C. (2015). The age-related emergence of cranial morphological variation. *Forensic Science International*, *251*, 220.e1–220.e2220. https://doi.org/10.1016/j.forsciint.2015.03.030

Wright, R. (2010). *Guide to using the CRANID programs Cr6bInd: For linear and nearest neighbours discriminant analysis*. http://www.box.net/shared/static/qyaq6thdds.pdf.

Wright, R. V. S. (1992). Correlation between cranial form and geography in Homo sapiens: CRANID - A computer program for forensic and other applications. *Archaeology in Oceania*, *27*(3), 128–134. http://www.jstor.org/stable/40386951

Yagain, V. K., Pai, S. R., Kalthur, S. G., Chethan, P., & Hemalatha, I. (2012). Study of cephalic index in Indian students. *International Journal of Morphology*, *30*(1), 125–129.

Yang, W., Chen, J., Shen, W., Hu, B., Chang, Q., Lyu, K., Pan, Q., Li, H., & Zhang, Y. (2019a). Cranial measurement analysis of 406 infants aged 0 to 6 months in Chongqing. *International Journal of Pediatrics*, *6*, 687–691.

Yang, W., Liu, X., Wang, K., Hu, J., Geng, G., & Feng, J. (2019b). Sex determination of three-dimensional skull based on improved backpropagation neural network. *Computational and Mathematical Methods in Medicine*, *2019*, 9163547. https://doi.org/10.1155/2019/9163547

Yang, W., Zhou, M., Zhang, P., Geng, G., Liu, X., & Zhang, H. (2020). Skull sex estimation based on wavelet transform and Fourier transform. *BioMed Research International*, *2020*, 8608209. https://doi.org/10.1155/2020/8608209

Yong, R. Y. Y., Lee, L. K. H., & Yap, E. (2006). Y-chromosome STR haplotype diversity in three ethnic populations in Singapore. *Forensic Science International*, *159*(2–3), 244–257. https://doi.org/10.1016/j.forsciint.2005.05.010

Yusof, A. (2007). *Craniofacial growth changes in Malaysian Malay children and young adults: A cross-sectional 3-dimensional CT study* [Doctoral dissertation, University of Adelaide]. https://hdl.handle.net/2440/39388

Zaafrane, M., Ben Khelil, M., Naccache, I., Ezzedine, E., Savall, F., Telmon, N., Mnif, N., & Hamdoun, M. (2018). Sex determination of a Tunisian population by CT scan analysis of the skull. *International Journal of Legal Medicine*, *132*(3), 853–862. https://doi.org/10.1007/s00414-017-1688-1

Zhang, D., Wang, S., Li, J., & Zhou, Y. (2018). Novel method of constructing a stable reference frame for 3-dimensional cephalometric analysis. *American Journal of Orthodontics and Dentofacial Orthopedics*, *154*(3), 397–404. https://doi.org/10.1016/j.ajodo.2017.11.038

Zhang, Y., & Schepartz, L. A. (2021). Three-dimensional geometric morphometric studies of modern human occipital variation. *PloS One*, *16*(1), e0245445. https://doi.org/10.1371/journal.pone.0245445

Zhou, L., Pan, S., Wang, J., & Vasilakos, A. V. (2017). Machine learning on big data: Opportunities and challenges. *Neurocomputing*, *237*, 350–361. https://doi.org/10.1016/j.neucom.2017.01.026

Zöller, M.-A., & Huber, M. F. (2021). Benchmark and survey of automated machine learning frameworks. *Journal of Artificial Intelligence Research*, *70*, 409–472. https://doi.org/10.48550/arXiv.1904.12054