CRANIAL METRIC AND NONMETRIC VARITION IN SOUTHEAST MEXICO AND GUATEMALA: IMPLICATIONS FOR POPULATION AFFINITY ASSESSMENT IN THE UNITED STATES

By

Kelly Rae Kamnikar

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Anthropology—Doctor of Philosophy

2022

ABSTRACT

CRANIAL METRIC AND NONMETRIC VARITION IN SOUTHEAST MEXICO AND GUATEMALA: IMPLICATIONS FOR POPULATION AFFINITY ASSESSMENT IN THE UNITED STATES

By

Kelly Rae Kamnikar

The scientific identification of unknown human skeletal remains in forensic contexts relies heavily on the estimation of demographic parameters (i.e., sex, age, stature, and population affinity). Population affinity, or the likelihood of group relatedness to a defined population of a decedent, can be estimated using measurements and observations from the cranial and postcranial skeleton. These estimations may be less accurate among populations which have been pooled together based on convention. Latin American individuals—with geographic origins widely distributed throughout Central and South America—are broadly pooled together under the blanket term Hispanic with little regard for the immense cultural and biological diversity represented by these groups. Consequently, forensic anthropologists may be unintentionally disregarding genetic diversity, population structure, and population history and their impact on the formation and morphology of these groups. The purpose of this dissertation is to investigate variation in craniofacial morphology and develop population affinity models for Latin American groups using cranial metric and nonmetric data. The intent is to move beyond a single classification level (i.e., Hispanic) to more refined levels based on geographic origins (e.g., Guatemala, Southeast Mexico).

The broad category of Hispanic was adopted by forensic anthropologists in large part because it is still used in medicolegal death investigations in the U.S. to describe individuals with familial origins in Latin America, Spain, and the Caribbean (U.S. Census Bureau 2021). Since the term Hispanic does not narrow down the region of origin for unidentified human remains, it is uninformative for identification and repatriation purposes, particularly regarding forensic investigations along the southern U.S. border. In this context, population affinity estimation benefits from refinement of a broad category to a more focused, population-level group. Craniometric and cranial macromorphoscopic (MMS) data are collected from samples in Guatemala City, Guatemala and Mérida, Mexico—with strong support from the forensic anthropologists in these countries—to capture aspects of skeletal variation associated with these regions. Biological distance and population affinity models are assessed and comparative data from other Latin American and U.S. populations are used to assess how well these model skeletal variation. Biological distance analysis demonstrates that Latin American populations, including the Meridian and Guatemala sample are distinct. Classification models obtain varying accuracy rates; the combined craniometric and cranial MMS model had the highest classification accuracy (70.7%). This study provides further support for the refinement of this broad category and is important for future investigations involved in identification efforts along the U.S.-Mexico border.

This dissertation is dedicated to my family and my support system. Thank you.

ACKNOWLEDGEMENTS

I would like to thank my advisor and dissertation committee chair, Dr. Joseph Hefner for his guidance and advice throughout the dissertation process, and for his unwavering trust and support in my decisions to shape my future. Thank you for supporting my ideas and wishes, and for helping this project come to fruition. I would like to thank my committee members, Dr. Elizabeth Drexler, Dr. Todd Fenton, Dr. Kate Spradley, and Dr. Gabriel Wrobel. Each of them provided me with an opportunity to explore various aspects of Anthropology related to my research that have greatly broadened my views and shaped my as an Anthropologist. I appreciate the discussion we have had on different subjects and the new avenues of research that were presented to me through them. A very big thank you to Dr. Nicholas Herrmann for advice and advising over the years. I am grateful to the institutions and medical examiners offices who opened their doors for me to collect data for this project: Dr. Vera Tiesler, Monica Rodriguez, and Julio Chi-Keb at the Universidad Autónoma de la Yucatán for their assistance and access to the individuals housed in the UADY collection, and for showing me treasures in the beautiful city of Mérida, Mexico; and Dr. Zarina Guzman, Dr. Carlos Rodas, and Dr. Elmar Gonzalez for allowing me to participate in activities and data collection in the Anthropology Unit at the Instituto Nacional de Ciencias Forenses in Guatemala City, Guatemala. I want to thank Daniel Jiménez, Myrna Díaz, Carlotta Díaz, and Celeste Pereira for accepting me as if I were another lab member at the INACIF and making each visit to Guatemala very special. I would like to thank the Center for Latin American and Caribbean Studies at Michigan State University for funding through the Tinker Foundation and the College of Social Sciences at Michigan State University for funding through the Corey Endowment and the Research Scholars Award. Lastly, I want to thank the MSU FAL members for their support, advice, and coding help during this project. Thanks for making the MSU FAL a really great place to work and thrive.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

One of the main objectives of a biological anthropological analysis is to estimate demographic variables from skeletal remains for the purpose of answering questions about the skeletal assemblage or individual. In forensic analysis, where questions are aimed at identification of the individual, these variables include sex, age, stature, and population affinity. Population affinity aims to understand the geographic origin of a person through the comparison of skeletal features to populational reference groups (Pilloud & Hefner 2016). Differentiation of these skeletal features is due to human genetic variation, which is shaped by microevolutionary processes (genetic drift, gene flow, natural selection, and mutation) acting on our genome (Relethford & Harding 2001). These forces are influenced by cultural and environmental variables (Goodman & Leatherman 1998; Leatherman & Goodman 2020; Stanford et al. 2011:118). Within population genetics modeling, evolutionary forces are examined in relationship to allele frequencies, finding higher levels of variation within any given population rather than between populations (Relethford & Harding 2001). In human groups, using phenotypic correlates, the same principle applies. Human variation is often displayed as a cline rather than distinct boundaries between populations, often correlating to geography and the environment to produce patterns of variation (Relethford & Harding 2001). Skeletal markers useful for examining human variation are tied to neutral genomic variation (Reyes-Centeno & Hefner 2019). Biological anthropologists use patterned variation in skeletal markers to reconstruct population history and to understand group relatedness in humans living in the past and in the present. In forensics, this variation is compared to social classifiers of identity, which is useful for the U.S. given the population history of the country (Sauer 1992). However, social identity and phenotypic data are not 1:1 correlates, and must be contextualized while acknowledging inherent biases (Michael et al. 2021).

While population affinity estimates are useful for several groups in the U.S., often population affinity estimation is more difficult for Latin American populations. Classification matrices and graphics often place Latin American groups in an intermediate position due to genetic admixture from continental, parental groups (Dudzik & Jantz 2016; Spradley 2016a). Because some craniometric and cranial MMS variables correspond to selectively neutral genetic traits (Relethford 2009; Reyes-Centeno & Hefner 2019), population history can be reconstructed (Relethford 2009). However, we must consider the impact of other variables, like human migrations, complex relationship networks based on culture, colonization by European groups, and the forced migration of African slaves, that contribute to genetic, and consequently skeletal variation in Latin America. For population affinity estimation using skeletal variables, several causes may account for lower classification rates, including: 1) erroneously grouping diverse populations under the term Hispanic; 2) a lack of reference data from geographical areas falling within this classification label; and, 3) a poor understanding of the range of human variation in Latin American populations (Spradley 2016a: 242). This is especially problematic as the U.S. demographic boasts Hispanics as the second largest population group and the ongoing humanitarian crisis at the U.S.-Mexico border.

In general, reference samples used to investigate biological variation in Latin American groups is limited, which severely hampers our understanding of human variation in Latin America. Studies examining variation note differences, skeletally, among populations within Latin America. However, the issue is that our limited understanding of variation is broadly applied to Latin American groups, which is problematic. This project will add to the current reference data available for Latin American populations and provide preliminary data to fill some reference sample gaps. Through continued research, we can gain a better understanding of the range of human variation, potential regional patterns of variation in Latin America, limitations associated with current methodology, and identify actionable paths for understanding human variation in this diverse group. The data from this project is best suited for use and reuse in comparative studies. The continuous addition of new sample data allows for exploration of variation using different labeling systems under new hypotheses. An understanding of variation within these samples can be useful for future studies investigating population affinity and identification for Latin American individuals within the U.S. criminal justice system and the humanitarian crisis at the border between Mexico and the U.S.

The purpose of this dissertation is to address the previously mentioned issues with population affinity in Latin American populations. To do this, I will: 1) investigate human variation in samples from Latin America and 2) test the ability and accuracy of population affinity models for Latin American derived populations using different methodological approaches and comparative sample data.

Research Design and Research Questions

This dissertation is guided by the <u>overarching research question</u>: *Is there significant craniofacial variability among Latin American populations?* Three sub-questions seek to understand human variation within Latin American samples as they relate to biological distance, population structure, and population affinity estimates in forensic anthropology. <u>Research question one</u> asks: *Are there significant differences in craniofacial variability across sex and population labels in samples from Guatemala City, Guatemala and Mérida, Mexico?* Craniometric and cranial MMS data are collected from two Latin American reference samples from Mérida, Mexico and Guatemala City, Guatemala, all of which are currently underrepresented in forensic reference databases. Trait and variable correlations are used to understand relationships between the variables and overarching labels like sex and population.

I am using these two data types, craniometric and cranial MMS, because they are heritable, meaning passed on genetically, and they correspond to selectively neutral genetic traits (Relethford 1994; 2009; Reyes-Centeno & Hefner 2019). Most importantly, previous research has demonstrated that combining metric and nonmetric data better captures skeletal morphology than either data type alone (Spiros & Hefner 2019; Maier 2019).

Research question two asks "*What is the relationship between Latin American groups using craniometric and cranial MMS data?*" Here, I will examine the relationship between craniometric and cranial MMS data as they relate to population structure in each of the Latin American samples. Research demonstrates the utility of craniometric data to reconstruct population history, which translates into population affinity estimation modeling (Relethford 2009). Preliminary research into cranial MMS data suggest a correlation between traits and genomic data, which implies a correspondence between population history and certain cranial MMS traits (Reyes-Centeno & Hefner 2019). The two previously described samples are compared to samples of identified Mexican and Guatemalan migrants using factor and distance analysis to identify relationships.

Research question three asks: *"Can craniometric and cranial MMS data be used to make predictions about population affinity?"* This question aims to understand if the two data types can make predictions regarding population affinity for the samples tested. Population affinity estimation models are created using a combined craniometric/cranial MMS approach within a machine learning method (MLM) classification framework of artificial neural network analysis (aNN). This type of modeling is appropriate for this research as this method uses both categorical and continuous data. Together, these research questions will provide more nuanced information about human variation and population structure in two samples from geographically proximate locations in Latin America, and allow assessment of classification models using each data type and in combination to demonstrate the applicability of this research.

Human Skeletal Variation and Heritability

Biological evolution grounds studies in human variation and is the foundational theory under which that variation is studied. Under Darwinian Theory and the Modern Synthesis, the main drivers of evolution are mutation, gene flow, genetic drift, and natural selection; however, other factors like symbiotic relationships between organisms, epigenetics at the DNA level, and internal cell control mechanisms can further shape evolutionary trajectories (Corning 2020). These evolutionary drivers depend on and co-vary with the environment, including natural, biological, and socio-cultural conditions. While natural selection relies on an organism's fitness as a main driver of change, most variation at the molecular level (DNA) does not affect fitness (Duret 2008). Under Neutral Genetic Theory, most evolutionary change occurs at the molecular level and results from genetic drift and mutation acting on genetic material that is selectively neutral (Kimura 1991). These neutral genetic traits correspond to craniofacial variables on the skeleton (Relethford 2009; Reyes-Centeno & Hefner 2019). This theoretical perspective can be used to explain the variation reflected in the skeleton and provides the framework for methodology used to estimate demographic variables from the skeleton (Boyd & Boyd 2018). The influence of culture on biology can be addressed using a biocultural framework. A biocultural framework examines the impact of culture and the natural environment on biological material (Goodman & Leatherman 1998). Local and global cultural forces, like larger political events, act in combination with human agency to produce biological variation (Leatherman & Goodman 2019). For example, disparities in access to resources have demonstrated biological effects on human groups (Klaus 2012; Soler & Beatrice 2018). These overarching cultural drivers are important to consider when conducting research on contemporary populations. Samples used in this dissertation include individuals from Mexico and Central America, a region where humans have been responding to cultural variability in pre- and post-contact societies for centuries. Notably, the social, economic, and political behaviors during pre-contact and colonization, coupled with contemporary economic and political dynamics in the study region, all contribute to contemporary human genetic and phenotypic variation. We know that cultural factors can influence biology (Leatherman & Goodman 2020), specifically genetics, which can directly impact the expression of craniofacial morphology.

Variation in cranial morphology is used to infer population structure and group relatedness in biological anthropology studies. Relethford (2010) notes that although environmental effects can impact cranial morphology the underlying genetic structure of a population is not obliterated. The cranium is especially suited for these types of studies because of the high levels of heritability in cranial morphology (Adhikari et al. 2016; Relethford & Harpending 1994; Relethford 1994; Roseman & Weaver 2004). In bioarcheology and modern biodistance studies, levels of skeletal variation are compared using nonmetric and metric variables. Group relatedness is determined by applying statistical analysis to examine group relatedness from a population-level perspective. In forensic anthropology, data is compared from an individual (unknown skeleton) to a population (reference group data) using classification statistics to produce probabilistic estimates (Dunn et al. 2020; Hefner et al. 2016; Ousley 2016; Ousley & Jantz 2012; Spradley & Jantz 2016). The strength of classification of an unknown individual depends on available and appropriate reference data used in model construction (Spradley 2016a; 2021).

Population Affinity and Personal Identification

Ancestry, in forensic contexts, is a demographic parameter of the biological profile that aims to establish the geographic origin or ancestral affiliation of a set of skeletal remains (SWGANTH 2013). Estimation of this variable in U.S. forensic casework is possible for several reasons: 1) skeletal data is correlated with geographic information due to genetic heritability (Ousley et al. 2009); and, 2) the unique population history of the U.S., including voluntary and forced migration from populations geographically distant from each other over various periods (e.g., dispersal from Asia, European colonization, forced migration of Africans) (Sauer 1992). These migratory events, coupled with socio-cultural constructs dictate behavioral practices. U.S. legislation and cultural constructs have long contributed to assortative mating practices, maintaining underlying population structure reflected in skeletal morphology (Ousley et al. 2009; Sauer 1992). Because of these correlations, skeletal data can be used—with high levels of accuracy—to classify an unknown individual into one of several reference populations (SWGANTH 2013). In its current form, ancestry estimation evaluates population affinity or the likelihood of an individual to be included in a specific population in the way that the population is defined for research (Winburn & Algee-Hewitt 2021). Definitions can range from socially constructed labels and context, genetic ancestry, or some combination of biological and social variables. While the term ancestry is used in standardized documentation (SWGANTH 2013), I will use the term population affinity as it more accurately captures estimation of group affinity using evolutionary history and population structure (Winburn & Algee-Hewitt 2021). Population affinity accounts for biocultural variables that contribute to variation rather than broadly grouping people based on arbitrary, bureaucratic labels that often correspond to geo-political states or geopolitical regions. (See **Chapter 2: Human Variation in Populations Considered Hispanic** for more discussion on population affinity.)

Traditional U.S. classification categories are broad and stem from hierarchical ideas of biological race attributed in part to Samuel Morton (1799-1851). These categories result in an unknown skeleton to be classified into one of three continental groups: African, Asian, or European (Dewbury 2007). This is largely a product of a typological approach to human variation, which categorized all humans into one of three 'races' (Winburn & Algee-Hewitt 2021). Development of skeletal reference samples largely consisted of 19th century American Black and American White individuals further reifying the 3-group structure (Spradley & Weisensee 2017). Subsequently, analytical methods for estimating population affinity were developed using these available reference samples limiting methodological outcomes. Researchers recognized the need to expand reference data used in methodological development to account for variation (Hefner & Spradley 2018; Spiros 2019) and studies have estimated population affinity on a more refined level (Atkinson & Tallman 2020; Hefner & Byrnes 2020; Hefner et al. 2015; Hughes et al 2013; Kamnikar et al. 2021; Maier & George 2020; Spradley 2014a; 2016a). The recommended approach for population affinity estimation follows a broad to narrow classification, allowing for finer resolution as the data permit (Hefner & Spradley 2018). Key to a more refined approach is the availability and inclusion of appropriate reference data in modeling (Spradley 2016a).

To facilitate a positive identification between a set of unknown skeletal remains and a person's identity, forensic practitioners rely on the comparison of antemortem and postmortem data, which can involve the use of medical radiography, dental records, or DNA analysis (Hurst et al. 2013). However, in forensic work of suspected cases of underrepresented groups (i.e., migrants), these data types are often unavailable or there are limitations associated with access to medical records hindering one-to-one comparisons (Anderson 2008). In lieu of records, many presumptive identifications are made utilizing circumstantial evidence, such as descriptions of dental devices, tattoos, healed bony fractures, scars, body markings, the use of skull-photo superimposition, or the use of mitochondrial DNA via family reference samples (Anderson 2008; Birkby et al. 2008; Fenton et al. 2008). Estimation of the biological profile via skeletal remains can provide critical information for creating a short list of potential individuals, thus aiding in the identification process (Tersigni-Tarrant & Shirley 2013). In this context, it is critical to have data from diverse reference samples to estimate not only population affinity, but also sex and stature, as their accuracy usually hinges on population-specific methods (Garvin & Klales 2020; Spradley 2016b).

The Hispanic Label and Terminology

The term Hispanic as used in the U.S. legal system is based on a shared language—Spanish—and ideas of shared culture (Oboler 1995). Within the U.S. system of reporting and forensic casework, Hispanic reference samples refer to individuals from Spanish-speaking countries. This label disregards human variability due to culture, language, biology, environment, and history (Ross et al. 2004; 2014; Spradley 2016a) and is uninformative for use in forensic anthropology at the U.S.-Mexico border, where hundreds of migrants die each year crossing the borderlands clandestinely (Spradley 2014a; Spradley et al. 2019). In this research, I refer to samples that are both grouped under the broad term "Hispanic" according to geography and the collection origination – Mérida, Mexico and a morgue sample from Guatemala - and indicate in **Chapter 4: Materials and Methods** exactly where each sample comes from and who is included in them. I caution that even though geopolitical labels are used as population affinity descriptors, the samples I use in this research do not accurately capture human variation within a country or region, rather they serve as a starting point for exploring variation that can be used in combination with new data for future comparative studies. Importantly, labels used to describe populations or samples in this research may be different than reported identity or self-identification labels. Broadly, I refer to samples originating from Mexico, Central, and South America as Latin American. The label "Latin American" is used in the region by forensic practitioners and forensic anthropologists to describe populations with a common history of conquest by the Spanish and Portuguese empires, which include groups in Latin America (Daniel Jiménez personal communication). For example, the professional organization of practicing forensic anthropologists in Mexico, Central, and South America is called the Asociación Latinoamericano de Antropología Forense [The Latin American Association of Forensic Anthropology; https://www.alafforense.org/es/], in which they collectively refer to Mexico and countries within Central and South America as Latin America.

Population Affinity in Latin American-derived Samples

Statistics in the National Missing and Unidentified Persons System, NamUs, for Hispanic individuals identify 3,237 active unidentified and 2,982 open missing persons' cases (NamUs 2020). In the context of migration, Arizona reported 2,816 migration-related deaths since 2000 (of which, approximately 53% are unidentified) (PCOME 2017), with more than 3,000 open missing persons' cases (Colibrí 2018). The figures are also high in Texas with approximately 2,655 documented migrant death cases in South Texas from 1990-2020 (Leutert et al. 2020). Of the cases recovered, approximately 300 unidentified migrant cases are curated at Texas State University (Spradley et al. 2019). Furthermore, Latin Americans live in patterned pockets within specific areas of the U.S. (i.e., Salvadorans in California and Texas, Venezuelans in Florida, etc.) (Noe-Bustamante et al. 2019a, b). An understanding of human variation as it relates to geographic origin could be useful for identification within U.S. forensic casework. Identifications are hindered by several factors, including limitations or inaccuracies of current methodology (Kimmerle et al. 2010).

Population affinity estimation models that capture nuances of population structure in the region could allow for more targeted analyses. As classification accuracy depends on available comparative reference data, a major limitation in refining the Hispanic category is a lack of available reference data from Latin America (Spradley 2014b; 2016a). Despite a large, geographically diverse sample of individuals in the Forensic Databank (FDB) (Jantz & Moore-Jansen 1988) and the Macromorphoscopic Databank (MaMD) (Hefner 2018), samples for Latin American individuals or individuals with Latin American heritage are largely driven by skeletal data collected from forensic contexts (Hefner 2018; Jantz & Ousley 2005; Spradley 2021). The FDB contains craniometric data for Latin American populations from four main sources: mostly unidentified migrants (Southwest [SW] Hispanic sample); identified individuals from forensic case work in the U.S.; and, an indigenous highland Mayan sample from Guatemala. An additional resource used for classification analysis, FORDISC (Ousley & Jantz 2012) contains these samples and two cemetery samples from Mexico, bringing the total reference data to approximately 480 individuals (Spradley & Weisensee 2013). The SW Hispanic sample comprises craniometric data collected during postmortem examination at the Pima County Office of the Medical Examiner (PCOME) in Tucson, Arizona from identified border crossers (Anderson & Spradley 2016; Tise 2014). As such, most individuals in the SW Hispanic sample are from Mexico. Forensic case data from identified individuals sometimes is submitted to the FDB and available for research. Craniometric data available from an indigenous Maya highland Guatemalan sample are composed of Guatemalan Civil War genocide victims (Spradley et al. 2008). These individuals are part of specific cultural groups targeted by military forces during the Guatemalan Civil War. The additional two reference samples in FORDISC from Mexico are made up of identified individuals from two cemeteries in Mexico City and Mérida, which are in different geographic locations and were subject to different historical cultural events. Reference data for Latin American populations in the MaMD derive from four major sources but are more limited than the FDB. Cranial MMS data were collected from the PCOME, La Verbena Cemetery in Guatemala, and the Operation Identification samples from Texas State University (Hefner 2018). The PCOME cranial MMS sample contains small sample sizes (n<10) for individuals from different regions in Mexico and Central America. It is important to note that it is difficult for researchers to access certain areas in Mexico and Central America due to current socio-political conditions and regional control of violent groups. Without sufficient reference data, estimation of population affinity in Latin American derived groups is difficult, at best, and prone to error. Small samples may skew results and inadequately capture the full range of variation present in Latin American populations. As Mexicans and Guatemalans comprise two of the Top 10 groups in the U.S. and two of the top four migrant origination countries (Eschbach et al.1999; U.S. Census Bureau 2019), reference samples from these countries are imperative for identification efforts.

Organization of Chapters

The dissertation is organized into seven chapters. The first chapter, **Introduction**, serves to introduce the dissertation and related goals. The chapter starts with a presentation of the research questions and research design, which is followed by a brief discussion of human variation and heritability in biological anthropology, which links to population affinity in forensic anthropology. Finally, I discuss terminology employed in the dissertation and limitations with estimating population affinity in Latin American-derived individuals.

Chapter 2, **Human Variation in Populations Considered Hispanic**, discusses cultural and biological sources of variation for the Latin American samples used in this research. I describe differences between the terms 'ancestry' and 'population affinity' and their use and misuse in biological anthropology. Next, I present a summary of findings from skeletal variation studies in Latin American populations and describe sources of variation that have impacted the samples used in this research. These events include cultural sources of variation like the Maya Culture, Spanish colonization, and the Caste War in the Yucatán, and biological fountains of variation like migration events and genetic variation. Finally, I conclude with expectations of variation among the Latin American samples used in this research and acknowledge sample biases.

Chapter 3, **Contemporary Migration to the U.S. from Mexico and Central America** discusses the origins of migration to the U.S. from Mexico and Central America, identifying sociopolitical and economic factors that initiated and perpetuated migration and changes over time in motivations. I identify U.S. policy that directly impacted migration routes, contributing to death of migrants as they tried to cross. I finish with a summary of theoretical perspectives used to describe the humanitarian crisis at the border in sociocultural and forensic anthropology, and how social inequality diffuses into forensic work, having implications for identification and repatriation.

In Chapter 4, **Materials and Methods**, I describe the samples and methodology used in this project. I collected data from two Latin American samples -- the Xoclán Cemetery in Mérida, Mexico and the Instituto Nacional de Ciencias Forenses (INACIF) in Guatemala City, Guatemala. Craniometric and cranial MMS data were collected using a 3D Microscribe Digitizer, the software 3Skull, and the program MMS v.1.61. Using biodistance analysis, I identify relationships between the Latin American samples and comparative reference data. I then investigate group membership using classification via machine learning models.

Chapter 5, **Results**, presents the results of my analyses. Biodistance analyses results are reported as tables and graphics, and the results of the classification methods as well as model performance are discussed. These results are extrapolated to my research questions and theoretical perspectives in the **Discussion** (Chapter 6). I tie these results to other studies on Latin American skeletal samples, biological and cultural influences, and social theory. Additionally, chapter 6 describes the broader impacts of research. This research provides matched populational data to serve as a starting point to investigate human variation in Guatemala and Mérida, Mexico. The code used in my approach is available on my GitHub page for future research and advancement of methodology used in identification efforts.

CHAPTER 2: HUMAN VARIATION IN POPULATIONS CONSIDERED HISPANIC

Human variation in the phenotypic composition of human beings is attributed to a combination of intrinsic (e.g., hormones, genotype) and extrinsic factors (e.g., climate, environment, culture) acting on a population (Agarwal 2016; Agarwal & Beauchesne 2010; Roseman & Auerbach 2015). Using a biocultural approach, we can examine these forces in tandem with cultural variables (e.g., political, economic, and social structure) to understand variability in global populations (Leatherman & Goodman 2019). Cultural variables are important to consider for the impact they have on population structure and local group variation, which are present in any research sample used in biological anthropology.

This research uses data from skeletal samples in Guatemala (the Instituto Nacional de Ciencias Forenses [INACIF]) and Mérida, Mexico (Universidad Autónoma de la Yucatán [UADY]), geopolitical groups which are characterized under the Hispanic heading in current U.S.-based forensic identification frameworks (Anderson 2008; Murray et al. 2018; United States Census Bureau 2022). Research demonstrates that this broad group label can be useful for population affinity estimates in forensic anthropology, but fails to provide specific information for missing persons cases (Spradley et al. 2008; Spradley 2021; Ross et al. 2014). Within this broad classification, research in skeletal and genetic variation show high diversity (Ibarra-Rivera et al. 2008; Rubi-Castellanos et al. 2009; Ross et al. 2014; Spradley 2014a), largely due to different historical and cultural events acting on each of these samples in specific ways to produce variability. To understand the presence and persistence of variation within the samples used in this research, I document different cultural and historical processes that may contribute to variation in each sample.

Human Variation and Population Affinity

Human variation in skeletal morphology is recruited in forensic anthropology to estimate a person's ancestry, capitalizing on the non-zero correlation between social race and geographic origin in the U.S. (Sauer 1992). However, ancestry estimation in forensic anthropology traditionally utilizes skeletal samples as proxies for populations, reinforcing the idea that certain racial types are associated with geography, whether this association is purposeful or not (Ross & Pilloud 2021). Samples are often conflated as representatives of continental or regional variation, and researchers often do not consider the impact of population structure, individual populational histories, and population-specific cultural factors on human variation (i.e., classification studies). This contributes to the oversimplification of human variation in ancestry estimation modeling like classification studies. Recent scholarship recommends that forensic anthropologists situate research on human variation as skeletal tissue relates to smaller populations rather than ancestry categories (Winburn & Algee-Hewitt 2021). The term population affinity better aligns with what forensic anthropologists aim to estimate, using models that consider evolutionary history and population structure instead of ancestry, which is seen as a correlate for race (Ross & Pilloud 2021; Winburn & Algee-Hewitt 2021). Rather, race is a social construct with direct biological consequences that can impact a person's health and well-being, but race has no correlate with global patterns of diversity (Gravlee 2009). Human variation must be explained along with reasons why said variation exists as it relates to population history and biocultural variables within a framework of evolutionary theory (Ross & Pilloud 2021).

Prior to identifying morphological skeletal variables (e.g., cranial MMS traits, interlandmark distances), research should aim to understand why such variables may or may not manifest, define what a population is relative to the research project, and understand limitations with grouping variables (Winburn & Algee-Hewitt 2021). Ross and Pilloud (2021) suggest using the definition by Sneath and Sokal (1973) that a population is a "group that shares some commonality based on phenetic similarities without a phylogenetic assumption, such as a deme, cultural factors, etc." (page #). These cultural factors can become entwined with biology leading to significant differences between populations. The remainder of this chapter will discuss sources of cultural and biological variation that impact the INACIF and UADY samples used in this study, emphasizing that variation is a very complex process and is constantly in motion.

Skeletal Variation in Latin America

Because of high heritability between the genome and the craniofacial variables (Relethford 1994), we expect groups considered Hispanic to exhibit skeletal variation. In 2008, Spradley and colleagues recommended a reevaluation of methods used to estimate population affinity in Latin American (Mexico, Central and South American) populations. They called for an update to forensic methodology originally developed using American Black and American White samples from U.S. collections compiled in the 19th century. This spurred several studies on variation in Latin American populations, which critique the use of the term 'Hispanic' to define populations from Latin America (Algee-Hewitt 2018; Dudzik 2019; Hefner et al. 2015; Hughes et al. 2013; Kamnikar et al. 2021; Monsalve & Hefner 2016; Ross et al. 2014; Spradley 2014a; 2016a; 2021; Spradley et al. 2008; Tise et al. 2014). As the cranium is a popular skeletal element used in estimating population affinity (Dunn et al. 2020), many of these studies focused on craniometric and morphoscopic-based differences in craniofacial form.

Tise and colleagues (2014) examined craniometric differences among four samples considered Hispanic (Mexican, Indigenous highland Guatemalan, Cuban, and Puerto Rican samples), American Blacks, and American Whites, finding considerable differences among the

Hispanic samples, with the greatest dissimilarity between the Puerto Rican and Indigenous highland Guatemalan samples. Sample clusters with similar cranial morphology (e.g., Puerto Rican + American White, American Black + Cuban, and Mexican + Indigenous Highland Guatemalan) reflect population histories and migration events within and between the Caribbean Islands and Mesoamerica, indicating the importance of understanding population structure when interpreting results (Tise et al. 2014). Using 3D spatial data in a geometric morphometric model, Ross and colleagues (2014) found differences between samples considered Hispanic from Cuba, Ecuador, Panama, and Mexico, noting low amounts of Amerindian contributions on the Cuban sample compared with other samples. Hefner and colleagues (2015) examined MMS trait variation between a Guatemalan and three other samples (SW Hispanic, American Black, American White) (Hefner et al. 2015). They found significant variation in trait frequency distribution between the Guatemalan and SW Hispanic sample. Research explored MMS trait and craniometric variation within a Colombian sample, and among other, comparative samples, finding very little intra-regional variation within the Colombian sample, despite the heavy use of socially-designated racial categories in Antioquia (see Monsalve & Hefner 2016). When compared to the non-Colombian data sets, more nuanced details of population structure were illuminated. Colombians in Antioquia exhibit a close morphological relationship to American Whites and other Hispanic groups, consistent with European colonization and population isolation in the area (Kamnikar et al. 2021; Monsalve & Hefner 2016).

Due to a lack of reference data and sample information for modern Latin American populations, many studies utilized unidentified and identified migrant data collected during skeletal analysis from the Operation Identification (OpID) program at Texas State University and the Pima County Office of the Medical Examiner (PCOME) in Tucson, Arizona. With identified migrant data, Hughes and colleagues (2013) used craniometric data from a PCOME sample to explore if cranial variation mirrored a European-Indigenous genetic admixture gradient. They found cranial variation coincided with genetic results, as group centroids more closely aligned with parent continental reference samples. Focusing on population affinity in an unknown migrant sample, Spradley (2014a) identified differences between migrants recovered from southern Texas and Arizona, a contemporary anatomical sample from in the School of Medicine from the National Autonomous University of Mexico in Mexico City, and an indigenous Guatemalan sample from the Fundación de Antropología Forense de Guatemala. The Mexican sample and Arizona migrants were very similar, indicating that the Arizona migrant sample most likely comprised migrants from Mexico. The Texas migrant sample was different from all comparative samples, likely indicating these individuals are not represented in current reference databanks.

Many U.S.-based skeletal studies examining variation among Latin American populations use samples of relatively small size. Many of these samples represent migrants identified in medical examiner's offices or during some other medicolegal death investigations. These samples are often the only data available from these migrant-originating countries for various social, political, or legal reasons. For example, members of some countries do not see body donation as a viable alternative to burial (Winburn et al. 2020). Therefore, all available data—even datasets with small sample sizes—should be used. This is particularly true for samples derived from various Latin American populations; although minority groups in the U.S., they are over-represented in medicolegal death investigations due to systematic and institutional racism, the cycle of poverty, and the inequality of identification (Goad 2020). These datasets can be used to create a starting point to develop our understanding of skeletal human variation in Latin America, which can drive future research related to population affinity and identification (Spradley 2021; Winburn et al.

2020). Research must include a discussion on sample bias and limitations associated with small samples and recognize how these impact results.

Cultural & Environmental Sources of Sample Variation

Location & Composition

The UADY sample comes from Mérida, Mexico, the capital city of the Yucatán state in Mexico. Individuals within this sample come from the Xoclán Cemetery, and overwhelmingly comprise individuals of Maya descent, born between 1900 and 1990 (Chi-Keb et al. 2013). Most of the individuals in the collection lived in Mérida proper or the surrounding rural towns. Individuals in the cemetery are continuously excavated by the university and stored at the Facultad de Ciencias Antropológicas if the next of kin cannot continue to pay burial fees after a two-year grace period (Chi-Keb et al. 2013).

Individuals in the INACIF sample come from forensic casework at the INACIF's Morgue Metropolitana in Guatemala City, Guatemala. While Guatemala City and the INACIF morgue are in the south-central part of the country, forensic casework comes from any department within Guatemala that requires anthropological analysis. Guatemala is a very diverse country with several Maya descendant groups, speaking 22 different Mayan languages (Translators Without Borders 2022). The sample from the INACIF does not capture the range of human variation within Guatemala's borders, but can be used as preliminary data for future research examining skeletal variation. Department of origin and demographic variables are recorded if known, and data are continuously collected by researchers at the INACIF. Therefore, as the sample grows, it can be reexamined and reassessed under the same and novel hypotheses regarding variation and population affinity. Research trends in the composition of unidentified morgue populations, like a portion of the INACIF sample require discussion. Unidentified individuals tend to be adult males and come from specific segments of society like at-risk groups and/or underrepresented minorities, resulting in a forensic population that is different from the general population (Kimmerle et al. 2010; Komar & Grivas 2008). This is a consideration for this project and future research using this sample.

Prehistoric Migration & Cultural Events

Archaeologically, the area which includes these samples (the country of Guatemala and the city of Mérida, Mexico) is referred to as Mesoamerica. This region includes the present-day geographic areas of central and southern Mexico, Guatemala, Belize, Honduras, and El Salvador. While many prehistoric Indigenous groups lived in the area, one of the most notable cultural groups in size and complexity is the Maya. The archaeological Maya presence in the region lasted for several periods over thousands of years (~1800 BC to 1500 AD) (Ibarra-Rivera et al. 2008; Sharer & Traxler 2006). Currently, Maya descendants still reside in areas of present day Central America.

In the Maya region, migration appeared to be a quotidian cultural experience that occurred across various time periods using inland and coastal migratory routes (Cucina 2014a; Miller-Wolf & Freiwald 2018; Ortega-Muñoz et al. 2019). Migrants were present in all societal levels, including elites and commoners (Ortega-Muñoz et al. 2019; Price et al. 2014). Archaeological and biological evidence indicates movements largely occurred within boundaries constructed by community ties with larger political centers. These centers were large complex sites, generally regarded as city states, or central polities. Each polity had an intricate relationship built on economics and power alliances with outlying communities and other city states (Cucina 2014b: v). Alliances were often extensions of these networks (Martin & Grube 2000). Because the ancient Maya society was stratified into a social and political hierarchy (Sharer & Traxler 2006), policies on assortative mating might dictate and direct gene flow within and among people. This might especially be the case for the elite Maya due to these economic and power relationships. Martin and Grube (2000:21) illustrate the complexity of site relationships in the Maya region based on hierarchical and kinship social components and their associated power dynamics.

Colonization by the Spanish

During colonization, boundaries, society, and populations were reorganized in the wake of the arrival of Europeans and Africans in the Americas. Some Maya populations fled from Spanish rule and culture, with several groups moving toward the central basin of the Yucatán, Belize and Guatemala (Rice & Rice 2005). Meanwhile, other groups capitalized on opportunities for gaining wealth and power in the areas of Spanish control (Alexander & Kepecs 2005). A more nuanced interpretation of colonial life in the region was elucidated from a large-scale study of a cemetery population (n=180) from the colonial town of San Francisco de Campeche (1540 A.D.). Located in the Yucatán Peninsula in Mexico, this was the first municipality established by the Spanish and served as the main shipping port. The city of Mérida, slightly more inland, was established as the capital of the region (Tiesler et al. 2010). Early inhabitants of San Francisco de Campeche included Indigenous Maya, Spaniards, and enslaved African. Africans were forcefully brought to the Yucatán region during conquest, which increased after the demographic collapse of Indigenous populations, for manual labor (Zabala 2010). Portuguese slavers dominated the slave trade and had well-established human-trafficking networks that extended deep into the African continent, so enslaved peoples were likely taken from several African countries (Zabala 2010). This implies variation and diversity among the African individuals brought to the Americas. If this were the situation, homogeneity among African populations in the Americans cannot and should not be assumed (Spradley 2006). Additionally, the colonizers brought disease to the New World (Ubelaker 1994). Smallpox took a devastating toll on the Indigenous populations (Tiesler et al.

2010). Demographic decline of the native population from foreign pathogens greatly reduced Indigenous genetic diversity and was a catalyst for cultural change in subsistence infrastructure (Alexander & Kepecs 2005; Ongaro et al. 2019) possibly leading to regional morphological variation.

Caste War and Social Tensions

Even though several states flourished under Spanish colonial rule, it created inequality and triggered social tension between populational groups. In the mid-1800s, several years after independence from Spain, inequality between social elites, many of whom were descendants of Spanish colonizers, and rural, mostly Indigenous populations remained commonplace in Mexico (Gabbert 2019). By 1848, tensions between social classes erupted into a large-scale conflict in the Yucatán Peninsula, referred to as the Caste War. While this conflict is often interpreted as a symbol of Maya resistance to colonial rule, the conflict was much more complex with Maya and non-Maya participants on both sides. Again, populations were reorganized as 'pacifist' Yucatec Maya groups fled south while revolutionary groups, like the Santa Cruz Maya remained embattled until the early 1900s (Cal 1983). The conflict was very violent, bloody, and nearly ousted the ruling class from the peninsula (Joseph 1985). Afterward, the ruling class enacted repressive social and political strategies to ensure their place was not challenged again. Yucatec Maya refugees fled to the northern part of the peninsula and neighboring Belize (Gabbert 2004). Many of the individuals in the UADY sample are Yucatec Maya and could include descendants of refugees from this conflict.

Contemporary Language and Culture

The Maya people are still present in the region numbering over seven million individuals (Ibarra-Rivera et al. 2008). Although generally described as a single culture, contemporary

Indigenous Maya groups are culturally and linguistically quite diverse. Today there are over 28 different Mayan dialects spoken in the region (Sharer & Traxler 2006), and the abundance of dialects is attributed to isolation by distance, conflict, migration, and political systems (Coe 1999). In parts of Mexico and Guatemala, cultural processes like ‘ladinization’ can also impact biology. Ladinization is the adoption a mix of native and European cultural elements like diet, dress, and language into the already present Indigenous culture; a process specific to Guatemala and adjacent regions in Mexico, Honduras, and El Salvador (Adams 1994). Several studies by Malina and colleagues (1981; 2008a; 2008b) and Little and colleagues (2006) on rural populations in Oaxaca demonstrate the significant impact cultural change can have on the physical bodies of people in rural communities in the region, specifically ladinization.

Biological Sources of Sample Variation

Genetic Variation in Latin America

Genetic analyses in the region characterize populations in Central and South America as having genetic components from three main source populations: West African, European, and Native American (Bryc et al. 2010; Rubi-Castellanos et al. 2009; Wang et al. 2008). The proportion of genetic material from these broad groups varies across geographic regions within the Americas, largely dependent on several factors, such as the size of Native American groups in the region prior to European contact, the rate of displacement of these groups by European settlers, the presence of enslaved Africans in the region, and the timing of arrival and size of these enslaved populations (Bryc et al. 2010). Ongaro and colleagues (2019) explain genetic variation as resulting from two processes: 1) a sharp decline in Indigenous populations due to genocide and disease, and 2) gene flow that occurred during and after European colonization. Independent analyses found varying proportions of genetic ancestry across the region; in all instances the African contribution is relatively low, apart from coastal and island populations near the Caribbean (Bryc et al. 2010; Wang et al. 2008). Rubi-Castellanos and colleagues (2009) examined the genetic make-up of a Mexican Mestizo sample and found a directional North-South gradient structure in which European ancestry varied inversely with Native American ancestry across the country from the north to the south. Almost all studies note significant variation of ancestry contributions on the sex chromosomes. The Y-chromosome contributions are almost exclusively European, while the X-chromosome contributions vary between Native American and African indicating gene flow between European males and women from Native American and African groups (Bryc et al. 2010; Ongaro et al. 2019; Wang et al. 2008).

A refined study on present day Maya in the Yucatán Peninsula, specifically in Mexico, and the Guatemalan Highlands describes genetic variation between Maya descended groups. The study included STR loci from Maya individuals in the states of Campeche and Yucatán (Mexico) and the Maya groups K'iche and Kakchikel in Guatemala (Ibarra-Rivera et al. 2008). Data from this study were also compared with previously published genetic data from Maya and non-Maya Indigenous groups in North America, El Salvador, and Colombia. Results indicate a higher genetic diversity in the Mexican Maya groups when compared to the Guatemalan Maya groups, suggesting more movement within the Yucatán and interactions with other groups in the region like the Olmec and other non-Maya groups. The Guatemalan highland sample showed less genetic diversity, which Ibarra-Rivera and colleagues (2008) attribute to limited gene flow from non-Maya groups. On a larger scale, Maya groups included in this study were more like Maya samples from El Salvador and less like non-Maya groups in the comparative dataset. This suggests Maya relationships based on culture were maintained despite large distances, conflicts, and changing political structures (Ibarra-Rivera et al. 2008). Because genetic analyses in the region demonstrate differing degrees of continental proportions and patterning among populations considered Hispanic and even the Maya, and cranial shape and form is highly heritable, the expectation is that variation should be visible cranially.

Expected Variation in Research Samples

This study contains biased samples towards a poor Maya-descended group in Mérida, Mexico and a forensic sample, likely composing individuals involved with organized crime in Guatemala (see **Chapter 4: Methods**). I reiterate that these samples are not representative of the Yucatán region or Mérida proper, nor the country of Guatemala, especially as the INACIF sample comes from individuals in all departments within the country. However, these data can be used to explore human skeletal variation. These data will be used to create and test population affinity models, identifying areas for improvement in future research. Hypotheses of expected variation and the reasons behind said variation are below:

Hypotheses & Expectations

H1: There will be measurable differences between the Guatemalan and UADY sample.

Highland groups may be represented in both the INACIF or Identified Guatemalan Migrant sample since both could include individuals from across the country. The UADY sample likely contains lowland Maya individuals based on historical events and location. Building on conclusions of Ibarra-Rivera and colleagues' (2008) genetic study that Maya-derived groups, or populations descending from the archaeological Maya are likely to be more like each other than non-Maya groups, I expect my samples with Maya descended groups to be more like each other than non-Maya samples. However, isolation of Maya highland populations from the lowland Maya populations could exacerbate craniofacial differences between the two groups. Additionally, many Guatemala Migrants come from the Western highlands (Grandin et al. 2011; Smith 2006) so I expect to see variation between my samples of the lowland (UADY) and highland Maya (potentially INACIF and Identified Guatemalan Migrants) groups. This research will test the significance of these differences to understand if they can be used meaningfully to understand human variation and serve as a starting point for future investigation into variation in Guatemala and the Yucatán Peninsula.

H2: There will be measurable differences between the Mexican migrant sample and the UADY sample.

Genetic and skeletal studies suggest high diversity within Mexico that is attributed to the different cultural and historical processes occurring in other regions of Mexico (Hughes et al. 2013; Rubi Castellanos et al. 2009). There are many Indigenous groups within Mexico and, during colonization, Europeans and enslaved Africans arrived at certain ports on the Atlantic and travelled inland. The variation of genetic contribution within Mexico largely depends on the size of Indigenous groups in the area, the arrival and number of Europeans, and the presence and number of enslaved people brought to the region (Bryc et al. 2010). The arrival of these groups on the East side of Mexico combined with gene flow between Europeans, Africans, and Indigenous populations, altered the population structure of groups in areas of heavy contact, so populations away from ports of European entry during colonization will be genetically and phenotypically different than the UADY sample. In fact, Mexican migrants to the U.S. do not typically come from Mérida or the Yucatán, but can originate from many regions in Mexico. According to data gathered from multiple sources, in 2004 – 2015, most migrants crossing the U.S. border from Mexico come from central and norther states, as well as Chiapas. Data from the Yucatán Peninsula indicated a very low number (n = 300) of Mexican migrants originating from the Yucatán state (Migration

Policy Institute 2022). I expect that individuals in the Identified Mexican Migrant sample to be different morphologically from the UADY sample.

CHAPTER 3: CONTEMPORARY MIGRATION TO THE U.S. FROM MEXICO AND CENTRAL AMERICA

Contemporary migration across the Mexico-U.S. border is a product of several decades of the flow of people, capital, and ideas within the region. Sassen (2011) likens migration in the region as movement along a chain, firmly implanted within the larger political, economic, and social structures of the U.S. and Mexico. Movement along the chain's links was well-established prior to the formation of the present-day geopolitical boundary, owing to migration's complexity. The movement of people and goods have adapted and responded to larger changes through time, transforming our current conceptualization of migration. In fact, change in this region is constant, shaping the *who*, *what*, *when*, *where*, and *why* of migration. Major shifts in political, economic, and social spheres are reflected in policy and social attitudes toward migration and migrants.

When characterizing migration into eras, scholars use perspective as the dependent variable on which to base a timetable. A U.S.-centric perspective often marks periods using changes in U.S. immigration policy and economic/foreign relationships with Mexico (Massey 2011). Alternatively, Mexican-centric perspectives divide migration epochs into time periods corresponding with economic policies and operations originating in the U.S., but ultimately implemented by Mexican elites (Gonzalez 2011). In both instances, the tendency is to focus on the economic/political milestones from the country of perspective. A comprehensive approach involves using several perspectives to characterize migration and understand migration as a 'genealogy' with multiple origins, branches, and overlapping histories (Overmyer-Velázquez 2011a).

This chapter will discuss the development of migration to the U.S. from Mexico and Central America. They are discussed separately as the roots of current migration developed from different causes, but ultimately merge in the forensic context of migrant deaths at the border. Next, I will identify theoretical perspectives from sociocultural anthropology that can be used to examine and explain migration and the death of migrants at the border. I will conclude with a summary of the forensic context at the border and a discussion of the challenges to identification of undocumented migrants.

Why Do People Migrate?

Over the last century, migration from Mexico was largely influenced by the relationship between the U.S. and Mexican economy and fluctuations in the economic market (Gonzalez 2011). After working in the U.S. many migrants returned to Mexico more financially stable. This encouraged other individuals to migrate for economic security. The pattern of working in the U.S. and returning to Mexico was cyclical in nature and continues into the 21st century (Overmyer-Velázquez 2011b). Migration to the U.S. from Central America initiated under different circumstances. Migrants from Central America primarily originate from El Salvador, Guatemala, and Honduras, three countries collectively referred to as the Northern Triangle. Here, migration is tied to a deep history of economic inequality, conflict, and violence (Martínez 2017a; Menjivar 1993). Inequality surrounding agricultural production, land ownership, and wealth sparked violence between the government and rural agriculturalists (Menjivar 1993), which, in El Salvador, contributed to a sixteen-year civil war that ended with a blanket 'peace for all' deal and impunity for many (Martínez 2017b). During the 1980s and 1990s, the Regan administration extended Cold War policies in the region, destabilizing 'Communist-like' governments and contributing to military-backed political coups d'états. These campaigns further exacerbated social inequality and led to extreme violence (Grandin et al. 2011). As a horrendous and on-going side effect, U.S. involvement in the region reinforced social inequalities, specifically against indigenous people born during the Spanish conquest and Colonial Period (Borger 2018). These inequalities had become intertwined with national identity during independence and reemerged during the 20th century (Paley 2018). Currently categorized as ‘weak states’, countries in the Northern Triangle are unable to provide core functions in security, capacity, and legitimacy for their citizens (Tyagi 2012) and are characterized by unregulated violence, low or stagnant economic growth, and some of the most impoverished people on the continent (Bialik 2019). Furthermore, the exportation to El Salvador of gangs originally formed in the U.S., like MS-13, has exacerbated the state's inability to provide safety for its citizens and operate free from corruption (Martínez 2017a). Essentially, after the peace agreement, criminal violence replaced political violence, which mimicked wartime brutality against the civilian population (Martínez 2017b).

In Guatemala, internal struggles initiated by a long, violent civil conflict and perpetuated by racial disparities culminated in emigration from the country (Jonas 2013). For 36 years, Guatemala was embroiled in the longest and arguably the most violent civil war in Central America (Jonas 2013). The Guatemalan Civil War (1960-1996) has been divided into two phases, each with different implications for emigration. During the first phase (1966-1968), political emigrants, largely professionals and middle-class elites, fled to nearby Mexico. During the second phase (1968-1996), the military junta shifted targets from larger metropolitan areas to the Western Highlands, which were populated by Indigenous Maya (Jonas 2013). The goal was to weaken the guerilla fighters by severing their rural community support (Grandin et al. 2011). Here, the military employed a scorched earth policy, intentionally targeting Indigenous Maya descendants and their communities with extreme and unconstrained violence. Over 600 massacres against the civilian population occurred in a span of two years with the most violent attacks occurring in northern Huehuetenango, Quiché, Rabinal, and Baja Verapaz (Grandin et al. 2011; Jiménez 2011). Under this violent campaign, approximately one million Maya were displaced and migrated to bordering Mexican towns, some continuing to the U.S. as *de facto* refugees (Grandin et al. 2011; Jonas 2013). The violence disrupted indigenous economies including agriculture, commerce, and trade (Grandin et al. 2011), which were further exacerbated by a series of natural disasters, including three hurricanes and an earthquake in the 1990s (Jonas 2013). Despite playing a critical role in the conflict and economic hardships in Guatemala, the U.S. has not granted Temporary Protected Status to Guatemalans, as they have for other nearby countries (El Salvador, Honduras, and Nicaragua) (Jonas 2013). Past and present emigration from Guatemala is the result of the combination of political and economic factors. Post-war economic and political conditions in Guatemala have not improved (Morrison & May 1994). The country is unable to care for and protect citizens, and the depressed labor market does not provide job security nor financial stability. A large wealth disparity separates the urban elite and rural agricultural laborers, essentially preventing upward mobility through structural barriers. Migration is often seen to overcome this hurdle for those who can finance the journey (Jonas 2013). Migration has also created economic opportunities for rural Indigenous Guatemalans. In Huehuetenango, migration to the U.S. has created an intense housing boom, in which social status is based on remittance economy (Grandin et al. 2011). Guatemala is still plagued by uncontrolled social violence, largely a remnant of the civil war. Violence exhibited by drug traffickers, organized crime, and clandestine paramilitary groups is largely reflective of the brutality exercised by the military during the conflict (Jiménez 2011) and can involve coercion of entire communities (Martínez 2017a). As many of the perpetrators of the violence during the Guatemalan Civil War were not held accountable, impunity plays into modern day violence (Jiménez 2011; Martínez 2017a). A cost benefit migration model showed that in all departments, or regions, of Guatemala, violence was an important factor influencing migration, especially during the 1970s and 1980s (Morrison & May 1994). Survey data also indicated a poor post-war economy was influential when considering migration (Jonas 2013).

Additionally, El Salvador, Honduras, and Guatemala sit in a strategic location along a drug corridor from Colombia to the U.S. where narcotraffickers and powerful drug families control all aspects of these states including politicians, judges, and police officers (Martínez 2017a; Martínez 2014). As these states are weak, they fail to protect their citizens from coercion into the drug trade, gangs, and violence. The homicide rate in El Salvador is more than 10 times that of the U.S. (OSAC 2020). The borders between El Salvador, Honduras, and Guatemala are especially dangerous, as they are completely controlled by narcotraffickers (Martínez 2014). With safety and poverty as driving issues, people are often forced to flee to the U.S. or stay in a violent and unstable country (Martínez 2017b).

A Very Brief History of Migration at the Southern U.S. border

The Mexico-U.S. border, in its current conceptualization, is relatively new. Prior to the mid-1800s, the Mexican territory included much of the western half of the North American continent, including most of the western states in the U.S. After the Mexican-American War (1846-1848), under the provisions of the *Treaty of Guadalupe Hidalgo* (1848), half of the Mexican territory, including the present day states of Arizona, California, Colorado, New Mexico, Nevada, Texas and Utah, were ceded to the U.S. (Overmyer-Velázquez 2011a). Movement of people in this area had been well-established prior to cessation of the territory, so the newly created geopolitical border only served to intensify migratory flows. This is particularly true between the states of Sonora and California and along the newly established Northeastern Mexican border and Texas (Mora-Torres 2011). Mora-Torres (2011) credits the California Gold Rush with the initiation of cyclical migration; early migrants returned home more financially stable than when they left, encouraging others to follow during the next cycle.

While the U.S. economy was flourishing, the Mexican economy was stagnant. The Porfirio government rule, under Porfirio Diaz (1876-1911) coincided with large-scale immigration to the Americas from Europe (Mora-Torres 2011). The political elites expected large numbers of immigrants to also come to Mexico, stimulating and expanding the flailing economy. Simultaneously, the elites viewed the indigenous labor force as inferior and damaging to economic prosperity, even though Mexico had the lowest wages on the continent. Taking advantage of proximity to the U.S. people in the northern Mexican border states, traveled to the U.S. where the same economic opportunities were much more financially prosperous. As European immigrants settled elsewhere, the Porfirio government opened Mexican borders to U.S. companies, relying on U.S. economic success to enhance the Mexican economy. U.S. investors flooded the market and grabbed land and resources. Gonzalez (2011) likens American involvement in the region as imperialistic, treating Mexico as an American colony. Diaz and other elites permitted this relationship through construction of American rail lines into the country. This action drastically altered the Mexican economy and migration, the ripple effects of which are still present today. Rail lines cut through indigenous farmland allowing American companies to extract precious resources like copper and silver. More notably, the railways disrupted the traditional farming lifeways indigenous people had practiced for centuries, uprooting them from the countryside and forcing them to migrate to cities for work. This shift restructured the country's demography as people migrated to northern Mexican states to work in American factories (Mora-Torres 2011). U.S. capitalist expansion caused the initial 'push', forcing laborers to move closer to the U.S., so, when the ‘pull’ for cheap, seasonal labor came from the U.S. economy, laborers were nearby (Gonzalez 2011).

The cycle of migration ebbed and flowed in sync with the demands of the U.S. economy. During times of economic downturn or depression, Mexicans living and working within the U.S. were deported *en mass* and immigration laws were enacted to restrict the entrance of foreign workers (Overmyer-Velázquez 2011b). The Bracero Program (1942-1964) had a large impact on Mexican migration to the U.S. The initial goals of the program were to accommodate the post-war labor shortage by contracting Mexican laborers, called *Braceros*, to work in the agricultural sector. During the program’s tenure, 4.6 million bracero contracts were active: the largest importation of foreign labor in U.S. history (Overmyer-Velázquez 2011b). The Bracero Program drastically impacted local communities in Mexico, as laborers sent remittances to their families, transforming the material and health conditions of their home communities (Malina et al 2008). At the same time the Bracero Program was importing labor, other legal initiatives were restricting foreign immigration. *Operation Wetback* (1950-1954) aimed to curb illegal immigration of Mexicans into the U.S.; however, many people who had migrated legally were deported. While the Bracero Program officially ended in 1964, undocumented migration to the U.S. continued, while policy aimed at restricting migration increased. In 1965, numerical limitations were placed on legal migration; however, with the cyclical nature of migration tied so tightly to the U.S. economy, clandestine migration continued (Massey et al 2014). To reduce undocumented migration, the Immigration Reform and Control Act (1988) penalized business that hired undocumented migrants and increased the budget for Border Patrol. Massey (2011) marks this event as the initiation of modern militarization of the border.

During the 1980s, discourse changed the border from a physical land boundary to a militarized zone of conflict, and the narrative surrounding migration changed from an economic issue to one of national security. Cloaked as policies protecting the state and its citizens from a foreign enemy, border security became a nation-building-tactic (Dunn 1996). Characterization of the southern international border as a *Low-Intensity Conflict* (LIC) zone allowed implementation and enforcement of stricter border laws with a militaristic flavor (De León 2015; Dunn 1996). These policies were carried out by border patrol agents, often with military training. Operations were, and remain, performed in conjunction with the military. Under LIC, action against a perceived threat to national security was allowed and condoned by the state. The word action is fluid depending on what the 'authority' deems a threat and the means necessary to quell the threat (Dunn 1996). In 1994, the North American Free Trade Agreement (NAFTA) was enacted, removing tariffs and other restrictions on agricultural products among Canada, Mexico, and the U.S. With government subsidies, American products rapidly overtook Mexican markets, displacing a vast number of Mexican laborers (Martínez et al. 2014). NAFTA significantly altered the economy of rural communities and forced individuals to leave in search of work. By the 1990s undocumented immigration to the U.S. was increasingly common, due to the continued labor demands of the U.S. economy and the prolonged economic crisis in Mexico (Overmyer-Velázquez 2011b).

Migration and border security changed drastically in the 1990s. In 1993, chief border patrol agent, Silvestre Reyes, was faced with complaints of border patrol agents harassing Latino citizens while searching for undocumented migrants in El Paso, TX. Reyes initiated a new tactic under *Operation Blockade*, where the city was flooded with border patrol agents, forcing migrants to cross away from metropolitan areas on the outskirts of the city. This strategy served to make illegal migration less visible, while creating a scenario where policing of undocumented migration was also out of sight (De León 2015). Touted as a success, this strategy was adopted by politicians in the Clinton administration, quickly spreading along the Southwest border. Wholly referred to as *Prevention Through Deterrence,* this tactic had two goals: 1) redistribute targeted resources (people and equipment) at specific border stations and 2) 'discourage' clandestine migration by shifting migration to rural, more dangerous routes away from urban centers (Martínez et al 2014; Eschbach et al. 2003). The argument centered on the ability of the Border Patrol to easily apprehend migrants. Using these tactics, initiatives like *Operation Gatekeeper* (1994) in San Diego, *Operation Safeguard* (1995) in southern Arizona, and *Operation Rio Grande* (1997) in South Texas targeted migration routes in urban areas, forcing people to choose more rural, remote crossing routes (De León 2015; Eschbach et al. 2003; Martínez et al 2014). Under PTD, the federal government boosted resources in these areas, increasing border patrol agent presence, technological resources, and constructing physical barriers or walls. Customs and Border Patrol (CBP) archives commend these initiatives as successful at reducing clandestine migration (CBP 2018); however, others argue the programs were ineffective (Eschbach et al. 2003; Kovic 2018), and only served to increase the number of dead along the border (De León 2015).

Slack and colleagues (2016) argue that previous administration border policies profoundly influence future policy decisions. Mimicking the PTD initiatives, the number of border agents doubled and tripled in some areas in the mid-2000s. Congress increased border security spending by millions of dollars with budgets in the low trillions (Slack et al. 2018). There are nine sectors along the southwest border, each guided by their own CBP culture. Slack and colleagues (2018) further argue that this culture emphasizes pain, suffering, and trauma and is used as a deterrent. While the actual policies and practice vary across regions, border patrol culture in each sector is linked using violence as an enforcement strategy. This culture is embodied by living, deceased, and disappeared migrants and can influence forensic investigations (Gocha et al. 2018; Slack et al. 2018; Spradley et al. 2019).

A Shift in Migrant Demography

Reflective of the history of migration along the U.S.-Mexico border, early migrant demography comprised young to middle-aged Mexican males with strong familial ties to migration within established migration networks (Massey et al. 2014). This demographic journeyed for economic reasons and were the target of the Bracero Program. After the program's dissolution in 1964, middle-aged Mexican males continued to migrate clandestinely for economic reasons. Changes in border policy effectively closed the border, curtailing cyclical economic migration. Migration from the Northern Triangle, albeit at a much smaller level, began in the 1980s and 1990s because of political violence and economic instability. In 2014, the number of Central Americans apprehended clandestinely crossing the border surpassed the number of apprehended Mexicans (Gonzalez-Barrera & Krogstad 2019). Individuals from Northern Triangle countries are more likely to be apprehended in Texas, reflecting migratory routes (Isacson et al. 2013). More recently, family unit apprehensions have outnumbered individual apprehensions (Gramlich & Noe-Bustamante 2019). Family units, or individuals traveling together that include a child under 18 years of age and a parent or legal guardian, outnumbered apprehensions for adults traveling alone and unaccompanied children (CBP 2018; Gramlich & Noe-Bustamante 2019). In 2018, 95% of family apprehensions comprised Salvadorans, Guatemalans, and Hondurans (Bialik 2019). In recent years, Guatemalan migration to the U.S. has included more women, even though the journey is much more dangerous (Jonas 2013).

Statistics from postmortem examinations collected from medical examiners offices reflect CBP apprehension data on migrant origins. Data from the Pima County Office of the Medical Examiner (PCOME) in Tucson, Arizona indicates most undocumented migrants that die in Arizona are from Mexico (Anderson 2008). Identification data from Operation Identification (OpID) at Texas State University also supports CBP evidence that Central Americans and Mexicans are crossing through Texas (Gocha et al. 2018). These examples show the utility of using CBP statistics in forensic research. Information on *who* is apprehended *where* can be compared to death data to examine migration routes used by different groups (Vogelsberg 2018) and inform methodological developments (Spradley 2014a). While this data is valuable, improvement in forensic methodology at the border is possible. One major hurdle is a lack of reference data for groups involved in migration (Spradley 2016a; 2021). Biological data used in migrant identification is lacking for Central American populations and large regions of Mexico. Unfortunately, data collection in these areas is often not an option as travel in the region is unstable.

Theoretical Perspectives on U.S.-Mexico Migration

Migration and the death toll from clandestine migrant crossing at the Mexico-U.S. border is a humanitarian crisis and has been categorized as a silent mass disaster (DeLeón 2015; Reineke 2016; Goldsmith & Reineke 2010; Martínez et al. 2014; Spradley et al. 2019; Spradley 2021). Several theoretical frameworks grounded in sociocultural theory examine the relationship of inequality and violence directed at marginalized groups in different contexts. Migration lies at the intersection of race, politics, the economy, and society, so it includes relationships that entail power and violence. These sociocultural lenses can be used to explore the migration crisis in more detail. In the Americas, migration is a product of our collective history, deeply embedded in the formation of current geopolitical power structures and national identities (Reineke 2016). A brief literature review identified Critical Race Theory (Crenshaw et al. 1995), Structural Violence (Galtung 1969), the State of Exception (Agamben 1998), and Necropolitics (Mbembe & Meintjes 2003) to be most prevalent when discussing violence and death at the U.S.-Mexico border. Looking at migration through these perspectives allows us to understand *why* and *how* deaths continue to accrue at the border and allow us to understand *where* and *how* forensic anthropology can change the narrative and facilitate the identification process.

Migrants clandestinely crossing the U.S.-Mexico border are living in the margins of the U.S. Here, the interaction between migrants and the State are constantly impacting relationships at the physical border. These margins are in a constant state of flux, continually (re)shaped through actions by the marginalized groups and reactions by the State/sovereign, or ruling body, which can often lead to state-sponsored violence and human rights abuses (Das & Poole 2004). Migrants are considered foundational to the creation and reinforcement of a U.S.-national identity narrative of who belongs and who does not (Reineke 2016). However, migrants are simultaneously excluded from invoking the identity they helped to create, reinforcing who does not belong (Das & Poole 2004). The reaction to those that do not belong, migrants, by the U.S. state has been a steady, militarized, and more restricting approach to border security (Dunn 1996). As the space for migration diminishes, migrants, narcotraffickers, and other groups participating in criminal activity are thrust into the same physical space and the distinctions between the groups are purposely blurred (Martínez 2017a). The close interaction of these groups with each other and governmental actors, like border patrol, continually (re)shapes and challenges the narrative of migration on both sides of the border.

Several scholars have used Structural Violence to describe indirect violence that is built into social structures (cultural, economic, religious, legal, and political), preventing people from meeting their needs (Farmer 2004; Farmer et al. 2006; Galtung 1969; Klaus 2012; Rylko-Bauer & Farmer 2016). This theoretical perspective has been used to contextualize border deaths (Kovic 2018) and describe the physical embodiment of marginalization (Beatrice & Soler 2016; Beatrice et al. 2021). Critical Race Theory (CRT) builds on structural violence, arguing that direct violence and systems of structural violence originate colonialism and imperialism. Violence here stems from racism and is directed to hurt people of color (Crenshaw et al. 1995). When speaking specifically about migration at the U.S.-Mexico border, Reineke (2016) argues that structural violence is not able to capture the specific historical context and social conditions that lead to migrant death, arguing that they are better explained using CRT due to the role that race and racism play in construction of barriers. CRT can extend into the realm of scientific investigation, where noncritical information plays a role in the outcome of scientific processes (Dror et al. 2021). A lack of population-specific methods for construction of the biological profile in migrant groups can potentially hinder identifications (Spradley 2008; Spradley 2016a). Through directed studies on migrant remains, like those conducted at OpID and the PCOME, results are generated to break down biases in methodology and subsequent identification.

Many sociocultural studies acknowledge that current policy and treatment of migrants and migrant remains stem from the colonial and racial past of the U.S. and European countries (Martínez et al. 2014; Reineke 2016). CRT can be used to understand the causes of migration (i.e., political and economic inequality, violence), the reaction to undocumented migration, and why people continue to clandestinely migrate to the U.S. (Reineke 2016). Economic migration from Mexico began in the 1800s, because of collaborative efforts between American companies and Mexican elites to exploit the country's resources for profit (Gonzalez 2011). The transition of indigenous farmers to a large, mobile cheap labor force was noted by visitors to Mexico's borderlands. Gonzalez (2011) describes written accounts of American tourists, academics, journalists, and missionaries describing Mexico and Indigenous laborers as “incapable of modernization without foreign assistance” (Gonzalez 2011:28). This statement exemplifies colonial hierarchical thought, where American travelers believed Indigenous Mexicans were inferior to North Americans (Europeans) and Mexican elites and wrote about them as such in written accounts of the country. In Central America, racism fueled economic inequality and the civil conflicts of the mid-20th century (Paley 2018). Regarding migration, Reineke (2016) explains how CRT describes the transformation of migrants into ‘illegals’ and creates targeted policies that have led to uncountable deaths and a generalized apathy toward them. The construction of who is allowed to be killed without repercussion ties into the State of Exception (discussed below). Additionally, CRT plays a role in the unequal policies regarding death investigation of suspected migrant cases. Once in the forensic realm, the Inequality of Identification highlights the uneven use of resources used in the death investigation process, arguing more expensive techniques can be used but are not because of the socio-economic or citizenship status of the deceased (Bartelink 2018; Spradley et al. 2019). Additionally, those charged with death investigation do not always adhere to legislated protocols in suspected migrant cases (Gocha et al. 2018; Spradley & Gocha 2020). Misinterpretation of the law in some South Texas counties resulted in no autopsy, no skeletal analysis, and consequently, no DNA samples were taken, leaving these individuals with no opportunity with even a chance to be identified (Gocha et al. 2018).

The State of Exception, as discussed by Agamben (1998), examines the sovereign’s power to allow life or death within a political sphere. The State of Exception describes a juridically empty space where the sovereign authorizes violence as a response to an exception. In this space, rights are suspended allowing for the emergence of *homo sacer*, a person who is allowed to be killed without punishment. Regarding migration, authors have described migrants as existing in a State of Exception (De León 2015; Vogt 2018). Migrants are not U.S. citizens and therefore not granted the same rights that come with citizenship. Over the years, the border has become militarized (Dunn 1996), and the language toward migrants reflects this militarization. Migrants are often spoken about as 'threats' or 'enemies of the state', creating a narrative that supports a manufactured emergency. Once threats are identified, the state can respond in the way it deems justifiable. Regarding migration, the state's responses have been purposely restricting migration routes to cross inhospitable terrain like the Sonoran Desert in Arizona with the knowledge that death will be a common side effect (GAO 2012; De León 2015). Instead of a display of direct physical violence, De León (2015) argues the state strategically uses the landscape to do the killing, intentionally out of view of the public, thus absolving the government of any wrongdoing. This practice extends beyond the U.S. to main transit routes within Mexico (Vogt 2018). Consequently, migration has overlapped with criminal enterprises in the same space (Martínez 2017a), further exacerbating bare life and exposing migrants to violence, especially along the journey from Central America.

The state's act of killing or allowing certain deaths to occur can be examined through another theoretical perspective of Necropolitics. This perspective builds on Foucault's concept of Biopower and describes the sovereign's ability to control who lives and who dies through various forms of power (Mbembe & Meintjes 2003). The state often uses politics to justify violence and protective action from a perceived enemy, therefore controlling life and death under this umbrella. Through restrictive policy directives touted as 'reducing' or 'curbing' migration, the state channels migrants through a dangerous, hostile environment, controlling *where*, and *how* migrants die (De León 2015; Martínez et al. 2014). The use of Necropolitics within the border region is not unique to the U.S. Magaña (2011) explains that the Mexican government employs Necropolitics to reaffirm its authority over the borderlands where it has lost control. Reineke (2016) argues that Necropolitics at the border is affected by racism, as race is the main factor in determining who is considered disposable by the state. De León (2015) includes Necroviolence, which is specific treatment of bodies that is meant to inflict violence through pain and suffering. Postmortem treatment is meant to offend the victim and the cultural group to which they belong (De León 2015:69). Examples of Necroviolence to migrant bodies include haphazardly piling the deceased into graves, disregarding cultural considerations for burial, not properly marking migrant graves for later identification, and placing trash or medical waste into the burials (Bemiss et al. 2020; Spradley & Gocha 2020). De León (2015) states the most egregious form of Necroviolence is the destruction and disappearance of a corpse. At the border, the geographical remoteness, harsh arid climate, and presence of scavengers can erase a body in a matter of days (De León 2015). Without physical evidence of their death, migrants are erased, and their families are suspended in a state of not knowing what happened. This can have long lasting emotional and mental effects on living family members, friends, and migrant communities (De León 2015; Reineke 2016).

Forensic Science Along the Border/Effect of PTD on Migration

The U.S. government labeled the PTD campaign a success: the program seemed to deter undocumented migration; however, the number of deaths along the border skyrocketed as a direct result of this policy (De León 2015; Parks et al. 2016). As migrant routes were pushed from metropolitan areas to more dangerous, rural areas, migrants were exposed to harsh environmental conditions and terrain, including intense desert heat. In Arizona, migration routes were directed toward the Sonoran Desert (De León 2015). The most common cause of death, as recorded by the

PCOME was directly related to high temperatures, classified as heat stroke or hyperthermia (Parks et al. 2016).

No standard practice for documenting migrant death on a large scale exists across all jurisdictions along the southern border (Gocha et al. 2018; Reineke 2016; Spradley et al. 2019). Reineke (2018:vi) states "[A] lack of an organized effort to count the dead, (and identification) indicates intentional ignorance and maintenance of certain blind spots on the part of the state". Furthermore, Texas does not have a centralized medical examiner system, leaving each jurisdiction responsible for death investigation (Spradley et al. 2019). To combat decentralization, two forensic institutions have reduced the gap between the missing and unidentified: the PCOME in Arizona and OpID in Texas. Prior to the year 2000, the PCOME received approximately 14 undocumented border crosser (UBC) cases per year. Since 2000, the PCOME reported a significant increase in migration-related deaths, which has fluctuated between 150 and 220 cases per year (Parks et al. 2016). In 2012, the number of migrant deaths in Texas surpassed those in Arizona; however, exact numeric data from Texas are unknown as death records are not kept in a centralized system. Furthermore, due to the large expanse of private land near the border, many deceased individuals are often never recovered (Kovic 2018; Spradley et al. 2019).

As a model of best practice, the PCOME adheres to protocol developed by Anderson and colleagues in the investigation of migrant remains, through the development of a UBC (undocumented border crosser) profile (Anderson & Parks 2008; Beatrice & Soler 2016; Birkby et al. 2008). UBC cases at the PCOME are treated with the same level of commitment as other forensic cases, within the bounds of methodology and available resources. In south Texas, several jurisdictional and bureaucratic hurdles challenge the identification and repatriation process (Gocha et al. 2018). To mitigate these challenges, OpID provides critical support in the identification process of migrants in south Texas counties. OpID takes charge of recovery, analysis, and DNA sample submission, all while liaising with external stakeholders. OpID even operates internationally with several non-governmental organizations working in migrants' home countries, which is outside of the purview of local law enforcement. To date, OpID has received 225 sets of unidentified remains from presumed migrants (Gocha et al. 2018). The forensic work undertaken by both the PCOME and OpID, importantly, counteracts decentralization and allows for a greater understanding of the number and magnitude of lives lost along the border, even though numbers represent an underestimation of the full scope of the crisis (Crossland 2013; Leutert et al. 2020; Soler & Beatrice 2018).

Despite the efforts by the PCOME and OpID, many individuals are still missing or remain unidentified. The South Texas Migrant Center reports 3,253 migrant deaths in southern Texas counties, but caution that this number is a gross underestimation due to the expanse of private land (Leutert et al. 2020). The Sheriff in Brooks County, Texas estimates that for every set of remains found, five are somewhere in the desert (Bemiss et al. 2020). Similarly, death data in Arizona is likely an underestimation due to the taphonomic effects on human tissue by the desert, literally disappearing bodies (De León 2015; Soler & Beatrice 2018). When bodies are recovered, the composition of the remains is different across geographic areas due to different taphonomic agents acting on the bones and soft tissue. Taphonomic analysis examines the postmortem processes that act on human remains and other organisms; specifically, in forensics, the events following death (Haglund & Sorg 1997). In Texas, when bodies of presumed migrants are recovered by law enforcement, they are typically buried in a designated location. As such, OpID exhumes complete bodies in varying states of decomposition. However, in Arizona, the desert heat, high temperatures, and scavenging often leads to extensive taphonomic damage in a short period of time, which affects the quality and completeness of the recovered skeletal elements and identification (Beck et al. 2015; Martínez et al. 2013). Data from the PCOME detail that 36% of border crosser cases that come through their office remain unidentified, which is, in part, a reflection of the destructiveness of the desert (PCOME 2017). Missing persons data collected from families of the missing by the Colibrí Center in Tucson report more than 3,500 open missing persons cases (Colibrí 2021; Reineke 2016). A cursory comparison indicates a larger number of missing persons cases compared to the number of individuals recovered, but even so, these are likely underreported.

The disconnect between recovered, identified, and missing persons at the border is likely reflective of several issues on both the forensic anthropology and reporting side. Spiros and Kamnikar (2021) note that cognitive biases within forensic methodology and reporting culture may influence who is identified and who is reported missing. In the case of missing migrants, researchers have identified several barriers that prevent families from reporting their loved ones to authorities (Gocha et al. 2018; 39). Some of these barriers include international status of family members and reporting and undocumented status of family or friends living within the U.S. For reporting a missing person outside of the U.S., consulates are involved in the process. Burnout and high staff turnover can lead to a lack of institutional knowledge (Gocha et al. 2018). The criminal justice literature also identifies underreporting in marginalized Hispanic communities due to distrust in the legal system, legal authorities, and immigration policy (Weitzer 2014). On the identification side, a lack of standardized protocol along the southern border, and shortfalls in forensic methodology can impact identification rates. For example, standard protocols in death investigation are not always followed leading to issues later in the process (Gocha et al. 2018). As the migrant demographic has shifted from Mexicans to include more Central Americans, one possibility could be the lack of reference data for non-Mexican individuals. Spradley (2016a) argues that a lack of understanding of Hispanic demographic, which is in part due to a lack of reference data for parent border crosser populations. In that same vein, research suggests that a reliance on the three-group model for ancestry estimation or a practitioner's will to be accurate instead of precise may hamper identification efforts (Spiros & Kamnikar 2021). These efforts are further complicated by the impact of taphonomy and damage to the remains and a lack of a centralized, international DNA databank for profile comparison (Spradley & Gocha 2021).

Conclusion

This dissertation directly impacts one of the challenges associated with the postmortem investigation of migrant deaths: more nuanced reference data for Latin American populations. This dissertation research aims to address the issues surrounding identification and forensic methodology with respect to population affinity estimation of Hispanic groups. Reference data from Latin America will contribute to the growing body of literature aimed at understanding human variation within this broad, diverse group.

CHAPTER 4: MATERIALS AND METHODS

Forensic anthropologists estimate the geographic origin of an individual using population-based approaches and skeletal data (SWGANTH 2013). A population affinity approach examines biological variation and its relationship to reference groups at a specifically defined level (Pilloud & Hefner 2016; Winburn & Algee-Hewitt 2021). While ancestry is the term currently employed in forensic anthropology reporting (SWGANTH 2013), population affinity more accurately describes what forensic anthropologists are trying to estimate, especially for migrant cases (Spradley 2021). Group variation under the heading Hispanic is poorly understood, due at least in part to a lack of comparative reference skeletal samples (Spradley 2016a; 2021). This project specifically addresses the lack of reference data for populations considered Hispanic by including data for two new reference samples.

Materials

Refining the Hispanic heading into more focused populational divisions requires data collection from multiple, diverse sources in Mexico, Central, and South America reflecting the history and culture of each individual. As accuracy in estimation of population affinity depends on the available reference data, adequate reference samples are required (Spradley 2016a; 2021). Currently, the limited samples in reference databanks are broadly applied to the entirety of countries within this region. No formal reference samples exist for individuals who consider themselves Ladino in either the FDB or MaMD, and there are no formal reference samples for cranial MMS trait data from Mexican populations. To address this gap, this research adds data from these underrepresented populations to supplement currently available reference data for groups considered Hispanic. The first phase of this project generates matched craniometric and cranial MMS reference data for two geographically proximate regions: 1) Mérida, Mexico in the

Yucatán Peninsula and 2) Guatemala (Table 4.1). Individuals from these countries importantly make up approximately 70% of the Hispanic population in the U.S. (Martinez & Castillo 2013) and include two of the top four sending countries for undocumented migration.

Craniometric and cranial MMS data are selected to assess craniofacial variation in relationship to population structure. These two data types are utilized in biological anthropology to answer questions surrounding group relatedness in cranial shape and form, including genetic inheritance and variation (Harvati & Weaver 2004; Relethford 1994; 2010; Roseman & Weaver 2004). Cranial MMS and craniometric data demonstrate variation corresponding to selective patterns in genetic variation (Betti et al. 2010; Relethford 2004; Reyes-Centeno & Hefner 2021). In biological anthropology, specifically in regard to biological distance and population affinity estimation, craniometric and cranial MMS data have: 1) demonstrated utility in geographic origin refinements beyond the continental level (Hefner & Byrnes 2020; Hefner et al., 2015; Kamnikar et al. 2021; Ross et al. 2014; Spradley 2014a; Tise et al., 2014), which make them ideal for studies aimed at refinement of broad groupings; 2) data collection methods are standardized and a variety of resources are available to guide practitioners (Dudzik & Kolatorowicz 2016; Fleischman & Crowder 2019; Langley et al. 2016; Plemons et al. n.d.; Hefner & Linde 2018); and 3) cranial MMS traits demonstrate low intra- and interobserver error between measurements and scoring when practitioners are trained prior to data collection (Kamnikar et al. 2018; Klales & Kenyhercz 2015). Furthermore, models using two biological data types are more accurate when establishing group membership (Maier 2019; Spiros & Hefner 2019).

Latin American Samples

The first reference sample includes individuals from the Mérida, in the Yucatán Peninsula region of Mexico. This sample is currently housed at the Universidad Autónoma de la Yucatan

(UADY) in Mérida (Chi Keb et al. 2013). These individuals were born in the 20th century, died in Mérida and surrounding communities, and were buried in the Xoclán Cemetery. Most individuals in this sample are from indigenous communities in and surrounding Mérida. After a two year period, if families are unable or unwilling to pay burial fees, remains are excavated by UADY and added to the sample. Craniometric (*n*=109) and cranial MMS (*n*=159) data were collected from the Xoclán Cemetery sample by the author and supplemented with craniometric data (*n*=59) previously collected by Dr. Kate Spradley of Texas State University (TXST) (Table 4.1).

The second reference sample is housed at the Instituto Nacional de Ciencias Forenses de Guatemala (INACIF) in Guatemala City and includes individuals recovered from various forensic contexts and likely includes individuals involved with organized crime in the country. The INACIF is the national forensic organization performing all medicolegal death investigations across the country. Because Guatemala is quite ethnically and culturally diverse, individuals in this sample come from several groups including ethnic Maya and Ladino groups. Investigation into similarities and differences among ethnic groups, specifically the Indigenous Maya and Ladino groups, is important and will be addressed herein. Craniometric (*n*=32) and cranial MMS (*n*=40) data were collected from the INACIF sample by the author.

Comparative Samples

Comparative samples of craniometric and cranial MMS data, from identified Guatemalan (*n*=12) and Mexican (*n*=24) migrants collected by the PCOME, Operation Identification (OpID) at TXST, and Macromorphoscopic (MaMD) Lab at Michigan State University are used. All individuals in these samples were identified through DNA analysis, which allows for the attachment of known demographic data corresponding to region of origin and sex to skeletal morphology. Case numbers from identified individuals with craniometric data are compared to case numbers included in the MaMD. All individuals with matched craniometric and cranial MMS data are selected for inclusion. A separate, unidentified migrant sample with matched data (*n*=155) from OpID and MaMD is included to explore relationships between known data and unknown individuals recovered from migration contexts.

Comparative reference samples are compiled from different sources to mimic the current U.S. demographic (U.S. Census Bureau 2019). Comparative samples include data from American Black, American White, and Thai samples. Data for the American Black and American White samples come from the Bass Donated Skeletal Collection at the University of Tennessee Knoxville. Individuals in this collection come from a body donation program in which demographic variables are known (Wilson et al. 2007). The Thai comparative dataset is from Khon Kaen University, Thailand. This collection constitutes modern Thai individuals who donated their bodies to the university (Techataweewan et al. 2017). Members of the MaMD Lab and the Khon Kaen Lab collected craniometric and cranial MMS data following standardized protocol. While Thai individuals or people with Thai heritage comprise roughly 350,000 people in the U.S. (Budiman 2021), this sample is included to test whether misclassification would occur between Hispanic groups and an Asian-derived group. Dudzik and Jantz (2016) addressed misclassification rates among groups under the broad Asian and Hispanic headings, finding that Thai males were the second least likely group for misclassification with a Hispanic male sample. Furthermore, distance scores between the two groups were intermediate compared to other Asian derived samples, supporting the use of the Thai data as a comparative dataset in this study. Table 4.1 provides sample information for each population group.

Table 4.1: Sample demographic of matched craniometric and cranial MMS datasets

Sample	*Population*	*M*	*F*	*Unknown*	*Total*
UADY	Merida, Mexico	114	54	–	168
INACIF	Broadly Guatemala	9	3	18	30
PCOME/OpID	Identified Mexican Migrants	22	2	–	24
PCOME/OpID	Identified Guatemalan Migrants	7	4	–	11
OpID	Unidentified Migrants	85	48	–	133
Bass Collection	American Black	32	6	–	38
Bass Collection	American White	46	25	–	71
Khon Kaen	Thai	150	111	–	261
				Total:	**736**

Methods

The second phase of this project analyses craniometric and cranial MMS data with the aim of identifying patterns and magnitudes of variation among the samples. Data are used to create classification models with other reference samples. Each biological data type is analyzed separately—craniometric, cranial MMS—and in conjunction. Analyses are performed separately on males and females, then pooled when appropriate.

Data Collection

Eighty-six cranial landmarks are collected from the Xoclán Cemetery and INACIF samples using a Microscribe® digitizer and the software 3Skull (v.1.76) (Ousley 2014). This program automatically calculates interlandmark distances (ILDs), or distances between cranial landmarks, while storing linear and coordinate cranial landmark data in Advantage Data Architect database. Data collected using 3Skull allow the user to include more measurements than the standard set of 24 ILDs. Expanded sets of ILDs have demonstrated higher accuracy when discriminating between diverse groups (Spradley & Jantz 2016) and have shown utility in population affinity estimates in migrant groups (Spradley 2014a; 2021). The ILDs used in this study overlap with data from all groups and are presented in Table 4.2. Seventeen cranial MMS traits are collected from the same samples using the MMS v1.61 program developed by Hefner and Ousley (2014) and are presented in Table 4.3. The MMS program contains standardized drawings and definitions for each character state, ensuring consistency in data collection.

Available demographic data (age, sex, birth location) are appended to all individuals after data collection. If the remains are unidentified, the individuals are categorized by the geopolitical country where the reference collection is located. Population structure does not necessarily conform to current geopolitical boundaries (Spradley 2021), but these labels are used as a first step in understanding variation.

Table 4.2: Interlandmark distances

Abbreviation	*Measurement*	*Abbreviation*	*Measurement*
GOL	maximum cranial length	XFB	maximum frontal breadth
BBH	basion-bregma height	ZYB	bizygomatic breadth
BNL	basion-nasion length	ASB	biasterionic breadth
XCB	maximum cranial breadth	OBH	orbit height
WFB	minimum frontal breadth	DKB	interorbital breadth
AUB	biauricular breadth	EKB	biorbital breadth
NLH	nasal height	FRC	frontal chord
NLB	nasal breadth	OCC	occipital chord
OBB	orbit breadth	MDH	mastoid height
PAC	parietal chord		

*Adapted from Fleischman & Crowder 2019; Langley et al. 2016; FORDISC 3.0 (help file).

Table 4.3: Cranial MMS traits

Abbreviation	*Trait*	*Character State*
ANS	Anterior nasal spine	1, 2, 3
INA	Inferior nasal aperture	1, 2, 3, 4, 5
IOB	Interorbital breadth	1, 2, 3
MT	Malar tubercle	0, 1, 2, 3
NAS	Nasal aperture shape	1, 2, 3
NAW	Nasal aperture width	1, 2, 3
NBC	Nasal bone contour	0, 1, 2, 3, 4
NFS	Nasofrontal suture	1,2,3,4
NBS	Nasal bone shape	1, 2, 3, 4
NO	Nasal overgrowth	0, 1
OBS	Orbital shape	1, 2, 3
PBD	Post bregmatic depression	0, 1
PZT	Posterior zygomatic tubercle	0, 1, 2, 3
SNS	Supranasal suture	0, 1, 2
TPS	Transverse palatine suture	1, 2, 3, 4
PS	Palate shape	1, 2, 3, 4
ZS	Zygomaticomaxillary suture	0, 1, 2

Research Question One

Descriptive Statistics and Preliminary Analysis

All statistical analyses are conducted in R (v. 4.0.2), a computational program freely available online (R Core Team 2018). Descriptive statistics are given for each sample used. These statistics provide a summary of the data and examine variability. The mean, standard deviation, maximum value, and minimum value are provided for the craniometric datasets, while frequency data are calculated using the '*psych*' package for each trait and character state in the cranial MMS datasets.

Each dataset is screened for errors and assessed for completeness. Missing data can be caused by antemortem trauma, pathology, postmortem damage, or taphonomy obscuring cranial landmarks or cranial MMS traits. Imputation offers a potential solution to problems associated with analysis and missing data, which have been tested with both data types. For craniometric data, Kenyhercz and Passalacqua (2016) recommend imputation if less than 50% missing data to maintain accuracy in classification. Kenyhercz and colleagues (2019) also recommend imputation for cranial MMS traits if the original dataset contains less than 50% missing. Data for craniometric and cranial MMS are imputed using the Multivariate Imputation by Chained Equations (MICE) approach (van Buuren & Groothuis-Oudshoorn 2011) in the *'mice'* package. The MICE method is highly flexible, allowing for the simultaneous imputation of binary, categorical, and continuous data. Within MICE, the predictive mean matching, or *pmm*, approach is favored here. Under the *pmm* method, imputation selects a random observation from the pool of observed values (by variable, in this case the population label) to replace a missing value (van Buuren & Groothuis-Oudshoorn 2011). This approach creates *n* number of datasets (five is the default) with imputed values. Next, the missing values are filled in from the generated dataset of choice using the *completeData* function, and the plausibility of values assessed using several plots. A significant benefit of this method is that imputed values are drawn from your dataset, preventing impossible or unrealistic values.

To address research question one, the relationship between craniometric and cranial MMS variables in each population group are investigated. This is to understand patterns of correlation among traits and for insight into potential impacts to the model, investigated later in research question three. In MLM models using cranial and postcranial MMS traits, Spiros and Hefner (2019) identified trait correlations within populational groups, noting models assuming trait independence should be applied with caution. Correlations between craniofacial variables (craniometric and cranial MMS) and population affinity labels are assessed using a polyserial correlation test in the *Polycor* package. This test measures associations between ordinal and numerical variables using a two-step process (Fox 2019; Lee et al. 1995), which is appropriate in assessing associations among craniometric (numerical), cranial MMS (ordinal) data and population-level labels. A polychoric correlation coefficient was calculated to identify inter-trait correlations among Latin American datasets using cranial MMS variables. as the method requires at least two of the same scores per character trait to calculate correlations. For a review of polychoric correlations using cranial MMS data, see Spiros and Hefner (2019). The correlation test indicates possible outcomes among variables that include: 1) a positive correlation, where lower character state values correlate to other lower character state values or higher character state values correlate to other higher character state values; 2) a negative correlation, where lower character state values correlate to higher character state values and vice versa; or 3) no correlation. An example of a positive correlation between two character state is an increase in projection for ANS (1<2<3) corresponds to a more sill-like projection in INA (1<2<3<4<5). An example of a negative correlation is an increase in width for IOB (1<2<3) correlates to a more rounded and smoother INA (5>4>3>2>1). The *cor* function is used to generate correlations using the craniometric data. Correlation plots for all metric variables in each Latin American sample are visualized. Positive correlations correspond to an increase in both ILDs, while negative correlations correspond to an increase in one ILD and a decrease in the other.

Next, to identify and assess the strength in relationships between cranial variables and population affinity labels, craniofacial data is assessed using the appropriate methods. For craniometric data, MANOVA is first used to assess significance between craniometric variables and population affinity labels. The MANOVA test assumes the data are normally distributed, so craniometric data are tested for normality using the Shapiro-Wilk test in the *mvnormtest* package. Next, an ANOVA is used to identify which craniometric variables are significant with population affinity labels. Cranial MMS data follows a different approach as the data are non-parametric. A Kruskal-Wallis test examines cranial MMS variables and population affinity labels for significance. To understand significant relationships among cranial MMS data, a pairwise comparison is performed with the Wilcoxon Rank Sum test.

Additionally, factor analysis of mixed data (FAMD) is used to understand the association between both qualitative and quantitative variables and labeling schemes used in analysis. This method assesses the data for patterns using principal component analysis. Results are presented graphically to describe variation within dimensions and the variable contributions to each dimension (Kassambara 2017:108). FAMD is used here to explore the data and identify patterns.

Research Question Two

Within Group Variation of Latin American Samples

As discussed in Chapter 3, biological distance examines the degree of group relatedness using underlying morphological variables from the skeleton that preserve population structure (Hefner et al. 2016). To address research question two, biological distance analyses focus on biological distance in geographically proximate samples.

Next, all data sets are assessed for similarity/dissimilarity using distance measures to understand the degree of relatedness among the samples and other populational reference groups. Populational distance analysis using craniometric data is achieved with the Mahalanobis Distance statistic. Distance analysis using cranial MMS data are analyzed according to the methodology described in Pink and colleagues (2016) and Go and Hefner (2020). Following protocol outlined in Go and Hefner (2020), cranial MMS traits that exhibit ordinal progression of character states (ANS, INA, PZT, PBD, NO, NAW, NBS, MT, IOB, and ZS) are dichotomized. Dr. Hefner and the author determined other trait dichotomizations. Sectioning points for dichotomization for each cranial MMS trait are listed in Table 4.4. Cranial MMS traits are transformed to binary variables with 0 as the low score and 1 as the high score for computational ease. Next, a distance matrix is calculated using Smith's Mean Measure of Divergence (MMD) in the *AnthropMMD* package (Santos 2018). The MMD is appropriate for categorical data like cranial MMS traits, converting frequency data to a numerical value, which indicates the level of similarity/dissimilarity (Harris & Sjøvold 2004; Pink et al. 2016). A larger numerical value indicates more dissimilarity between groups. See Pink and colleagues (2016) for a more detailed discussion of the mathematics involved in MMD. A Mantel test is used to test for significance. The craniometric and cranial MMS distance data are subject to a Procrustes analysis using the *smacof* package in R (Mair et al. 2021), which transforms the data so it could be visualized graphically in the same multivariate space.

Table 4.4: Sectioning points for cranial MMS data

Trait	*Sectioning Point*	*Trait*	*Sectioning Point*
ANS	1 | 2*	INA*	3 | 4
IOB	1 | 2	MT*	2 | 3
NAW	1 | 2*	NBC	1 | 2
NBS	1 | 2*	NO*	0 | 1
NFS	1 | 2	OBS	1 | 2
PBD*	0 | 1	PZT*	2 | 3
SPS (SNS)	0 | 1	TPS	1 | 2
PS	3 | 4	ZS	0 | 1
NAS	2 | 3		

Dichotomizations adopted from Go and Hefner (2020) are indicated with a (*).

Craniometric data for the Xoclán Cemetery, INACIF, identified Guatemalan migrants, and identified Mexican migrant samples are first subjected to Factor Analysis for Mixed Data (FAMD). This method is performed on the Latin American samples, first without and subsequently with the Unidentified Migrant sample. This method is useful for identifying patterns in datasets with mixed categorical and continuous variables while not prioritizing either type of variable over the other (Pagés 2004). Here, the means are centered and standard deviation set to 1, to remove any influence sex may have on measurements.

Research Question Three

Comparison of Cranial MMS and Craniometric Variation

This project uses the machine learning method (MLM)—Artificial Neural Networks (aNN)— to assess the classification power of craniometric and cranial MMS data for the Latin American and comparative samples. MLMs are computer intensive methods that learn from the data to arrive at the best classification outcome, in a process called tuning (Ousley 2016). MLMs do not require that data meet assumptions required of traditional classification statistics and they aim to avoid problems like overfitting by using more rigorous cross-validation methods (Hefner & Ousley 2014; Ousley 2016:204). Importantly, MLMs allow for the use of multiple data types within modeling. Research using MLMs show that combined biological data types have produced higher classification rates (Maier 2019; Spiros & Hefner 2019). A combination of craniometric and cranial MMS data demonstrate increased accuracy using RFM within a 3-group model structure (Hefner et al. 2014), but MLMs have yet to be explored for group refinement, including within the Hispanic category. To answer research question three, three classification models are created: an aNN model using only craniometric data, aNN models using only cranial MMS data, and aNN models using a combination of craniometric and cranial MMS data. This study assesses whether aNN accurately discriminates on a more refined level, past the Hispanic label, and assess which of the data types and combinations provide the most accurate results.

The aNN method is a type of neural network analysis inspired by neuronal functioning in human and animal brains (Liu 2020). Neural networks function by introducing several variables within your dataset that pass-through layers via nodes to arrive at an outcome based on patterns in the data. Each node represents a relatively simple operation that reorganizes the data as it moves to the next layer; however, the weights and connections between nodes and layers happen in a 'black box' and are difficult to interpret (Haykin 2009; Liu 2020). In aNN, random weights are assigned to each variable, in this case craniometric or cranial MMS, which generate multiple classification models, iterated over many repeats. The model with the best fit for the data is used. A train/test approach is used for building the aNN model, which is a type of cross validation where a proportion of the original dataset is reserved from model construction and used to test the formal model. Variable importance is modeled by identifying the strength of weighted connections between specific nodes of the model, as described in Beck (2013).

Results from each model are compared using the Matthew's Correlation Coefficient (MCC). The MCC measures classification accuracies between models and is better at assessing the accuracy in models with imbalanced samples (e.g., models that contain more cranial MMS data than craniometric data or models built with different numbers of populational groupings 3-group vs 6-group) (Chicco & Jurman 2020). For example, a 3-group classification model may have a higher accuracy than a 6-group classification model but assessing which of the models is doing a better job is accomplished using the MCC. Results from all the models are presented as a confusion matrix with values that range from -1 to +1 and speak to the strength of the observed and predicted classification values (Chicco 2017).

Limitations

Limitations for this study include travel restrictions, institutional protocol, and the presence of skeletal trauma. Reference samples from Mexico and Guatemala are not located within the U.S. Therefore, I traveled to Mexico in 2019 and Guatemala in 2020 as preliminary research trips to assess the collections and collect pilot data. Shortly after the 2020 trip to Guatemala, the COVID-

19 pandemic affected research globally. The shutdown effectively stopped all university related travel and prevented future travel. As skeletal remains are not stored indefinitely at the INACIF, all unidentified cases are stored for a period of six months, then if still unidentified, they are buried in a local cemetery according to INACIF protocol. This limits that amount of skeletal material available for analysis at any given time. The anthropologists at the INACIF are working with me to collect craniometric and cranial MMS data to amplify reference databases and for use in future research projects. Lastly, antemortem and perimortem trauma in skeletal specimens precludes data collection of craniometric data. Relatively few specimens in the Xoclán Cemetery collection exhibited cranial vault trauma, preventing data collection of craniometric landmarks. However, many of the cases at the INACIF exhibit perimortem trauma to the cranial vault, which affect the ability to collect craniometric landmark data and further reduced sample size.

CHAPTER 5: RESULTS

Missing Data and Imputation

Figure 5.1 shows the number of missing data by individuals and samples. The graphic shows that individuals with missing data generally have less than five variables missing per case for all samples.

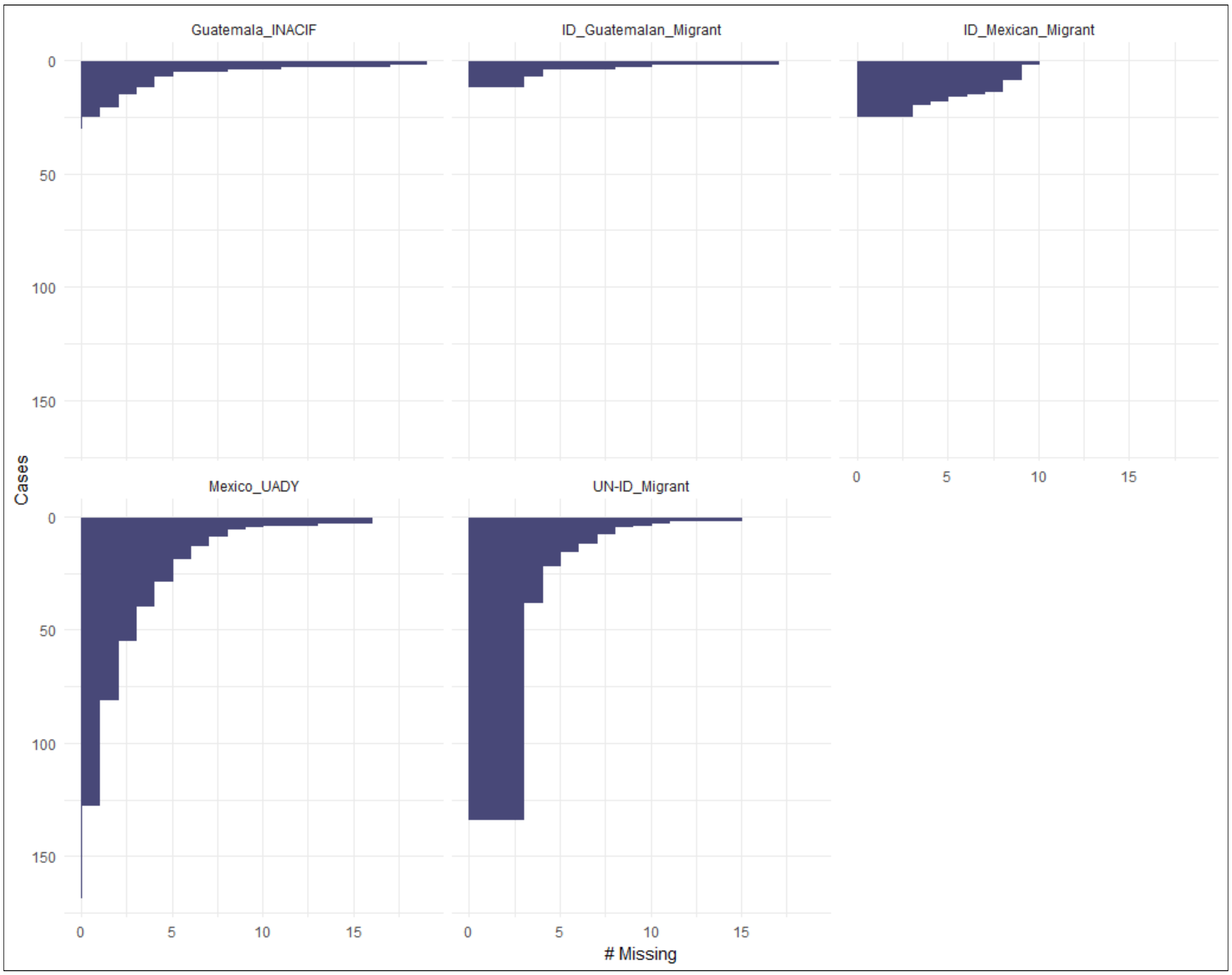

Figure 5.1: Missing data by individual and sample.

Figure 5.2 illustrates the variables with the highest percentage of missing data along with any patterns. The variables FOL, FOB, and UFBR are missing together for 134 individuals. This pattern is present in the Identified Guatemalan Migrant, Identified Mexican Migrant, and the Unidentified Migrant samples (Figure 5.3). The second most common missing variable is NO, which is the highest missing variable in the INACIF (~60%) and UADY (~50%) samples.

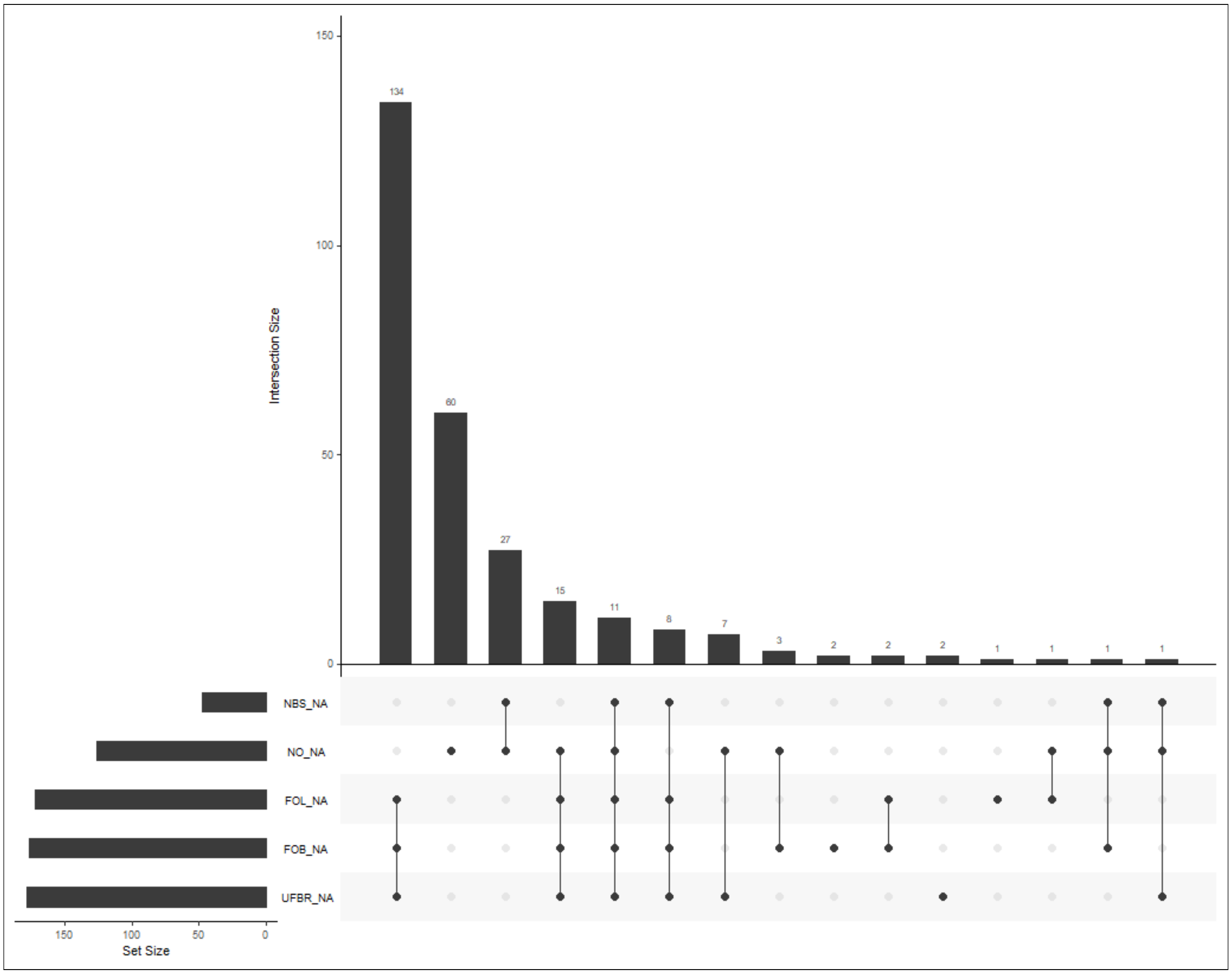

Figure 5.2: Highest frequency of missing data by variable for the Latin American samples

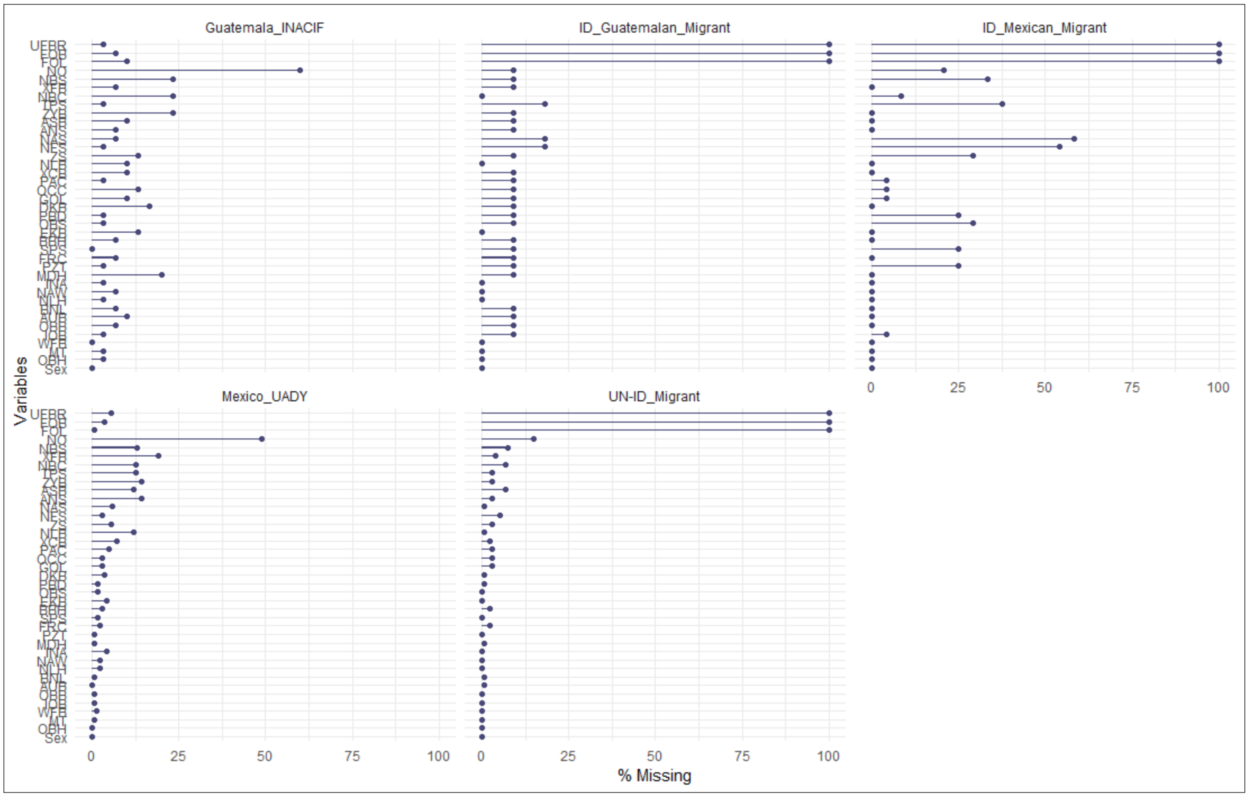

Figure 5.3: Percent missing data by population and variable

The mice method was set to complete five imputations before the original dataset was completed with final, imputed values where missing data once was. The algorithm isolated data to its specific column, therefore each predicted value is set by predictors specific to that column. The default was selected, so the measurement level available by variable (ILD or cranial MMS trait) were the limit for the imputed value.

Outlier Detection

The Cook's Distance identified 17 potential outliers in the metric data, 15 from the INACIF, Identified Guatemalan Migrant, and Identified Mexican Migrant samples; however, these individuals are retained to maximize the total sample size for these groups. The remaining two outliers were in the American Black sample, and—given the larger size of this sample—were removed from subsequent analysis.

Research Question One

To answer research question one, summary statistics, correlation tests, MANOVA, ANOVA, Kruskal-Wallis, and FAMD are used with data collected from the INACIF sample in Guatemala City, Guatemala and the UADY sample in Mérida, Mexico.

Summary Statistics

Descriptive statistics, prior to imputation, for cranial MMS data collected from the Latin American samples are provided in the appendix directly following this chapter. These include frequency distribution data for each trait and the dichotomization scheme used some of the subsequent analyses. Summary data for American Black, American White, and Thai samples are provided in Spiros & Hefner (2019) and Techataweewan et al. (2021).

Summary Metric Data by Population and Sex

Craniometric data for the Latin American samples (prior to any imputation of missing data) are summarized in the appendix directly following this chapter, by population and sex.

Cranial MMS Trait Correlations

The following figures and tables provide the Polychoric correlation coefficient for the cranial MMS traits by individual samples. These illustrative figures demonstrate the relative intertrait correlations of MMS data and follow (relatively closely) previously published results (see below). Figure 5.4 shows the inter-trait correlations among the INACIF sample. Significant positive correlations occurred between NAW and MT, NFS and NBC, PBD and NBC, IOB and TPS, PZT and SPS, ZS and NO, and TPS and IOB. Negative correlations occurred between NAW and INA, NAW and ANS, IOB and NO, MT and NO, ZS and IOB. Table 5.1 shows the correlation matrix for the INACIF sample.

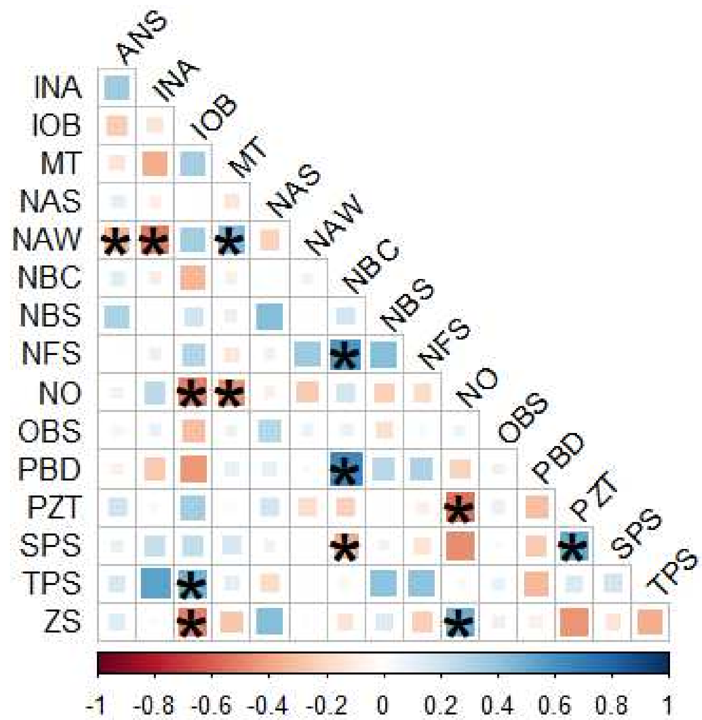

Figure 5.4: INACIF polychoric correlation values. *The (*) indicates significant values.*

Table 5.1: Correlation matrix for the INACIF polychoric correlations

	ANS	*INA*	*IOB*	*MT*	*NAS*	*NAW*	*NBC*	*NBS*	*NFS*	*NO*	*OBS*	*PBD*	*PZT*	*SPS*	*TPS*
INA	0.35														
IOB	-0.24	-0.14													
MT	-0.13	-0.37	0.35												
NAS	0.11	-0.09	-0.02	-0.13											
NAW	**-0.34**	**-0.52**	0.35	**0.44**	-0.23										
NBC	0.13	-0.11	-0.34	-0.08	0.01	0.06									
NBS	0.32	-0.02	0.19	0.09	0.42	0.01	0.18								
NFS	-0.01	0.09	0.29	-0.12	0.08	0.36	**0.59**	0.43							
NO	0.08	0.26	**-0.50**	**-0.45**	-0.08	-0.26	0.19	-0.23	-0.18						
OBS	-0.07	0.09	-0.30	0.08	0.27	0.09	0.09	-0.16	0.04	0.07					
PBD	-0.07	-0.27	-0.44	0.10	0.09	-0.03	**0.65**	0.28	0.31	-0.22	0.08				
PZT	0.20	0.04	0.35	-0.03	0.19	-0.17	-0.22	-0.03	-0.09	**-0.54**	0.10	-0.30			
SPS	0.09	0.24	0.24	0.18	0.08	-0.01	**-0.36**	-0.07	-0.15	-0.47	0.03	-0.25	**0.50**		
TPS	0.18	0.53	**0.47**	0.11	-0.19	0.03	-0.05	0.40	0.41	-0.05	0.10	-0.33	0.15	0.19	
ZS	0.13	-0.02	**-0.50**	-0.28	0.42	0.03	-0.14	0.16	-0.25	**0.49**	0.09	-0.07	-0.45	-0.14	-0.37

*significant values are bolded

Inter-trait correlations in the UADY sample are shown in Figure 5.5. A single significant negative correlation exists between NO and PBD. The correlation matrix for the UADY sample is presented in Table 5.2.

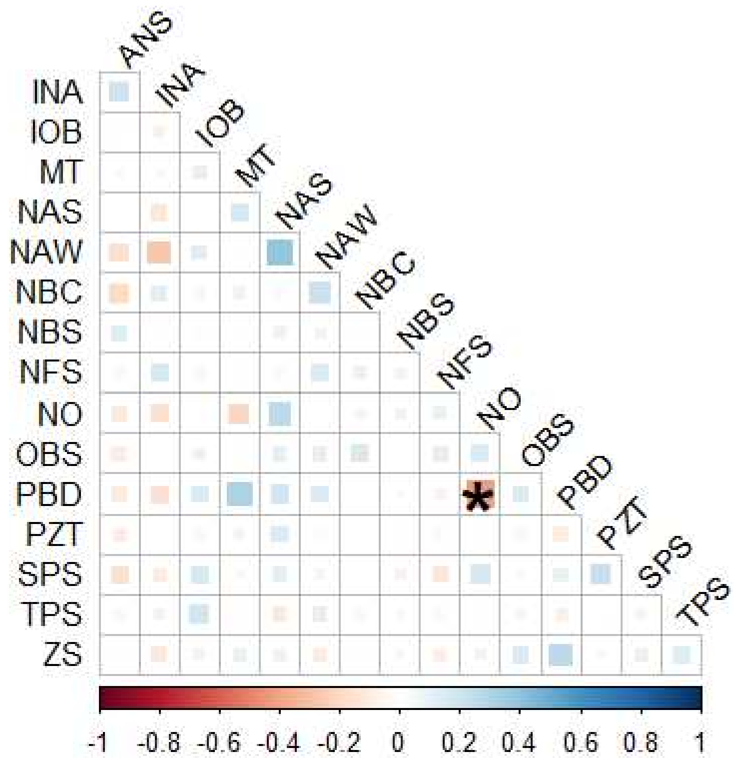

Figure 5.5: UADY polychoric correlation values. *The (*) indicates significant values.*

Table 5.2: Correlation matrix for the UADY polychoric correlations

	ANS	*INA*	*IOB*	*MT*	*NAS*	*NAW*	*NBC*	*NBS*	*NFS*	*NO*	*OBS*	*PBD*	*PZT*	*SPS*	*TPS*
INA	0.20														
IOB	-0.02	-0.07													
MT	0.03	-0.03	0.11												
NAS	-0.01	-0.13	0.00	0.17											
NAW	-0.17	-0.27	0.12	-0.01	0.39										
NBC	-0.20	0.12	0.06	0.07	0.03	0.21									
NBS	0.13	0.02	0.02	-0.01	0.07	-0.06	-0.02								
NFS	0.04	0.16	-0.06	0.03	0.04	0.15	0.08	0.07							
NO	-0.11	-0.17	0.01	-0.21	0.27	0.01	0.06	0.07	0.10						
OBS	-0.10	0.00	0.08	-0.01	0.10	0.11	0.16	-0.01	0.11	0.16					
PBD	-0.11	-0.16	0.15	0.32	0.18	0.14	-0.01	0.04	-0.07	**-0.42**	0.15				
PZT	-0.09	0.00	-0.04	-0.06	0.15	0.02	0.01	0.02	-0.03	0.02	-0.04	-0.10			
SPS	-0.16	-0.09	0.16	0.03	0.11	-0.06	-0.01	-0.07	-0.12	0.18	0.04	0.09	0.23		
TPS	0.05	0.07	0.18	0.01	-0.12	0.11	0.05	0.05	-0.03	0.03	0.07	-0.09	-0.01	-0.06	
ZS	-0.01	-0.13	0.06	0.10	0.07	-0.10	0.01	-0.04	-0.08	-0.06	0.15	0.26	-0.05	0.09	0.14

*significant values are bolded

Figure 5.6 shows the inter-trait correlations among the Identified Migrant sample. Significant positive correlations occur between ANS and INA, NBC and ANS, and PBD and PZT. Negative correlations occur between MT and ANS, MT and INA, MT and NBD, and IOB and OBS. Table 5.3 presents correlation coefficients for the Identified Migrant sample.

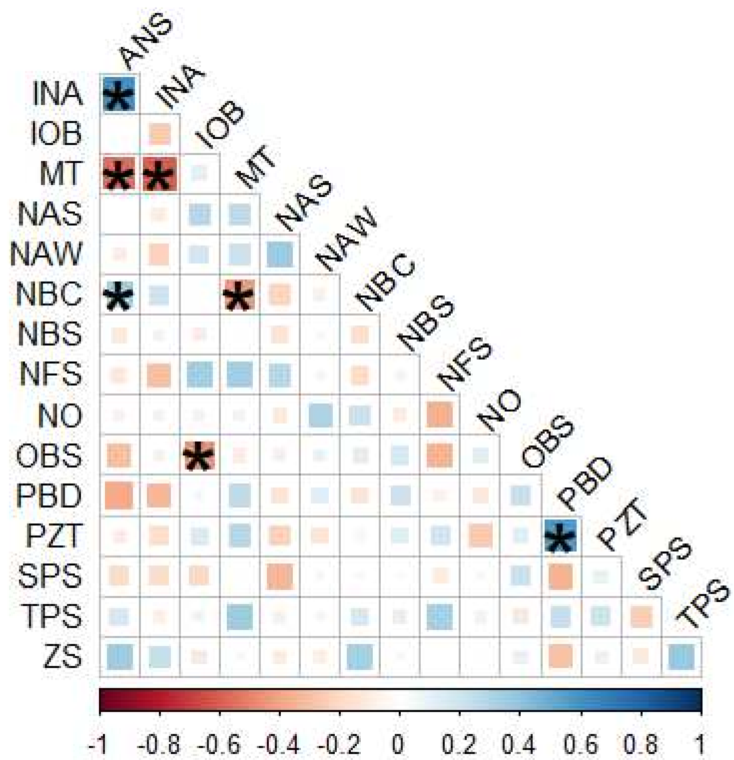

Figure 5.6: Identified Mexican Migrant polychoric correlation values. *The (*) indicates significant values.*

Table 5.3: Correlation matrix for the Identified Mexican Migrant polychoric correlations

	ANS	*INA*	*IOB*	*MT*	*NAS*	*NAW*	*NBC*	*NBS*	*NFS*	*NO*	*OBS*	*PBD*	*PZT*	*SPS*	*TPS*
INA	**0.62**														
IOB	0.00	-0.25													
MT	**-0.55**	**-0.61**	0.11												
NAS	0.01	-0.10	0.27	0.25											
NAW	-0.09	-0.23	0.19	0.22	0.36										
NBC	**0.35**	0.20	-0.02	**-0.44**	-0.21	-0.08									
NBS	-0.12	-0.07	-0.11	0.01	-0.15	0.03	-0.18								
NFS	-0.13	-0.30	0.34	0.35	0.27	-0.05	-0.18	0.05							
NO	-0.07	-0.07	-0.07	0.06	-0.11	0.30	0.22	-0.10	-0.35						
OBS	-0.31	-0.07	**-0.45**	-0.10	-0.09	0.08	0.13	0.18	-0.35	0.13					
PBD	-0.38	-0.33	0.05	0.25	-0.14	0.13	-0.14	0.22	-0.08	-0.13	0.23				
PZT	-0.09	-0.18	0.16	0.27	-0.22	-0.15	-0.05	0.14	0.19	-0.26	0.13	**0.57**			
SPS	-0.19	-0.18	-0.20	0.02	-0.34	-0.05	0.03	-0.03	-0.11	0.04	0.22	-0.34	0.10		
TPS	0.18	-0.09	-0.07	0.35	-0.07	-0.05	0.17	0.10	0.34	-0.08	-0.12	0.23	0.20	-0.24	
ZS	0.35	0.24	-0.11	0.04	-0.10	-0.08	0.33	0.04	0.00	-0.01	0.09	-0.30	0.09	-0.12	0.37

*significant values are bolded

No significant inter-trait correlations are noted for the Unidentified Migrant sample (Figure 5.7). Table 5.4 illustrates the correlation coefficients for the Unidentified Migrant sample.

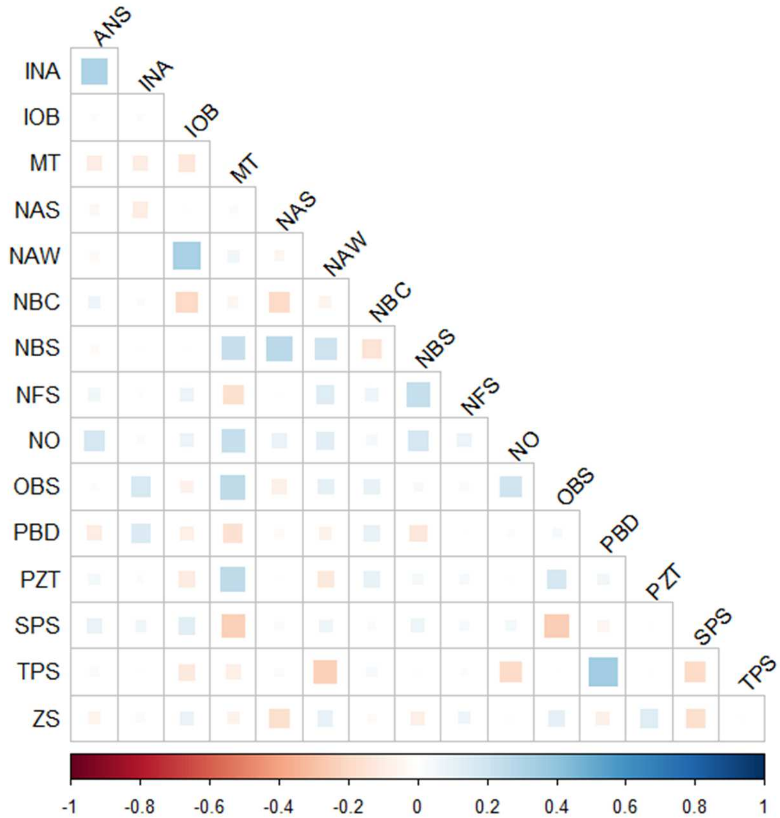

Figure 5.7: Unidentified Migrant polychoric correlation values.

Table 5.4: Correlation matrix for the Unidentified Migrant polychoric correlations

	ANS	*INA*	*IOB*	*MT*	*NAS*	*NAW*	*NBC*	*NBS*	*NFS*	*NO*	*OBS*	*PBD*	*PZT*	*SPS*	*TPS*
INA	0.30														
IOB	0.01	-0.01													
MT	-0.09	-0.09	-0.12												
NAS	-0.04	-0.10	0.00	0.03											
NAW	-0.03	0.00	0.33	0.06	-0.04										
NBC	0.07	0.02	-0.19	-0.05	-0.18	-0.06									
NBS	-0.02	-0.01	0.01	0.22	0.26	0.21	-0.15								
NFS	0.06	0.01	0.07	-0.17	0.01	0.14	0.06	0.24							
NO	0.17	-0.02	0.08	0.22	0.08	0.12	0.05	0.17	0.08						
OBS	-0.01	0.16	-0.07	0.26	-0.08	0.10	0.10	0.03	0.03	0.19					
PBD	-0.09	0.14	-0.08	-0.15	-0.03	-0.06	0.10	-0.12	0.00	-0.02	0.03				
PZT	0.05	-0.02	-0.11	0.26	0.00	-0.12	0.09	0.04	0.04	-0.01	0.17	0.05			
SPS	0.08	0.05	0.13	-0.23	0.03	0.07	0.03	0.07	0.04	0.05	-0.24	-0.05	-0.01		
TPS	0.03	0.00	-0.11	-0.09	0.02	-0.24	0.04	-0.01	-0.01	-0.19	0.00	0.34	0.00	-0.18	
ZS	-0.05	0.03	0.08	-0.06	-0.17	0.10	-0.03	-0.08	0.06	0.01	0.10	-0.08	0.14	-0.16	-0.01

*significant values are bolded

A polychoric correlation coefficient calculation is not possible for the Identified Guatemalan Migrant sample due to its small sample size (n = 11).

Craniometric Correlations

The Pearson correlation coefficient calculations are presented for each Latin American sample with metric data. Figure 5.8 shows the correlations among metric variables in the INACIF sample. Significant values exist between several length and breadth measurements. Individual correlation values are listed in Table 5.5.

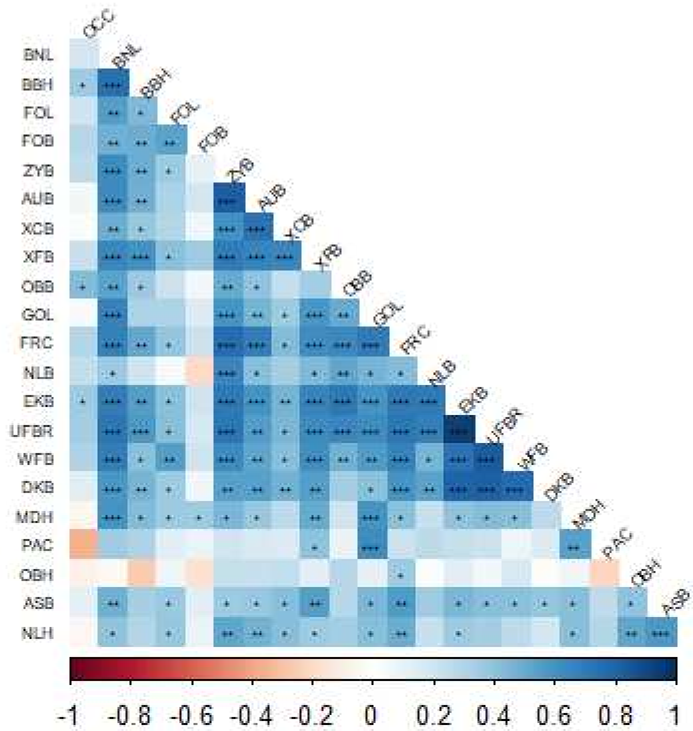

Figure 5.8: Correlation plot for craniometric variables (INACIF).

Table 5.5: Correlation matrix for craniometric data (INACIF)

	GOL	*BNL*	*BBH*	*XCB*	*XFB*	*WFB*	*ZYB*	*AUB*	*ASB*	*NLH*	*NLB*	*MDH*	*OBH*	*OBB*	*DKB*	*EKB*	*FRC*	*PAC*	*OCC*	*FOL*	*FOB*
BNL	**0.68**																				
BBH	0.31	**0.76**																			
XCB	**0.36**	**0.50**	**0.39**																		
XFB	**0.59**	**0.66**	**0.61**	**0.65**																	
WFB	**0.56**	**0.69**	**0.41**	**0.42**	**0.62**																
ZYB	**0.60**	**0.64**	**0.50**	**0.60**	**0.71**	**0.64**															
AUB	**0.47**	**0.63**	**0.49**	**0.75**	**0.66**	**0.52**	**0.83**														
ASB	**0.46**	**0.48**	0.34	**0.43**	**0.56**	**0.42**	**0.36**	**0.38**													
NLH	**0.43**	**0.40**	0.30	**0.42**	**0.37**	0.27	**0.53**	**0.48**	**0.58**												
NLB	**0.46**	**0.36**	0.20	0.34	**0.44**	**0.46**	**0.65**	**0.43**	**0.36**	0.24											
MDH	**0.60**	**0.60**	**0.43**	0.27	**0.51**	**0.42**	**0.45**	**0.38**	**0.43**	**0.42**	0.23										
OBH	0.12	-0.04	-0.25	0.25	0.13	0.17	0.24	0.22	**0.40**	**0.53**	0.02	0.07									
OBB	**0.51**	**0.53**	**0.37**	0.24	0.35	**0.47**	**0.50**	**0.40**	0.28	0.35	**0.53**	0.21	0.28								
DKB	**0.41**	**0.58**	**0.47**	**0.47**	**0.56**	**0.77**	**0.55**	**0.55**	**0.39**	0.19	**0.54**	0.27	-0.01	0.35							
EKB	**0.59**	**0.70**	**0.55**	**0.47**	**0.66**	**0.75**	**0.71**	**0.60**	**0.45**	**0.37**	**0.71**	**0.41**	0.13	**0.71**	**0.80**						
FRC	**0.71**	**0.69**	**0.52**	**0.43**	**0.62**	**0.67**	**0.76**	**0.70**	**0.56**	**0.50**	**0.44**	**0.42**	**0.36**	**0.65**	**0.59**	**0.71**					
PAC	**0.63**	**0.35**	0.31	0.15	**0.40**	0.08	0.20	0.17	0.25	0.29	0.26	**0.54**	-0.22	0.10	0.16	0.23	0.21				
OCC	0.04	0.19	**0.37**	0.01	0.23	0.30	0.25	0.06	0.11	-0.04	0.25	-0.05	-0.07	**0.43**	0.13	**0.37**	0.30	**-0.36**			
FOL	0.32	**0.56**	**0.45**	0.29	**0.41**	**0.55**	**0.36**	0.32	**0.37**	**0.39**	0.04	**0.37**	0.07	0.22	**0.37**	**0.42**	**0.39**	0.13	0.20		
FOB	0.14	**0.48**	**0.49**	0.06	0.35	0.21	0.11	0.19	0.13	0.09	-0.20	**0.37**	-0.16	0.05	0.06	0.21	0.21	0.09	0.28	**0.54**	
UFBR	**0.63**	**0.73**	**0.57**	**0.46**	**0.64**	**0.83**	**0.72**	**0.56**	**0.43**	0.35	**0.69**	**0.44**	0.05	**0.65**	**0.82**	**0.94**	**0.73**	0.22	0.34	**0.43**	0.17

*significant values are bolded

Figure 5.9 shows the correlations among metric variables in the UADY sample. Significant values exist between most cranial length and breadth measurements. The exception to this is OBH with NLB and DKB, which exhibit slight negative correlations. Individual correlation values are listed in Table 5.6.

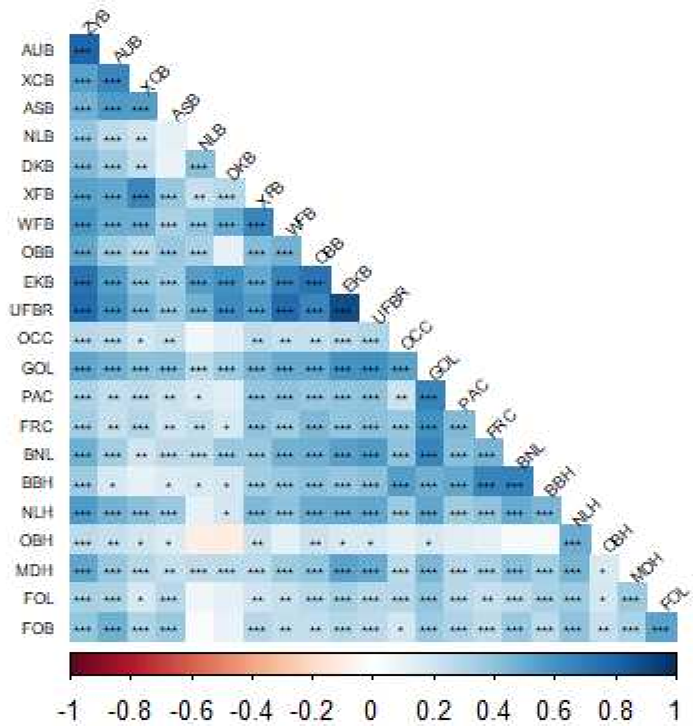

Figure 5.9: Correlation plot for craniometric variables (UADY).

Table 5.6: Correlation Matrix for craniometric data (UADY)

	GOL	*BNL*	*BBH*	*XCB*	*XFB*	*WFB*	*ZYB*	*AUB*	*ASB*	*NLH*	*NLB*	*MDH*	*OBH*	*OBB*	*DKB*	*EKB*	*FRC*	*PAC*	*OCC*	*FOL*	*FOB*
BNL	**0.66**																				
BBH	**0.48**	**0.68**																			
XCB	**0.42**	**0.20**	0.11																		
XFB	**0.46**	**0.38**	**0.35**	**0.66**																	
WFB	**0.50**	**0.48**	**0.39**	**0.49**	**0.67**																
ZYB	**0.53**	**0.49**	**0.36**	**0.55**	**0.53**	**0.58**															
AUB	**0.48**	**0.34**	**0.17**	**0.66**	**0.51**	**0.50**	**0.81**														
ASB	**0.42**	**0.26**	**0.15**	**0.55**	**0.39**	**0.33**	**0.47**	**0.56**													
NLH	**0.52**	**0.51**	**0.46**	**0.41**	**0.44**	**0.40**	**0.57**	**0.47**	**0.38**												
NLB	**0.27**	**0.26**	**0.17**	**0.20**	**0.24**	**0.39**	**0.39**	**0.27**	0.10	0.10											
MDH	**0.45**	**0.42**	**0.34**	**0.29**	**0.27**	**0.36**	**0.53**	**0.39**	**0.23**	**0.47**	**0.26**										
OBH	**0.18**	0.04	0.04	**0.19**	**0.21**	0.11	**0.26**	**0.23**	**0.15**	**0.49**	-0.10	**0.16**									
OBB	**0.51**	**0.48**	**0.41**	**0.28**	**0.44**	**0.48**	**0.53**	**0.35**	**0.38**	**0.50**	**0.33**	**0.40**	**0.22**								
DKB	**0.35**	**0.38**	**0.19**	**0.24**	**0.29**	**0.51**	**0.44**	**0.36**	0.09	**0.19**	**0.41**	**0.27**	-0.10	0.10							
EKB	**0.60**	**0.55**	**0.38**	**0.40**	**0.51**	**0.67**	**0.74**	**0.56**	**0.36**	**0.51**	**0.55**	**0.53**	**0.16**	**0.74**	**0.60**						
FRC	**0.61**	**0.47**	**0.66**	**0.30**	**0.41**	**0.36**	**0.35**	**0.24**	**0.21**	**0.41**	**0.24**	**0.33**	0.11	**0.44**	**0.18**	**0.39**					
PAC	**0.67**	**0.43**	**0.52**	**0.25**	**0.40**	**0.41**	**0.32**	**0.25**	**0.22**	**0.35**	**0.16**	**0.29**	0.14	**0.39**	0.14	**0.39**	**0.45**				
OCC	**0.53**	**0.40**	**0.56**	**0.19**	**0.25**	**0.23**	**0.29**	**0.26**	**0.23**	**0.36**	0.05	**0.26**	0.11	**0.24**	0.12	**0.26**	**0.39**	**0.21**			
FOL	**0.40**	**0.27**	**0.32**	**0.18**	**0.20**	**0.22**	**0.32**	**0.31**	**0.27**	**0.36**	0.07	**0.36**	**0.18**	**0.27**	0.12	**0.34**	**0.23**	**0.28**	**0.28**		
FOB	**0.37**	**0.38**	**0.27**	**0.31**	**0.29**	**0.25**	**0.39**	**0.49**	**0.37**	**0.40**	0.02	**0.30**	**0.24**	**0.25**	0.10	**0.28**	**0.26**	**0.30**	**0.18**	**0.53**	
UFBR	**0.61**	**0.56**	**0.42**	**0.46**	**0.56**	**0.78**	**0.77**	**0.60**	**0.39**	**0.52**	**0.48**	**0.50**	**0.18**	**0.66**	**0.62**	**0.89**	**0.42**	**0.40**	**0.30**	**0.27**	**0.26**

*significant values are bolded

Figure 5.10 shows the correlations among metric variables in the Identified Guatemalan Migrant sample. Significant values exist between cranial length and breadth measurements. There are a few negative correlations between OCC and FOB, PAC, then FOL and BBH, ASB, XCB, ZYB, AUB, PAC, FRC, and finally, DKB and OBH. Individual correlation values are listed in Table 5.7.

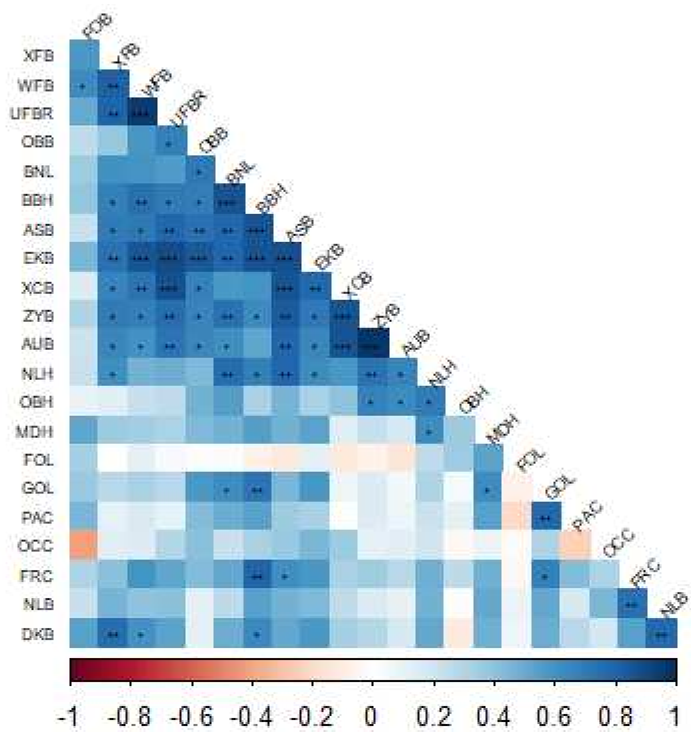

Figure 5.10: Correlation plot for craniometric variables (Identified Guatemalan Migrants).

Table 5.7: Correlation matrix for craniometric data (Identified Guatemalan Migrants)

	GOL	*BNL*	*BBH*	*XCB*	*XFB*	*WFB*	*ZYB*	*AUB*	*ASB*	*NLH*	*NLB*	*MDH*	*OBH*	*OBB*	*DKB*	*EKB*	*FRC*	*PAC*	*OCC*	*FOL*	*FOB*
BNL	**0.62**																				
BBH	**0.74**	**0.86**																			
XCB	0.08	0.58	0.59																		
XFB	0.28	0.60	**0.69**	**0.66**																	
WFB	0.32	0.60	**0.76**	**0.74**	**0.82**																
ZYB	0.15	**0.75**	**0.64**	**0.87**	**0.71**	**0.68**															
AUB	0.06	**0.61**	0.53	**0.87**	**0.65**	**0.60**	**0.97**														
ASB	0.46	**0.79**	**0.87**	**0.86**	**0.71**	**0.72**	**0.84**	**0.80**													
NLH	0.31	**0.76**	**0.69**	0.58	**0.62**	0.48	**0.75**	**0.66**	**0.79**												
NLB	0.51	0.27	0.55	0.25	0.47	0.40	0.15	0.11	0.48	0.48											
MDH	**0.62**	0.48	0.55	0.14	0.37	0.34	0.23	0.18	0.48	**0.63**	0.52										
OBH	0.05	0.56	0.33	0.41	0.13	0.23	**0.67**	**0.63**	0.47	**0.70**	-0.02	0.36									
OBB	0.57	**0.71**	**0.69**	**0.68**	0.39	0.59	**0.66**	**0.65**	**0.75**	0.44	0.12	0.44	0.47								
DKB	0.50	0.50	**0.65**	0.34	**0.76**	**0.61**	0.30	0.18	0.52	0.52	**0.79**	0.48	-0.11	0.12							
EKB	0.58	**0.79**	**0.87**	**0.79**	**0.75**	**0.86**	**0.72**	**0.65**	**0.87**	**0.62**	0.44	0.54	0.33	**0.85**	0.58						
FRC	**0.69**	0.52	**0.81**	0.32	0.42	0.59	0.35	0.28	**0.62**	0.49	**0.76**	0.49	0.26	0.44	0.55	0.59					
PAC	**0.79**	0.50	0.54	0.00	0.10	0.17	0.15	0.07	0.36	0.21	0.17	0.54	0.12	0.43	0.28	0.33	0.43				
OCC	0.31	0.23	0.33	0.36	0.13	0.14	0.10	0.13	0.37	0.20	0.45	0.08	-0.04	0.42	0.18	0.45	0.32	-0.23			
FOL	-0.07	0.02	-0.08	-0.12	0.00	0.12	-0.08	-0.14	-0.12	0.26	0.08	0.54	0.36	0.02	0.08	0.11	-0.04	-0.19	-0.03		
FOB	0.38	0.35	0.39	0.15	0.57	**0.64**	0.30	0.22	0.23	0.22	0.23	0.53	0.10	0.26	0.55	0.45	0.30	0.45	-0.42	0.34	
UFBR	0.28	0.56	**0.71**	**0.87**	**0.81**	**0.95**	**0.76**	**0.74**	**0.80**	0.49	0.41	0.31	0.26	**0.68**	0.53	**0.89**	0.53	0.10	0.29	0.04	0.51

*significant values are bolded

Figure 5.11 shows the correlations among metric variables in the Identified Mexican Migrant sample. Significant values exist between some cranial length and breadth measurements. There are more negative correlations, with one significant negative correlation between NLB and OCC. Individual correlation values are listed in Table 5.8.

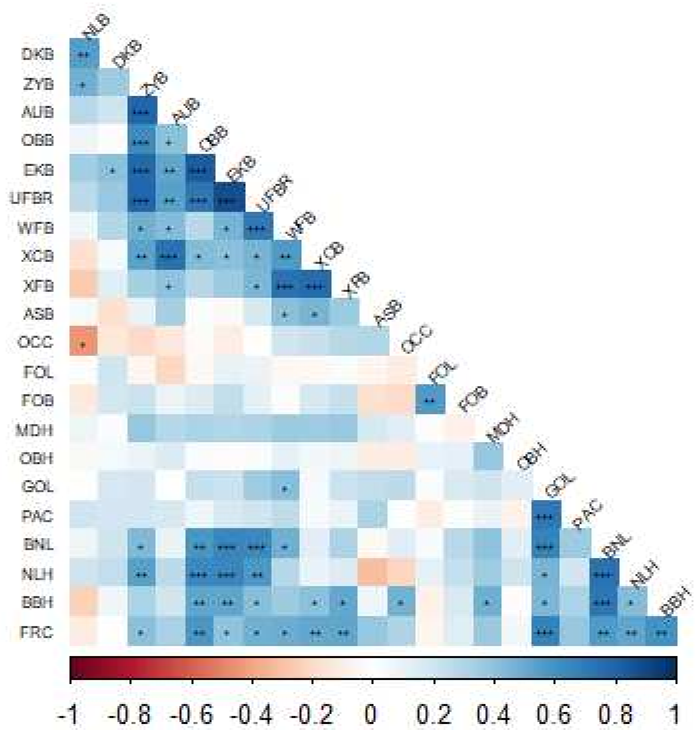

Figure 5.11: Correlation plot for craniometric variables (Identified Mexican Migrants).

Table 5.8: Correlation matrix for craniometric data (Identified Mexican Migrants)

	GOL	*BNL*	*BBH*	*XCB*	*XFB*	*WFB*	*ZYB*	*AUB*	*ASB*	*NLH*	*NLB*	*MDH*	*OBH*	*OBB*	*DKB*	*EKB*	*FRC*	*PAC*	*OCC*	*FOL*	*FOB*
BNL	**0.68**																				
BBH	**0.45**	**0.73**																			
XCB	0.05	0.13	**0.42**																		
XFB	0.22	0.32	**0.51**	**0.77**																	
WFB	**0.43**	**0.48**	0.37	**0.57**	**0.75**																
ZYB	0.19	0.45	0.31	**0.53**	0.34	**0.49**															
AUB	0.00	0.10	0.19	**0.74**	**0.41**	**0.44**	**0.79**														
ASB	0.24	-0.04	0.07	**0.46**	0.37	**0.41**	0.10	0.34													
NLH	**0.46**	**0.74**	**0.43**	0.09	0.15	0.29	**0.55**	0.27	-0.31												
NLB	0.00	0.08	-0.24	-0.15	-0.27	0.06	**0.49**	0.28	0.04	0.21											
MDH	0.20	0.40	**0.48**	0.36	0.37	0.37	0.39	0.29	0.18	0.36	0.08										
OBH	0.13	0.12	-0.02	0.06	0.08	0.10	0.08	0.15	-0.11	0.21	-0.03	0.39									
OBB	0.24	**0.59**	**0.52**	**0.43**	0.29	0.27	**0.64**	**0.42**	0.01	**0.64**	0.08	0.32	0.02								
DKB	0.17	0.20	0.07	0.04	0.13	0.29	0.36	0.21	-0.17	0.24	**0.55**	0.03	0.05	-0.02							
EKB	0.22	**0.65**	**0.52**	**0.42**	0.37	**0.47**	**0.78**	**0.53**	-0.04	**0.65**	0.35	0.30	-0.01	**0.84**	**0.41**						
FRC	**0.68**	**0.56**	**0.60**	**0.52**	0.52	**0.50**	**0.41**	0.33	0.39	**0.53**	-0.11	0.38	0.06	**0.58**	0.00	**0.41**					
PAC	**0.71**	0.36	0.32	0.04	0.09	0.26	0.18	0.17	0.32	0.20	0.19	0.09	-0.07	0.07	0.18	0.16	0.38				
OCC	0.27	0.12	**0.48**	0.22	0.29	0.19	-0.20	-0.13	0.29	-0.21	**-0.44**	0.12	-0.11	-0.03	-0.12	-0.09	0.31	0.01			
FOL	0.02	0.04	-0.07	-0.04	0.02	-0.03	-0.06	-0.21	-0.05	0.10	0.02	-0.02	0.09	-0.04	0.20	0.10	-0.05	-0.10	-0.10		
FOB	0.17	0.28	0.17	0.18	0.22	-0.01	0.22	0.09	-0.17	0.24	-0.11	-0.08	0.10	0.15	0.18	0.25	0.14	0.02	-0.20	**0.57**	
UFBR	0.35	**0.64**	**0.43**	**0.46**	**0.45**	**0.72**	**0.79**	**0.55**	0.18	**0.58**	0.27	0.32	0.03	**0.72**	0.38	**0.88**	**0.49**	0.21	0.00	0.09	0.11

*significant values are bolded

Figure 5.12 shows the correlations among metric variables in the Unidentified Migrant sample. Highly significant values exist between many of the cranial length and breadth measurements. There are three slightly negative correlations, with one significant negative correlation between DKB and OBH. Individual correlation values are listed in Table 5.9.

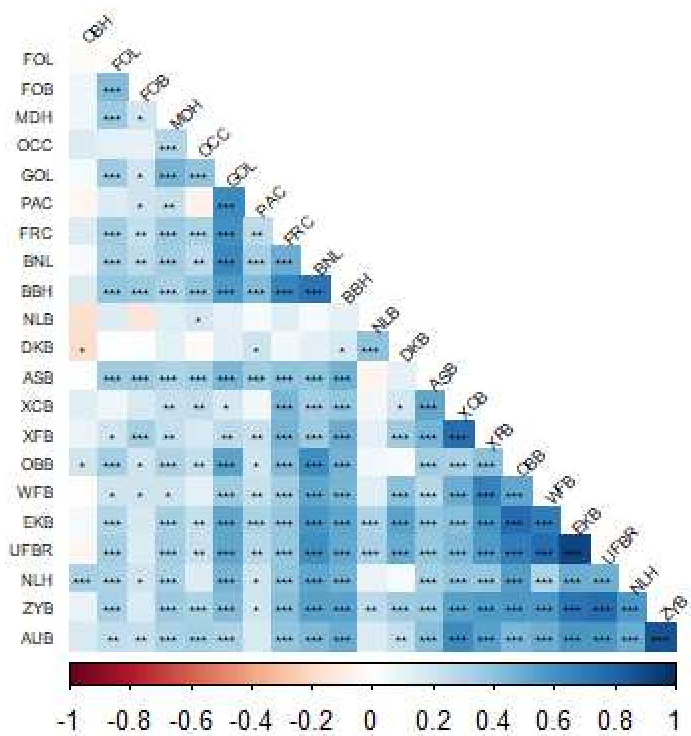

Figure 5.12: Correlation plot for craniometric variables (Unidentified Migrants).

Table 5.9: Correlation matrix for craniometric data (Unidentified Migrants)

	GOL	*BNL*	*BBH*	*XCB*	*XFB*	*WFB*	*ZYB*	*AUB*	*ASB*	*NLH*	*NLB*	*MDH*	*OBH*	*OBB*	*DKB*	*EKB*	*FRC*	*PAC*	*OCC*	*FOL*	*FOB*
BNL	**0.65**																				
BBH	**0.56**	**0.74**																			
XCB	**0.17**	**0.34**	**0.39**																		
XFB	**0.27**	**0.37**	**0.50**	**0.76**																	
WFB	**0.35**	**0.52**	**0.46**	**0.49**	**0.67**																
ZYB	**0.44**	**0.55**	**0.53**	**0.55**	**0.56**	**0.57**															
AUB	**0.32**	**0.45**	**0.50**	**0.64**	**0.55**	**0.50**	**0.86**														
ASB	**0.45**	**0.40**	**0.48**	**0.53**	**0.46**	**0.28**	**0.38**	**0.42**													
NLH	**0.47**	**0.53**	**0.48**	**0.39**	**0.38**	**0.31**	**0.57**	**0.49**	**0.34**												
NLB	0.11	0.03	0.10	0.06	0.12	0.12	**0.26**	0.16	-0.06	0.09											
MDH	**0.47**	**0.34**	**0.32**	**0.22**	**0.24**	**0.20**	**0.37**	**0.32**	**0.31**	**0.32**	0.13										
OBH	0.04	0.04	0.15	0.14	0.09	-0.01	0.10	**0.17**	0.00	**0.33**	-0.16	0.06									
OBB	**0.54**	**0.62**	**0.48**	**0.32**	**0.42**	**0.53**	**0.58**	**0.46**	**0.35**	**0.55**	0.06	**0.31**	**0.19**								
DKB	0.13	0.12	**0.17**	**0.18**	**0.36**	**0.41**	**0.36**	**0.24**	0.11	0.04	**0.40**	0.10	**-0.17**	0.03							
EKB	**0.52**	**0.58**	**0.52**	**0.43**	**0.59**	**0.71**	**0.74**	**0.59**	**0.39**	**0.50**	**0.30**	**0.33**	0.04	**0.77**	**0.53**						
FRC	**0.61**	**0.49**	**0.67**	**0.45**	**0.46**	**0.34**	**0.43**	**0.40**	**0.41**	**0.40**	0.12	**0.36**	0.14	**0.43**	0.06	**0.39**					
PAC	**0.63**	**0.33**	**0.43**	0.05	**0.23**	**0.24**	**0.19**	0.16	**0.36**	**0.21**	0.04	**0.24**	-0.05	**0.21**	**0.21**	**0.32**	**0.26**				
OCC	**0.41**	**0.27**	**0.39**	**0.24**	**0.17**	0.11	**0.32**	**0.33**	**0.33**	0.15	**0.18**	**0.29**	0.14	**0.23**	-0.03	**0.24**	**0.31**	-0.07			
FOL	**0.35**	**0.33**	**0.39**	0.08	**0.20**	**0.20**	**0.33**	**0.24**	**0.35**	**0.32**	0.14	**0.36**	-0.02	**0.34**	0.01	**0.30**	**0.35**	0.15	0.12		
FOB	**0.21**	**0.27**	**0.36**	0.16	**0.33**	**0.20**	0.16	**0.23**	**0.35**	0.21	-0.14	**0.20**	0.06	**0.20**	0.00	**0.17**	**0.24**	**0.20**	0.10	**0.44**	
UFBR	**0.52**	**0.61**	**0.51**	**0.45**	**0.57**	**0.76**	**0.77**	**0.58**	**0.38**	**0.49**	**0.33**	**0.34**	-0.04	**0.69**	**0.50**	**0.91**	**0.39**	**0.28**	**0.23**	**0.34**	0.14

*significant values are bolded

Variable Comparison

Using the population and sex variables independently, a MANOVA test identifies significant differences between population (*p* = *<0.001*) and ILDs and sex (*p* = *<0.001*) and ILDs at the ($p < 0.001$) for the identified Latin American samples. An ANOVA test identified specific ILDs where these differences occur by sex and population affinity. Among populations, significant differences exist at the following ILDs: GOL, BNL, BBH, XCB, XFB, ZYB, AUB, MDH, OBH, OBB, FRC, PAC, and OCC. These ILDs include a wide array of breadth and height measurements. For sex, significant differences among the data exist at the following ILDs: GOL, BNL, BBH, XCB, XFB, WFB, ZYB, AUB, ASB, NLH, MDH, OBH, OBB, DKB, EKB, FRC, PAC, and OCC. Again, these are a combination of breadth and height measurements across the midfacial skeleton and vault.

When sex and population affinity are tested together, ANOVA identifies significant differences in sex and population affinity ($p < 0.001$) (Tables 5.10 and 5.11). A Tukey Two-Way test identifies significant differences between males and females and unidentified individuals and females. However, no significant differences are noted between males and unidentified individuals.

Table 5.10: ANOVA values by sex for metric data

	Male	*Female*
Female	**0.00**	–
Unknown	0.30	**0.00**

*significant values are bolded

For population affinity, the ANOVA identifies significant differences between the Identified Mexican Migrant and the INACIF sample (*p = 0.021*), the Identified Mexican Migrant and the Identified Guatemalan Migrant samples (*p = 0.036*), and the UADY and Identified Mexican Migrant samples (*p = 0.00012*). Notably, there are no significant differences between the Identified Guatemalan, UADY, and INACIF samples.

Table 5.11: ANOVA values by population for metric data

	INACIF (Guatemala)	*Identified Guatemalan Migrants*	*Identified Mexican Migrants*
Identified Guatemalan Migrants	0.95	–	–
Identified Mexican Migrants	**0.02**	**0.04**	–
UADY (Mexico)	0.86	0.99	**1.27×10^{-4}**

*significant values are bolded

An ANOVA test on the craniometric variables against the interaction of population and sex indicated significant differences ($p = 2 \times 10^{-16}$).

A series of Kruskal-Wallis tests identified significance among cranial MMS data and the variables of population and sex. All cranial MMS variables, except for PBD, are significantly different across the Latin American samples. Five cranial MMS traits are significantly different for sex, including Table 5.12 illustrates the p-values for each trait and variable tested.

Table 5.12: P-values for Kruskal-Wallis test on cranial MMS variables

	Population	*Sex*
ANS	**5.3×10^{-13}**	0.041
INA	**$<2.2 \times 10^{-16}$**	**0.008**
IOB	**$<2.2 \times 10^{-16}$**	0.556
MT	**$<2.2 \times 10^{-16}$**	0.079
NAS	**$<2.2 \times 10^{-16}$**	0.080
NAW	**$<2.2 \times 10^{-16}$**	0.066
NBS	**7.3×10^{-8}**	0.072
NBC	**$<2.2 \times 10^{-16}$**	0.002
NFS	**2.6×10^{-8}**	**5.1×10^{-6}**
NO	**$<2.2 \times 10^{-16}$**	0.063
OBS	**$<2.2 \times 10^{-16}$**	0.084
PBD	0.003	**0.007**
PZT	**6.9×10^{-7}**	**5.0×10^{-4}**
SPS	**1.5×10^{-5}**	0.088
TPS	**$<2.2 \times 10^{-16}$**	0.052
ZS	**7.4×10^{-7}**	**1.6×10^{-5}**

*significant values are bolded

Data Mining

Factor Analysis for Mixed Data (FAMD) is performed on the Latin American samples, first without and subsequently with the Unidentified Migrant sample. This examined patterns according to population affinity labels within the datasets. Using only data from the known Latin American samples, variation can be explained with five dimensions. The first dimension captures 21% of the variance, while the second dimension captures approximately 18% of the variance (Figure 5.13).

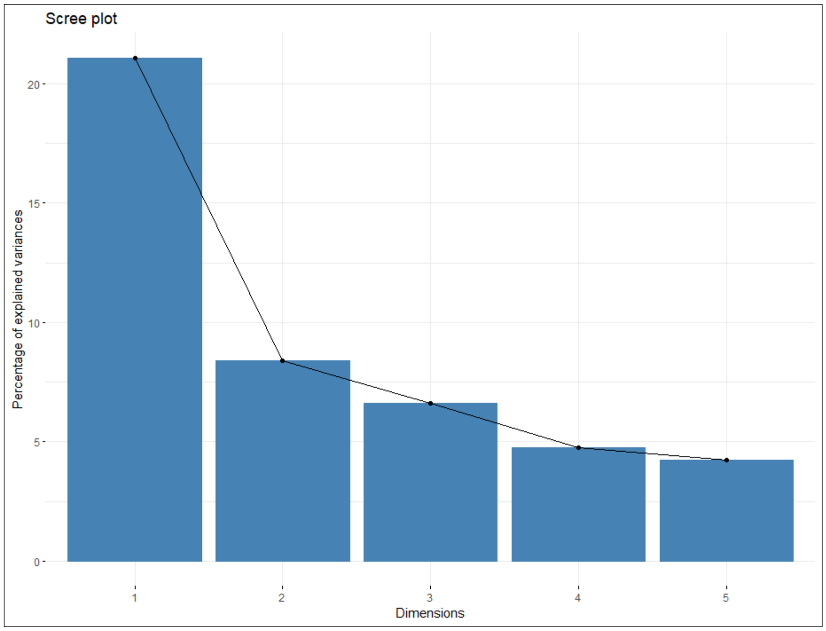

Figure 5.13: Scree plot from FAMD of identified Latin American samples

The main variables used to separate groups are presented in Figures 5.14 and 5.15. The most important variables in Dimension 1 are metric and include EKB, ZYB, OCL, WFB, BNL, OBB, XFB, MLB, NLH, FRC, MDH, BBH, SCB, PAC, and ASB (Figure 5.14). The most important variables contributing to group separation in dimension two are: Population and a combination of metric and cranial MMS variables (BBH, XCB, OCC, AUB, FRC and ANS) (Figure 5.15).

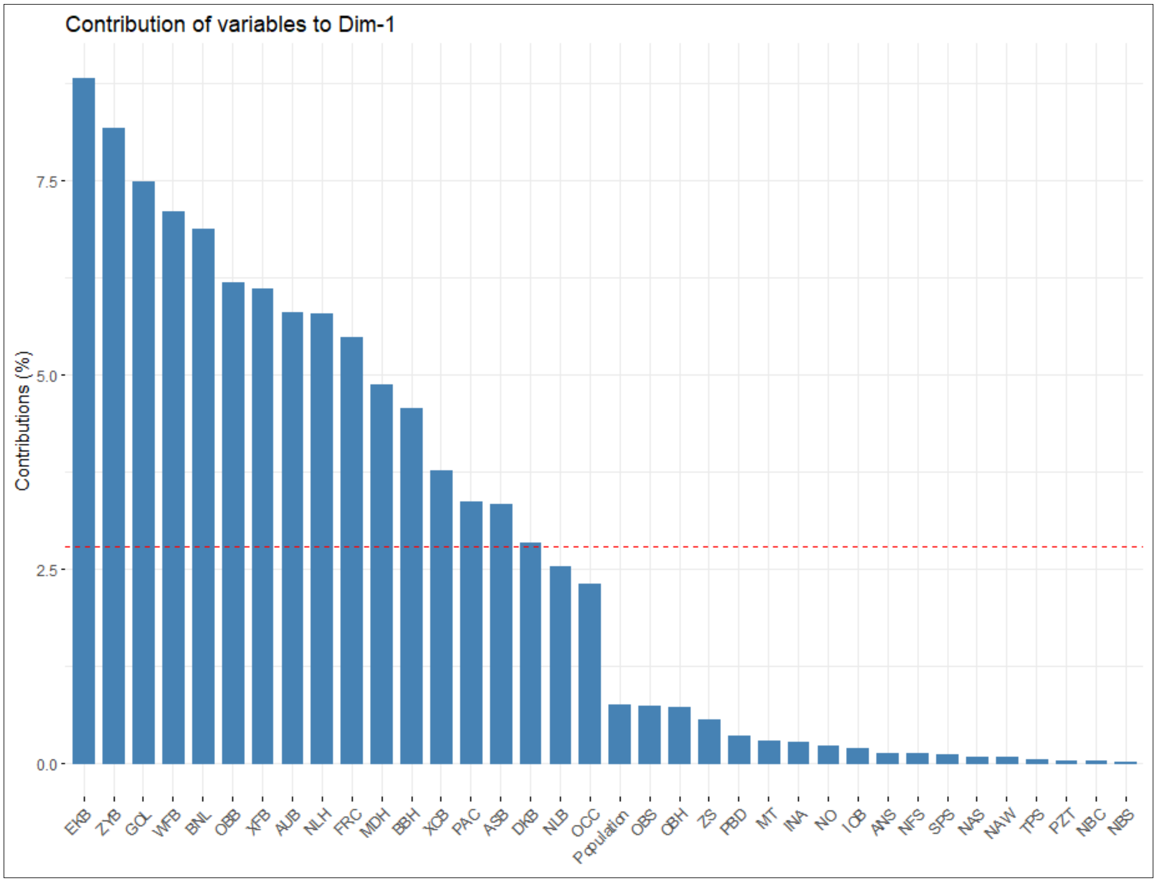

Figure 5.14: Variable contribution for dimension one (identified Latin American samples)

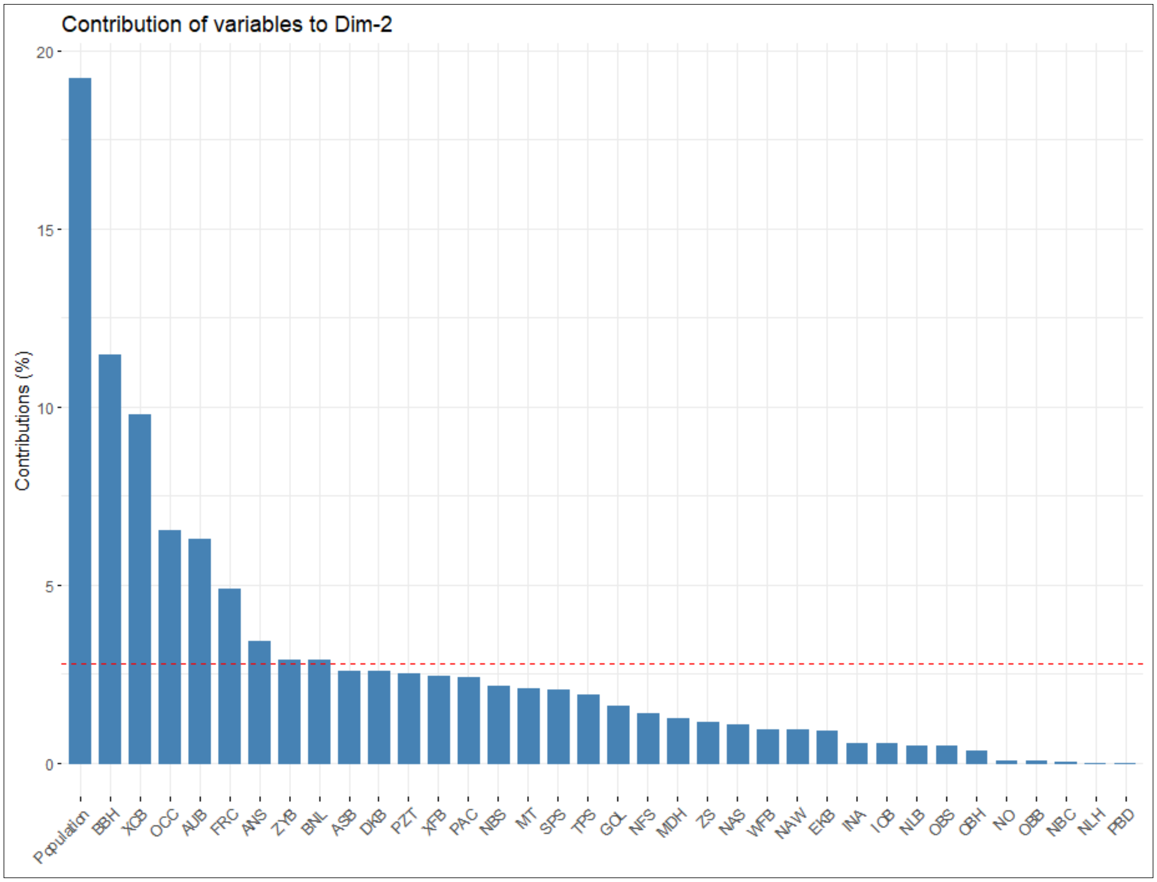

Figure 5.15: Variable contribution for dimension two (identified Latin American samples).

Data points for each individual are plotted and color-coded by population affinity (Figure 5.16). Dimension one isolates the UADY sample from the other Latin American samples. The INACIF, the Identified Guatemalan Migrant, and the Identified Mexican Migrant sample exhibit overlap with each other.

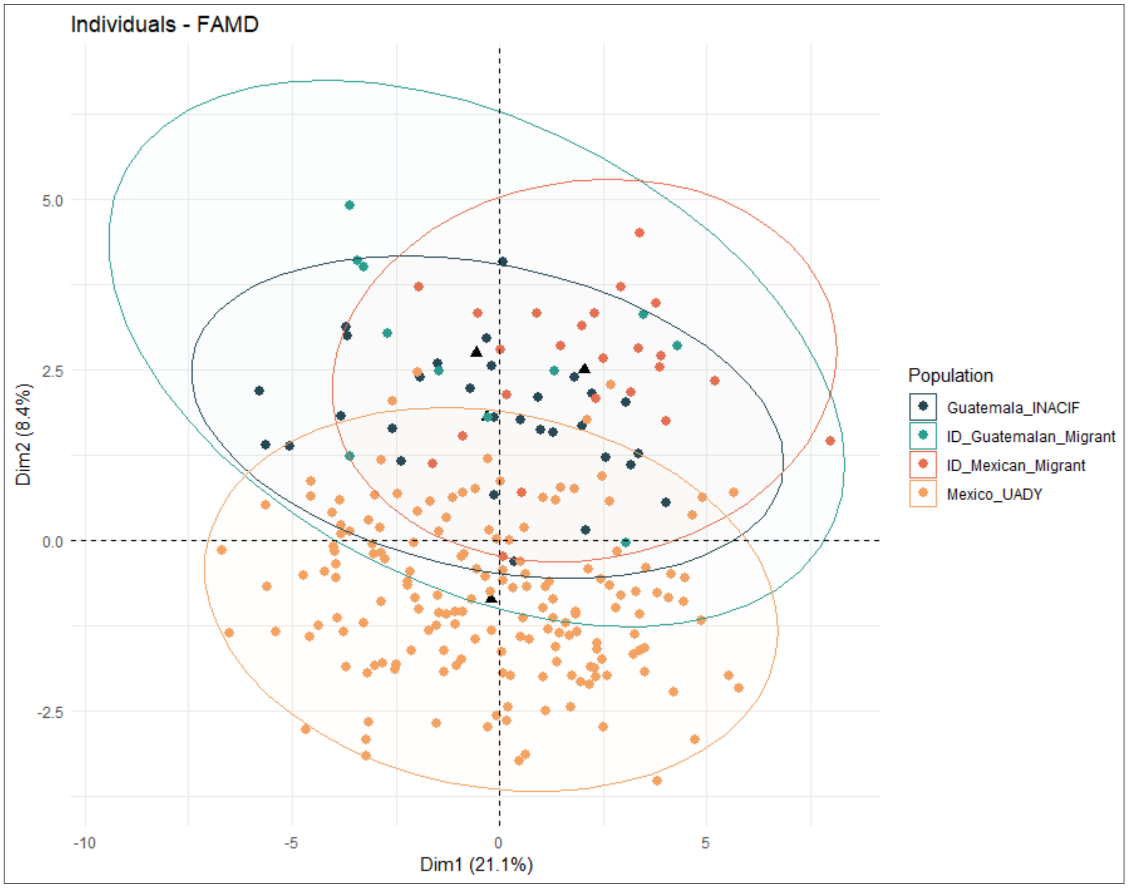

Figure 5.16: FAMD plot of Identified Latin American samples.

FAMD analysis is performed again, but this time with the Unidentified Migrant Sample is included. The first dimension captures approximately 20% of the variation across the dataset (Figure 5.17). Eigen values indicate that 30% of the variation is captured in the first two dimensions. The driving variables contributing to group separation in Dimensions 1 and 2 are presented in (Figure 5.18 and 5.19). The main variables in Dimension 1 separating the dataset into smaller clusters are metric (EKB, ZYB, GOL, BNL, WFB, OBB, XFB, FRC, NLH, AUB, BBH, MDH, ASP, and XCB). While the main separating variables in Dimension 2 are Population, metric variables (XCB, BBH, AUB, OCC, ZYB, FRC, DKB), and cranial MMS variables (IOB, PZT, TPS, MT, NFS).

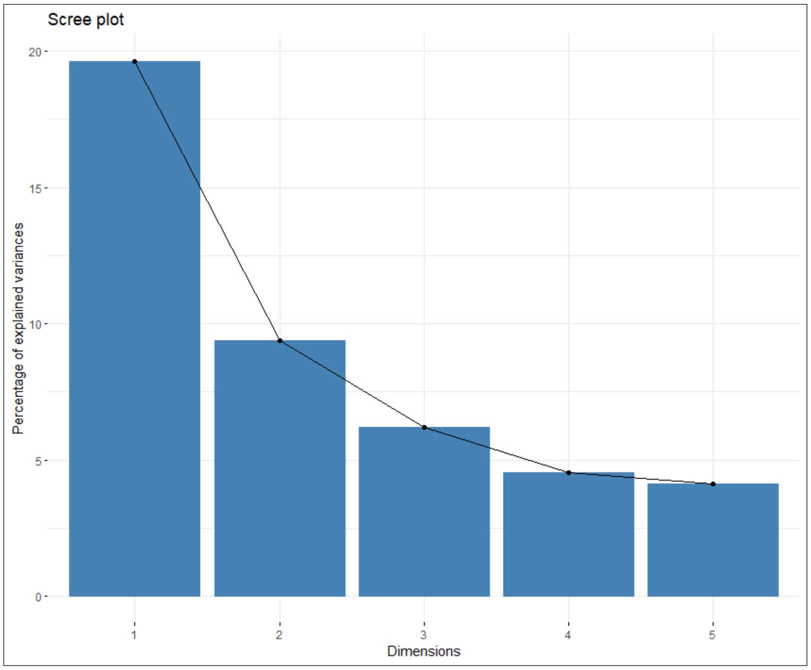

Figure 5.17: Scree plot from FAMD of all Latin American samples

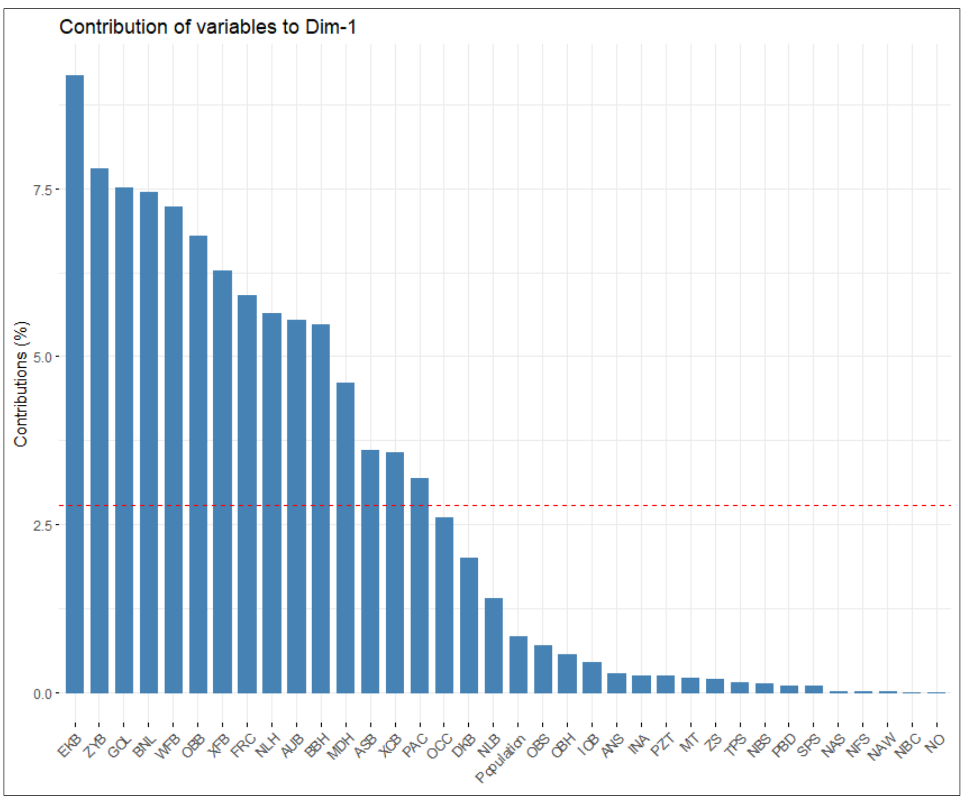

Figure 5.18: Variable contribution for dimension one (all Latin American samples).

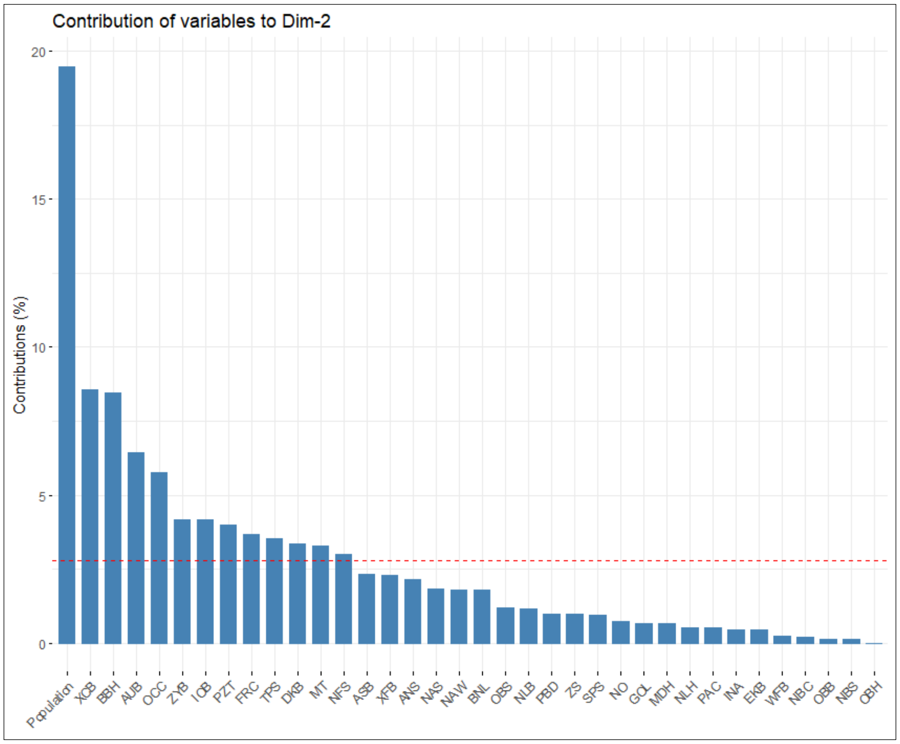

Figure 5.19: Variable contribution for dimension two (all Latin American samples).

Data points for each individual are plotted and color-coded by population affinity labels (Figure 5.20). The x-axis separates the majority of the UADY and Unidentified Migrant sample. The Identified Guatemalan Migrants and the Identified Mexican Migrants cluster within the Unidentified Migrant sample, while the INACIF sample overlaps all groups.

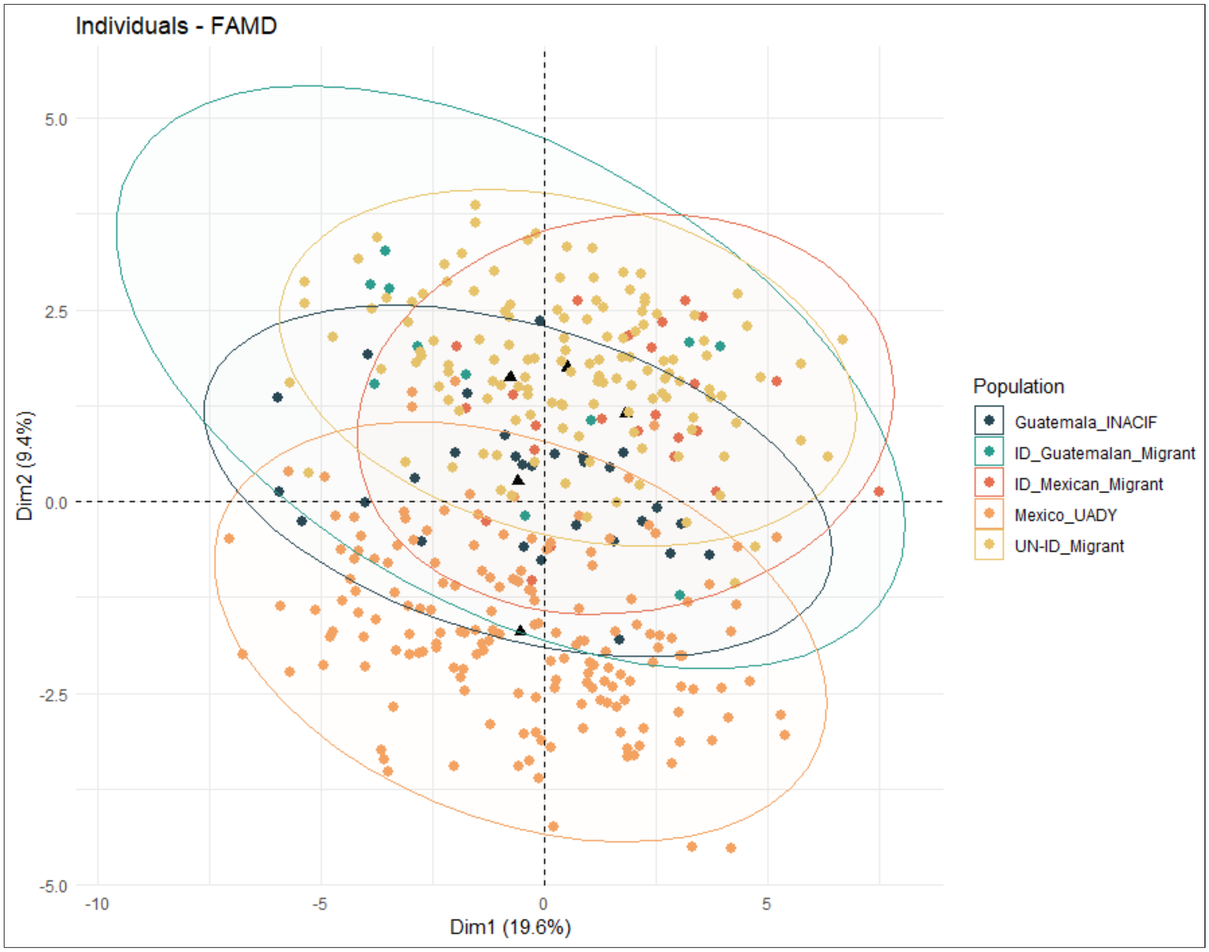

Figure 5.20: FAMD plot of all Latin American samples.

Research Question Two

To answer research question two, I created a Mahalanobis Distance matrix with craniometric data, a Smith's Mean Measure of Divergence matrix with the cranial MMS data, and used a Procrustes transformation to place the two matrices in the same multivariate space.

Mahalanobis Distance

Mahalanobis distance (MD) is calculated on the craniometric measurements in each sample to indicate levels of similarity and dissimilarity among samples. The first set of distances are calculated using the identified migrant samples (Mexican and Guatemalan), the UADY sample, and the INACIF sample. Results are visualized graphically (Figure 5.21) and presented as a dissimilarity matrix (Table 5.13). Distance measures indicate the Guatemalan groups (INACIF and Identified Guatemalan Migrants) are the most similar. The UADY sample is closer in multivariate space and more similar to the Guatemalan samples. The largest distance is between the two Mexican derived samples, the Identified Mexican Migrant sample and the UADY sample.

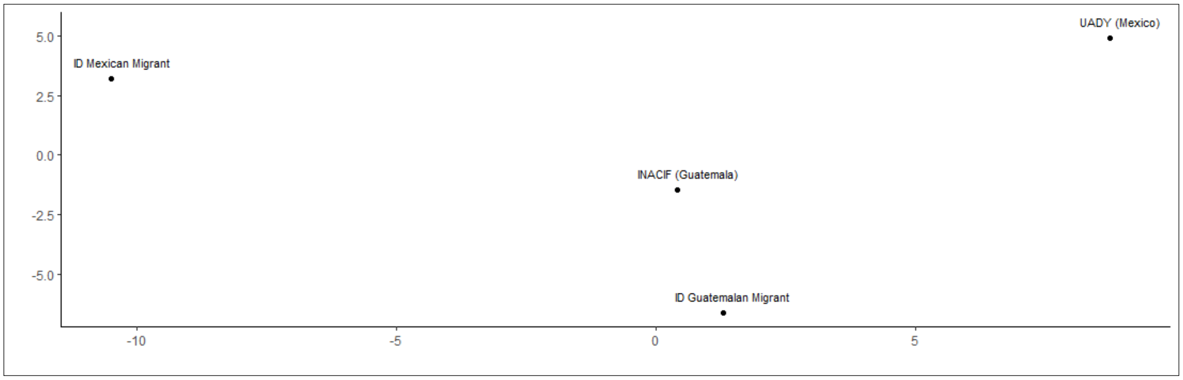

Figure 5.21: 2D scatterplot of Mahalanobis distance (identified Latin American samples).

Table 5.13: Mahalanobis distance (identified Latin American samples)

	INACIF (Guatemala)	*Identified Guatemalan Migrants*	*Identified Mexican Migrants*
Identified Guatemalan Migrants	9.85	–	–
Identified Mexican Migrants	13.54	15.02	–
UADY (Mexico)	12.34	13.20	18.75

A second set of distance measures are calculated on all samples including the Unidentified Migrant sample. Results are illustrated as a 2-dimensional scatterplot in Figure 5.22 and presented as a dissimilarity matrix in Table 5.14. The Unidentified Migrant sample is similar to the INACIF sample, but lies partway between the INACIF and Identified Mexican Migrant sample in multivariate space. The Guatemalan samples (INACIF and Identified Guatemalan Migrants) are most unlike the UADY then Identified Mexican Migrant samples. The UADY sample is most unlike the other samples.

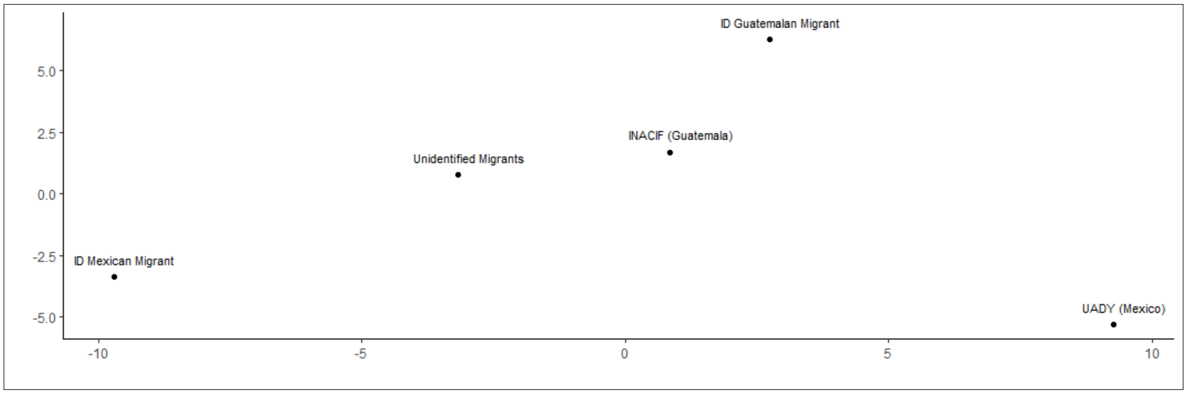

Figure 5.22: 2D scatterplot of Mahalanobis distance (including Unidentified Migrants)

Table 5.14: Mahalanobis distance (including the Unknown Migrant sample)

	INACIF (Guatemala)	*Identified Guatemalan Migrants*	*Identified Mexican Migrants*	*UADY (Mexico)*
Identified Guatemalan Migrants	9.87	–	–	–
Identified Mexican Migrants	13.47	15.40	–	–
UADY (Mexico)	12.55	12.95	18.41	–
Unidentified Migrants	7.78	9.41	8.87	13.69

Mean Measure of Divergence

Smith's MMD is calculated on the 16 cranial MMS traits: ANS, INA, PZT, PBD, NO, NAW, NBS, MT, and IOB. Only one nonpolymorphic trait, ZS is excluded. Frequency data for the dichotomized cranial MMS traits are listed in the appendix in Tables 5A.1-5A.16.

Figure 5.23 shows a 2-dimensional scatterplot based on MMD results for the identified Latin American samples. This scatterplot illustrates that all samples exhibit relative dissimilarity. The Identified Mexican Migrant and Identified Guatemalan Migrant samples appear to be more similar to each other, than the UADY or INACIF samples.

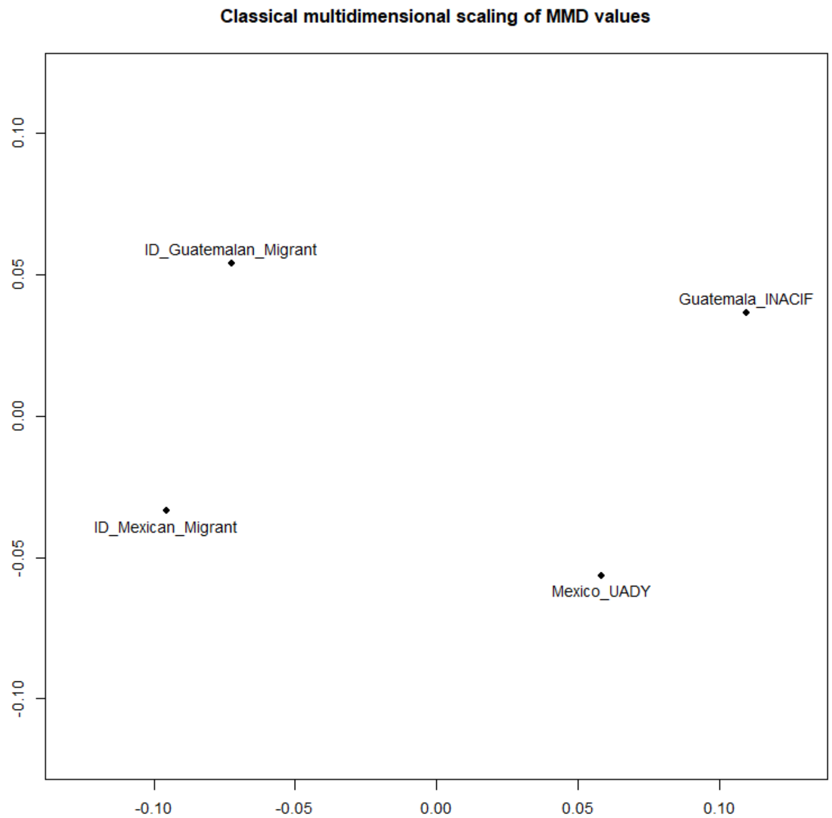

Figure 5.23: 2D scatterplot of MMD (identified Latin American samples).

The similarity/dissimilarity matrix for the identified Latin American samples is shown in Table 5.15. Values that are significant at the (*p = 0.05*) level are bolded. Most groups do not exhibit high dissimilarity scores to other groups. The UADY sample is the least similar to the Identified Guatemalan Migrant sample and is most similar to the INACIF sample. The Identified Guatemalan Migrant sample is most similar to the Identified Mexican Migrant sample, but is least similar to the INACIF then UADY samples. Finally, the INACIF sample is the least similar to the Identified Mexican Migrant sample, then the Identified Guatemalan Migrant sample, and is most similar to the UADY sample. Table 5.16 describes the variables in order of their discriminating power. For group separation with the analytical samples, TPS is the trait most useful, followed by ZS. The least useful traits for group discrimination are SPS and NAW.

Table 5.15: MMD dissimilarity matrix for cranial MMS variables (identified)

	INACIF (Guatemala)	*Identified Guatemalan Migrants*	*Identified Mexican Migrants*
INACIF (Guatemala)	–	–	–
Identified Guatemalan Migrants	**0.170**	–	–
Identified Mexican Migrants	**0.217**	0.052	–
UADY (Mexico)	**0.082**	**0.171**	**0.136**

*bolded values are statistically significant at (p = 0.05)

Table 5.16: Variable importance in MMD

Trait	*Overall MD*
TPS	4.11
ZS	2.65
NO	1.59
PZT	1.30
NBC	1.03
PBD	0.95
ANS	0.69
NAS	0.58
NFS	0.48
OBS	0.43
MT	0.13
IOB	0.02
INA	-0.08
NBS	-0.13
SPS	-0.19
NAW	-0.32

Figure 5.24 shows a 2-dimensional scatterplot based on MMD results with all samples including the Unidentified Migrant sample. This scatterplot illustrates that the Unidentified Migrant sample is most like the UADY sample. The next nearest similarity is the INACIF sample, followed by the Identified Mexican Migrant sample. The Identified Guatemalan Migrant sample appears to be the most dissimilar to all samples.

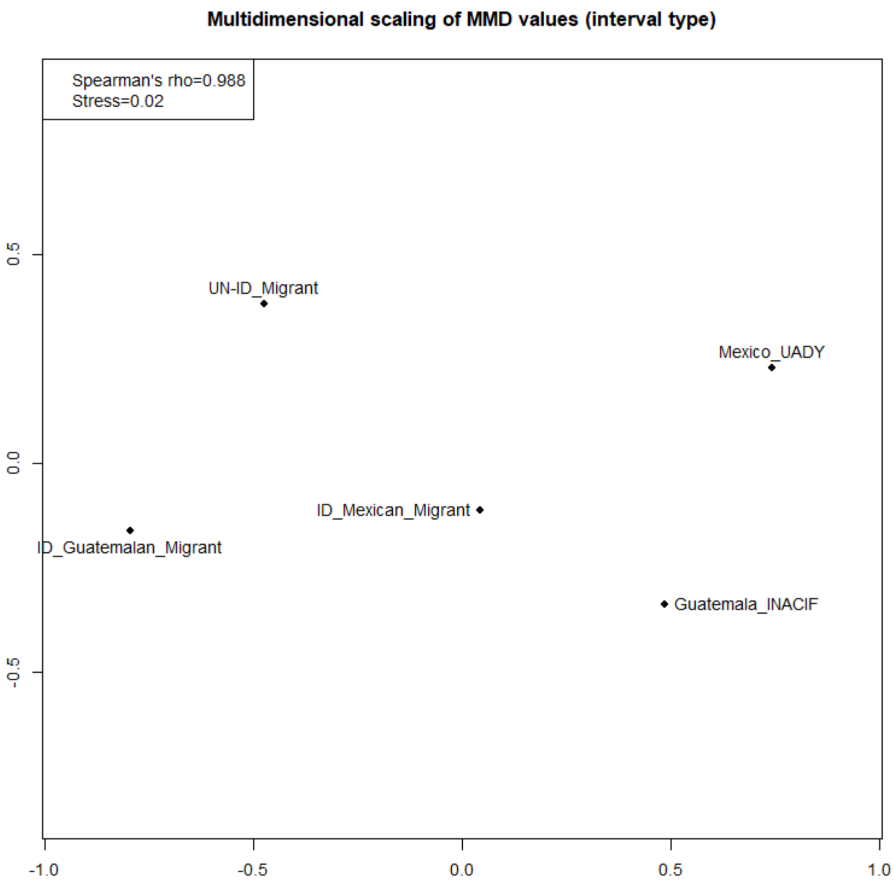

Figure 5.24: 2D scatterplot of MMD (all Latin American samples).

The similarity/dissimilarity matrix for all Latin American samples, including the Unidentified Migrant sample, is shown in Table 5.17. Values that are significant at the ($p = 0.05$) level are presented in bold text. Most groups do not exhibit high dissimilarity scores to other groups. The Identified Mexican sample does not show strong dissimilarities toward any groups, and is similar to all samples in this research according to the MMD. The UADY sample is the least similar to the Identified Guatemalan Migrant sample, then Unidentified Migrant sample, but is more similar to the INACIF and Mexican Migrant sample. The Identified Guatemalan Migrant sample is most similar to the Unidentified Migrant sample, then the Mexican Migrant sample, but is least similar to the UADY then INACIF samples. Finally, the INACIF sample is the least similar to the UADY sample, then the Guatemalan Migrant sample, the Unidentified Migrant sample, and most similar to the Mexican Migrant sample.

Table 5.17: MMD dissimilarity matrix for cranial MMS variables (all)

	INACIF (Guatemala)	*Identified Guatemalan Migrants*	*Identified Mexican Migrants*	*UADY (Mexico)*
INACIF (Guatemala)	–			
Identified Guatemalan Migrants	**0.215**	–		
Identified Mexican Migrants	0.005	0.085	–	
UADY (Mexico)	**0.039**	**0.287**	**0.069**	–
Unidentified Migrants	**0.176**	0.039	**0.066**	0.195

*bolded values = statistically significant at (p = 0.05)

Table 5.20 describes the variables in order of their discriminating power. For group separation with these samples, IOB is the most useful trait, followed by NO. The least useful traits for group discrimination are INA and PZT.

Table 5.18: Variable importance in MMD

Trait	*Overall MD*
IOB	3.16
NO	2.25
NBS	1.46
MT	1.39
PBD	1.23
NAW	1.06
ANS	0.80
INA	-0.15
PZT	-0.62

Procrustes Transformation

The MMD matrix is transformed to the same space as the Mahalanobis dissimilarity matrix. Figure 5.25 displays the transformation plot of the cranial MMS and craniometric variables by sample. The craniometric and cranial MMS data for each sample are near to each other.

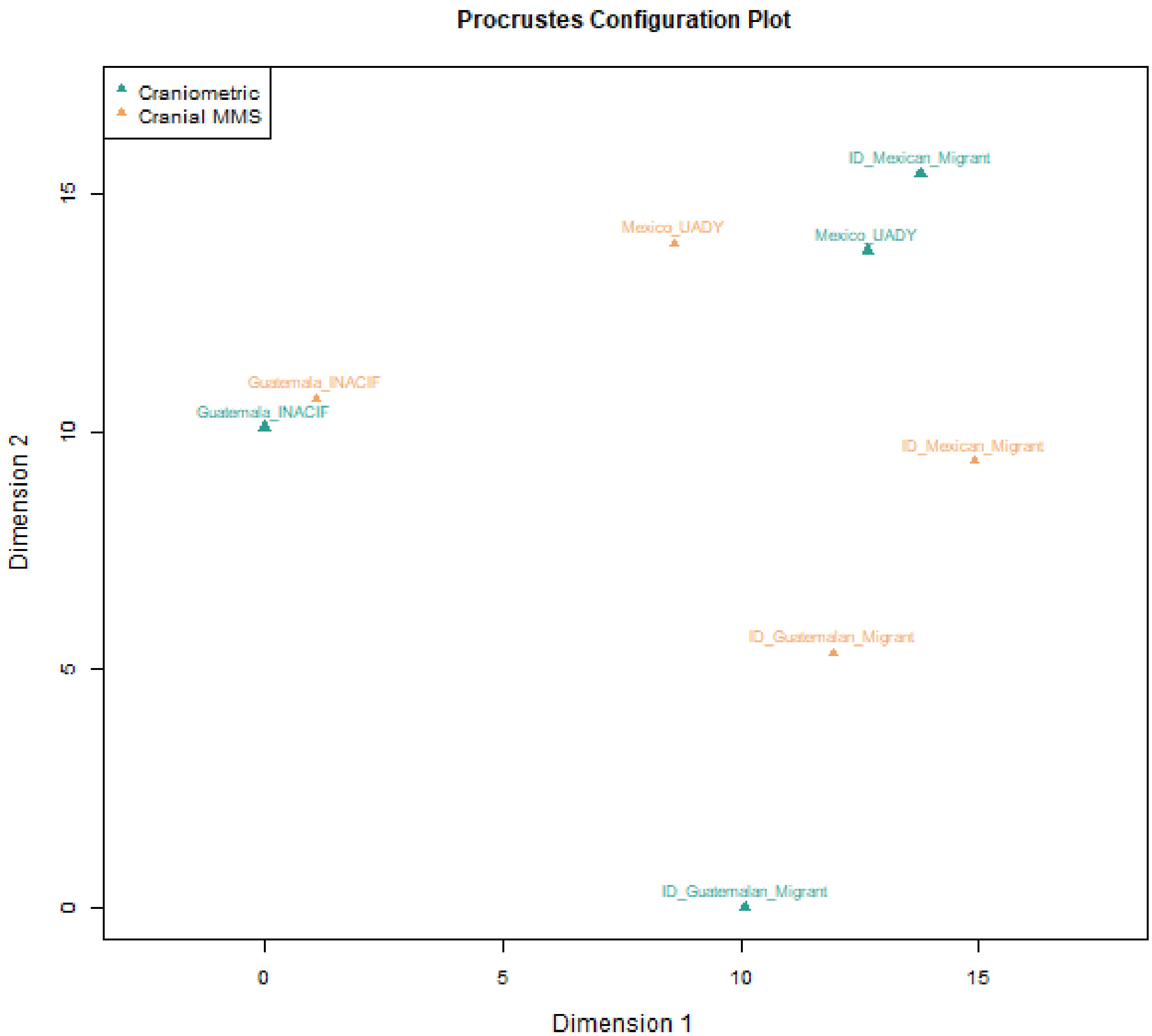

Figure 5.25: Procrustes transformation plot.

A Mantel Test on the dissimilarity matrices from the Mahalanobis distance and MMD analyses is graphically represented in Figure 5.26. The p-value ($p = 0.349$) indicates that the matrices are linearly correlated with each other. The vertical line in Figure 5.26 shows the observed z-statistic.

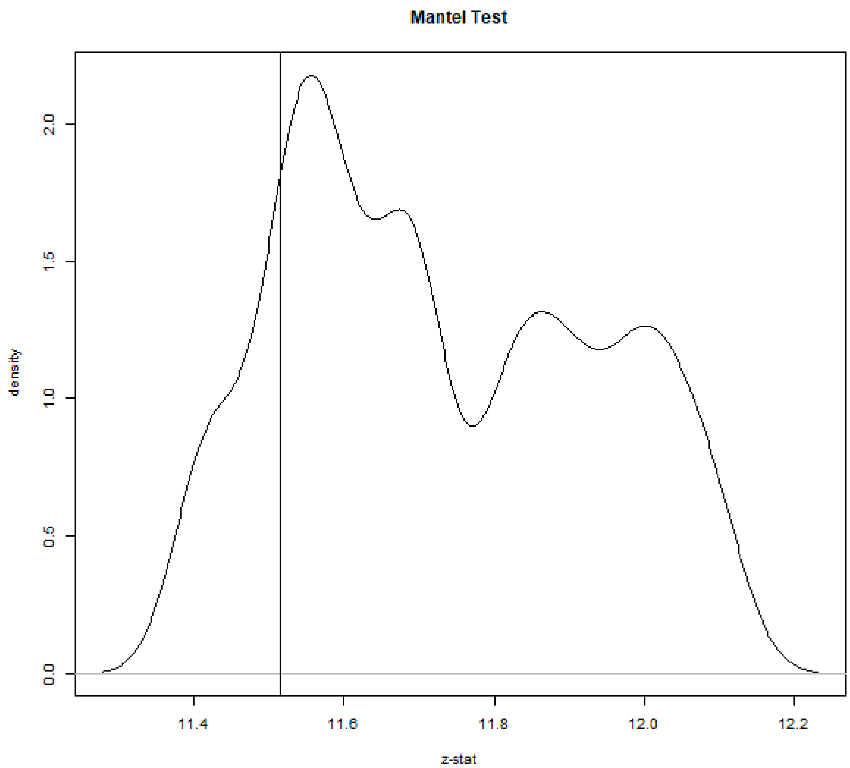

Figure 5.26: Mantel test results.

Research Question Three

To answer research question three, I created different classification models using aNN Prior to modeling, the metric data were centered to remove any sex influence on ILDs. Next, the data were divided into train and test sets. The training data comprise a random 70% (*n = 417*) of the original sample; the test data comprise the remaining 30% (*n = 180*). The training and testing samples are presented by the data type used in each model in Table 5.19.

Table 5.19: Train/test datasets for modeling

	American Black	*American White*	*INACIF*	*Identified Guatemalan Migrants*	*Identified Mexican Migrants*	*Thai*	*UADY*	***Total***
Train (MMS)	23	43	17	8	17	121	188	**417**
Test (MMS)	11	26	13	3	7	47	73	**180**
Train (metric)	25	43	22	9	13	118	187	**417**
Test (metric)	9	26	8	2	11	50	74	**180**
Train (metric + MMS)	23	49	19	8	15	123	179	**417**
Test (metric + MMS)	11	20	11	3	9	45	82	**181**

Artificial Neural Networks

Prior to creating and testing the aNN models, the number of hidden layers selected is determined. Figures 5.27, 5.28, and 5.29 illustrate a conservative approach where the optimal value is selected to avoid overfitting the model. Seven hidden layers were selected for the craniometric model, achieving stability without overfitting the data and providing overly optimistic results. A threshold value of eight is the ideal value for the cranial MMS aNN model, as the model exhibits stability and the lowest group CCR is above 25%. With a size value of nine, the cranial MMS aNN model deteriorates markedly for the Identified Guatemala Migrant group to 12% CCR, before increasing to 75% at a size value of ten. A threshold value of four is the ideal value for the combined craniometric and cranial MMS model, because the model is stable at this value and quickly jumps to 100% CCR for all groups at threshold values of six and above.

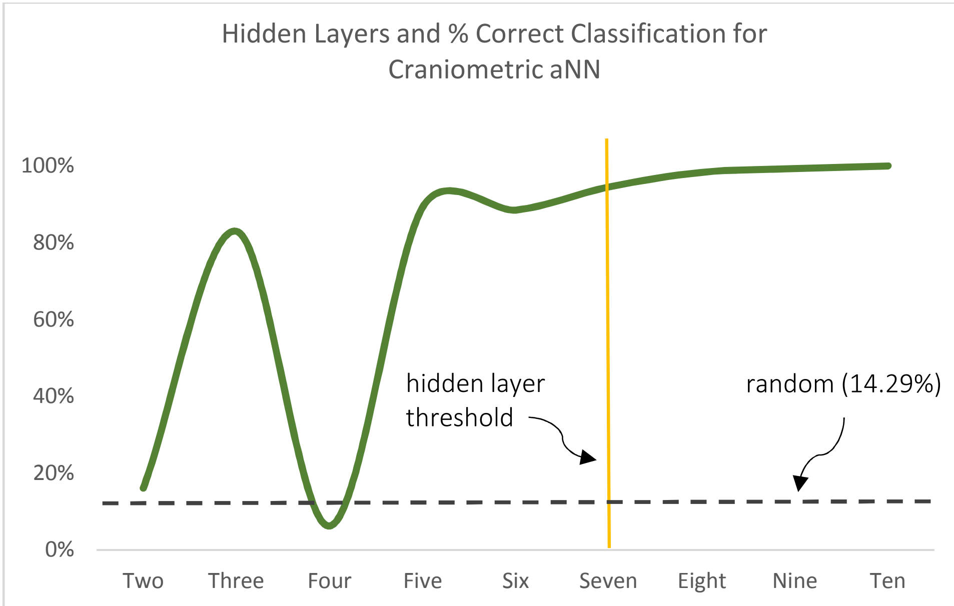

Figure 5.27: Threshold value for craniometric model.

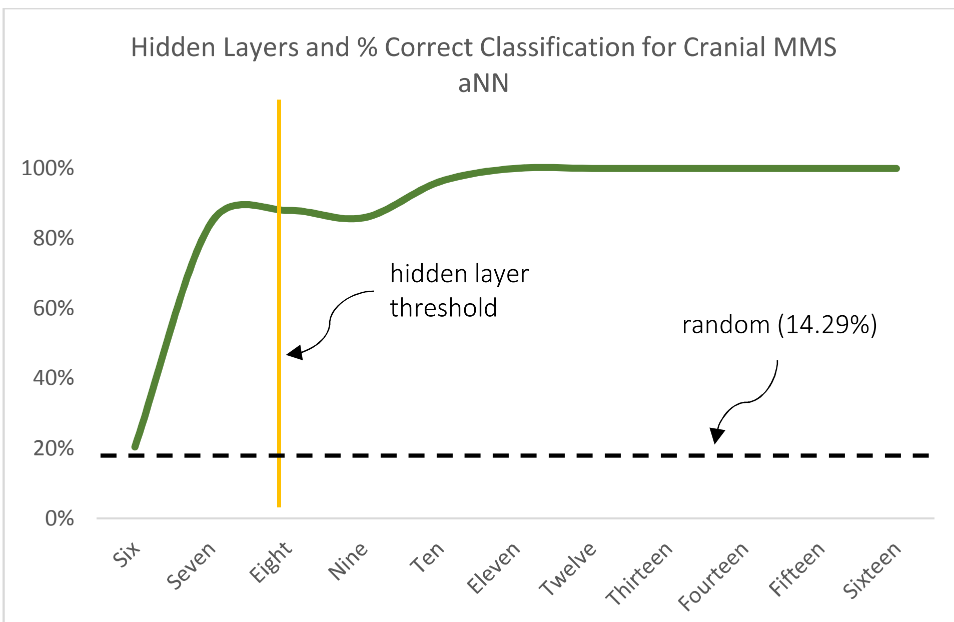

Figure 5.28: Threshold value for cranial MMS model.

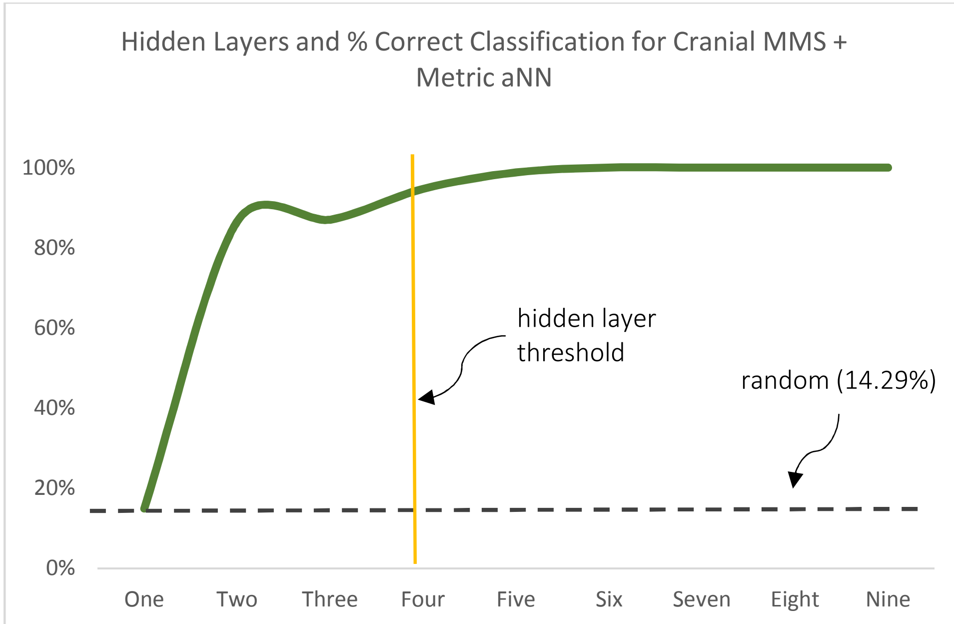

Figure 5.29: Threshold value for cranial MMS + craniometric model.

Tables 5.20-5.22 show the confusion matrices for the train data used to build each of the aNN models. CCR data for each sample and an overall model CCR are presented. The cranial MMS only model and craniometric only model both show overall CCRs greater than 90.0%. The combined cranial MMS and craniometric model exhibits a CCR of 94.7%. Individual classifications are lowest across single variable models for the Identified Mexican Migrants with a CCR of 56.3% (metric) and 52.9% (cranial MMS). In the combined model, the CCR increases to 73.3% for this group. The INACIF sample performs well in both single variable models (>80.9%), and classifies everyone correctly in the combined model. All samples remain relatively stable across all models, with a general pattern of an overall higher classification rate in the combined model. Even the worst CCRs are higher than chance (14.3%).

Table 5.20: Confusion matrix for training dataset for the craniometric model

	American Black	*American White*	*INACIF*	*Identified Guatemalan Migrants*	*Identified Mexican Migrants*	*Thai*	*UADY*	*% CCR*
American Black	22	3	0	0	0	0	0	88.0
American White	0	43	0	0	0	0	0	100.0
INACIF	0	0	20	2	0	0	0	90.9
Identified Guatemalan Migrants	0	0	0	9	0	0	0	100.0
Identified Mexican Migrants	0	3	0	0	10	0	0	76.9
Thailand	0	0	0	0	0	187	0	100.0
UADY	0	1	0	0	0	0	117	99.2
							Total:	**97.8**

Table 5.21: Confusion matrix for training dataset for the cranial MMS model

	American Black	*American White*	*INACIF*	*Identified Guatemalan Migrants*	*Identified Mexican Migrants*	*Thai*	*UADY*	*% CCR*
American Black	21	0	0	0	0	2	0	91.3
American White	0	39	0	0	0	0	4	90.7
INACIF	0	1	16	0	0	0	1	94.1
Identified Guatemalan Migrants	0	2	1	6	0	0	1	75.0
Identified Mexican Migrants	0	1	2	0	9	0	5	52.9
Thailand	1	1	0	0	0	183	3	97.3
UADY	0	1	5	0	0	8	107	88.4
							Total:	**91.4**

Table 5.22: Confusion matrix for training dataset for the combined model

	American Black	*American White*	*INACIF*	*Identified Guatemalan Migrants*	*Identified Mexican Migrants*	*Thai*	*UADY*	*% CCR*
American Black	23	0	0	0	0	0	0	100.0
American White	5	44	0	0	0	0	0	89.8
INACIF	0	0	19	0	0	0	0	100.0
Identified Guatemalan Migrants	0	0	0	5	0	1	2	62.5
Identified Mexican Migrants	0	1	0	0	11	0	3	73.3
Thailand	0	1	0	0	0	178	0	99.4
UADY	0	2	4	0	0	0	178	92.7
							Total:	**94.7**

Variable importance is assessed to show which variables contribute the most to each model. Variable importance for the metric only model is shown in Figure 5.30. The variables that contribute the most to the model are FRC, XFB, WFB, GOL, ZYB, OBB, XCB, AUB, OBH, PAC, MDH, and BNL.

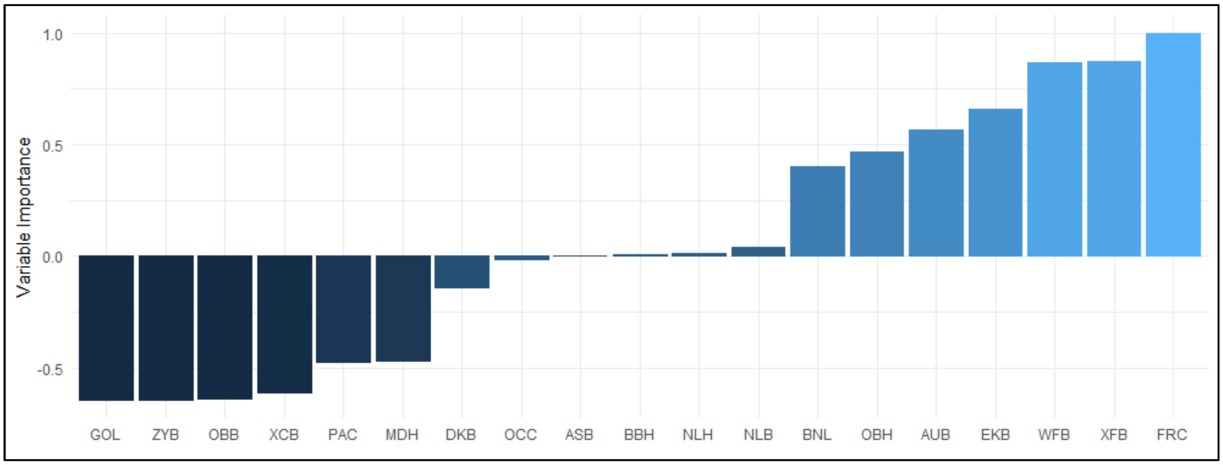

Figure 5.30: Variable importance graph for craniometric model.

Figure 5.31 shows the variable importance for the cranial MMS only model. The variables that contribute the most to the model are NBS, INA, PZT, ANS, NAW, NFS, PBD, MT, and OBS.

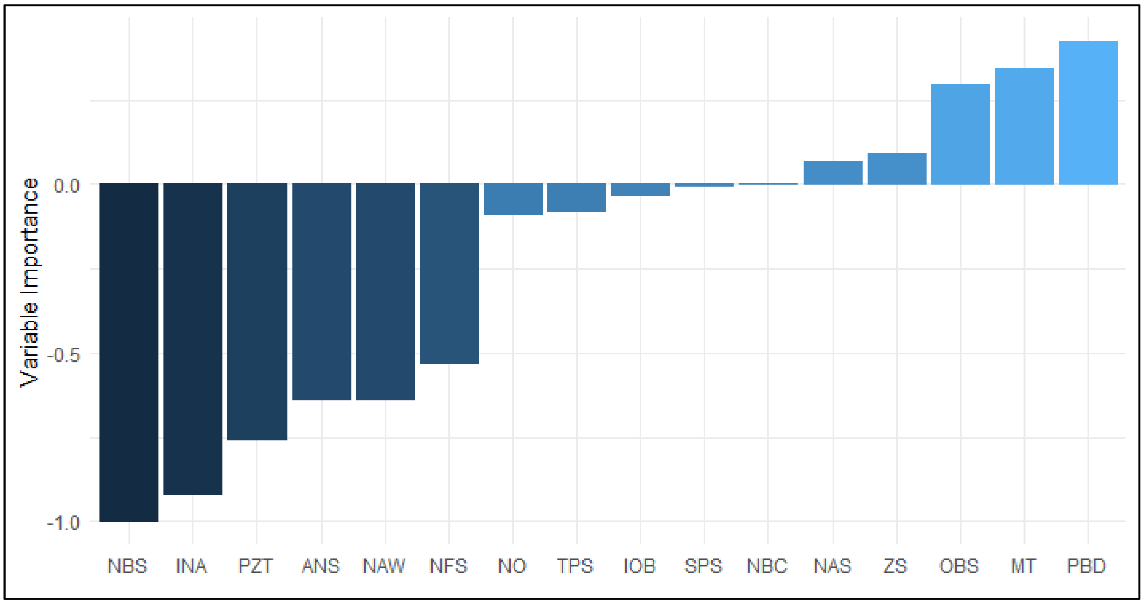

Figure 5.31: Variable importance graph for cranial MMS model.

Finally, Figure 5.32 shows the variable importance graphic for the craniometric and cranial MMS variable model. The variables most impacting the model are DKB, GOL, OBH, NAW, NAS, XCB, OBB, NBC, PAC, ZS, NLB, AUB, MT, EKB, SPS, ANS, NBS, FRC, PZT, INA, XFB, BNL, and PBD.

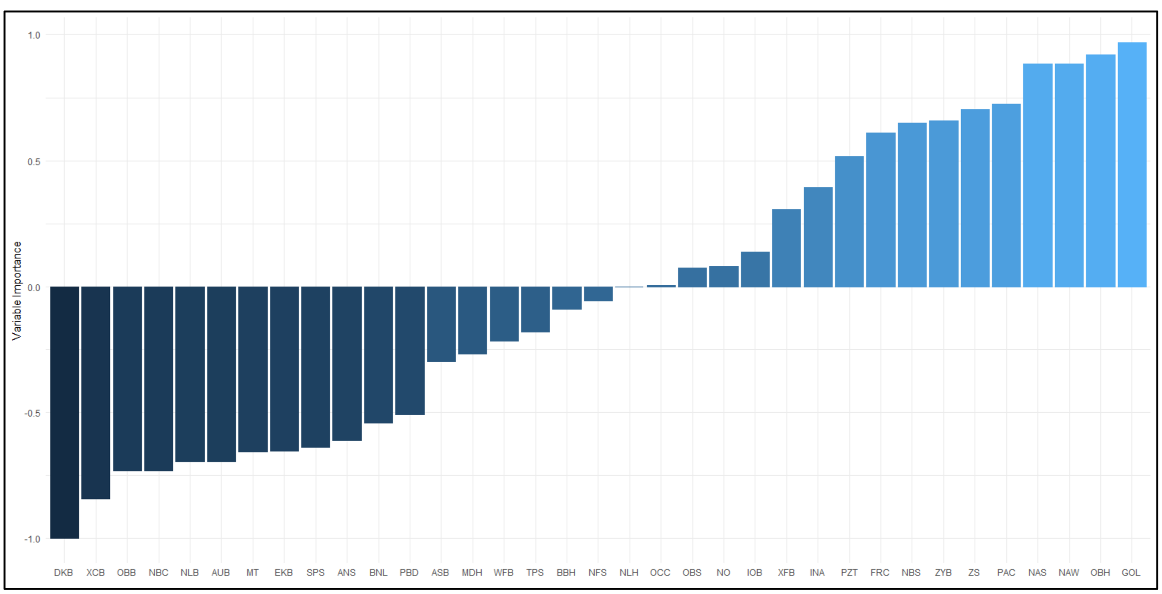

Figure 5.32: Variable importance graph for combined model

Tables 5.23-5.25 show the classification matrices for the testing datasets for each of the aNN models. CCR data for each test sample and the overall model CCR are presented in the tables. The cranial MMS only model shows a CCR of 54.4%, which is the lowest of the three models. The craniometric only aNN model shows a CCR of 66.1%, and the combined model shows the highest CCR at 70.7%. Despite overall classification rates higher than 50.0%, CCR data for specific individual samples is low. For example, the craniometric only and cranial MMS only models do not provide a single correct classification for individuals in the Identified Guatemalan Sample and Identified Migrant Samples. Furthermore, in the combined metric and cranial MMS model, there are no correct classifications for the individuals in the Identified Mexican Migrant sample. Interestingly, the American White sample decreases in accuracy for the combined cranial MMS and craniometric model, misclassifying individuals in the American Black, Thai, UADY, and INACIF samples. The CCR for all samples, except the Identified Mexican Migrant samples, perform better than chance (14.3%) allocations in the combined craniometric and cranial MMS model. However, this is not true for the Identified Migrant samples in the craniometric only model, and the Identified Migrant and INACIF samples in the cranial MMS only model.

Table 5.23: Classification matrix for testing data for the craniometric model

	American Black	*American White*	*INACIF*	*Identified Guatemalan Migrants*	*Identified Mexican Migrants*	*Thai*	*UADY*	*% CCR*
American Black	7	1	0	0	1	0	0	77.8
American White	3	20	1	0	1	0	1	76.9
INACIF	1	0	2	0	1	2	2	25.0
Identified Guatemalan Migrants	1	0	0	0	0	1	0	0.0
Identified Mexican Migrants	2	1	2	0	0	5	1	0.0
Thailand	1	1	3	3	0	57	9	77.0
UADY	1	4	5	1	0	6	33	66.0
							Total:	**66.1**

Table 5.24: Classification matrix for testing dataset for the cranial MMS model

	American Black	*American White*	*INACIF*	*Identified Guatemalan Migrants*	*Identified Mexican Migrants*	*Thai*	*UADY*	*% CCR*
American Black	2	2	0	0	0	4	2	18.2
American White	2	14	2	1	0	0	7	53.9
INACIF	1	0	1	1	0	6	4	7.7
Identified Guatemalan Migrants	0	1	0	0	0	1	1	0.0
Identified Mexican Migrants	0	3	0	0	0	2	2	0.0
Thailand	3	1	3	2	2	56	6	76.7
UADY	1	5	3	0	0	11	25	53.2
							Total:	**54.4**

Table 5.25: Classification matrix for testing dataset for the combined model

	American Black	*American White*	*INACIF*	*Identified Guatemalan Migrants*	*Identified Mexican Migrants*	*Thai*	*UADY*	*% CCR*
American Black	7	1	1	0	1	1	0	63.6
American White	6	7	1	0	0	3	3	35.0
INACIF	1	0	4	0	0	3	3	36.4
Identified Guatemalan Migrants	0	0	0	1	0	1	1	33.3
Identified Mexican Migrants	1	2	0	0	0	4	2	0.0
Thai	0	0	1	1	0	77	3	93.9
UADY	0	3	3	0	0	7	32	71.1
							Total:	**70.7**

Model Selection

Overall model percentages for correct classification of the test data are presented in Table 5.26. However, the Matthew's Correlation Coefficient statistics identify the best performing model based on sample size and the results of confusion matrix categories (true positives, false negatives, true negatives, and false positives).

Table 5.26: Classification rates by model

	Craniometric + cranial MMS	*Craniometric only*	*Cranial MMS only*
aNN	70.7%	66.1%	54.4%

Matthew's Correlation Coefficient

Matthew's Correlation Coefficient is calculated from the test data for each model. The results are compared across models to asses which models perform the best. Overall, the combined craniometric + cranial MMS model perform better than each of the models based on only one data type, craniometric or cranial MMS. Each model is compared to each other and the value for the MCC listed in Table 5.27.

Table 5.27: MCC values for each testing model

Cranial MMS only	0.37
Craniometric only	0.54
Craniometric + cranial MMS	0.58

Exploratory Analyses

Combined Latin American Sample

Individuals from the INACIF and Identified Migrant samples (Guatemalan and Mexican) were modeled together within the aNN framework to understand classification rates on a pooled sample. The model uses both craniometric (centered) and cranial MMS data. A hidden layer threshold value of four is chosen for training the model (Figure 5.33).

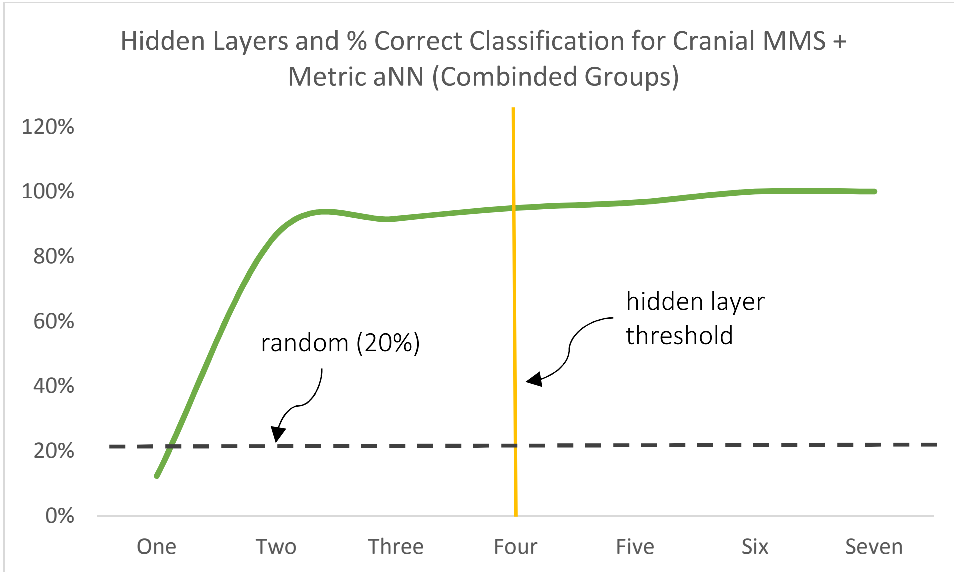

Figure 5.33: Threshold value for combined model using an exploratory pooled dataset.

Classification accuracies for the training model are shown in Table 5.28. The training model correctly classified four of the samples at ~93% or higher. The American Black sample does not perform as well with a correct classification rate of 60.9%. American Black individuals are exclusively classified as American White in this sample and model. The INACIF + Identified Migrant sample classifies well, with one individual misclassifying as American Whites, and two individuals misclassifying as the UADY sample.

Table 5.28: Classification matrix for training dataset with five groups (combined model)						
	American Black	*American White*	*INACIF + Identified Migrants*	*Thai*	*UADY*	*% CCR*
American Black	14	9	0	0	0	60.9
American White	0	47	0	2	0	95.9
INACIF + Identified Migrants	0	1	39	0	2	92.9
Thailand	0	0	1	178	0	99.4
UADY	1	0	4	1	117	95.1
					Total:	**94.9**

Results of the testing dataset on the model are shown in table 5.29. The overall classification rate is 74.6%. The lowest classification accuracy is for the American Black sample, in which no individual correctly classified. The most misclassifications for these individuals occur in the American White sample (*n* = 8) then the INACIF + Identified Migrant samples (*n* = 2), then the UADY (*n* = 1). The INACIF + Identified Migrant samples correctly classifies at 26.1%, which is just above random allocation (20.0%). Most misclassifications occur as American White (*n* = 5), Thai (*n* = 5) and UADY (*n* = 6). The American White, Thai, and UADY samples all show classification rates above 70.0%.

Table 5.29: Classification matrix for test dataset with five groups (combined model)

	American Black	*American White*	*INACIF + Identified Migrants*	*Thai*	*UADY*	*% CCR*
American Black	0	8	2	0	1	0.0
American White	0	14	1	5	0	70.0
INACIF + Identified Migrants	1	5	6	5	6	26.1
Thailand	0	1	1	75	5	91.5
UADY	0	1	1	3	40	88.9
					Total:	**74.6**

Unidentified Migrant Sample

Data from the Unidentified Migrant Sample is tested on the combined cranial MMS and craniometric models. These individuals have complete datasets, and the means are centered for metric data. Classification data for the aNN model are listed in Tables 5.30. Most of the Unidentified Migrant sample classifies as UADY ($n = 55$), followed by American Black ($n = 22$), INACIF ($n = 18$), American White ($n = 16$), Thai ($n = 15$), then Identified Guatemalan Migrants (n = 4) and Identified Mexican Migrants ($n = 3$). It is impossible to determine accuracy as the individuals in the Unidentified Migrant sample are unknowns. However, based on previously published data and statistics, the unidentified individuals housed at PCOME and OpID are likely from Mexico, Guatemala, El Salvador, and Honduras.

Table 5.30: Classification matrix for the Unidentified Migrant sample (combined model)							
	American Black	*American White*	*INACIF*	*Identified Guatemalan Migrants*	*Identified Mexican Migrants*	*Thai*	*UADY*
Unidentified Migrant Sample	22	16	18	4	3	15	55

Incomplete Cases

A second exploratory analysis is done using incomplete data from the INACIF, UADY, and the migrant samples where only craniometric or only cranial MMS data are available. This scenario emulates situations in forensic practice where remains are damaged or incomplete for several reasons including damage due to trauma or the environment. Additionally, cranial MMS traits and craniometrics are not collected where antemortem trauma has altered the shape of the bone (i.e., a previously broken nose, evidence of cranial surgery or healed cranial vault trauma). Damage from the environment is more often associated with migration contexts in Arizona due to the extreme temperatures in remote locations where migrant remains are often found, and the short amount of time that extreme temperatures and carnivore activity can impact skeletal remains (De León 2015). Table 5.31 illustrates the number of individuals with each type of data present for the test samples used.

Table 5.31: Summary of exploratory incomplete data				
	Metric Only	*MMS Only*	*Metric and MMS*	**Total:**
Identified Mexican Migrant	0	6	0	**6**
INACIF	0	8	0	**8**
UADY	6	10	0	**16**
Unidentified Migrant	0	39	0	**39**

The incomplete data are tested in each of the models, in order to understand classification patterns with craniometric or cranial MMS data only. In the combined craniometric and cranial MMS model, no single individual classified in any of the groups. Classification rates from individuals with only craniometric data available, which are UADY individuals (*n* = *6*) are presented in Table 5.32. The models classify only 33.3% of the total sample correctly, and one individual is not classified.

Table 5.32 Classification matrix using the exploratory data for the craniometric model

	American Black	*American White*	*INACIF*	*Identified Guatemalan Migrants*	*Identified Mexican Migrants*	*Thai*	*UADY*	*CCR*
UADY	1	1	0	0	0	1	2	33.3%

The individuals with only cranial MMS data show varied classification rates (Table 5.33). The entire INACIF (*n* = *8*) sample, and many individuals in the Identified Mexican Migrant (*n* = *6*), and UADY *(n* = *7*) samples do not classify. Of those that do, one Identified Mexican Migrant classified as UADY, which is not incorrect, but may not be accurate since specific region of origin data is unavailable. Of the UADY individuals that do classify (*n* = *2*), they correctly classify.

Table 5.33: Classification matrix using the exploratory data for the cranial MMS model

	American Black	*American White*	*INACIF*	*Identified Guatemalan Migrants*	*Identified Mexican Migrants*	*Thai*	*UADY*	*CCR*
INACIF	0	0	0	0	0	0	0	0.0%
Identified Mexican Migrant	0	0	0	0	0	0	1	Unknown
UADY	0	0	0	0	0	0	2	12.5%
Unidentified Migrant	0	7	2	0	1	9	9	Unknown

APPENDIX

Summary frequency data for each cranial MMS trait are listed in tables 5A.1-5A.16.

Table 5A.1: Frequency distribution of anterior nasal spine (ANS)

		INACIF (Guatemala) (n=37)		*Identified Guatemalan Migrants (n=10)*		*Identified Mexican Migrants (nu=24)*		*Unidentified Migrants (n=129)*		*UADY (Mexico) (n=136)*	
Character State	*0/1*	n	%	n	%	n	%	n	%	n	%
1	*0*	8	*21.6*	2	*20.0*	8	*33.3*	42	*32.6*	82	*60.3*
2	*1*	26	*70.3*	7	*70.0*	12	*50.0*	62	*48.1*	49	*36.0*
3	*1*	3	*81.1*	1	*10.0*	4	*16.7*	25	*19.3*	5	*3.7*

Table 5A.2: Frequency distribution of inferior nasal aperture (INA)

		INACIF (Guatemala) (n=38)		*Identified Guatemalan Migrants (n=11)*		*Identified Mexican Migrants (n=24)*		*Unidentified Migrants (n=133)*		*UADY (Mexico) (n=153)*	
Character State	*0/1*	n	%	n	%	n	%	n	%	n	%
1	*0*	2	*5.3*	0	*0.0*	2	*8.3*	23	*17.3*	8	*5.2*
2	*0*	6	*15.8*	2	*18.2*	4	*16.6*	31	*23.3*	33	*21.6*
3	*1*	12	*31.6*	6	*54.5*	11	*45.8*	58	*43.6*	78	*50.9*
4	*1*	15	*39.5*	3	*27.3*	6	*25.0*	14	*10.5*	24	*15.7*
5	*1*	3	*7.9*	0	*0.0*	1	*41.6*	7	*5.3*	10	*6.5*

Table 5A.3: Frequency distribution of inter-orbital breadth (IOB)

		INACIF (Guatemala) (n=36)		*Identified Guatemalan Migrants (n=10)*		*Identified Mexican Migrants (n=23)*		*Unidentified Migrants (n=133)*		*UADY (Mexico) (n=159)*	
Character State	*0/1*	n	%	n	%	n	%	n	%	n	%
1	*0*	18	*50.0*	3	*30.0*	11	*47.8*	22	*16.5*	83	*52.2*
2	*0*	18	*50.0*	6	*60.0*	10	*43.5*	62	*46.6*	71	*44.7*
3	*1*	0	*0.0*	1	*10.0*	2	*8.7*	49	*36.8*	5	*3.1*

Table 5A.4: Frequency distribution of malar tubercle (MT)

		INACIF (Guatemala) (n=38)		*Identified Guatemalan Migrants (n=11)*		*Identified Mexican Migrants (n=24)*		*Unidentified Migrants (n=133)*		*UADY (Mexico) (n=159)*	
Character State	*0/1*	n	%	n	%	n	%	n	%	n	%
0	*0*	8	*21.1*	0	*0.0*	4	*16.7*	22	*16.5*	53	*34.5*
1	*1*	23	*60.5*	8	*72.7*	10	*41.7*	67	*50.4*	101	*63.5*
2	*1*	7	*18.4*	3	*27.3*	8	*33.3*	41	*30.8*	5	*3.1*
3	*1*	0	*0.0*	0	*0.0*	2	*8.3*	3	*2.3*	0	*0.0*

Table 5A.5: Frequency distribution of nasal aperture shape (NAS)

		INACIF (Guatemala) (n=35)		*Identified Guatemalan Migrants (n=9)*		*Identified Mexican Migrants (n=10)*		*Unidentified Migrants (n=132)*		*UADY (Mexico) (n=150)*	
Character State	*0/1*	n	%	n	%	n	%	n	%	n	%
1	*0*	30	*81.0*	7	*77.8*	6	*60.0*	96	*72.7*	129	*86.0*
2	*1*	0	*0.0*	0	*0.0*	0	*0.0*	0	*0.0*	12	*8.0*
3	*1*	5	*13.5*	2	*22.2*	4	*40.0*	36	*27.3*	9	*6.0*

Table 5A.6: Frequency distribution of nasal aperture width (NAW)

		INACIF (Guatemala) (n=35)		*Identified Guatemalan Migrants (n=11)*		*Identified Mexican Migrants (n=24)*		*Unidentified Migrants (n=133)*		*UADY (Mexico) (n=155)*	
Character State	*0/1*	n	%	n	%	n	%	n	%	n	%
1	*0*	10	*28.6*	2	*18.2*	5	*20.8*	24	*18.1*	49	*31.6*
2	*0*	25	*71.4*	7	*63.6*	15	*62.5*	93	*69.9*	105	*67.7*
3	*1*	0	*0.0*	2	*18.2*	4	*16.7*	16	*12.0*	1	*0.6*

Table 5A.7: Frequency distribution of nasal bone contour (NBC)

		INACIF (Guatemala) (n=26)		*Identified Guatemalan Migrants (n=11)*		*Identified Mexican Migrants (n=22)*		*Unidentified Migrants (n=124)*		*UADY (Mexico) (n=141)*	
Character State	*0/1*	*n*	*%*	*n*	*%*	*n*	*%*	*n*	*%*	*n*	*%*
0	*0*	0	*0.0*	4	*36.4*	0	*0.0*	16	*12.9*	5	*3.5*
1	*0*	25	*96.2*	3	*27.3*	9	*8*	46	*37.1*	99	*70.2*
2	*1*	0	*0.0*	0	*0.0*	1	*4.5*	7	*5.6*	0	*0.0*
3	*1*	1	*3.8*	4	*36.4*	9	*40.9*	46	*37.1*	25	*17.7*
4	*1*	0	*0.0*	0	*0.0*	3	*13.6*	9	*7.3*	12	*8.5*

Table 5A.8: Frequency distribution of nasal bone shape (NBS)

		INACIF (Guatemala) (n=31)		*Identified Guatemalan Migrants (n=10)*		*Identified Mexican Migrants (n=16)*		*Unidentified Migrants (n=123)*		*UADY (Mexico) (n=140)*	
Character State	*0/1*	*n*	*%*	*n*	*%*	*n*	*%*	*n*	*%*	*n*	*%*
1	*0*	1	*3.2*	1	*10.0*	2	*12.5*	20	*16.3*	1	*0.7*
2	*0*	29	*93.6*	7	*70.0*	12	*75.0*	96	*78.0*	138	*98.6*
3	*0*	1	*3.2*	1	*10.0*	2	*12.5*	5	*4.1*	1	*0.7*
4	*1*	0	*0.0*	1	*10.0*	0	*0.0*	2	*1.6*	0	*0.0*

Table 5A.9: Frequency distribution of nasal overgrowth (NO)

		INACIF (Guatemala) (n=18)		*Identified Guatemalan Migrants (n=10)*		*Identified Mexican Migrants (n=19)*		*Unidentified Migrants (n=113)*		*UADY (Mexico) (n=76)*	
Character State	*0/1*	*n*	*%*	*n*	*%*	*n*	*%*	*n*	*%*	*n*	*%*
0	*0*	11	*61.1*	8	*80.0*	6	*31.6*	75	*66.4*	41	*53.9*
1	*1*	7	*38.9*	2	*20.0*	13	*68.4*	38	*33.6*	35	*46.1*

Table 5A.10: Frequency distribution of orbit shape (OBS)

		INACIF (Guatemala) (n=39)		*Identified Guatemalan Migrants (n=10)*		*Identified Mexican Migrants (n=17)*		*Unidentified Migrants (n=133)*		*UADY (Mexico) (n=161)*	
Character State	*0/1*	*n*	*%*	*n*	*%*	*n*	*%*	*n*	*%*	*n*	*%*
1	*0*	6	*15.4*	5	*50.0*	8	*47.1*	96	*72.2*	47	*29.2*
2	*0*	31	*79.5*	3	*30.0*	8	*47.1*	28	*21.1*	110	*68.3*
3	*1*	2	*5.1*	2	*20.0*	1	*6.8*	9	*6.7*	4	*2.5*

Table 5A.11: Frequency distribution of post bregmatic depression (PBD)

		INACIF (Guatemala) (n=37)		*Identified Guatemalan Migrants (n=10)*		*Identified Mexican Migrants (n=18)*		*Unidentified Migrants (n=132)*		*UADY (Mexico) (n=160)*	
Character State	*0/1*	*n*	*%*	*n*	*%*	*n*	*%*	*n*	*%*	*n*	*%*
0	*0*	34	*91.9*	6	*60.0*	14	*77.7*	89	*67.4*	130	*81.3*
1	*1*	3	*8.1*	4	*40.0*	4	*22.3*	43	*32.6*	30	*18.7*

Table 5A.12: Frequency distribution of posterior zygomatic tubercle (PZT)

		INACIF (Guatemala) (n=39)		*Identified Guatemalan Migrants (n=10)*		*Identified Mexican Migrants (n=18)*		*Unidentified Migrants (n=133)*		*UADY (Mexico) (n=162)*	
Character State	*0/1*	*n*	*%*	*n*	*%*	*n*	*%*	*n*	*%*	*n*	*%*
0	*0*	0	*0.0*	0	*0.0*	0	*0.0*	1	*0.8*	3	*1.9*
1	*1*	25	*64.1*	6	*60.0*	5	*27.8*	45	*33.8*	109	*67.3*
2	*1*	13	*33.3*	1	*10.0*	7	*38.9*	65	*48.9*	42	*25.9*
3	*1*	1	*2.6*	3	*30.0*	6	*33.3*	22	*16.5*	8	*4.9*

Table 5A.13: Frequency distribution of superior nasal suture (SNS)

		INACIF (Guatemala) (n=38)		*Identified Guatemalan Migrants (n=10)*		*Identified Mexican Migrants (n=18)*		*Unidentified Migrants (n=133)*		*UADY (Mexico) (n=156)*	
Character State	*0/1*	*n*	*%*	*n*	*%*	*n*	*%*	*n*	*%*	*n*	*%*
0	*0*	9	*23.6*	1	*10.0*	0	*0.0*	10	*7.5*	11	*7.1*
1	*1*	16	*42.1*	2	*20.0*	10	*55.6*	51	*38.3*	36	*23.1*
2	*1*	13	*34.2*	7	*70.0*	8	*44.4*	72	*54.1*	109	*69.9*

Table 5A.14: Frequency distribution of transverse palatine suture (TPS)

		INACIF (Guatemala) (n=36)		*Identified Guatemalan Migrants (n=9)*		*Identified Mexican Migrants (n=15)*		*Unidentified Migrants (n=129)*		*UADY (Mexico) (n=137)*	
Character State	*0/1*	*n*	*%*	*n*	*%*	*n*	*%*	*n*	*%*	*n*	*%*
1	*0*	25	*69.4*	0	*0.0*	3	*20.0*	44	*34.1*	90	*65.7*
2	*1*	11	*30.6*	8	*88.9*	10	*66.7*	66	*51.2*	41	*29.9*
3	*1*	0	*0.0*	1	*11.1*	2	*13.3*	12	*9.3*	6	*4.4*
4	*1*	0	*0.0*	0	*0.0*	0	*0.0*	7	*5.4*	0	*0.0*

Table 5A.15: Frequency distribution of zygomaticomaxillary suture course (ZS)

		INACIF (Guatemala) (n=34)		*Identified Guatemalan Migrants (n=10)*		*Identified Mexican Migrants (n=17)*		*Unidentified Migrants (n=129)*		*UADY (Mexico) (n=148)*	
Character State	*0/1*	*n*	*%*	*n*	*%*	*n*	*%*	*n*	*%*	*n*	*%*
0	*0*	28	*82.4*	8	*80.0*	7	*41.2*	88	*68.2*	67	*45.3*
1	*1*	6	*17.6*	2	*20.0*	8	*47.1*	26	*20.2*	64	*43.2*
2	*1*	0	*0.0*	0	*0.0*	2	*11.7*	15	*11.6*	17	*11.5*

Table 5A.16: Frequency distribution of palate shape (PS)

		INACIF (Guatemala) (n=34)		*Identified Guatemalan Migrants (n=10)*		*Identified Mexican Migrants (n=17)*		*Unidentified Migrants (n=129)*		*UADY (Mexico) (n=148)*	
Character State	*0/1*	*n*	*%*	*n*	*%*	*n*	*%*	*n*	*%*	*n*	*%*
1	*0*	28	*82.4*	8	*80.0*	7	*41.2*	88	*68.2*	67	*45.3*
2	*0*	6	*17.6*	2	*20.0*	8	*47.1*	26	*20.2*	64	*43.2*
3	*0*	0	*0.0*	0	*0.0*	2	*11.7*	15	*11.6*	17	*11.5*
4	*1*	0	*0.0*	0	*0.0*	0	*0.0*	0	*0.0*	0	*0.0*

Summary craniometric data are listed in tables 5A.17-5A.21.

Table 5A.17: Descriptive statistics for craniometric data (INACIF)

	Females						Males					
ILD	*n*	*Mean (mm)*	*SD*	*Min (mm)*	*Max (mm)*	*SE*	*n*	*Mean (mm)*	*SD*	*Min (mm)*	*Max (mm)*	*SE*
GOL	3	163	2.0	161	165	1.2	8	175	5.8	165	184	2.0
BNL	3	89	3.2	87	93	1.9	8	98	4.7	88	104	1.7
BBH	3	119	4.7	114	123	2.7	8	135	5.9	127	144	2.1
XCB	3	135	1.5	134	137	0.9	9	136	5.0	128	144	1.7
XFB	3	106	1.2	105	107	0.7	9	115	4.2	110	121	1.4
WFB	3	89	3.2	87	93	1.9	9	93	3.8	85	97	1.3
ZYB	3	122	2.5	119	124	1.5	7	130	5.1	125	139	1.9
AUB	3	117	1.0	116	118	0.6	8	122	4.8	115	130	1.7
ASB	3	104	3.1	101	107	1.8	8	110	7.1	100	121	2.5
NLH	3	49	0.6	48	49	0.3	9	53	3.5	47	59	1.2
NLB	3	23	2.3	22	26	1.3	9	25	2.6	20	29	0.9
MDH	2	21	5.7	17	25	4.0	8	27	3.7	19	31	1.3
OBH	3	36	0.6	36	37	0.3	9	36	2.2	32	40	0.8
OBB	3	37	1.5	35	38	0.9	9	39	1.6	37	42	0.5
DKB	3	20	2.1	18	22	1.2	6	21	1.9	18	24	0.8
EKB	2	87	0.0	87	87	0.0	8	96	2.5	92	99	0.9
FRC	3	99	2.7	97	102	1.5	9	110	4.8	103	116	1.6
PAC	3	102	7.0	94	107	4.0	9	111	7.3	103	126	2.4
OCC	3	91	3.2	89	95	1.9	8	98	4.2	93	105	1.5
FOL	3	33	1.0	32	34	0.6	8	36	2.6	32	39	0.9
FOB	2	26	1.4	25	27	1.0	8	30	1.8	28	32	0.6
UFBR	3	95	4.0	93	100	2.3	8	103	3.6	98	108	1.3

Table 5A.18: Descriptive statistics for craniometric data (Identified Guatemalan Migrants)

	Females						Males					
ILD	*n*	*Mean (mm)*	*SD*	*Min (mm)*	*Max (mm)*	*SE*	*n*	*Mean (mm)*	*SD*	*Min (mm)*	*Max (mm)*	*SE*
GOL	4	173	1.5	172	175	0.8	6	176	8.2	167	185	3.3
BNL	4	91	1.4	89	92	0.7	6	96	3.8	88	98	1.5
BBH	4	127	0.6	127	128	0.3	6	133	4.8	125	137	1.9
XCB	4	130	4.1	125	135	2.1	6	140	5.8	134	150	2.4
XFB	4	106	4.3	100	109	2.1	6	115	5.0	110	122	2.0
WFB	4	88	1.9	87	91	0.9	7	92	4.9	87	99	1.9
ZYB	4	120	2.5	119	124	1.3	6	133	6.2	126	144	2.5
AUB	4	115	3.6	113	121	1.8	6	126	4.7	123	135	1.9
ASB	4	101	2.5	98	104	1.3	6	111	4.8	104	116	1.9
NLH	4	46	2.6	44	49	1.3	7	53	2.3	50	56	0.9
NLB	4	23	0.9	23	25	0.5	7	26	2.3	22	29	0.9
MDH	4	26	3.7	21	30	1.9	6	31	3.5	27	37	1.4
OBH	4	35	1.8	33	37	0.9	7	37	1.9	34	39	0.7
OBB	4	38	1.0	38	40	0.5	6	41	2.1	38	43	0.9
DKB	4	19	0.8	19	20	0.4	6	22	2.6	19	26	1.1
EKB	4	92	0.8	91	93	0.4	7	97	4.9	91	103	1.8
FRC	4	109	1.9	107	111	0.9	6	111	3.5	105	115	1.5
PAC	4	109	6.5	103	118	3.3	6	112	7.9	102	120	3.3
OCC	4	96	7.5	86	104	3.8	6	97	5.4	91	104	2.2

Table 5A.19: Descriptive statistics for craniometric data (UADY)

	Females						Males					
ILD	*n*	*Mean (mm)*	*SD*	*Min (mm)*	*Max (mm)*	*SE*	*n*	*Mean (mm)*	*SD*	*Min (mm)*	*Max (mm)*	*SE*
GOL	53	166	7.7	153	185	1.0	110	175	7.3	150	197	0.7
BNL	54	90	3.9	84	101	0.5	113	97	5.0	85	114	0.5
BBH	53	120	6.8	109	141	0.9	110	125	7.2	108	145	0.7
XCB	50	137	5.4	125	151	0.8	106	144	5.9	130	160	0.6
XFB	41	113	4.6	104	125	0.7	95	117	4.9	104	131	0.5
WFB	54	89	4.3	81	101	0.6	112	93	4.9	82	104	0.5
ZYB	48	124	4.0	116	137	0.6	96	133	4.9	117	143	0.5
AUB	54	120	4.5	109	132	0.6	114	127	5.0	111	139	0.5
ASB	46	108	5.0	100	120	0.7	102	112	6.6	86	125	0.7
NLH	54	48	3.1	42	55	0.4	110	52	3.3	44	62	0.3
NLB	48	24	1.8	22	30	0.3	100	25	2.1	20	30	0.2
MDH	54	24	2.7	19	29	0.4	113	28	3.3	20	38	0.3
OBH	54	34	1.8	31	38	0.2	114	35	2.0	30	40	0.2
OBB	54	38	1.7	35	42	0.2	113	40	1.8	35	44	0.2
DKB	52	20	2.0	16	26	0.3	110	22	2.1	17	27	0.2
EKB	52	93	2.9	87	101	0.4	109	97	3.6	88	105	0.3
FRC	53	103	6.1	90	118	0.8	111	105	6.2	89	126	0.6
PAC	52	105	7.6	91	125	1.1	108	109	7.2	92	126	0.7
OCC	53	89	5.5	79	102	0.8	110	93	5.6	80	106	0.5
FOL	54	33	1.8	29	38	0.2	113	35	2.5	29	42	0.2
FOB	54	28	1.8	25	33	0.3	108	30	2.1	25	36	0.2
UFBR	59	99	3.4	92	107	0.5	110	104	4.1	93	112	0.4

Table 5A.20: Descriptive statistics for craniometric data (Identified Mexican Migrants)

	Females						Males					
ILD	*n*	*Mean (mm)*	*SD*	*Min (mm)*	*Max (mm)*	*SE*	*n*	*Mean (mm)*	*SD*	*Min (mm)*	*Max (mm)*	*SE*
GOL	2	160	9.9	153	167	7.0	21	180	6.9	163	194	1.5
BNL	2	91	3.5	89	94	2.5	22	102	4.8	93	110	1.0
BBH	2	127	3.5	125	130	2.5	22	136	4.7	127	145	1.0
XCB	2	138	7.8	133	144	5.5	22	139	6.6	128	156	1.4
XFB	2	117	4.2	114	120	3.0	22	116	5.6	108	126	1.2
WFB	2	94	3.5	92	97	2.5	22	93	4.9	87	105	1.0
ZYB	2	129	0.7	129	130	0.5	22	132	4.2	125	144	0.9
AUB	2	124	2.1	123	126	1.5	22	125	3.9	116	134	0.8
ASB	2	108	10.6	101	116	7.5	22	110	6.0	99	126	1.3
NLH	2	48	2.1	47	50	1.5	22	52	2.5	47	57	0.5
NLB	2	26	1.4	25	27	1.0	22	25	2.8	21	31	0.4
MDH	2	27	0.7	27	28	0.5	22	30	2.7	25	37	0.6
OBH	2	34	2.8	32	36	2.0	22	35	2.1	32	39	0.4
OBB	2	38	0.7	38	39	0.5	22	40	1.8	38	45	0.4
DKB	2	21	0.7	21	22	0.5	22	20	2.4	17	26	0.5
EKB	2	95	2.1	94	97	1.5	22	97	4.2	91	109	0.9
FRC	2	99	4.2	96	102	3.0	22	112	5.4	101	122	1.1
PAC	2	101	0.7	101	102	0.5	21	111	6.3	102	126	1.4
OCC	2	92	7.1	87	97	5.0	21	98	5.2	92	110	1.1

Table 5A.21: Descriptive statistics for craniometric data (Unidentified Migrants)												
	Females						Males					
ILD	*n*	*Mean (mm)*	*SD*	*Min (mm)*	*Max (mm)*	*SE*	*n*	*Mean (mm)*	*SD*	*Min (mm)*	*Max (mm)*	*SE*
GOL	47	167	5.9	154	179	0.8	82	178	6.6	163	198	0.7
BNL	48	93	4.7	82	102	0.7	84	100	4.5	89	110	0.5
BBH	48	128	5.7	118	140	0.8	82	136	4.7	124	147	0.5
XCB	48	136	5.1	127	149	0.7	82	139	6.2	127	159	0.7
XFB	47	112	5.0	103	125	0.7	81	116	5.6	104	128	0.6
WFB	48	91	4.0	80	99	0.6	85	94	4.6	85	106	0.5
ZYB	47	124	4.6	111	133	0.7	82	130	4.7	119	141	0.5
AUB	48	119	4.7	108	129	0.7	84	123	4.9	110	133	0.5
ASB	45	107	5.5	97	120	0.8	79	111	5.7	98	129	0.6
NLH	48	48	2.3	44	53	0.3	85	52	2.7	46	58	0.3
NLB	48	24	1.8	21	29	0.3	84	24	2.1	20	30	0.2
MDH	48	26	3.5	20	34	0.5	84	29	2.6	25	38	0.3
OBH	48	35	1.8	29	38	0.3	85	35	2.0	31	44	0.2
OBB	48	38	1.7	35	42	0.2	85	40	1.8	36	45	0.2
DKB	47	20	2.3	16	25	0.3	85	20	2.0	16	25	0.2
EKB	48	94	3.7	85	103	0.5	85	97	3.8	88	107	0.4
FRC	48	106	3.4	98	113	0.5	82	111	4.2	99	122	0.5
PAC	47	106	7.2	89	119	1.0	82	111	7.4	97	135	0.8
OCC	47	96	6.1	84	109	0.9	82	99	6.4	87	118	0.7

CHAPTER 6: DISCUSSION

This dissertation aims to distinguish craniofacial variability among Latin American populations using craniometric and cranial MMS data. Additionally, differences in craniofacial variability are used to generate and test population affinity estimation models. Included in this chapter are a discussion of the results related to the research questions, , limitations in the study, and the broader implications of this research.

Research Question One

Research question one asked "*Are significant differences in craniofacial variability across sex and population labels in Latin American populations?*".

Relationships Among Sex and Population Affinity

The only sample that exhibits multiple positive and negative correlations among cranial MMS traits is the INACIF sample. This could be due an effect of the small size, and the fact that individuals come from all over the country, expressing a wide range of human variation. Guatemala is a very diverse country with several ethnic and cultural groups, each with different cultural and historical variables impacting human variation. For example, there are Maya descended groups throughout the country in the highlands and lowlands, European-descended individuals in larger municipalities, African-descended individuals along the east coast, and individuals with combinations of these population affinities. Therefore, it is likely that human variation is not homogenous among this sample.

When testing for differences between sex with MANOVA and ANOVA, significant differences are identified using ILDs between males and females, and females and unknowns. Since the INACIF sample is the only dataset with individuals of undetermined sex, this indicates that the unknown sex individuals are likely males, which tend to be overrepresented in forensic samples (Komar & Grivas 2008). The ANOVA test identifies specific ILDs where differences occurred for sex. Results indicate that several variables across the facial skeleton and cranial vault change according to male or female. This is also true for cranial MMS traits, as the Kruskal-Wallis test identifies several traits that are significantly different among the sexes. These results suggest sexual dimorphism is present within the samples and can be used in research focused on sex estimation.

The MANOVA and Kruskal-Wallis test results report that most ILDs and cranial MMS traits are different among the Latin American samples tested using population affinity as a variable. This indicates promise in using craniometric and cranial MMS data to differentiate among groups on a more refined level. The MANOVA indicates significant differences between the Identified Mexican Migrant sample and the INACIF sample, the Identified Guatemalan Migrant sample and the Identified Mexican Migrant sample, and the Identified Mexican Migrant sample and the UADY sample using craniometric measurements. This indicates that population structure of these groups is different, despite being grouped under the term 'Hispanic' or even Mexican or Guatemalan.

Factor Analysis for Mixed Data

The variables are examined using FAMD and the population affinity label. Most of the variation among the samples is captured by craniometric measurements and few cranial MMS variables. This is true when comparing only identified sample data, and identified and unidentified sample data. However, the cranial MMS variables identified as important for group separation when the Unidentified Migrant sample is included. These variables include ANS, IOB, PZT, TPS, MT, and NFS, the majority of which are related to neutral genomic variation (Reyes Centeno & Hefner 2021). This indicates that more cranial MMS variables are useful for group separation of the Unidentified Migrant sample, which could point to reference populations not included in this research as present in the Unidentified Migrant sample. The FAMD graphics (Figures 5.22 and 5.26) illustrate the overlap of the INACIF, Identified Guatemalan Migrant, and Identified Mexican Migrant samples, while the UADY sample is separated along the *x*-axis in both graphics. When the Unidentified Migrant sample is added to the analysis, it overlaps with the INACIF, Identified Guatemalan and Mexican Migrants, and Unidentified Migrant samples. Again, this illustrates that the Unidentified Migrant sample likely contains individuals from Guatemala and migrant sending regions of Mexico. This also suggest that the UADY sample is different from the other samples used in this research, and may be more beneficial to use alone in analyses rather than combined with other Latin American samples.

Research Question Two

Research question two asked "*What is the relationship between Latin American groups using craniometric and cranial MMS data?*"

Biological Distance and Group Similarity

The general expectations for biological distance analyses are that the Guatemalan samples to cluster more closely together based on similar population structure and that the Guatemalan and UADY samples would cluster more based on a similar cultural and biological history. The Guatemalan and UADY samples likely include individuals of Maya descent and have a similar history of invasion and colonization by the Spanish.

The Guatemalan samples, the Identified Guatemalan Migrants and the INACIF sample, graphically display closer together based on their similarities to each other more than the other samples used in this research. As expected, this is likely related to a similar biological and cultural evolutionary history. However, the UADY sample is spatially isolated from the Guatemalan samples in distance scatterplots with both metric and cranial MMS data. This could indicate the impact on population structure by specific historical or cultural events that occurred in more recent history, like the Caste War, and/or it could indicate a high level of variability within the Guatemalan INACIF and Identified Guatemalan Migrant samples. Either way, the data suggest that the UADY sample not be combined with Guatemalan samples when performing a refined analysis. When comparing the two Mexican-derived samples, UADY and Identified Mexican Migrants, the samples are quite spatially distant from each other using both ILDs and cranial MMS data. This could be due to the likelihood that most migrants are coming from regions other than the Yucatán, which is where individuals in the UADY sample derive. Data from the Migration Policy Institute (2022), suggest only 0.3% of emigrants came from the Yucatán in 2015. The results suggest that the UADY and Identified Mexican Migrant samples should not be pooled in a refined analysis.

When the Unidentified Migrant sample is added to the distance analysis, it is more similar to the INACIF sample with metric data and to the Identified Guatemalan and Identified Mexican Migrant samples with the cranial MMS data. In fact, in the distance scatterplot using metric data, the Unidentified Migrant sample is in between the Identified Mexican Migrant and INACIF samples, but slightly closer to the INACIF sample. This indicates the Unidentified Migrant sample is likely composed of individuals from Guatemala and Mexico or other samples not included in this dissertation. Future research to parse out relationships could include highland Guatemala sample data collected from the FAFG to test if there are significant differences between highland Maya groups compared to the forensic and migrant samples from Guatemala, as the relative separation and isolation of highland groups over time could have influenced craniofacial variation.

The Procrustes plot of both distance matrices indicate relative similarity among data types for each group. In this plot (Figure 5.25) it is easier to visualize the relationships of each sample to each other using both data types. Craniometric data for the Identified Mexican Migrants do plot relatively closer to the UADY sample, while the cranial MMS data are more distant. This could result from the patterned missing data described earlier for cranial MMS variables. Importantly, the Procrustes plot demonstrates that craniometric and cranial MMS data are expressing relatively similar, they are saying the same thing about cranial shape and form in relation to populational groups.

Research Question Three

Research question three asked "*Can craniometric and cranial MMS data be used to make predictions about population affinity?*" To answer this question, aNN modeling was employed, which is appropriate for the data types used in the study.

Interpretation of Classification Results

Expectations for classification modeling include correct classification of the test samples within their respective samples based on studies supporting regional variation in Mexico and Guatemala (Helgeson 2019; Spradley 2014a; 2021). Additionally, unidentified migrants are expected to classify within the Identified Guatemalan Migrant, Identified Mexican Migrant, UADY, and/or INACIF samples. I expected the combined craniometric and cranial MMS trait models to perform the best (have the highest CCR%) based on results from previous studies (Spiros & Hefner 2019; Maier 2019). There is the expectation that cranial MMS traits will perform well in classification models, as they are designed to be used on complete and incomplete or fragmented crania.

The overall classification rate are as follows: the combined model is 70%, the craniometric model is 66%, and the cranial MMS model is 54%. These rates are comparable to other research using classification modeling to understand group membership in Latin American samples using more refined levels of analysis (Hefner et al. 2015; Kamnikar et al. 2021). Despite these classification rates, not all samples perform well within the models. The UADY sample classifies between 53%-71% across all models with most misclassifications occurring in the American White, INACIF, and Thai samples; thus, performing well as a stand-alone dataset. However, samples from Mexico and Guatemala produce classification rates ranging from 0% to 36%.

Examining sample composition and misclassification patterns of data collected from Guatemala and Mexico might help clarify the reasons behind lower classification rates. In this research, the INACIF and Identified Guatemalan Migrant samples produce classification rates using the combined model between 33%-36%, which are higher than random allocation (14%). The Identified Mexican Migrant sample produces a 0% classification rates with most misclassifications occurring within the American Black, American White, Thai, and UADY samples. Similar rates were generated when refining a Hispanic sample into Mexican and Guatemalan subsamples for comparison to Colombian data (Kamnikar et al. 2021). In that study, the authors attribute low misclassification to small sample sizes coming from a broad range of populations within geopolitical states, which could be true for this research as well. In the craniometric only model, the Identified Guatemalan and Mexican Migrant sample classify at 0%. Misclassifications for the both migrant samples occur in the American Black, American White, INACIF, Thai, and UADY samples. Non-migrant samples did misclassify within the Identified Guatemalan Migrant sample, which include Thai (n=3) and UADY (n=1). The INACIF sample classify at 25%, with misclassifications in the American Black, Identified Mexican Migrants, Thai and UADY samples. This pattern supports the idea that the INACIF sample likely comprises individuals from several ethnic groups within Guatemala, reinforcing the idea that an understanding of specific origination regions within a diverse country like Guatemala is important for understanding patterns of variation. Misclassifications for the UADY sample occur in all groups except the Identified Mexican Migrant sample, indicating dissimilarity between the two Mexican samples. In a separate study, a SW Hispanic sample, comprised mostly of Mexican individuals from migration contexts, misclassified as American White (Hefner et al. 2015). Understanding where Mexican migrants originate in Mexico, and that the Yucatán is not a large migrant sending department help clarify why certain misclassifications occur.

Several studies examine genetic variation within migrants and Latin American populations (Algee-Hewitt et al. 2018; Hughes et al. 2019; New et al. 2021). They demonstrate that differences in genetic structure have an impact on population structure and phenotypic variation. For example, misclassifications of the UADY and Identified Mexican Migrant samples within other, reference samples offer additional support of a genetic gradient across Mexico as described by Rubi-Castellanos and colleagues (2009). Examining the genetic make-up of Mexican Mestizo populations, they found a directional, north-south gradient in which European ancestry varied inversely with Native American ancestry across three regions (north/west, central, and southeast). Hughes and colleagues (2013) mirrored this research with craniometric data and found cranial variation coincided with the genetic gradient. This could explain why individuals in the UADY sample did not misclassify as the Identified Mexican Migrant sample, and why Mexican migrants misclassify as American White. As a higher proportion of Amerindian affinity, both genetically and craniometrically, is higher in the Yucatán, it is imperative to parse out differences and similarities to other Amerind-derived groups present in Guatemala, El Salvador, and Honduras.

General conclusions from the classification model results suggest that as individual samples, the Identified Migrant and INACIF samples did not perform well. However, the UADY sample classified with high accuracy rates across all models. Previous research recognizes differences among Guatemalan and Mexican samples (Hefner et al. 2015; Helgeson 2019; Spradley 2021); however, classification accuracies are generally lower when compared to other, comparative samples (Kamnikar et al. 2021). This evidence, along with the biological distance results from this study demonstrate and support differences among populations in the same region. The issue lies in using these groupings and labels within classification modeling. Because the samples that are available for identified migrants are small, they are often grouped by geopolitical unit, which is likely not appropriate for Mexican and Guatemalan groups. Given that biological distance analyses demonstrate differential patterning in craniofacial form across Mexico and among Maya-derived and non-Maya groups, one would expect classification models to perform equally as well in exploiting these differences. This research shows that samples with an adequate size that contain individuals from the same or very similar social and geographical populations, like UADY, can be useful in classification modeling. However, the challenge in forensic work is to find adequate samples where all individuals have similar population structures, histories, and cultural factors to use in human variation studies. While these variables are considered in research, but the nature of sample procuration is often difficult. The INACIF and migrant samples are random as they represent forensic cases and migrant cases that are recovered and identified, which is not always the case for individuals in these samples. They are random samples of populations and contain individuals with differing population structure and histories. Because of this, the INACIF sample is more useful as an exploratory dataset as it currently stands rather than a baseline of variation for Guatemala.

For forensic work on population affinity of unknown individuals, these results support research on the refinement of the Hispanic category, but only if the data allow. This study demonstrates that the data from UADY can be used as a reference data sets individually for population affinity modeling. However, the small samples of migrant data and data from the INACIF are not yet sufficient to use as stand-along reference samples for classification on a more refined scale. Continued data collection and hypothesis testing is recommended.

Interpretation of Exploratory Analyses

Exploratory analysis focused on combining the INACIF and Identified migrant samples to assess classification accuracies on a pooled Latin American dataset and testing the Unidentified Migrant sample and incomplete datasets in the models.

Pooled Latin American Data

The classification rate for the combined model using a pooled dataset (INACIF + Identified Guatemalan Migrants + Identified Mexican Migrants) produced an overall classification accuracy of 74%. Interestingly the classification accuracy for the pooled Guatemalan and Mexican sample performs worse when pooled, with a CCR of 26%. This is slightly above random allocation (20%). Misclassifications of the pooled dataset occurred in all other reference groups. Additionally, the American Black sample shows a classification rate of 0%. These two pieces of information suggests the model may be overfitting the data or inappropriate for use with a pooled sample. The pooled sample could be too small with a high amount of diversity to allow for any meaningful patterns to emerge. The UADY sample performed well with a classification rate of ~90%. This supports the conclusion that the UADY sample can be used as reference data for comparative studies examining biological distance and population structure in Mexico and within Central America.

Unidentified Migrant Data

Individuals in the Unidentified Migrant sample come from PCOME and OpID at TXST. While we do not have region of origin data for these individuals, information from classification models could provide insight into population affinity. Biological distance analyses (MD and MMD) suggest that the Unidentified Migrant sample is partway between the Guatemalan samples and Identified Mexican Migrant sample, suggesting that unidentified individuals resemble a population structure similar to both groups. Most unidentified individuals classify as UADY (n = 55), American Black (n = 22), INACIF (n = 18), American White (n = 16) and Thai (n = 15). Few classify as Identified Guatemalan Migrants (n = 4) or Identified Mexican Migrants (n = 3). These results demonstrate a large amount of variability within the Unidentified Migrant sample. Testing each case on a discrete level would provide more insight into specific population affinity information for each individual.

Incomplete Data and Modeling

Several individuals from the Identified Mexican Migrant, INACIF, UADY, and Unidentified Migrant samples have only craniometric or cranial MMS data available for various reasons. This is often the case in forensic work as preservation, trauma, or a combination of both may affect the ability to collect ILDs or cranial MMS traits. To emulate these practical scenarios, these data are run through each model to assess classification. Interestingly, no individuals classify in the combined model when one data type is present, suggesting that the model utilized should depend on the data available in each case. This is important to know for active casework in order to select the most appropriate model for use depending on the data available.

UADY sample data is available for the craniometric model. The classification rate is 33% with misclassifications in the American Black, American White, and Thai group. UADY, Identified Mexican Migrant, and Unidentified migrant data is available for the cranial MMS model. The UADY sample classifies both individuals as UADY, the Identified Mexican Migrant also classifies as UADY. Specific region of origin, apart from Mexico as the country of origin, is not available to assess the accuracy of these results. The Unidentified Migrant Sample classifies within American White, INACIF, Identified Mexican Migrants, Thai, and UADY. These results make sense as Mexican migrants have often classified as American White (Hefner et al. 2015), and unidentified migrants can be from anywhere.

Limitations

Missing Data and Impact on Modeling

Most of the samples have low amounts of overall missing data. However, the INACIF and UADY samples exhibit patterned missingness caused by antemortem and perimortem trauma. Specifically, missing data is the highest for the cranial MMS trait, NO, in the INACIF (60%) and UADY (48.8%) samples. While NO is missing for females, the majority of missing NO values occurs for males in both samples. Missing values for this trait are largely due to nasal trauma, and may be a product of interpersonal violence and socioeconomic status; although any conclusions must be corroborated with historical documentation of behavior related to violence that could produce these specific patterns of trauma (de la Cova 2012). Importantly, traits using the nasal area are often exhibit evolutionary significance (Reyes-Centeno & Hefner 2021), useful for population affinity models (Hefner et al. 2014; 2015). Hefner and colleagues (2014) found NAS to be the most important variable for distinguishing population affinity in their RFM models of craniometric and cranial MMS data for an American Black, American White, and SW Hispanic samples. In another study, Hefner and colleagues (2015) used cranial MMS traits to differentiate between a Guatemalan and SW Hispanic sample. Here, NO was significantly different among the two samples, indicating its likely influence on support vector machine (SVM) modeling to estimate population affinity. They found a lower incidence of NO (score of 1) in the Guatemalan sample; however, as SVM is a black-box method, an understanding of individual trait contribution to the model is unknown (Hefner et al. 2015). This study is no exception as nasal-derived traits are important variables for the aNN models which use cranial MMS only data and the combination of metric + cranial MMS trait data. Therefore, missing data in the INACIF and UADY samples, especially resulting from damage to the nasal area, could affect the overall accuracy of the models given the importance of nasal-derived variables in population affinity research.

Sample Biases

As outlined in Materials and Methods, each sample used in the research from Latin America comes with a set of biases. These biases are built into each sample and stem from sample composition and regional/local social and power dynamics surrounding the inclusion and exclusion of individuals in each sample. Winburn and colleagues (2020) note a disconnect between the forensic population and osteological collections available for research in the U.S. They note that many osteological collections are biased toward White or European-descended individuals and are accompanied by documentation of demographic parameters, while most individuals in forensic casework have origins from non-European countries, do not contain complete documentation, and are likely in medical examiner office collections or donated by medical examiners rather than next-of-kin. As Gravlee (2009) pointed out, social inequality in marginalized individuals can present on the physical body, which Winburn and colleagues (2020) argue can limit our understanding of skeletal variation and hinder the ability of our methods to estimate aspects of the biological profile. Specifically, they argue a lack of non-White individuals in osteological research collections can exacerbate this issue. This argument is especially relevant at the U.S.-Mexico border, where migrants likely come from marginalized groups and display skeletal indicators of structural violence (Beatrice et al. 2021). Therefore, an understanding of the limitations and biases within the sample and how they relate to potential migrants is an important consideration. The UADY sample is biased toward low socioeconomic status, Maya-descended individuals who lived in or near Mérida (Chi Keb et al. 2013). These individuals are included in the sample as family members cannot afford or choose not to pay burial fees. It is difficult to know if individuals specifically from Mérida are migrating clandestinely to the U.S as migration patterns may change. Migration statistics indicate approximately 300 individuals from the Yucatán State accounting for 0.3% of migrants to the U.S. (Migration Policy Institute 2022). The INACIF sample, on the other hand, is a forensic sample from Guatemala. This sample contains any individual requiring anthropological analysis (i.e., biological profile, trauma, etc.) from the country. While Guatemalans comprise approximately 1,111,000 migrants or 29.4% of migrants to the U.S., the exact statistics for outmigration per department within the country are unclear. Many migrants come from the Huehuetenango department in the west of the country as evidenced by the remittance economy described in Chapter 2. Again, it is unclear of the INACIF forensic sample captures variation of migrants from Guatemala. But, because no modern, osteological collections exist in Guatemala, this is a first step in exploring variation and can inform as to what the next steps should include. Apart from social and cultural factors surrounding body donation and skeletal sampling, narcotrafficker and organized crime syndicates are active in many regions of Mexico and Guatemala (Martínez 2014). Therefore, the idea of visiting local skeletal assemblages may not be viable or safe for local and international researchers. These factors severely impact and shape the sample regarding who is included and who is excluded, which affects methodological development.

Impact of Small Sample Size

An additional limitation for this project was the initiation and continuation of the COVID-19 pandemic. Because of the pandemic, the author was not permitted to travel internationally as much as would have been possible without governmental and health organization limitations. As the UADY sample data was collected in early 2019 and by Dr. Spradley, years prior to the pandemic, this sample data contains a robust size of over 100 individuals for craniometric and cranial MMS data. However, the INACIF sample size is much smaller. During data collection trips to Guatemala, the author collected all available data, but was limited to active casework currently in the lab. The institution's policy of reburial of unknowns after a short period of time constrains the number of skeletonized individuals in the lab at any given time. The INACIF staff were very accommodating and supportive of the data collection effort and requested specialized training to continue data collection for their own research. During the pandemic, INACIF forensic anthropologists participated in virtual trainings on the use of craniometric and cranial MMS data for their casework. When travel restrictions lifted in the summer of 2021, the author returned to collect more data and provide hands-on training of data collection methods. In time, the INACIF will have a robust dataset that can be reanalyzed under these research questions or utilized in other types of investigations.

With these limitations in mind, the outlier test identified several outliers from the INACIF, Identified Guatemalan Migrant, and Identified Mexican Migrant samples. I argue that these individuals may not actually be outliers. They could be identified as outliers because these samples are small and individuals can come from any region within Guatemala and Mexico. The issue with this is that each region, and even local communities within regions, have different population histories and different cultural impacts acting on biology and human variation. The results of this study show that more data is needed to illuminate baselines for variation in unique regions and communities. Removal of outliers may be premature as our understanding of human variation within the INACIF and identified migrant data is extremely limited.

Broader Impact

While the reference data collected here aim to provide a more nuanced picture of variation in Latin America and to enhance current biological profile methodology already being applied to forensic contexts in the U.S., results indicate that more data and more analyses are needed. Modeling created and used on test samples in this research performed will on the UADY sample, but did not for the INACIF and identified migrant samples. The UADY sample can be used as a stand-along sample for more refined population affinity estimation analyses. However, more research and investigation into human variation in Guatemala and Mexico is needed in order to be used as stand-along datasets in forensic casework. This research provides a good foundation; however, as is the case in many skeletal studies, more data is preferred, as well as, reference data from other countries whose citizens are involved in international migration, like Honduras and El Salvador.

Despite the biases built into the samples, this study provides matched populational data for a cemetery sample from Merida, Mexico and a forensic sample from Guatemala. Matched data is especially important as it allows researchers to collectively utilize separate methodological approaches (Hefner et al. 2014b; Spiros & Hefner 2019). Additionally, data collected for this dissertation partially fills a significant gap in reference data for Latin American populations. Furthermore, the author is working with anthropologists at the INACIF who have incorporated craniometric and cranial MMS data collection into their casework protocol to create datasets for future investigations like this one and research relevant to the medico-legal system in Guatemala.

In time, there should be a large dataset to reassess the hypotheses and questions presented here with a larger sample set and assess the impact this may have on current results. Additionally, craniometric and cranial MMS reference data collected in this study can be used to investigate secular change and variation within modern and archaeological populations in the region.

Conclusions

This research demonstrates that Latin American samples from Mexico and Guatemala are different from each other. The Identified Guatemalan and Identified Mexican Migrant samples, INACIF, and UADY samples are different from each other in biological distance analyses Specifically, differences in the UADY and Guatemalan samples are present in several distance measures, even though individuals in both samples are descended from Maya. This shows that local population histories and historical events can impact population structure. Classification accuracies are not as clear cut as the biological distance results. The UADY sample classifies well; however, the Identified Migrant and INACIF samples do not. They likely require the addition of more samples and specific geographic information attached to each individual for further testing. Additionally, this research demonstrated that matched craniometric and cranial MMS data from individuals produce similar results in biological distance analyses. Collectively, these data types capture different aspects of cranial morphology used to understand group relatedness.

Implications for forensic science then challenge us to think about which categories we use in classification and population affinity language in our analyses. Forensic anthropologists have identified the issues with the term Hispanic (Kamnikar et al. 2021; Ross et al. 2014; Spradley 2016a; Spradley et al. 2008). However, with more refined analyses, it is likely we fall into the same problem using geopolitical terms that homogenize countries and populations, like Mexican and Guatemalan. Hefner and Spradley (2018) advocate for a broad to narrow analytical approach that depends on the nature of the analytical outcomes (i.e., migrant identification vs U.S. forensic casework). This research intends to push further into what we consider 'narrow' by refining not only Hispanic, but Guatemalan and Mexican sample data. Here, I demonstrate that the UADY sample can be used as a distinct group in refined classification models. However, as the samples currently stand, Identified Migrant data from Mexico and Guatemala are limited. For these datasets, where we know there are differences in evolutionary histories and population structure, we can add clarifiers to the samples like 'migrant' or 'Maya' (Spradley 2021). These labels should have the flexibility to be refined and modified to fit cultural ideas of taxonomies while still demonstrating biological variation and maintaining viability in forensic research (Edgar & Ousley 2022). The INACIF sample has more limitations due to sample composition described previously, but shows promise as the sample grows and can be reanalyzed.

REFERENCES

REFERENCES

Adams RN. Guatemalan ladinization and history. *The Americas* 1994;50:527-543.

Adhikari K, Fuentes-Guajardo M, Quinto-Sánchez M, Mendoza-Revilla J, Chacón-Duque JC, Acuña-Alonzo V, et al. A genome-wide association scan implicates DCHS2, RUNX2, GLI3, PAX1 and EDAR in human facial variation. *Nature communications* 2016;7(1): 1-11.

Agamben G. *Homo Sacer: Sovereign Power and Bare Life*. Stanford: Stanford University Press; 1998.

Agarwal SC. Bone morphologies and histories: Life course approaches in Bioarcheology. *Yearbook of Physical Anthropology* 2016;159:S130-S149.

Agarwal SC, Beauchesne P. It is not carved in bone. Development and plasticity of the aged skeleton. In: Agarwal SC, Glencross BA, eds. *Social Bioarcheology.* Blackwell Publishing LTD.;2011:312-332.

Alexander RT, Kepecs S. The Postclassic to Spanish-era transition in Mesoamerica: An introduction. In: Kepecs S, Alexander R, eds. *Postclassic to Spanish-Era Transition in Mesoamerica: Archaeological Perspectives*. Albuquerque: University of New Mexico Press; 2005:1-

Algee-Hewett BFB. Population inference from contemporary American craniometrics. *American Journal of Physical Anthropology* 2016;160:604-624.

Algee-Hewitt BFB, Hughes CE, Anderson BE. Temporal, geographic and identification trends in craniometric estimates of ancestry for persons of Latin American origin. *Forensic Anthropology* 2018;1(1): 4.

Anderson BE. Identifying the dead: Methods utilized by the Pima County (Arizona) office of the medical examiner for undocumented border crossers: 2001-2006. *Journal of Forensic Sciences* 2008;53(1):8–15.

Anderson BE, Parks BO. Symposium on border crossing deaths: Introduction. *Journal of Forensic Sciences* 2008;53(1):6-7.

Anderson BE, Spradley MK. The role of the anthropologist in the identification of migrant remains in the American Southwest. *Academic Forensic Pathology* 2016;6(3):432-438.

Atkinson ML, Tallman SD. Nonmetric cranial trait variation and ancestry estimation in Asian and Asian-derived groups. *Journal of Forensic Sciences* 2020; 65(3):692-706.

Bartelink EJ. Identifying difference: Forensic methods and the uneven playing field of repatriation. In: Latham KE, O'Daniel AJ, eds. *Sociopolitics of Migrant Death and Repatriation: Perspectives from Forensic Science*. Cham: Springer; 2018: 129-142.

Beatrice JS, Soler A. Skeletal indicators of stress: A component of the biocultural profile of undocumented migrants in southern Arizona. *Journal of Forensic Sciences* 2016;61(5).

Beatrice JS, Soler A, Reineke RC, Martínez DE. Skeletal evidence of structural violence among undocumented migrants from Mexico and Central America. *American Journal of Biological Anthropology* 2021;176(4):584-605.

Beck M. Variable importance in neural networks. https://beckmw.wordpress.com/2013/08/12/variable-importance-in-neural-networks/. Created on March 4, 2013. Accessed on February 2, 2022.

Beck J, Ostericher I, Sollish G, De Leon J. Animal scavenging and scattering and the implications for documenting the deaths of undocumented border crossers in the Sonoran Desert. *Journal of Forensic Sciences* 2015;60:S11-S20.

Bemiss J, Molomot L, Bricca J. *Missing in Brooks County*. [film]. ITVS: United States; 2020.

Betti L, Balloux F, Hanihara T, Manica A. The relative role of drift and selection in shaping the human skull. *American Journal of Physical Anthropology* 2010;141:76-82.

Bialik K., Border apprehensions increased in 2018 - especially for migrant families. Pew Research Center: Washington, D.C. https://www.pewresearch.org/fact-tank/2019/01/16/border-apprehensions-of-migrant-families-have-risen-substantially-so-far-in-2018/ Created January 16, 2019. Accessed April 18, 2019.

Birkby WH, Fenton TW, Anderson BE. Identifying Southwest Hispanics using nonmetric traits and the cultural profile. *Journal of Forensic Sciences* 2008;53(1):29-33.

Borger J. Fleeing a hell the US helped create: Why Central Americans journey north. The Guardian: Washington D.C.; 2018.

Boyd CC, Boyd DC. The theoretical and scientific foundations of forensic anthropology. In: Boyd CC, Boyd DC, eds. *Forensic Anthropology: Theoretical Framework and Scientific Basis*. 1st ed. Hoboken: John Wiley & Sons; 2018:1-20.

Bryc K, Velez C, Karafet T, Moreno-Estrada A, Reynolds A, Auton A, et al. Genome-wide patterns of population structure and admixture among Hispanic/Latino populations. *PNAS* 2010;107(Supplement 2): 8954-8961.

Budiman A. Thai in the U.S. Fact Sheet. Pew Research Center. 2021. Accessed at: https://www.pewresearch.org/social-trends/fact-sheet/asian-americans-thai-in-the-u-s/

Cal AE. *Anglo Maya contact in norther Belize: A Study of British Policy Toward the Maya During the Caste War of Yucatan*, 1847-1872 [Master's Thesis] University of Calgary: Calgary, AB; 1983.

Chi-Keb JR, Albertos-González VM, Ortega-Muñoz A, Tiesler VG. A new reference collection of documented human skeletons from Mérida, Yucatan, Mexico. *HOMO* 2013;64(5):366-376.

Chicco D. Ten quick tips for machine learning in computational biology. *BioData Mining* 2017;10:35.

Chicco D, Jurman G. The advantages of the Matthews Correlation Coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genomics* 2020;21:6.

Coe MD. *Breaking the Maya Code*, Second Edition. Thames & Hudson; 1999.

Colibrí Center for Human Rights. Human Rights Crisis; 2018. Retrieved from: http://www.colibricenter.org/wp-content/uploads/2018/07/Press-Kit-PDF-English.pdf

Colibrí Center for Human Rights (Colibrí). https://colibricenter.org/about/. Accessed April 25, 2021.

Corning PA. Beyond the modern synthesis: A framework for a more inclusive biological synthesis. *Progress in Biophysics and Molecular Biology* 2020;153:5-12.

Crenshaw K, Gotanda N, Peller G, Thomas K. *Critical Race Theory. The Key Writings that Formed the Movement*. New York: The New Press; 1995.

Crossland Z. Evidential regimes of forensic archaeology. *Annual Review of Anthropology* 2013;42:121-137.

Cucina A. Mayan lowlands: The analysis of dental morphological traits. In: Cucina A, ed. *Archaeology and Bioarchaeology of Population Movement among the Prehispanic Maya*. Cham: Springer; 2014a:71-83.

Cucina A. Introduction. In: Cucina A, ed. *Archaeology and Bioarchaeology of Population Movement among the Prehispanic Maya*. Cham: Springer; 2014b: v

Das V, Poole D. Anthropology in the margins of the state. *PoLAR: Political and Legal Anthropology Review.* 2004;30(1):140-144.

de la Cova, C. Cultural patterns of trauma among 19th-century-born males in cadaver collections. *American Anthropologist* 2012;112(4):589-606.

De León J. *The Land of Open Graves: Living and Dying on the Migrant Trail.* University of California Press; 2015.

Dewbury A. The American School and scientific racism in early American Anthropology. *Histories of Anthropology Annual* 2007;3121-147.

Dror I, Melinek J, Arden JL, Kukucka J, Hawkins S, Carter J, Atherton DS. Cognitive bias in forensic pathology decisions. *Journal of Forensic Sciences* 2021;00:1-7.

Dudzik B. Examining cranial morphology of Asian and Hispanic populations using geometric morphometrics for ancestry estimation. *Forensic Anthropology* 2019;2(4): 304-315.

Dudzik B, Kolatorowicz A. Craniometric data analysis and estimation of biodistance. In: Pilloud MA, Hefner JT, eds. *Biological Distance Analysis: Forensic and Bioarchaeological Perspectives.* San Diego: Elsevier; 2016:36-60.

Dudzik B, Jantz RL. Misclassifications of Hispanics using Fordisc 3.1: Comparing cranial morphology in Asian and Hispanic populations. *Journal of Forensic Sciences* 2016;61(5):1311-1318.

Dunn RR, Spiros MC, Kamnikar KR, Plemons AM, Hefner JT. Ancestry estimation in forensic anthropology: A review. *WIREs Forensic Science* 2020; e1369.

Dunn TJ. *The Militarization of the U.S.-Mexico border, 1978-1992: Low-Intensity Conflict Doctrine Comes Home.* Austin: CMAS Books, University of Texas; 1996.

Duret L. Neutral Theory: The null hypothesis of molecular evolution. *Nature Education* 2008;1(1):218.

Edgar H, Ousley SD. Testing the homogeneity of "White": Dental morphology in Americans and Australians of European descent. *Forensic Anthropology* 2021; 4(4):161-170.

Eschbach K, Hagan J, Rodriguez N, Hernandez-Leon R, Bailey S. Death at the border. *International Migration Review* 1999;33(2):430-454.

Eschbach K, Hagen J, Rodriguez N. Deaths during undocumented migration: Trends and policy implications in the new area of homeland security. In Defense of the Alien. Center for Migration Studies of New York, Inc. 2003; 26: 37-52.

Farmer PE. An anthropology of structural violence. *Current Anthropology* 2004;45(3):305-325.

Farmer PE, Nizeye B, Stulac S, Keshavjee S. Structural violence and clinical medicine. *PLOS Medicine* 2006;3(10): 1686-1691.

Fenton TW, Heard AN, Sauer NJ. Skull□photo superimposition and border deaths: identification through exclusion and the failure to exclude. *Journal of Forensic Sciences* 2008;53(1):34-40.

Fleischman JM, Crowder CM. *Standard Operating Procedure for MicroScribe 3-Dimensional Digitizer and Craniometric Data*. Harris County Institute of Forensic Sciences, Forensic Anthropology Division: Houston; 2019.

Fox J. 'Polycor' 2019. https://cran.r-project.org/web/packages/polycor/polycor.pdf

Gabbert W. Of friends and foes: The Caste War and ethnicity in Yucatan. *Journal of Latin American Anthropology* 2004;9(1):90-118.

Gabbert W. *Violence and The Caste War of Yucatán*. Cambridge University Press: Cambridge UK; 2019.

Galtung J. Violence, peace, and peace research. *Journal of Peace Research* 1969;6:167-191.

Garvin HM, Klales AR. Adult skeletal sex estimation and global standardization. In: Parra RC, Zapico SC, Ubelaker DH. *Forensic Science and Humanitarian Action: Interacting with the Dead and the Living.* John Wiley & Sons Ltd.; 2020:199-209.

Go MC, Hefner JT. Morphoscopic ancestry estimates in Filipino crania using multivariate probit regression models. *American Journal of Physical Anthropology* 2020; 172(3):386-401.

Goad G. Expanding humanitarian forensic action: An approach to U.S. cold cases. *Forensic Anthropology* 2020;3(1):50-58.

Gocha TP, Spradley MK, Strand R. Bodies in limbo: Issues in identification and repatriation of migrant remains in South Texas. In: Latham KE, O'Daniel AJ, eds. *Sociopolitics of Migrant Death and Repatriation: Perspectives from Forensic Science*. Cham: Springer. 2018:143-156.

Gonzalez GG. Mexican labor migration, 1876-1924. In: Overmyer- Velázquez M, ed. *Beyond la Frontera: The History of Mexico-U.S. Migration.* Oxford: Oxford University Press: 2011:28-50.

Goldsmith RR, Reineke R. Border Deaths and Federal Immigration Enforcement. *NACLA Report on the Americas* 2010;43(6):48-49.

Gonzalez-Barrera A, Krogstad JM, What we know about illegal immigration from Mexico, in Fact Tank: News in the Numbers. 2019, Pew Research Center: Washington, D.C.

Goodman AH, Leatherman TL. Traversing the chasm between biology and culture: an introduction. In: Goodman AH, Leatherman TL, eds. *Building a New Biocultural Synthesis: Political-Economic Perspectives on Human Biology*. Ann Arbor: University of Michigan Press; 1998:3-41.

Government Accountability Office (GAO). Border Patrol: Key Elements of New Strategic Plan Not Yet in Place to Inform Border Security Status and Resource Needs. Report to Congressional Requesters; 2012. www.gao.gov/assets/660/650730.pdf

Gramlich, J. and L. Noe-Bustamante, What's happening at the U.S.-Mexico border in 6 charts, in Fact Tank: News in the Numbers. 2019, Pew Research Center: Washington, D.C.

Grandin G, Levenson DT, Oglesby E. *The Guatemala Reader: History, Culture, Politics*. Durham: Duke University Press; 2011.

Gravlee CC. How race becomes biology: Embodiment of social inequality. *American Journal of Physical Anthropology* 2009;139(1):47-57.

Haglund WD, Sorg MH. Method and theory of forensic taphonomy research. In: Sorg MH, Haglund WD, eds. *Forensic Taphonomy: The Postmortem Fate of Human Remains*. Boca Raton: CRC Press; 1997:13-25.

Harvati K, Weaver TD. Human cranial anatomy and the differential preservation of population history and climate signatures. *The Anatomical Record* 2006;288:1225-1233.

Harris EF, Sjøvold T. Calculation of Smith's Mean Measure of Divergence for intergroup comparisons using nonmetric data. *Dental Anthropology Journal* 2004;17(3):83-93.

Haykin S. *Neural Networks and Learning Machines. Third Edition*. New York: Pearson; 2009.

Hefner JT. The macromorphoscopic databank. *American Journal of Physical Anthropology* 2018;166(4):994-1004.

Hefner JT, Byrnes JF. Ancestry estimation using craniometric and macromorphoscopic data: analytical feasibility and 21st century transnationalism. In: Garvin-Elling H, Langley NR, eds. *Case Studies in Forensic Anthropology: Bonified Skeletons*. Boca Raton: CRC Press; 2020.

Hefner JT, Linde KC. Atlas of Human Cranial Macromorphoscopic traits. New York, NY: Elsevier; 2018.

Hefner JT, Ousley SD. Statistical classification methods for estimating ancestry using morphoscopic traits. *Journal of Forensic Sciences* 2014a;59(4):883-890.

Hefner JT, Pilloud MA, Black CJ, Anderson BE. Morphoscopic trait expression in "Hispanic" populations. *Journal of Forensic Sciences* 2015;60(5):1135-1139.

Hefner JT, Pilloud MA, Buikstra JE, Vogelsberg CCM. A brief history of biological distance analysis. In: Pilloud MA, Hefner JT, eds. *Biological Distance Analysis: Forensic and Bioarchaeological Perspectives.* San Diego: Elsevier; 2016:3-22.

Hefner JT, Spradley MK. Ancestry/forensic applications. In: Trevathan W, ed. *The International Encyclopedia of Biological Anthropology.* Hoboken: John Wiley & Sons, Inc.; 2018.

Hefner JT, Spradley MK, Anderson BE. Ancestry assessment using random forest modeling. *Journal of Forensic Sciences.* 2014; 59(3):583-589.

Helgeson, KE. *Geographic Origin Estimation of Latin American Individuals Using Craniometric Data*. [Master thesis] San Marcos: Texas State University; 2019.

Hughes CE, Tise ML, Trammell LH, Anderson BE. Cranial morphological variation among contemporary Mexicans: Regional trends, ancestral affinities, and genetic comparisons. *American Journal of Physical Anthropology* 2013;151(4):506-517.

Hughes CE, Dudzik B, Algee-Hewitt BFB, Jones A, Anderson BE. Understanding (mis)classification trends of Latin Americans in Fordisc 3.1: Incorporating cranial morphology, microgeographic origin, and admixture proportions for interpretation. *Journal of Forensic Sciences* 2019;54(2):353-366.

Hurst CV, Soler A, Fenton TW. Personal identification in forensic anthropology. In: Siegel JA, Saukko PJ, eds. *Encyclopedia of Forensic Sciences, 2nd ed. Vol. 1*. Waltham: Academic Press; 2013:68-75.

Ibarra-Rivera L, Mirabal S, Regueiro MM, Herrera RJ. Delineating genetic relationships among the Maya. *American Journal of Physical Anthropology* 2008;135:329-347.

Isacson A, Meyer M, Davis A. Border Security and Migration: A report from Arizona. Washington Office on Latin America: Washington, D.C.; 2013.

Jantz RL, Moore-Jansen PH. Database for Forensic Anthropology in the United States, 1962-1991. University of Tennessee: Knoxville;1988.

Jantz RL, Ousley SD. FORDISC 3: Computerized forensic discriminant functions. Version 3.1. Knoxville, TN: The University of Tenneessee; 2005.

Jiménez D. *Contexto Histórico y Desarrollo de la Antropología Forense en Guatemala (1954-2011)*. [Masters Thesis] Guatemala City: Universidad de San Carlos de Guatemala; 2011.

Jonas S. Guatemalan migration in times of civil war and post-war challenges. *Migration Policy Institute* 2013.

Joseph GM. From Caste War to Class War: The historiography of modern Yucatán (c. 1750-1940). *Hispanic American Historical Review* 1985;65:111-134.

Kamnikar KR, Hefner JT, Monsalve T, Bernal Florez L. Craniometric variation in a regional sample from Antioquia, Medellin, Colombia: Implications for forensic work in the Americas. *Forensic Anthropology* 2021; 5(3):239-250.

Kamnikar KR, Plemons AM, Hefner JT. Intraobserver error in macromorphoscopic trait data. *Journal of Forensic Sciences* 2018;63(2):361–70.

Kassambara A. Practical Guide to Principal Component Methods in R. First Edition. STHDA; 2017.

Kenyhercz M, Passalacqua NV, Hefner JT. Missing data imputation using morphoscopic traits and their performance in the estimation of ancestry. *Forensic Anthropology* 2019; 2(3):1-11.

Kimmerle EH, Falsetti A, Ross AH. Immigrants, undocumented workers, runaways, transients and the homeless: Towards contextual identification among unidentified decedents. *Forensic Science Policy & Management: An International Journal* 2010;1(4):178-186.

Kimura M. The neutral theory of molecular evolution: A review of recent evidence. *The Japanese Journal of Genetics* 1991; 6(4):367-386.

Klales AR, Kenyhercz MW. Morphological assessment of ancestry using cranial macromorphoscopics. *Journal of Forensic Sciences* 2015;60(1):13–20.

Klaus HD. The bioarcheology of structural violence: A theoretical model and a case study. In: Martin DL, Harrod RP, Perez VR, eds. *The Bioarcheology of Violence*. Gainesville: The University of Florida Press; 2012:29-62.

Komar DA, Grivas C. Manufactured populations: What do contemporary reference skeletal collections represent? A comparative study using the Maxwell Museum documented collection. *American Journal of Physical Anthropology* 2008;137:224-233.

Kovic C. Naming state crimes, naming the dead: Immigration policy and 'the new disappeared; in the United States and Mexico. In: Lathan KE, O'Daniel AJ, eds. *Sociopolitics of Migrant Death and Repatriation. Perspectives from Forensic Science.* Cham: Springer. 2018;129-141.

Langley NR, Meadows Jantz L, Ousley SD, Jantz RL, Milner G. *Data Collection Procedures for Forensic Skeletal Material 2.0*. University of Tennessee: Knoxville; 2016.

Leatherman T, Goodman A. Building on the biocultural syntheses: 20 years and still expanding. *American Journal of Human Biology* 2020;32(4): e23360.

Lee SY, Poon WY, Bentler PM. A two-stage estimation of structural equation models with continuous and polytomous variables. *British Journal of Mathematical and Statistical Psychology* 1995;48(2):339-358.

Leutert S, Lee S, Rossi V. *Migrant Deaths in South Texas*. Strauss Center for International Security and Law: The University of Texas at Austin; 2020.

Little BB, Buschang PH, Peña Reyes ME, Tan SK, Malina RM. Craniofacial dimensions in children in rural Oaxaca, southern Mexico: secular change, 1968–2000. *American Journal of Physical Anthropology* 2006;131(1):127-136.

Liu W. *Neural Networks. Foundation Entries*. London: SAGE Publications; 2020.

Magaña R. Dead bodies: The deadly display of Mexican border politics. In: Mascia-Lees FE, ed. *A Companion to the Anthropology of the Body and Embodiment,* First Edition. Blackwell Publishing Ltd.;2011:157.

Maier C. Evaluating mixed-methods models for the estimation of ancestry from skeletal remains. *Forensic Anthropology* 2019;2(1):45-56.

Maier CA, George RL. Examining differences in presumed migrants from Texas and Arizona using cranial and dental data. *Forensic Anthropology* 2020;3(1): 17-28.

Mair P, Groenen PJF, de Leeuw J. More on multidimensional scaling in R: smacof version 2. *Journal of Statistical Software* 2021; early view.

Malina RM, Himes JH, Custom Stepic S, Gutierrez Lopez F, Buschang PH. Growth of rural and urban children in the Valley of Oaxaca, Mexico. *American Journal of Physical Anthropology* 1981;54(3):327-336.

Malina RM, Peña Reyes ME, Little BB. Secular change in the growth status of urban and rural schoolchildren aged 6-13 years in Oaxaca, southern Mexico. *Annals of Human Biology* 2008a;35(5): 475-489.

Malina RM, Peña Reyes ME, Little BB. Epidemiological transition in an isolated indigenous community in the Valley of Oaxaca, Mexico. *American Journal of Physical Anthropology* 2008b; 137: 69-81.

Martin S, Grube N. *Chronicle of the Maya Kings and Queens: Deciphering the Dynasties of the Ancient Maya*. New York: Thames and Hudson; 2000.

Martínez DE, Reineke RC, Rubio-Goldsmith R, Parks BO. Structural violence and migrant deaths in Southern Arizona: Data from the Pima County Office of the Medical Examiner, 1990-2013. *Journal on Migration and Human Security* 2014;2:257-286.

Martinez M, Castillo M. Study: Hispanics show increasing cultural, economic, and social diversity. CNN; September 16, 2013.

Martínez O. *The Beast: Riding the Rails and Dodging Narcos on the Migrant Trail*. New York: Verso; 2014.

Martínez O. *A History of Violence: Living and Dying in Central America*. New York: Verso; 2017a.

Martínez O. How not to assemble a country: In El Salvador, the legacies of violence persist and intensify. *NACLA Report on the Americas* 2017b;49(2):139-144.

Massey DS. The past and future of Mexico-U.S. migration. In: Overmyer- Velázquez M, ed. *Beyond la Frontera: The History of Mexico-U.S. Migration*. Oxford: Oxford University Press; 2011:251-266.

Massey DS, Durand J, Pren KA. Explaining undocumented migration to the U.S. *International Migration Review* 2014;48(4): 1028-1061.

Mbembe A, Meintjes L. Necropolitics. *Public Culture* 2003;15(1):11-40.

Menjivar C. History, economy and politics: Macro and micro-level factors in the recent Salvadorian migration to the US. *Journal of Refugee Studies* 1993;6(4):350-371.

Michael A, Bengston J, Blat S. Genes, race, ancestry, and identity in Forensic Anthropology: Historical perspectives and contemporary concerns. *Forensic Genomics* 2021;1(2):41-46.

Migration Policy Institute. Accessed: January 10, 2022. https://www.migrationpolicy.org/programs/data-hub/charts/origins-mexican-migrants-united-states-mexican-state-residence-number-and

Miller-Wolf KA, Freiwald C. Re-interpreting ancient Maya mobility: A strontium isotope baseline for Western Honduras. *Journal of Archaeological Science: Reports* 2018; 20:799-807.

Monsalve T, Hefner JT. Macromorphoscopic trait expression in a cranial sample from Medellin, Colombia. *Forensic Science International* 2016;266:574.e571-574.e578.

Mora-Torres J. "Los de casa se van, los de fuera no vienen": The first Mexican immigrants 1848-1900. In: Overmyer- Velázquez M, ed. *Beyond la Frontera: The History of Mexico-U.S. Migration*. Oxford: Oxford University Press; 2011:28-50.

Morrison AR, May RA. Escape from terror: Violence and migration in post-revolutionary Guatemala. *Latin American Research Review* 1994;29(2):111-132.

Murray EA, Anderson BE, Clark SC, Hanzlick RL. The history and use of the National Missing and Unidentified Persons System (NamUS) in the identification of unknown persons. In: Latham KE, Bartelink EJ, Finnegan M, eds. *New Perspectives in Forensic Human Skeletal Identification.* San Diego: Academic Press;2018:115-126.

National Missing and Unidentified Persons System (NamUs). www.namus.gov. Created 2016. Updated 2020. Accessed March 20, 2020.

New BT, Algee-Hewitt BFB, Spradley MK, Fehren-Schmitz L, Hughes CE, Anderson BE, Jasinski ME, Arciszewska J, Grażyna Z, Szargut M, Cytacka S, Ossowski A. Comparing genetic variation among Latin American immigrants: Implications for forensic casework in the Arizona- and Texas-Mexico borderlands. *Human Biology* 2021;93(1):33-50.

Noe-Bustamante L, Flores A, Shah S. Facts on Hispanics of Salvadoran origin in the United States, 2017. Pew Research Center, September 16, 2019a

Noe-Bustamante L, Flores A, Shah S. Facts on Hispanics of Venezuelan origin in the United States, 2017. Pew Research Center, September 16, 2019b

Oboler S. *Ethnic Labels, Latino Lives: Identity and Politics of (Re)presentation in the United States.* University of Minnesota Press; 1995.

Ongaro L, Scliar MO, Flores R, Raveane A, Marnetto D, Sarno S, et al. The genomic impact of European colonization on the Americas. *Current Biology* 2019; 29:3974-3986.

Ortega-Muñoz A, Price TD, Burton JH, Cucina A. Population movements and identify in Postclassic Yucatan. Bioarchaeological analysis of human remains from the East Coast of the Yucatan peninsula. *Journal of Archaeological Sciences: Reports* 2019;23:490-500.

Ousley SD. 3skull [computer program]. Windows version 1.76, 2014. Retrieved from: http://math.mercyhurst.edu/~sousley/Software/

Ousley SD. Forensic classification and biodistance in the 21st century: The rise of learning machines. In: Pilloud MA, Hefner JT, eds. *Biological Distance Analysis: Forensic and Bioarchaeological Perspectives*. Boca Raton: CRC Press; 2016:197-212

Ousley SD, Jantz RL. Fordisc 3 and statistical methods for estimating sex and ancestry. In: Dirkmaat D, ed. *A Companion to Forensic Anthropology*. Chinchester: Wiley-Blackwell; 2012:311-329.

Ousley SD, Jantz RL, Freid D. Understanding race and human variation: Why forensic anthropologists are good at identifying race. *American Journal of Physical Anthropology* 2009;139(1):68-76.

Overseas Security Advisory Council (OSAC). El Salvador 2020 Crime & Safety Report. https://www.osac.gov/Country/ElSalvador/Content/Detail/Report/b4884604-977e-49c7-9e4a-1855725d032e. Accessed: March 10, 2021.

Overmyer-Velázquez, M., Introduction. In: Overmyer- Velázquez M, ed. *Beyond la Frontera: The History of Mexico-U.S. Migration.* Oxford: Oxford University Press; 2011a:xix-xlv.

Overmyer- Velázquez M. *Beyond la Frontera: The history of Mexico-US Migration*. Oxford: Oxford University Press; 2011b.

Pagés J. Analyse factorielle de donnees mixtes *Revue Statistique Appliquee* 2004;4(93-111).

Paley D. Capitalism and crisis in Central America. In: Latham KE, O'Daniel AJ, eds. *Sociopolitics of Migrant Death and Repatriation: Perspectives from Forensic Science*. Cham: Springer; 2018:25-37.

Parks BO, Peters ED, Porterfield Cy, Winston D, Anderson BE. Border migrant deaths that the Pima County Medical Examiner's Office. In: Rubio-Goldsmith R, Fernandez C, Finch JK, Masterson-Algar A, eds. *Migrant Deaths in the Arizona Desert: La Vida No Vale Nada*. Tucson: The University of Arizona Press; 2016:120-131.

Pilloud MA, Hefner JT. *Biological Distance Analysis. Forensic and Bioarchaeological Perspectives.* Boca Raton: CRC Press; 2016.

Pima County Office of the Medical Examiner (PCOME). Annual Report 2017. Tucson; 2017.

Pink CM, Maier C, Pilloud MA, Hefner JT. Cranial nonmetric and morphoscopic data sets. In: Pilloud MA, Hefner JT. *Biological Distance Analysis. Forensic and Bioarchaeological Perspectives.* Boca Raton: CRC Press; 2016:91-107.

Plemons AM, Goots AC, Kamnikar KR. n.d. www.locatelambda.org

Price TD, Burton JH, Fullagar PD, Wright LE, Buikstra JE. Tiesler V. Strontium isotopes and the study of human mobility among the ancient Maya. In: Cucina A, ed. *Archaeology and Bioarcheology of Population movement among the Prehispanic Maya*. Cham: Springer; 2014:119-132.

R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing, 2018.

Reineke R. *Naming the Dead: Identification and Ambiguity along the U.S.-Mexico Border*. [PhD dissertation]. University of Arizona, Tucson; 2016.

Reineke R. Foreword. In: Latham KE, O'Daniel AJ, eds. *Sociopolitics of Migrant Death and Repatriation: Perspectives from Forensic Science*. Cham: Springer; 2018:v-viii.

Relethford JH. Craniometric variation among modern human populations. *American Journal of Physical Anthropology* 1994:95:53-62

Relethford JH. Population-specific deviations of global human craniometrics variation from a neutral model. *American Journal of Physical Anthropology* 2010;142(1):105-111.

Relethford JH. Race and global patterns of phenotypic variation. *American Journal of Physical Anthropology* 2009;139:16-22.

Relethford JH, Harding RM. Population genetics of modern human evolution. *Encyclopedia of Life Sciences*. John Wiley & Sons, LTD; 2001.

Relethford JH, Harpending HC Craniometric variation, genetic theory, and modern human origins. *American Journal of Physical Anthropology* 1994; 95:249-270.

Reyes-Centeno H, Hefner JT. Evolution of cranial macromorphoscopic trait variation in modern humans. Poster presented at the Human Biology Meeting virtual conference 2019.

Rice DS, Rice PM. Sixteenth- and seventeenth-century Maya political geography in central Petén, Guatemala. In: Kepecs S, Alexander RT, eds. *The Postclassic to Spanish-Era Transition in Mesoamerica. Archaeological Perspectives.* University of New Mexico Press: Albuquerque. 2005;139-160.

Roseman CC, Auerbach BM. Ecogeography, genetics, and the evolution of human body form. *Journal of Human Evolution* 2015;78:80-90.

Roseman CC, Weaver TD. Multivariate apportionment of global human craniometric diversity. *American Journal of Physical Anthropology* 2004;125(3):257-263.

Ross AH, Pilloud M. The need to incorporate human variation and evolutionary theory in forensic anthropology: A call for reform. *American Journal of Physical Anthropology* 2021;176:672-683.

Ross AH, Slice D, Ubelaker DH. Population affinities of Hispanic crania. In: Berg GE, Ta'ala SC, eds. *Biological Affinity in Forensic Identification of Human Skeletal Remains: Beyond Black and White*. Boca Raton: CRC Press; 2014:155-164.

Ross AH, Slice DE, Ubelaker DH, Falsetti AB. Population affinities of 19th century Cuban crania: Implications for identification criteria in South Florida Cuban Americans. *Journal of Forensic Sciences* 2004;49(1):1–6.

Rubi-Castellanos R, Martínez-Cortés G, Francisco Muñoz-Valle J, González-Martín A, Cerda-Flores RM, Anaya-Palafox M, et al. Pre-Hispanic Mesoamerican demography approximates the present-day ancestry of Mestizos throughout the territory of Mexico. *American Journal of Physical Anthropology* 2009;139:284-294.

Rylko-Bauer B, Farmer PE. Structural violence, poverty, and social suffering. In: Brady D, Burton L, eds. *The Oxford Handbook of the Social Science of Poverty*. Oxford: Oxford University Press; 2016

Santos F. AnthropMMD: an R package with a graphical user interface for the mean measure of divergence. *American Journal of Physical Anthropology* 2018;165(1):200-205.

Sassen S. Weaponized fences and novel borderings? The beginning of a new history. In: Overmyer- Velázquez M, ed. *Beyond la Frontera: The History of Mexico-U.S. Migration*. Oxford: Oxford University Press: 2011:xi-xiii.

Sauer NJ. Forensic anthropology and the concept of race: if races don't exist, why are forensic anthropologists so good at identifying them? *Social Science & Medicine* 1992;34(2):107-111.

Scientific Working Group for Forensic Anthropology (SWGANTH). Ancestry Assessment. 2013. Retrieved from: http://www.swganth.org/.

Sharer RJ, Traxler LP. *The Ancient Maya*. Stanford: Stanford University Press; 2006.

Slack J, Martínez DE, Lee AE, Whiteford S. The geography of border militarization: Violence, death, and health in Mexico and the United States. *Journal of Latin American Geography* 2016;15(1):7-32.

Smith J. Guatemala: Economic migrants replace political refugees. Migration Policy Institute. Created on April 1, 2006. Accessed on January 27, 2022. https://www.migrationpolicy.org/print/4613

Soler A, Beatrice JS. Expanding the role of forensic anthropology in a humanitarian crisis: An example from the USA-Mexico border. In: Latham KE, O'Daniel AJ, eds. *Sociopolitics of Migrant Death and Repatriation: Perspectives from Forensic Science*. Cham: Springer; 2018:115-128.

Spiros MC. Standardization of postcranial nonmetric traits and their utility in ancestry analysis. *Forensic Anthropology* 2019;2(1):29-44.

Spiros MC, Hefner JT. Ancestry estimation using cranial and postcranial macromorphoscopic traits. *Journal of Forensic Sciences* 2019;65:921-929.

Spiros MC, Kamnikar KR. Reporting biases between missing persons and unidentified persons in the U.S. Paper presented at the 73rd annual meeting of the American Academy of Forensic Sciences, February 15-19, 2021; Virtual Meeting.

Spradley MK. *Biological Anthropological Aspects of the Africa Diaspora; Geographic Origins, Secular Trends, and Plastic Versus Genetic Influences Utilizing Craniometric Data*. [PhD Dissertation]. Knoxville: The University of Tennessee; 2006.

Spradley MK. Biological distance, migrants, and reference group selection in forensic anthropology. In: Pilloud M, Hefner JT, eds. *Biological Distance Analysis Forensic and Bioarchaeological Perspectives*. San Diego: Academic Press; 2016a:231-244.

Spradley MK. Metric ancestry estimation from the postcranial skeleton. In: Berg GE, Ta'ala SC, eds. *Biological Affinity in Forensic Identification of Human Skeletal Remains: Beyond Black and White*. Boca Raton: CRC Press; 2014b:83-94.

Spradley MK. Metric methods for the biological profile in Forensic Anthropology: Sex, Ancestry, and Stature. *Academic Forensic Pathology* 2016b;6(3):391-399.

Spradley MK. Toward estimating geographic origin of migrant remains along the United States-Mexico border. *Annals of Anthropological Practice* 2014a;38(1):101-110.

Spradley MK. Use of craniometric data to facilitate migrant identifications at the United States/Mexico border. *American Journal of Physical Anthropology* 2021:175:486-496.

Spradley MK, Gocha TP. Migrant deaths along the Texas/Mexico border: A collaborative approach to forensic identification of human remains. In: Parra RC, Zapico SC, Ubelaker DH, eds. *Forensic Science and Humanitarian Action: Interacting with the Dead and the Living*. John Wiley & Sons Ltd.; 2020:537-548.

Spradley MK, Herrmann NP, Siegert CB, McDaneld CP. Identifying migrant remains in South Texas: policy and practice. *Forensic Sciences Research* 2019;4(1):60-68.

Spradley MK, Jantz RL. Ancestry estimation in forensic anthropology: Geometric morphometric versus standard and nonstandard interlandmark distances. *Journal of Forensic Sciences* 2016;61:892-898.

Spradley MK, Jantz RL, Robinson A, Peccerelli F. Demographic change and forensic identification: Problems in metric identification of Hispanic skeletons. *Journal of Forensic Sciences* 2008;53:21-28.

Spradley MK, Weisensee K. Why do forensic anthropologists estimate ancestry, and why is it so controversial? In: Tersigni-Tarrant MA, Shirley NR. *Forensic Anthropology: An Introduction.* Boca Raton: CRC Press; 2013:231-243.

Sneath PHA, Sokal RR. Numerical taxonomy: the principles and practice of numerical classification. W.H. Freeman; 1973.

Stanford C, Allen JS, Antón SC. *Biological Anthropology*. Third Edition. Pearson; 2011.

Techataweewan N, Hefner JT, Freas L, Surachotmongkhon N, Benchawattananon R, Tayles N. Metric sexual dimorphism of the skull in Thailand. *Forensic Science International: Reports, 4* 2021; 100236.

Techataweewan N, Tuamsuk P, Toomsan Y, Woraputtaporn W, Prachaney P, Tayles N. A large modern Southeast Asian human skeletal collection from Thailand. *Forensic Science International* 2017;278:406.e1-406.e6.

Tersigni-Tarrant MT, Shirley N. *Forensic Anthropology: An Introduction*. Boca Raton: CRC Press; 2013.

Tise ME. *Craniometric Ancestry Proportions among Groups Considered Hispanic: Genetic Biological Variation, Sex-Based Asymmetry, and Forensic Applications*. [PhD dissertation]. Tampa: University of South Florida; 2014.

Tise ME, Kimmerle EH, Spradley MK. Craniometric variation of diverse populations in Florida: Identification challenges within a border state. *Annals of Anthropological Practice* 2014;38:111-123.

Tiesler V, Zabala P, Cucina A. *Natives, Europeans, and Africans in Colonial Campeche: History and Archaeology*. Gainesville: University Press of Florida; 2010.

Translators Without Borders 2022. https://translatorswithoutborders.org/mayan-languages-of-guatemala-interactive-en/. Retrieved Jan 07, 2022.

Tyagi J. Weak states. *The Wiley-Blackwell Encyclopedia of Globalization*; 2012.

Ubelaker DH. The biological impact of European contact in Ecuador. In: Larsen CS, Milner GR, eds. *In the Wake of Contact: Biological Responses to Conquest*. Wiley-Liss, Inc.; 1994:147-160.

U.S. Census Bureau https://www.census.gov/quickfacts/fact/table/US/RHI725218. Created 2019. Accessed March 20, 2020.

U.S. Census Bureau. Accessed: April 10, 2021. https://www.census.gov/quickfacts/fact/note/US/RHI725219#:~:text=for%20racial%20categories.-,Definition,%E2%80%A2Puerto%20Rican

van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software* 2011;45.

Vogelsberg CCM. *Identification of Deceased Border Crossers: Investigating Spatial and Skeletal Attributes of Migrant Deaths*. [PhD Dissertation]. Michigan State University: East Lansing; 2018.

Vogt WA. Loss, uncertainty, and actin: ethnographic encounters with families of the missing in the Central America-Mexico-United States corridor. In: Latham KE, O'Daniel AJ, eds. *Sociopolitics of Migrant Death and Repatriation: Perspectives from Forensic Science*. Cham: Springer; 2018:53-66.

Wang S, Ray N, Rojas W, Parra MV, Bedoya G, Gallo C, et al. Geographic patterns of genome admixture in Latin American Mestizos. *PLoS Genetics* 2008;4(3).

Weitzer R. The puzzling neglect of Hispanic Americans in research on police–citizen relations. *Ethnic and Racial Studies* 2014;37(11):1995-2013.

Wilson RJ, Algee-Hewitt B, Jantz LM. Demographic trends within the Forensic Anthropology Center's body donation program. Paper presented at the Annual Meeting of American Association of Physical Anthropologists; 2007:252.

Winburn AP, Algee-Hewitt BFB. Evaluating population affinity estimates in forensic anthropology: Insights from the forensic anthropology database for assessing methods accuracy (FADAMA). *Journal of Forensic Sciences* 2021;1-10.

Winburn AP, Jennings AL, Steadman DW, DiGangi EA. Ancestral diversity in skeletal collections: Perspectives on African American body donation. *Forensic Anthropology* 2020; 1-12.

Zabala P. The African presence in the Yucatan: Sixteenth and seventeenth centuries. In: Tiesler V, Zabala P, Cucina A, eds. *Natives, Europeans, and Africans in Colonial Campeche: History and Archaeology*. Gainesville: University Press of Florida; 2010: Chapter 8.