2023

Temporal change in nonmetric traits of indigenous peoples of the American southwest

BOSTON UNIVERSITY

ARAM V. CHOBANIAN & EDWARD AVEDISIAN SCHOOL OF MEDICINE

Thesis

TEMPORAL CHANGE IN NONMETRIC TRAITS OF INDIGENOUS PEOPLES OF THE AMERICAN SOUTHWEST

by

MANDI M. BEAUVAIS

B.A., Utica College, 2017

Submitted in partial fulfillment of the

requirements for the degree of

Master of Science

2023

Approved by

First Reader

Sean D. Tallman, Ph.D.
Assistant Professor of Anatomy and Neurobiology

Second Reader

Tara L. Moore, Ph.D.
Associate Professor of Anatomy and Neurobiology

.

ACKNOWLEDGMENTS

Thank you to Eli, my husband, for his never-ending support and encouragement. Through coursework, setbacks, distance, a PCS, changes to your own career, and loss, your support for me never wavered. Without you, I may have never finished this thesis. I don't know that you'll ever know how much I truly appreciate everything that you do.

I would like to thank the New Mexico Decedent Image Database for accepting my research request and allowing me to use their data for this study. I would also like to extend a huge thank you to Megan L. Atkinson for allowing me to use her AMNH data when the pandemic prevented me from accessing a skeletal collection in person. Without your graciousness, I would not have been able to continue with this research.

I would also like to thank my advisor, Dr. Sean D. Tallman, who will have my everlasting gratitude for his help and patience throughout this endeavor.

Additional thanks to the Oneida Indian Nation of New York for the scholarship provided, which made attending graduate school possible.

TEMPORAL CHANGE IN NONMETRIC TRAITS OF INDIGENOUS PEOPLES OF THE AMERICAN SOUTHWEST

MANDI M. BEAUVAIS

ABSTRACT

This study seeks to distinguish the presence of secular change in contemporary nonmetric trait expression and explores the efficacy of morphological ancestry estimation with regards to North American indigenous groups. The study was conducted using cranial and mandibular trait data collected from two indigenous samples originating from the American Southwest: a pre-Contact sample from the American Museum of Natural History (n=150) and a modern sample from the New Mexico Decedent Image Database (n=100). To observe the modern sample, the RadiAnt DICOM viewer program was used to transform CT scan stacks into 3D models.

Pearson's chi-square analyses were used to assess the presence of statistically significant difference between nonmetric trait expression between the two samples. The analyses produced significant p-values (≤0.05) in 22 of the 24 observed traits. Using binary logistic regression equations, four models were developed to assess which traits contributed significantly to predicting group membership: Model 1 combined cranial and mandibular traits, Model 2 used seven of Hefner's (2009) 11 macromorphoscopic traits, Model 3 used only cranial traits, and Model 4 used only mandibular traits.

To evaluate the efficacy of an extant ancestry estimation method on indigenous sample, data from both the pre-Contact and modern samples were entered into the hefneR decision support system. The results indicate that the hefneR algorithm does not produce

reliable ancestry estimates for either pre-Contact or modern indigenous samples; the samples received 14% and 11% correct classification, respectively. This study demonstrates that secular change has affected nonmetric trait expression in indigenous groups and that pre-Contact samples should not be used as proxies for modern populations.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AMF ..... Accessory mandibular foramen

AMNH .....American Museum of Natural History

ANS .....Anterior nasal spine

ARP .....Ascending ramus profile

ARS .....Ascending ramus shape

ASUDAS .....Arizona State University Dental Anthropology System

CON .....constant

CP .....Chin prominence

CS .....Chin shape

CT ..... Computed Tomography

DAS .....Dental arcade shape

DICOM .....Digital Imaging and Communications in Medicine

DNA ..... Deoxyribonucleic acid

FDB .....Forensic Data Bank

GE .....Gonial eversion

GMR .....Gonial muscle attachment ridging

IBM SPSS .....International Business Machines Statistical Package for Social Sciences

INA .....Inferior nasal aperture

IOB .....Interorbital breadth

JPAC .....Joint POW/MIA Accounting Command

KLN .....Keeling

LPRI ..... Location of ramus inversion

MB .....Mylohyoid bridging

MBS .....Mandibular border shape

MF .....Number of mental foramen

MMS .....Macromorphoscopic Databank

MT .....Mandibular tori

MTF .....Malar tubercle form

NA ..... Native American

NAF .....Native American females

NAM .....Native American males

NAS .....Nasal aperture shape

NAW .....Nasal aperture width

NBC .....Nasal bone contour

NBS .....Nasal bone shape

NFS .....Nasofrontal suture

NMDID ..... New Mexico Decedent Image Database

NO ..... Nasal overgrowth

OMI .....Office of the Medical Investigator

OS .....Orbital shape

OSSA .....Optimized Summed Scored Attributes

PACS .....Picture Archiving and Communication System

PBD ..... Postbregmatic depression

PCOME .....Pima County Office of the Medical Examiner

PI .....Platymeric Index

PRG .....Prognathism

PRI .....Posterior ramus inversion

PZT .....Posterior zygomatic tubercle

SF ..... Suture form

SNS .....Supranasal suture

SPNS .....Supranasal suture

UNM .....University of New Mexico

US .....United States

WB..... Wormian bones

ZMS .....Zygomaticomaxillary suture

ZP ..... Zygomatic projection

CHAPTER ONE

Introduction

When skeletal remains are encountered, a forensic anthropologist is often consulted to aid in law enforcement investigations. During these consultations, the forensic anthropologist's two main objectives are to confirm that the remains in question are human osteological material and to develop the biological profile. The biological profile consists of estimations of age, ancestry/population affinity, sex, and stature. In addition to the biological profile, the anthropologist will also describe any other informative aspects of the remains, including the presence of pathological conditions, evidence of antemortem, perimortem, and postmortem trauma, and the presence of any medical implants. This information is gathered and synthesized to provide the most information about the unknown individual for use in the identification process.

Ancestry estimation may be the most extensively researched and most heavily debated aspect of the biological profile (Sauer 1992; Kennedy 1995; Little and Kennedy 2010; Bethard and DiGangi 2020; DiGangi and Bethard 2021; Ross and Williams 2021; Stull et al. 2021). Since its inception, the field of biological anthropology has worked to document and understand human variation, both within and between human groups. Then called physical anthropology, the field developed in the United States during the early 19th century; at the time, studies of human variation were primarily concerned with methods of racial classification and the development of a racial hierarchy (Little and Kennedy 2010). Early studies relied heavily on the biological race concept, which was based on the belief that human groups could be differentiated into immutable "races,"

each of which was biologically separate and unrelated to the others. As variation between these "races" were explored, the documented differences were used to draw conclusions about the intellectual capacity and potential for criminality. In these studies, white Europeans were centered as an ideal reference; they were thought to be the pinnacle of intellectualism and inherently physically superior to other races. These studies provided scientific support for beliefs of European superiority, the American institution of slavery, and the eugenics movement (Little and Kennedy 2010; Fuentes et al. 2019). As a subfield of biological anthropology, forensic anthropology's ancestry estimation practices have not remained unaffected by the field's history.

The field of biological anthropology has since rejected the biological race concept and has mostly, but not entirely, moved away from the racial terminology used in those early studies (Little and Kennedy 2010; Ross and Williams 2021). Modern biological anthropologists strive to distance contemporary studies from the field's racial past without ignoring the lingering effects of historic studies. Contemporary practitioners of biological anthropology have turned away from their predecessors' previous attempts to delineate hard boundaries between broad geographic populations (Little and Kennedy 2010). Instead, recent research focuses on topics of population specific biological variation, exploring the relationship between population history and human variation, quantifying their conclusions with statistical analyses, and addressing the relationship between population affinity estimates and social race (Birkby et al. 2008; Hefner 2007; Hefner 2009; Hefner et al. 2012; Hurst 2012; Hefner et al. 2015; Tallman 2016; Spiros and Hefner 2019; Atkinson and Tallman 2020; Go and Hefner 2020; Kilroy et al. 2020).

The continued refinement of ancestry estimation methodologies allows forensic anthropologists to associate individuals with more specific populations, potentially closer to the group an individual would have associated themselves with in life. Forensic anthropologists understand that the results of their ancestry estimations reflect a statistical probability based on the individual's similarity to a sample population; ancestry estimates are not absolute. While the topic of race has not completely faded, it is now understood that variation between human groups cannot be attributed to deterministic biological factors and that race is a social construct; the influence that social race exerts on an individual's lived experience is the result of societal factors, not biology (Little and Kennedy 2010).

An individual's ancestry can be estimated using either metric or nonmetric characteristics of the skeleton. Nonmetric traits, also called morphological or discrete traits, are physical characteristics of the skeleton whose expression varies due to a multitude of factors. Nonmetric traits have been documented in association with ancestral groups since the inception of the biological anthropological field (Chambellan 1883). Early anthropologists considered the presence of certain traits to be indicative of an individual's ancestry; in practice, trait lists were used to differentiate skeletal remains into "races." These "races" were initially broad and geographically associated, with many different populations combined under one term. For example, Asian groups and Indigenous peoples have historically been grouped together under the umbrella term "Mongoloid" (Rhine 1990). Contemporary ancestry estimation has moved away from the trait list approach and forensic anthropologists now use statistical approaches to

investigating ancestry. In the last few decades, studies lumping together Asian and Indigenous peoples have dwindled and population-specific studies have become more prevalent (Tallman 2016; Atkinson and Tallman 2020). Studies with Indigenous samples have overwhelmingly used archaeologically derived or historic Indigenous samples as proxies for modern Indigenous peoples, which is problematic as it functionally ignores the potential of change in their nonmetric trait expression. There is currently a paucity of literature exploring the reliability using ancestry estimation methods developed on archaeological samples on modern Indigenous peoples.

Rationale of the current study

The relevance of the present study is based on two factors: the importance of ancestry estimation to the creation of an informative biological profile and the disproportionate number of violent crimes committed against the Indigenous people of the United States. While the necessity of ancestry estimation is currently a topic of debate (Bethard and DiGangi 2020; DiGangi and Bethard 2021), there are valid reasons to continue to revisit and refine the extant methods (Stull et al. 2021). While the biological profile is not the only, or even the best, vehicle for the identification of an unknown individual, there are cases where it is the best source of information available. Since the advancement of modern technology, identification via DNA sample has become the gold standard of identification. Despite modern reliance on DNA technology, there are circumstances where DNA samples cannot be acquired or, even if it is, there is no

familial reference sample to compare it to. The same challenges can thwart attempts to use fingerprints or dental records for identification. Abandoning the study or advancement of biological profile methodologies while they maintain the potential for beneficial use seems neglectful.

It has been estimated that rates of violence on reservations are up to 10 times higher than the national average (Bureau of Indian Affairs 2023). Statistics gathered by the Center for Disease Control indicate that the homicide rate for Native women living on reservations is 10 times higher than the national average and that homicide is the third leading cause of death for Native women in the United States. In 2016, the National Crime Information Center reported that there were 5,712 reports of missing American Indian and Alaskan Native women and girls, though only 116 of these cases were input into the National Missing and Unidentified Persons System (NamUS); this number does not represent a reliable total for the number of Indigenous women who go missing or are murdered each year, as it has been reported that indigenous women are more likely to be listed as "unidentified remains" in the database than women of other groups (Hawes et al. 2022). Additionally, researchers have found that Indigenous women are often misclassified as Hispanic, Asian, or another non-Indigenous group on missing persons forms (Salam 2019). Native American men experience high rates of victimization as well; the National Institute for Justice found that 81.6% of American Indian and Alaskan Native men reported having experienced violence in their lifetime (National Institute for Justice 2022). As of June 2023, there are more than 800 open cases of missing Indigenous persons and over 180 unidentified Indigenous persons cases on NamUS

(NamUS 2023). The Bureau of Indian Affairs estimates that roughly 4,200 cases of missing and murdered Indigenous people currently remain unsolved (Bureau of Indian Affairs 2022).

Modern ancestry estimation methods are based on information gathered from samples of known individuals; ancestry estimation can be performed using metrical and nonmetric analyses. These methods have been developed since the field's inception, with contemporary methods focusing on creating reliable, replicable methodologies using statistical analyses and standardized methods (Ousley and Jantz 2005; Hefner 2009; Spradley 2016; Tallman 2016; Byrnes et al. 2017; Hefner and Linde 2018; Scott et al. 2018; Pilloud et al. 2019. Replicability of methods has been a focus of recent studies, since the Daubert ruling created stricter guidelines for expert witness testimony in criminal cases. Until recently, studies seeking to develop better methods of ancestry estimation have predominately focused on Black and white American populations; work with other major populations has begun in the last decade (Hurst 2012; Tise et al. 2014; Tallman 2016; Spradley et al. 2016; Atkinson and Tallman 2020; Herrera and Tallman 2019; Kilroy et al. 2020). Currently, there is a paucity of scholarship focused on modernizing of Native American ancestry estimation techniques and the extant standards are based on data collected from archaeological samples. Given the knowledge that nonmetric trait expression can change over time (Jantz and Meadows Jantz 2000; Martin and Danforth 2009; Kilroy et al. 2020), it is unlikely that the archaeological samples used to develop nonmetric ancestry estimation for Indigenous North American populations continue accurately represent modern populations. With thousands of unsolved cases of

missing and murdered Indigenous people (Bureau of Indian Affairs 2022), the use of theoretically inaccurate ancestry estimation methods has the potential to hinder the identification process.

Hypotheses

The purpose of this study is to establish the existence of dissimilarities in cranial and mandibular trait frequencies between two temporally distinct but biogeographically similar Indigenous samples and determine how it influences ancestry estimation efforts. Nonmetric ancestry estimation approaches continue to rely on archaeologically derived Indigenous samples and there is currently a lack of research exploring any disparities between historic and modern Indigenous nonmetric trait expression. To address these concerns, two hypotheses will be tested:

1. Statistically significant differences in cranial and mandibular nonmetric trait expression exist between pre-Contact and modern Indigenous individuals.
2. Differences in cranial and mandibular nonmetric trait expression and frequencies exist between pre-Contact and modern Indigenous individuals, causing currently employed ancestry estimation tools to misclassify modern Indigenous individuals into other ancestry groups.

If these hypotheses are supported, pre-Contact Indigenous samples should not be considered representative of contemporary Indigenous individuals and the validity of extant methods should be retested against modern Indigenous populations.

Terminology

There are a multitude of terms that have been used to describe the Indigenous peoples of North America, including "American Indian," "First Nations," and "Native American." It must be acknowledged that Indigenous groups may not identify with any of these terms; some Indigenous people choose to forego pan-Indigenous labels entirely, instead preferring to identify with the name of their tribe. Through her own lived experience as an Indigenous person, the author understands that each of these labels have their own deficiencies, and some are more easily recognized by non-Indigenous groups than others (e.g., tribal or band associations). Though the author recognizes that the Indigenous peoples of North America, are not a homogenous group that can be easily defined by a generalizing, pan-Indigenous label, the terms "Native American" and "Indigenous" will be used in the present study to discuss these groups in a broad manner.

While these terms, specifically "Native American," may not be labels that North American Indigenous peoples identify with, they are commonly used on official documents that require racial and ethnic information, such as census forms, tax documents, and government issued identification. The United States alone is home to more than 574 federally recognized tribes, with many more tribes currently fighting for federal recognition (usa.gov 2022). Each Indigenous tribe has their own unique culture, tribal interests, face their own challenges, and have unique population histories; the

author's use of a general term to refer to these people is not intended to ignore or undermine these differences but only to use concise, consistent terminology.

Organization of Chapters

Chapter Two of this study discusses the evolution of ancestry estimation practices in the field of forensic anthropology, beginning with the typological race concept that was characteristic of early physical anthropological theory and concluding with contemporary studies focused on population-specific biological variation and explorations of the reliability of extant ancestry estimation methods. Chapter Three describes the composition of the pre-Contact and modern samples, including a brief introduction to the institutions housing their parent collections. The chapter also includes information about the programs used to render the three-dimensional models and perform the statistical analyses. Chapter Four presents the results of the statistical analyses performed, including frequency distributions, chi-square analyses, binary logistic regression equations, ancestry estimates developed using the hefneR algorithm, and intraobserver error rates. Chapter Five discusses the results of the statistical analyses, their implications, and issues encountered during the course of the study. The chapter also includes a brief discussion of the recent discourse regarding the problematic nature of ancestry estimation practices and its potential removal from the forensic anthropological toolset. Lastly, Chapter Six briefly summarizes the previous chapters and discusses the broader implications of this study, including how current methods of

ancestry estimation may negatively affect Indigenous communities. The final chapter concludes with a brief discussion of the potential utility of digital skeletal collections to preserve data and improve accessibility for future research opportunities.

CHAPTER TWO

Previous Research

This chapter discusses the history and development of ancestry estimation practices, addressing the racial dogmas that drove early research and highlighting how contemporary biological anthropologists approach ancestry estimation study. Because ancestry estimation continues to be a focal point for biological anthropological research, research concerned with population histories and the adaptation of extant methods for new technologies are highlighted. A discussion of the relevant literature regarding the effect of secular change on human biological variation is also presented.

Physical anthropology and race

Human variation has been a topic of interest since geographically separate human groups began to encounter one another in earnest; attempts to categorize human groups began long before the field of physical anthropology had fully synthesized. Early scholars divided the human species into races, numbering as few as two and as many as fifteen (Morton 1839); four to five races were common. An Enlightenment period physician, Linnaeus, was one of the earliest scholars to attempt to differentiate and name modern humans. Linnaeus is known as the father of taxonomy for his creation of binomial nomenclature, or the modern system of naming organisms. In his work *Systema Naturae* (1758) that detailed his taxonomic system, Linnaeus defined the modern human species and four subcategories, according to their geographic origin: *Homo sapiens asiaticus,*

Homo sapiens europaeus, Homo sapiens afer, and *Homo sapiens americanus*. The subspecies were characterized by Linnaeus, providing the basis for prejudices and stereotypes. Though the Latin subspecies names did not continue into later works, Linnaeus' four broad, geographically associated races did.

Johann Blumenbach, a German anthropologist and physician, wrote his treatise *On the Natural Variety of Mankind* (1775), which agreed with Linnaeus' previously described four races but shifted the geographic boundaries for each and added a fifth. Blumenbach's treatise is an early example of scientific racism (Hamm 2019), as it positions white Europeans as the pinnacle of beauty and civility while describing other groups as degenerated versions of humanity. He attempted to find a reason for this degeneration, suggesting that environmental factors could be the cause. Some of the traits discussed include skin color; hair color and texture; face and head shape; and eye color and shape. Additionally, Blumenbach attempted to define how these traits would present in the "hybrid" children born of interracial unions, describing the expected presentation up to three generations (Blumenbach et al. 1865). Scientific racism and Christian creationism can be found in subsequent scientific literature and was prevalent in early physical anthropological thought.

During the early 19th century, anthropologists were divided between two main schools of thought: monogenism and polygenism. Monogenesists, like Blumenbach, believed that the Christian god created one race of human beings, most closely resembling white Europeans, and that all other races were offshoots of the original template (Blumenbach et al. 1865). Polygenist theory asserted that each race was created

separately from each other and that they had their own distinct, immutable characteristics (Little and Kennedy 2010). A well-known polygenist of the time, Samuel G. Morton, believed that human "races" were different species whose intellectual abilities could be determined through cranial capacity analysis. To justify his claims, Morton acquired a large collection of human skulls from around the world and determined their intracranial volume; those found to have larger cranial capacities were thought to be more intelligent, as he believed that larger skulls housed larger brains. Morton developed an intellectual hierarchy based on his research, which ranked white Europeans as the most intelligent and found females to be inherently inferior to males. The belief that all other races were intrinsically less intelligent later became the basis for continued scientific racism, including the justification for American slavery. While the reliability of his data collection process has been questioned (Gould 1981; Weisberg 2014), his research was very influential in the formation of the field; Morton was called the "Father of American Physical Anthropology," by 20th century anthropologist Ales Hrdlička (Hrdlička 1919).

The creator of the *American Journal of Physical Anthropology*, the American Association of Physical Anthropologists, and the Smithsonian's Division of Physical Anthropology, Ales Hrdlička, was also integral to the development of the American field of physical anthropology (Little and Kennedy 2010). Hrdlička, a Bohemian born immigrant to America, began his anthropological career at the Paris Anthropological Institute; he was so heavily influenced by his time there that he became determined to replicate the school in the United States. While the field was flourishing in European countries, there were very few American scholars who held degrees in physical

anthropology. Many who performed work that fell under the purview of a physical anthropologist were medical doctors, professional anatomists, or ethnographers. While Hrdlička would never manage to recreate his French alma mater in the United States, his efforts were fruitful in other ways. The creation of the *American Journal of Physical Anthropology* provided a centralized route of research dissemination, separated physical anthropological work from the more generalized body of anthropological publications, and set standards for research publication. During the creation of the American Association of Physical Anthropologists, the founding members created the objectives that members should work toward, including the standardization of methods and instruction, the preparation of textbooks and other teaching materials, cooperation between physical anthropologists and other researchers, and dedication to the advancement of worthy students. These achievements synthesized the field of physical anthropology in America and helped prime it for future generations of students (Little and Kennedy 2010).

Hrdlička's scholarship, like other scholars of the early 20th century, was affected by the biological race concept, the eugenics movement, and other sociopolitical ideas of the time (Little and Kennedy 2010). Hrdlička's views on such topics have been debated by those who discussed him posthumously and there seems to be evidence of changing views throughout his career (Little and Kennedy 2010). His early anthropometric research focused on determining criminal predisposition or mental inferiority through physical characteristics of white and Black American orphans, though he did not find evidence to support his belief in the existence of a "physical degenerate" (Hrdlička 1898).

He moved away from exploring the difference between Black and white Americans, focusing instead on the Indigenous peoples of the Americas (Hrdlička 1917). His research on how Indigenous peoples came to populate the Americas led to the Bering Strait theory; the idea that the Indigenous people of the Americas were the descendants of ancient people who crossed the Bering Strait from Asia was not initially popular. The theory's basis in the evolution and migration of humans out of Africa directly opposed the traditional idea that people of European descent were biologically separate and superior to the other races. Hrdlička's research became less focused on biological race, instead he concentrated on the idea of the "American type" and the change that occurred in the descendants of the "Old American" (Hrdlička 1925). While Hrdlička's belief in the plasticity of the human body and his position on the peopling of the Americas, Hrdlička's views on nonwhite groups cannot be heralded as progressive. While his efforts helped found the American field of physical anthropology, the contemporary conceptualization of biological race and prejudices against non-white populations have echoed since its founding.

Ancestry estimation using nonmetric traits

Morphological trait variation of the human skull has been described by anthropologists since the late 19th century. Chambellan's (1883) study of suture variation is thought to have been the first publication on the anthropological potential of nonmetric traits, while Dwight (1890) may have been the first to publish on their use in the creation

of the biological profile (Dwight 1890; Berry and Berry 1967). Since that time, anthropologists have continued to develop and refine nonmetric methods of ancestry estimation (Hooton 1930; Rhine 1990; Gill 1998; Hefner 2009; Hefner and Ousley 2014; Hefner and Linde 2018; Tallman 2016; Atkinson and Tallman 2019; Go and Hefner 2020). One of the earliest attempts to delineate ancestry via nonmetric trait expression, the creation of the Harvard Blanks, was performed by Harvard professor of Physical Anthropology, Earnest Hooton (Brues 1990; Rhine 1990; Hefner 2007; 2009). The Harvard Blanks, also called the Harvard Lists, were suites of morphological cranial characteristics though to be indicative of a single race. Not only were these lists used as a foundation for subsequent studies of ancestry estimation but, in his position as a professor, Hooton was able to propagate his ideas on human variability to his students.

Alice Brues, a student of Hooton's, honored her academic forefather through her work on human variation, racial classification, and, specifically, nonmetric traits (Brues 1958; Brues 1990). As one of the first generations of physical anthropologists to identify themselves as a forensic anthropologist, Brues' scholarship contains topics not dissimilar to the field's modern publications. Her article, *Identification of Skeletal Remains* (Brues 1958), discusses the elements of the biological profile and presents the methodologies of the time. During the discussion of "race determination," Brues differentiates three types of racial characteristics that can be found in modern humans: those that only appear in soft tissue (e.g., skin color, eye color and shape), features visible in both hard and soft tissue (e.g., nasal root and bridge contour), and those that are evident in the skull, alone (e.g., the form of the inferior border of the nasal aperture and cranial sutures). The

recognition of these traits is discussed as a method of racial determination; certain forms of these skull-exclusive traits were thought to correspond with a certain race. The article briefly mentions admixture as a potential factor in differing trait expression, stating that an unidentified individual with intermediate features may be interpreted as "mixed." To contextualize this interpretation further, the article suggests that anthropologists consider the racial constitution of the area the remains were recovered from to further narrow down potential social identities (Brues 1958). The topic of the biological profile continued to be a focus of subsequent publications, with anthropologists researching new ways to improve its accuracy for forensic casework.

Rhine (1990) assessed 45 cranial and mandibular morphological traits, providing lists and frequencies expected to be found in four broad groups: Anglo, Black, Hispanic, and Indian. The traits included had mostly been adapted from Hooton's early attempts to record nonmetric traits for ancestry estimation purposes. He defined each trait and described possible character states, though there was no true scoring system involved. In addition to these lists, Rhine (1990) published diagrams of the "typical" morphological presentation of American Caucasoid, Southwestern Mongoloid, and American Black crania, each with an associated trait list. To address individuals whose characteristics did not correspond with previously held expectations, Rhine stated that the samples were "not unmixed," suggesting that admixture caused the differences (Rhine 1990). This study can be criticized due to the exclusion of an Asian sample, the minute sample sizes, use of casts as portions of samples, and its typological nature. Using samples titled "Caucasoid," "Mongoloid," and "Negroid," Rhine (1990) grouped dissimilar populations

according to overarching geographic association. Despite these critiques, these trait lists, and the associated diagrams, have been influential in the study of human biological variation. Contemporary biological anthropologists understand that such broad groupings cannot address the range of human variation between and within the populations and conclusions drawn from such research are too general to be meaningful. In the last two decades, the field of biological anthropology has undergone a shift in paradigm regarding ancestral grouping and the terminology associated with it.

Contemporary nonmetric ancestry estimation

Since the turn of the century, there has been significant change to the way ancestry estimation research occurs. In the last two decades, anthropologists have focused on ancestry estimations that are reliable and produced through statistical analyses, population-specific studies, and the exploration of new technologies (Hefner 2009; Hefner and Ousley 2014; Hefner and Linde 2018; Tallman 2016; Scott et al. 2018; Pilloud et al. 2019; Atkinson and Tallman 2020). In an effort to develop an empirical method of ancestry estimation using morphological traits, Hefner and Ousley (2014) presented the Optimized Summed Scored Attributes (OSSA) method, which used dichotomized trait scores from six nonmetric traits (ANS, INA IOB, NAW, NBC, and PBD) to differentiate American Black and white individuals. To create dichotomized scores from ordinal systems containing more than two possible character states, scores most closely associated with Black individuals were assigned scores of 0 and those most

closely associated with white individuals were assigned a score of 1. Once all traits have been scored and all scores had been converted to their corresponding OSSA binary, the sum of the scores is calculated and the OSSA score is determined. This method produces an OSSA score ranging from 0 to 6; the sectioning point between the two possible ancestry estimates is set at 3, with scores of 3 or less indicating American Black ancestry and scores of 4 or more indicating American white ancestry. While this method attempted to develop ancestry estimates with statistical backing, several key issues diminish its applicability. Perhaps most notably, the OSSA method can only differentiate between two ancestry possibilities; if this technique is used to analyze an individual outside of the sample populations, the individual will receive an incorrect ancestry estimate of either American Black or white. This makes the method ill-suited for use on the general population, as its use would inevitably produce incorrect ancestry estimates. Hefner and Ousley (2014) addresses this concern, stating that OSSA is most applicable in situations where ancestry of most remains assessed are either American Black or white individuals, such as at JPAC (Hefner and Ousley 2014). Another flaw is that the method requires all six traits to be present for the analysis to produce an ancestry estimate. This prevents its use on fragmentary or partial crania, which are often recovered in forensic casework. Additionally, the OSSA method ignores the possibility of multivariate relationships between traits, as all six traits are weighed the same in the analysis. Although this method's weaknesses greatly restrict its use, Hefner and Ousley (2014) do point out that scoring nonmetric traits is more suitable than the trait list approach and facilitates the statistical analyses required by the Daubert criteria.

Two publications, Hefner (2009) and Hefner and Linde (2018), have aided in the production of reliable ancestry estimation results through the standardization of nonmetric trait observation. Hefner's (2009) article, and the subsequent volume produced by Hefner and Linde (2018), provide clear directions for trait scoring, detailed character state descriptions, and images to aid in visualization. These works help bridge the gap between observers with varying levels of experience working with skeletal material; accessible character state descriptions and illustrations provide guidance for the inexperienced and a centralized source for score confirmation for the experienced. The need for this type of standardizing material was supported by Kamnikar et al.'s (2017) exploration of short- and long-term intra-observer error, which implored the use of clearly conveyed trait nomenclature and highlighted the importance of thorough training on morphoscopic traits. Hefner's (2009) study was one of the first to attempt to explain the variation between and within multiple ancestral groups using statistical analysis of trait frequencies. Subsequent researchers have continued to explore the role of morphoscopic traits in ancestry estimation (Tallman 2016; Atkinson and Tallman 2020; Spiros and Hefner 2019; Go and Hefner 2020).

Another study that propagates the benefits of multivariate statistical analysis in ancestry estimation was performed by Spiros and Hefner (2019). Their research explored the accuracy of ancestry estimation that employed both cranial and postcranial morphological traits. Spiros and Hefner (2019) demonstrated the value of postcranial skeletal traits due to the previous literature's tendency to focus primarily on cranial morphoscopic traits. The study assessed eight cranial and 11 postcranial traits on

American Black and American white individuals from the Terry and Bass skeletal collections. The results indicated that the combined scoring of cranial and postcranial traits had a classification accuracy of 88-92%. This accuracy rate is significantly higher than using only cranial or postcranial traits, which were 80-85% and 77.6-81.6% respectively. The authors encourage the use of a multivariate approach to ancestry estimation, illustrating the increase of accurate classification.

Go and Hefner (2020) focused on the application of the probit regression model to the differentiation of Filipino individuals from American Black, American white, Asian, and Hispanic populations. Probit regression analysis was chosen for the study because it outperforms logit regression analysis when the traits examined are few or not normally distributed within the sample. Additionally, previous research had failed to test this method. The Filipino sample was obtained from two osteological collections, the Smithsonian National Museum of Natural History's historic collection and an archaeological collection recovered in Manila. All other sample populations were taken from the Macromorphoscopic Databank (MMS); the MMS data collection program was created to complement Hefner's (2009) method and facilitates the collection and analysis of macromorphoscopic data while minimizing observer error (Hefner 2018). The samples were compared through analysis of nine morphoscopic traits, which were scored using the ordinal systems detailed by Hefner (2007; 2009) and Hefner et al. (2012). The study found that the probit regression analysis most often classified Filipino individuals into the Asian group. Go and Hefner (2020) addressed this, stating that probit analysis works best when the test sample is reflected in the training samples. This study highlights the

necessity of appropriate reference samples and illustrates the viability of a previously untested method of statistical analysis.

In the last few decades, dental morphology has become a successful approach to assess population affinity. While dental morphological variation had previously been documented in association with ancestry estimation (Hrdlička 1920; Rhine 1990; Scott et al. 2018), subsequent researchers have developed a formal statistical framework through which meaningful conclusions can be drawn (Edgar 2013; Pilloud et al. 2014; Pilloud et al. 2018; Scott et al. 2018; Pilloud et al. 2019). Like the development of cranial and mandibular nonmetric trait ancestry estimation, the use of dental morphology in the forensic context has advanced through the culmination of much scholarship (Scott and Turner 1997; Turner et al. 1991; Edgar and Lease 2007; Hanihara 2008; Edgar 2013; Edgar 2017; Pilloud et al. 2018; Pilloud et al. 2019). The use of dental morphology in ancestry estimation has several advantages that cranial and mandibular traits do not possess: teeth are often better preserved than the other portions of the skull, researchers have outlined the heritability of dental morphological traits, and dental morphology is relatively stable throughout an individual's lifetime and is more closely related to genetics than skeletal plasticity (Scott et al. 2018). Efforts to regulate dental nonmetric trait scoring systems have culminated in the Arizona State University Dental Anthropology System (ASUDAS), which has standardized ordinal scoring systems and has made physical reproductions of each trait, depicting the associated character states (Turner et al. 1991). The codification of dental morphology and increased availability of trait casts supports consistent data collection between researchers, making dental

morphological ancestry estimation accessible to a wider range of biological anthropologists. Recent research has advocated for the incorporation of dental morphological assessment to the biological profile (Scott et al. 2018; Pilloud et al. 2019).

Scott et al. (2018) presented their web-based application that uses dental morphological trait data to statistically estimate ancestry. The application, rASUDAS, allows the user to input trait score for 21 independent nonmetric traits and compares it to reference samples representing seven biogeographic regions: American Arctic & Northeast Siberia, Native American, East Asia, Southeast Asian and Polynesia, Australo-Melanesia and Micronesia, Sub-Saharan Africa, and Western Eurasia. The application uses a naïve Bayesian classifier algorithm to develop posterior probabilities regarding individual group membership. Though tests of the application yielded moderate classification accuracies (51.8%-72.7%), the authors posit that accuracy could increase if the reference sample was refined. Scott et al. (2018) did not recommend that rASUDAS be used as the sole method contributing to an ancestry estimate but did suggest that the application could be used to complement the results of other methods. Scott et al.'s (2018) rASUDAS is currently available on the web platform, Osteomics (d'Oliveira Coelho et al. 2020; Osteomics 2023).

Osteomics is a free web platform that hosts decisions support systems for biological anthropologists (d'Oliveira Coelho et al. 2020; Osteomics 2023). The website offers decision support systems for both forensics and paleoecology; the forensics section is divided into subsections which include biogeography, sex diagnosis, age at death, and body parameters. Another biogeography prediction algorithm, named hefneR, uses

cranial nonmetric trait scores to produce an ancestry estimate (d'Oliveira Coelho et al. 2020; Osteomics 2023). Developers João d'Oliveira Coelho and David Navega used Hefner's (2009) dataset to create their algorithm; hefneR produces ancestry estimates using a simple Bayesian classifier. The user can input trait scores for the 11 nonmetric traits described in Hefner (2009): anterior nasal spine, interorbital breadth, malar tubercle, nasal aperture width, nasal bone contour, nasal overgrowth, postbregmatic depression, supranasal suture, transverse palatine suture, and zygomaticomaxillary suture. If scores are missing for any of the traits, selecting "missing" for that trait is also an option. Once all data is input, the user must then transition to the Analysis tab, verify their inputs, and select "Process Data." The ancestry estimate is made available on the Results tab, which provides the probability of an individual's group membership in each of the four reference samples: African, American Indian, Asian, or European. The Batch Processing option allows the user to submit data for multiple individuals at once (d'Oliveira Coelho et al. 2020; Osteomics 2023). While this tool may be useful as a supplemental validation of ancestry estimation conclusions, the reference samples have not been augmented beyond Hefner's (2009) dataset, which used archaeological and historic collections to create his indigenous sample.

Perhaps less studied but still relevant are ancestry estimation methodologies that focus on portions of the skeleton other than the cranium and dentition are techniques that use the mandible and femur (Wescott 2005; Berg 2008). While this study does not utilize postcranial methods of ancestry estimation, Wescott's (2005) method does use morphological characteristics to differentiate American Black and white populations

from Native Americans. Wescott (2005) examined the utility of the subtrochanteric morphology of the proximal femur morphology for ancestry estimation. To assess this, Wescott (2005) collected measurement data from five groups: Native American, Polynesian, "Hispanic," American Black, and American white. The four measurements observed, femur maximum length, vertical head diameter, subtrochanteric anteroposterior diameter, and subtrochanteric mediolateral diameter were used to calculate the platymeric index and examine the relationship between femur size and morphology. Femur subtrochanteric shape is evaluated using the Platymeric index (PI), which is calculated by dividing the subtrochanteric anteroposterior diameter by the mediolateral diameter and multiplying the result by 100. A PI of less than 84.9 is platymeric (wider mediolaterally), a PI between 85-99.9 is eurymeric (approximately equal width mediolaterally and anteroposteriorly), and a PI greater than 100 is stenomeric (wider anteroposteriorly) (Wescott 2005). Wescott (2005) critiqued previous studies for using samples that treated Native American groups as homogenous and subdivided the Native American sample according to their region and typical terrain encountered, and subsistence strategies. This allowed for an examination of how biomechanics influence femur morphology. The study found that femur subtrochanteric shape can be used to differentiate American Black and white individuals from Native Americans, supporting the work of previous researchers who asserted the morphology of the proximal femur can be used for ancestry estimation purposes (Gill 1990; 2001). Wescott (2005) found that the femora of Native Americans were on average more platymeric than those of American Black and white individuals. Similar to other studies of ancestry estimation methods that include a Native American

sample, the reference sample used in the study was largely composed of the remains of prehistoric and historic Native Americans from the Great Plains. He concluded, while the method worked, his Native American sample may have inflated the accuracy as Native American groups from the Great Plains were generally the most platymeric of all observed groups (Wescott 2005).

Berg's (2008) study examined both metric and nonmetric traits of the mandible for use in the development of the biological profile. To explore their potential, Berg (2008) collected data for 11 mandibular measurements and 7 morphoscopic traits; the study used 17 world samples, including American Black and white, "Hispanic," pre- and protohistoric Native Americans, and Southeast and Northeast Asian samples. While most of the observed morphological traits were analyzed according to the extant literature, Berg (2008) expanded several trait definitions and scoring categories. Additionally, he created graphics depicting the various character states for the traits used. The study found that both mandibular metrics and morphoscopic traits can be used to estimate sex and ancestry in a variety of world populations. While metrical analyses did produce higher rates of accuracy than morphological analyses, the study demonstrated that nonmetric traits produced reliable sex and ancestry estimates. Berg's (2008) work with morphoscopic traits of the human mandible are relevant to the present study because four of the seven observed traits were also studied here; chin shape, lower border of the mandible (mandibular border shape in the present study), ascending ramus shape, and gonial angle flare (described as gonial eversion here).

Population-specific studies

While early studies of human biological variation did briefly discuss ancestral groups other than American Black and white populations, most studies focused on the differences between those two groups. Though this could be, in part, due to the availability of large skeletal collections comprised of American Black and white individuals, investigations into other ancestral groups were largely tangential. Population-specific studies focused on groups other than American Black and white began to become more prevalent in the 2000s (Birkby et al. 2008; Hurst 2012; Hefner et al. 2015; Tallman 2016; Atkinson and Tallman 2020).

To address the complexities of increasing numbers of unidentified decedents who died while attempting to cross the US-Mexico border, it was necessary to develop new methods of differentiating foreign nationals and US citizens (Birkby et al. 2008). In their study, Birkby et al. (2008) detail how the Pima County Office of the Medical Examiner (PCOME) have identified a suite of nonmetric traits that, in combination with a specific cultural profile, can be used to distinguish the remains of undocumented border crossers from deceased United States citizens. This task is particularly important as the remains of nearly 750 individuals were recovered from the areas surrounding the US-Mexico border in a five-year span, with investigations resulting in the positive identification of 75% of individuals. The differentiation between US citizens from the American Southwest and foreign national border crossers is difficult because the two groups are likely to be similar, given their potential to have shared population histories. Referencing Rhine (1990), Birkby et al. (2008) highlights the past use of the "Hispanic" label as an umbrella

term for certain populations and how it fails to address the nuance of the various groups that fall under that term. The term “Southwest Hispanics” (as opposed to those in other areas such as the Southeastern United States or the Caribbean) was used to describe individuals with skeletal traits from both European and Indigenous groups. The PCOME developed a suite of nonmetric skeletal traits to identify “Southwest Hispanics” from their extant trait lists for European and Southwest Indigenous groups. To help separate “Southwest Hispanic” US citizens and undocumented border crossers, a cultural profile was developed involving geographic context, personal effects, the condition of the individual’s dentition, and cultural accoutrements. Birkby et al. (2008) reported that the individuals identified as undocumented border crossers were generally young adults between 20-30 years of age from low-income areas of Mexico and other Latin American countries. These individuals also were generally of shorter stature than their peers in the US and had experienced worse dental health during their lifetime. In particular, individuals identified as undocumented border crossers were found to have higher rates of dental caries without amalgam restoration, tooth crowns that had been destroyed by caries, and enamel hypoplasias (Birkby et al. 2008). The remains of undocumented border crossers were most likely to be recovered from areas identified as migrant corridors and be in possession of items such as water bottles, foreign currencies, and cultural items such as amulets for protection or items with religious icon imagery. This study illustrated the need for population-specific identification methodologies, which has continued to be a topic of interest in the field of forensic anthropology (Birkby et al. 2008).

Hurst's (2012) study explored the viability of nonmetric traits to differentiate Hispanic, European American, and African American populations. The work resulted in the addition of two morphological traits, improvement upon extant character state illustrations, and the development of new illustrations where they had previously been missed. It was determined that Birkby et al.'s (2008) suite of traits is viable for ancestry estimation, and she posits that assessing a larger number of traits increases prediction accuracy. Hurst highlighted the problematic use of "Hispanic" in ancestry estimation studies, pointing out that the term refers to a shared-language group and has no basis in biological population affinity. It is suggested that biological anthropologists should avoid using colloquial terminology and instead use language specific to ancestral groups (Hurst 2012). This perspective is echoed in Hefner et al. (2015), who also explored the use of nonmetric traits in differentiating Hispanic individuals from other ancestral groupings.

Hefner et al. (2015) used cranial nonmetric traits to differentiate individuals of known ancestry; the individuals were from Southwestern Hispanic, Guatemalan, American Black, and American white samples. Using nonmetric analyses, the authors found that they could distinguish between individuals from the Southwestern Hispanic and Guatemalan samples. Additionally, they found that combining the two groups resulted in higher accuracy rates. This conflicts with the findings that indicate that six out of eight traits examined present differently in the Guatemalan and southwest Hispanic samples. Hefner et al. (2015) suggest that the increased accuracy that resulted from combining the two samples is likely due to inefficiency of the machine learning algorithm used in the study and is not related to population history similarities. The study

illustrated that groups that have historically been classified into the same broad category can be separated using morphological traits.

Further investigation into population-specific variation has continued in groups other than American Black and white groups; Tallman (2016) and Atkinson and Tallman (2020) focused on the investigation of skeletal variation in Asian and Asian-derived groups. Tallman's (2016) study demonstrated that nonmetric traits can be used to differentiate between groups that have historically been classified broadly as Asian. The work's findings determined that it is possible to morphologically discern between Japanese, Thai, and Native American groups. These conclusions were supported by Atkinson and Tallman's (2020) study, which sought to use nonmetric traits to differentiate modern Thai and Japanese individuals with pre-Contact Native Americans. This addressed the field's previous attitude toward Native American and Asian groups, positing that the two groups should not be combined during such studies, as the two groups would present differently due to their separate population histories. The analysis concluded that modern Thai and Japanese individuals and pre-Contact Native American populations have different nonmetric trait frequencies and, therefore, can be individuated.

Metrical ancestry estimation

Discriminant function analyses have been employed for ancestry estimation in the forensic anthropological context since the 1960s (Giles and Elliot 1962; Ayers et al. 1990) but were difficult to use because they required all measurements to be present to

produce results. Improvements on the Giles and Elliot (1962) formulae inspired the creation of the FORDISC computer program (Jantz and Ousley 2005; Ousley and Jantz 2013). FORDISC is currently the primary analytical tool used to estimate ancestry using cranial and postcranial skeletal measurements (Jantz and Ousley 2005; Ousley and Jantz 2013). Using customized linear discriminant function analyses, the program uses up to 24 available measurements to compare an unknown individual to up to 13 reference groups. The cranial reference groups currently available in FORDISC are American Black, American Indian, American white, Hispanic, and Japanese males and females, and Chinese, Guatemalan, and Vietnamese males. Howells' (1973, 1989, 1996) craniometric dataset was also incorporated in the FORDISC 3.0 update, which provides data from 28 groups worldwide. The postcranial reference samples are limited to American Black and white males and females. Much of the reference data comes from measurements taken on known individuals compiled in the Forensic Data Bank (FDB). The FDB contains demographic information on more than 2400 individuals, with nearly 900 cases of definitive sex and ancestry information, and at least 625 positively identified individuals. The FDB currently accepts data contributions from forensic anthropologists regarding individuals positively identified through their forensic casework (Jantz and Ousley 2005).

The program's ancestry estimation output provides posterior and typicality probabilities based on the measurements used in the analysis (Jantz and Ousley 2005). Posterior probabilities inform about the individual's probability of membership in each of the reference groups in the analysis, based on the similarity of measurements. The posterior probabilities sum to 1, with the highest posterior probability indicating group

membership. The typicality probability represents the likelihood of the individual's group membership based on the reference groups' average variability. Through the calculation of absolute distances, an individual's typicality probability demonstrates how much they differ from the reference groups' average morphology. Typicality probabilities of less than 0.05 (5%) indicate dubious likelihood of group membership or potential measurement error (Jantz and Ousley 2005).

Although the FORDISC program provides statistical backing to metrical ancestry estimation, its authors did caution users that there are several inherent weaknesses to be wary of. A major concern about the program's reliability for ancestry estimation relates to its limited reference samples (Jantz and Ousley 2005). Although the reference samples have been augmented through periodic updates, eight of 13 reference groups contain less than 100 individuals. FORDISC's current "American Indian" sample is severely lacking; the sample is comprised of a total of 91 individuals, 59 males and 32 females (Jantz and Ousley 2005; Dudzik and Jantz 2016). Of the 91 individuals in the sample, only 20 individuals are from a modern forensic context. To supplement the forensic cases, the rest of the sample consists of Indigenous individuals from the mid-to-late 19th century. Most of the forensic cases originate in the American Southwest, while the mid-to-19th century individuals are mainly from the Great Plains (Jantz and Ousley 2005; Dudzik and Jantz 2016). Additionally, Jantz and Ousley (2005) noted that secular change may affect FORDISC's ability to reliability estimate ancestry and highlighted the importance of reference samples that are temporally representative of the unknown individuals the program is used on (Jantz and Ousley 2005).

Several critiques about the use of the FORDISC program for ancestry estimation have been highlighted in subsequent research. Williams et al. (2005) noted that the ancestry estimation generated by FORDISC may vary depending on the sex of the reference samples included in the analysis. Similarly, the number of variables included in the analysis may either increase (Hubbe and Neves 2017) or decrease (Jantz and Ousley 2005) the accuracy of an ancestry estimation. Elliot and Collard (2009) discuss several reasons that the use of FORDISC for ancestry estimation is problematic, including their findings that the program's ancestry estimation capabilities are only useful when the remains are complete or nearly complete. Complete and nearly complete skeletal remains tend to be less common in archaeological and forensic contexts, which reduces the program's applicability. Additionally, ancestry estimates are only reliable when the unknown individual belongs to a population represented in the reference samples. Because FORDISC is based on discriminant function analyses, the program will always produce an ancestry estimation, even if an individual is not represented by a reference group. With such a limited number of reference groups, FORDISC is unlikely to be effective for all unidentified individuals (Elliot and Collard 2009).

Further, there are concerns regarding the reliability of the Howells craniometric data due to the strategy used to select individuals for the sample (Elliot and Collard 2009). Howells did not use a randomized sampling method, instead he chose individuals he considered to have the most "typical" presentation for each group and intentionally excluded individuals he deemed to be outliers (Howells 1989; Elliot and Collard 2009). This results in low levels of overlap in sample groups, making the dataset less ideal for

use in FORDISC, as discriminant function analyses rely on overlap between samples (Elliot and Collard 2009). Additionally, the variability between sexes present in the Howells dataset may not be reliable, as Howells excluded individuals whose sex he could not be certain of. This likely resulted in an exaggeration of dimorphic characteristics in the sample, as any individuals with indeterminate or mixed characteristics would not have been observed (Elliot and Collard 2009). Despite its issues, the use of FORDISC metrical ancestry estimation is prevalent in the forensic anthropological context.

Secular change and human biological variation

While the present research is not specifically a secular change study due to the large-scale time depth and inability to ensure population continuity, studies of secular change have demonstrated the effect of time on skeletal morphology for decades (Angel 1976; Angel 1982; Meadows Jantz and Jantz 1999; Jantz and Meadows Jantz 2000; Martin and Danforth 2009; Spradley et al. 2016; Kilroy et al. 2020). In their study of skeletal variation in American Black and white individuals, Angel (1976) found that there was an increase in overall head and body size since the colonial era. These changes were attributed to better nutrition and less exposure to illness, but the samples were too small to draw truly meaningful conclusions (Angel 1976). In a later study of secular change in the cranial base height of American Black and white individuals, Angel (1982) found that cranial base height change is correlated with body size increase, which was also attributed to socially influenced factors such as nutrition and health. Specifically, Angel

(1982) suggested that suboptimal nutrition during childhood may exaggerate the effects of secular change in skeletal presentation. Meadows Jantz and Jantz (1999) focused on allometric change of the postcranial skeleton in American Black and white individuals. The study found evidence of the effect of secular change on upper and lower limb length, with the lower limbs being altered to a lesser degree. This change in skeletal morphology was thought to be linked to fetal development conditions, which differed from previous discussions of socioeconomic influence. Jantz and Meadows Jantz (2000) collected data on five craniofacial variables to assess secular change in the cranial morphology of 19th century and 20th century American Black and white individuals. They found that the cranium had undergone change, with the dimensions of the vault expression the most difference while the facial skeleton exhibited little secular change. While there was variation in the extent of the change exhibited in these areas, both the facial skeleton and the cranial vault were found to have become more narrow and higher, generally speaking. These shifts in morphological presentation were also attributed to changes in nutrition and health, to which Jantz and Meadows Jantz (2000) described the modern American as having an overly nutritious diet and having a lifestyle so sedentary that it had become a health problem. Martin and Danforth's (2009) investigation into secular change in the mandible observed 1002 adult mandibles from the Hamann-Todd and Terry collections, collecting data on ten standardized measurements. The study found a two-fold trend in mandibular morphology: the mandibular body had become more gracile, while the mandible was becoming longer, overall. These changes were consistent with general

trends demonstrated in previous studies of the cranium. Martin and Danforth (2009) also attributed these changes to nutrition, health, and other lifestyle factors.

In a more recent population-specific study, Spradley et al. (2016) used craniometrics to determine if secular change had occurred in modern Mexican migrants and if those changes can be used to differentiate between historic Hispanics and recent Mexican decedents. The historic Hispanic sample (n=112) was derived from a predominately Mexican cemetery in Tucson, Arizona and the recent migrant sample (n=216) was comprised of the remains of decedents recovered along the US-Mexican border. Data were collected for a total of 82 craniofacial interlandmark distances measuring the size of the facial skeleton and the cranial vault. While secular change was only demonstrated in a few variables, some overarching differences in cranial morphology were evident. The historic Hispanic sample was found to have generally larger dimensions than those found in recent Mexican migrants. Additionally, as previous discussions (Birkby et al. 2008; Hurst 2012), the study highlighted the issues regarding the term "Hispanic" as a population label, pointing out that its use results in many dissimilar populations being forced into a singular category.

Kilroy et al.'s (2020) study is particularly relevant to the present research, as it was the first study to document secular change in nonmetric trait expression. Data were collected for 23 cranial and mandibular nonmetric traits in European American individuals. The European American sample (n=1130) was comprised of individuals from the Hamann-Todd Skeletal Collection and the William M. Bass Donated Skeletal Collection. The sample was divided into six birth-year cohorts and Pearson's chi-square,

along with other statistical analyses, was used to evaluate variations in trait frequencies between the groups. The study documented statistically significant change in the expression of 11 traits, six of which were examined in the present study. The most marked alteration occurred around the turn of the century, which Kilroy et al. (2020) theorized could have been the result of increased industrialization, lifestyle changes, and a decrease in nutritional stress. The study demonstrated that nonmetric trait frequencies can undergo statistically significant change in relatively short time spans. If microevolutionary forces have altered the expression of nonmetric traits in the Indigenous populations of the American Southwest, the present research will document its effect and explore how it influences ancestry estimation.

Currently, there have been no studies of secular change in Native Americans and, in general, modern Native Americans have been neglected by forensic anthropological research in favor of archaeologically derived or museum-based samples. While this may be partially due to the availability of these samples, the lack of data on the effect of secular change in modern Native Americans reduces the reliability of both nonmetric and metric methods of ancestry estimation. The Indigenous peoples of North America have not been exempted from changes in lifestyle, nutrition, and other forces highlighted by these authors as causes of secular change (Angel 1976; Angel 1982; Meadows Jantz and Jantz 1999; Jantz and Meadows Jantz 2000; Martin and Danforth 2009; Spradley et al. 2016; Kilroy et al. 2020); the effects of these changes deserve investigation.

Three-dimensional reproduction of skeletal material and the biological profile

In the wake of continuously advancing technologies, the field of forensic anthropology has recently begun to explore the utility of computer programs in the development of the biological profile. Among these programs, software that enables the user to access virtual, three-dimensional renderings of the human skeleton have been a topic of interest (Decker et al. 2011; Garcia de Leon Valenzuela 2014; Bertoglio et al. 2020). Such studies are of particular importance, as the current research employs a computer program which renders CT scan stacks into a three-dimensional model. Decker et al.'s (2011) study investigated the viability of using CT scans to assess metric and nonmetric traits of the adult pelvis for sex estimation. The study utilized Mimics software to render the DICOM data into a 3D reproduction, which was used to score morphoscopic traits. The metric analysis was performed with an osteometric toolkit built into the Mimics software package. The metric data gathered was also entered into FORDISC 3.0 to assess how the study's results compare to the field's standard. The results demonstrated that the use of 3D images developed from CT scans can be accurately used to estimate sex and were found to be more reliable than FORDISC 3.0. When scored and measured by students, the accuracy of correct sex classification remained high. This is particularly valuable as CT scans allow researchers to view skeletal features without needing to physically manipulate them or remove soft tissue (Decker et al. 2011).

Garcia de Leon Valenzuela's (2014) study tested the reliability of a biological profile developed from 3D imagery of skeletal remains. Using CT scans taken of

individuals from the Boston University (School of Medicine) skeletal collection, 3D models of the innominate bones were generated. Employing both the skeletal remains and the corresponding digital models, the study applied nonmetric age and sex estimation methods; approaches outlined by Phenice (1969) and Brooks and Suchey (1990) were used. Additional observers with varying levels of experience in the anthropological field applied the same methods to a subset of the sample to test for interobserver reliability. Garcia de Leon Valenzuela (2014) found that the results of the age and sex estimations were the most reliable when performed by the observer with the most experience working with digital models, the author. Higher levels of experience with anthropological methods did not necessarily correspond with higher levels of result reliability; observers found it difficult to translate methods they learned on physical remains to digital models. The author suggested that standardized scanning procedures focused on the optimization of 3D image resolution and the provision of 3D model training could result in better rates of reliability (Garcia de Leon Valenzuela 2014).

Bertoglio et al.'s (2020) study indicated that the efficiency of the 3D rendering software used can influence morphological trait analysis. The authors utilized the open-source DICOM viewing software Slicer 4.10.1. The 3D images rendered in Slicer were considered reasonable replicas yet only five of nine cranial bones observed were perfectly reproduced. When adjustments were made to the program's settings, seven of nine cranial bones were perfectly reproduced. Additionally, the 3D images only conveyed 60% of bone defects present on the physical bone, with the majority of loss occurring on small bone defects. While morphological trait assessment was accurate in cranial elements that

were perfectly reproduced, the bones that were imperfectly rendered presented difficulties for observers (Bertoglio et al. 2020). The two studies indicate that 3D rendering of CT scans is viable but careful consideration must be taken when choosing a DICOM viewing software, especially when assessing small bony structures or areas composed of a thin layer of compact bone (Decker et al. 2011; Bertoglio et al. 2020).

Questioning ancestry estimation

The field's recent discourse questioning the effectiveness of ancestry estimation and highlighting issues that may arise from its inclusion in the biological profile raises legitimate concerns. In their letter to the editor of the *Journal of Forensic Sciences*, Bethard and DiGangi (2020) discuss the uncertainty surrounding how well extant statistical methods reflect the contemporary population structure of the United States. This concern is further strengthened, as it is noted that there is currently no existing means to evaluate how extant methods interpret intersectional identities. Ancestry estimation methods employed by forensic anthropologists tend to produce results that correspond to one social race; assigning a singular social race to an unknown individual can be problematic as a multitude of sociocultural factors influence the social perception of race and ethnicity and cannot be described with skeletal remains alone. Additionally, the majority of the morphoscopic traits used to create an ancestry estimate do not have associated studies to provide heritability data, nor do they have a corresponding craniometric measurement. Without the backing of heritability data or craniometrics,

nonmetric trait analysis for ancestry estimation continues to stand on its typological past (DiGangi and Bethard 2021).

Previous attempts to assess the efficiency of forensic anthropologically produced ancestry estimates against known decedents cases have been criticized as circular logic (Thomas et al. 2017; DiGangi and Bethard 2021), as the likelihood of correctly identifying an unknown individual increases with an accurate biological profile. Further, such a study does not have the capacity to evaluate how incorrect ancestry estimates affect identification efforts in unresolved cases nor can they address the potential influence of ancestry estimates resulting in "unknown" or multiple ancestry possibilities (DiGangi and Bethard 2021). Reliance on the four main geographically derived ancestral populations as potential ancestry estimation results continues to be characteristic of nonmetric and metrical methods alike, though the boundaries dividing these groups have been criticized as being linked to colonialism and geopolitics (Fuentes 2020; Tsai et al. 2020).

DiGangi and Bethard (2020; 2021) suggest removing ancestry estimation as a facet of the biological profile, listing not only deficiencies in the methodologies but also ancestry estimation's ties to forensic anthropology's racist past. The authors acknowledge that discontinuing ancestry estimation practices would not be without logistical complications but stress that the field has not done enough work to address the connection between ancestry estimation and the biological race concept and state that contributing to a scientifically racist paradigm is incompatible with the overarching goals of the field.

In a response to Bethard and DiGangi (2020), Stull et al. (2021) argued that forensic anthropologists accept that race is a social construct and are highly sensitive to the negative impact the field has had on the public's perception of race and how it has been used to reinforce concepts of racial hierarchies in the past. The authors also acknowledge that science can reflect the preconceptions of its practitioners, stating that these biases must constantly be recognized and addressed to mitigate their influence on the field. Further, Stull et al. (2021) contend that the field has made efforts to move beyond its racist past, as current research focuses on more nuanced patterns of human biological variation and statistically based analyses.

Stull et al. (2021) continue to stress that while the field understands that the extent of human variation cannot be accurately discussed using these labels, an individual's self-identified and perceived social race does affect all facets of their life. These labels are not only used by the public but, are also included as an identifying factor of missing persons profiles produced by law enforcement agencies. In their efforts to aid in identification processes, forensic anthropologists relay the results of skeletal assessments as probable social race groups. The authors question Bethard and DiGangi's (2020) suggestion to eliminate ancestry estimation from the biological profile, pointing out that some techniques utilized by forensic anthropologists are population-specific and are meant to be implemented after ancestry has been estimated. Furthermore, while DiGangi and Bethard (2020) specifically condemn ancestry estimation methods that assess cranial morphology due to their typological origins, the authors reiterate that early anthropological goals were largely typological and methods born of these pursuits,

including morphological and metrical assessments of the cranium, were originally used to reinforce racial hierarchies.

Stull et al. (2021) state that addressing systematic issues, like Missing White Woman Syndrome (Sommers 2016; Slakoff and Fradella 2019), requires collaborative efforts from multiple communities, including medicolegal and law enforcement agencies. The dogmas and policies that currently enable biases to influence medicolegal casework are too deep-seated to be mitigated by the elimination of ancestry estimation alone. The authors posit that efforts are currently being made by many forensic anthropologists to provide continuing education to medicolegal groups and engage with the public to address these issues. The authors acknowledge that there is still more work to be done regarding the field's typological origins and its past handling of ancestry research topics but feel that eliminating ancestry estimation methods is not the solution to these problems. Stull et al. (2021) outline a path forward, calling upon practitioners of forensic anthropology to continue to address biases, educate other communities on the complexities of human biological variation and population histories, reject research based on race science, and continue to openly discuss the relationship between biological anthropology and systemic racism.

CHAPTER THREE

Materials and Methods

This chapter describes the skeletal samples, data collection methods, and statistical analyses used in the present research. The composition of the skeletal samples is discussed first, followed by a brief summary of the DICOM viewing program used to create the 3D renderings and the nonmetric scoring methods. Lastly, the statistical analyses used to address the hypotheses outlined in Chapter One are discussed.

Pre-Contact sample

To facilitate the exploration of nonmetric trait variation in pre-Contact and modern Indigenous peoples, 24 traits were examined in two geographically similar yet temporally distant samples. Between the two samples, trait data were collected from a total of 250 individuals. The data for the pre-Contact sample were generously provided by Atkinson and Tallman (2019), whose research involved the scoring of nonmetric traits in 150 individuals curated at the American Museum of Natural History (AMNH) in New York. The AMNH osteological collection is comprised of both archaeological and contemporary materials, representing more than 12,000 individuals recovered from over 50 countries (AMNH 2022). In their research, Atkinson and Tallman (2019) chose individuals that were recovered from archaeological sites located in the Four Corners region of the American Southwest; the sample chosen was broadly representative of Ancestral Puebloans.

For the purpose of their research, the American Southwest was defined as an area encompassing New Mexico, Colorado, Utah, and Arizona (Atkinson and Tallman 2019). To assemble a sample representing Ancestral Puebloans, Atkinson and Tallman (2019) selected individuals recovered from the Pueblo Bonito, Pueblo San Cristobal, Aztec Ruins, Mitten Rock, La Plata Valley, Grand Gulch, Canyon del Muerto, and Canyon de Chelley archaeological sites. These archaeological sites temporally range between the Early Basketmaker (1500 BC – 50 AD) and Pueblo III (1150-1350 AD) time periods. Though the individuals chosen are defined as "pre-Contact," it is important to note that the sample is not homogenous; it is comprised of temporally, culturally, and tribally diverse individuals. The pre-Contact sample presented here should not be taken to be representative of all Indigenous peoples that lived in the pre-Contact American Southwest. This sample was selected as the pre-Contact sample in the present study for two reasons: the individuals included in the sample generally share a geographic origin with the modern sample and the data collected by Atkinson and Tallman (2019) includes scores for all 24 traits considered in this study.

Modern sample

The modern sample consisted of 100 individuals whose nonmetric traits were scored via 3D images generated from CT scans. The scans were obtained from the New Mexico Decedent Image Database (NMDID), which is associated with both the New Mexico Office of the Medical Investigator (OMI) and the University of New Mexico

(UNM). In 2010, the OMI received a National Institute for Justice (NIJ) grant to fund the incorporation of CT scans into the investigation of cases that were suspected to have certain causes of death, including blunt-force trauma, gunshot wounds, drug intoxication, and childhood trauma (Daneshvari Berry and Edgar 2017, 2019). The collaborative efforts of the OMI and UNM's departments of pathology and radiology departments created the Center for Forensic Imaging, where all the forensic imagery in NMDID is generated. Upon determining that whole body scanning of decedents was valuable to the death investigation process, the practice was made permanent at the OMI. Forensic imagery continues to be collected for approximately 90% of all decedents processed by the OMI, amounting to over 2500 cases per year. The database stores all images produced alongside the associated metadata logged by the OMI their investigation. To facilitate research, the database has been modified to make the data accessible through a free, online application process. All images and metadata available for use in research have been anonymized (NMDID 2022).

To select individuals for the modern sample, demographic parameters were established, and a search of the available metadata returned cases that met those criteria. The search criteria stipulated that each individual be aged between 18 and 99 years, have been assigned "Native American" as their race, an ethnicity of "Not Hispanic, Latino, or Middle Eastern," and a Natural manner of death. Individuals under the age of 18 were not included in the sample to avoid the possibility of individuals whose death was related to genetic conditions and childhood trauma. Given the restrictions of the search parameters, the individuals in the sample were roughly estimated to have been born

between 1910 and 2005. This range is approximate due to the unidentified status of some individuals. Additionally, any individuals who were found to have sustained extensive injuries or defects to the facial bones were also excluded to avoid scoring complications.

Of the cases that met the criteria summarized above, 100 individuals were selected to be scored: 50 males and 50 females. It is important to note that the racial, ethnic, and age categorization is based on data collected by the OMI; this information was determined by decedent identification, Next of Kin, or OMI estimation. Once the research application had been approved, NMDID provided access to the requested case files in Digital Imaging and Communications in Medicine (DICOM) file form. DICOM files are the international standard for storing and transmitting digitized medical imagery (Medixant 2021).

Data collection software

DICOM files are transmitted in compressed stacks of 2D images. To reconfigure the DICOM files into a viable format, DICOM viewing software was required; Medixant's RadiAnt program was selected. RadiAnt DICOM viewer, version 2021.1 (64-bit) is a PACS DICOM viewer that was originally published for use by medical professionals but has since been expanded for a wider audience (Medixant 2021). The program allows the user to view DICOM files individually, as a series, or render them into a 3D model. The 3D volume rendering feature can reconstruct a stack of multiplanar CT slices into a cohesive model. The quality of the 3D model depends on the number of

scan slices produced for each area, spatial resolution, image contrast and noise, and the presence of surgical, medical, or dental artifacts (Medixant 2021). Cases that produced models that did not have the detail necessary to accurately observe multiple traits were not scored and removed from the study.

The RadiAnt software package has preprogrammed 3D volume rendering presets, enabling the user to easily toggle between layers (Angio, Airways, Bones and Skin (1-3), Bones B/W, and Skin B/W) and model orientation. The Angio and Bones B/W presets were the most relevant to the present research (Medixant 2021). Observing the mandibular traits required the use of the scalpel tool, which can be used to delineate areas to remove from the 3D rendering. The scalpel tool was also used to view traits that were difficult to observe and remove portions of the reconstruction that were unnecessarily reproduced (i.e., ossified cartilage, vascular structures, and artifacts). The measurement tool aided in the observation of traits with character state definitions that include metrical ranges, which include malar tubercle (MT), nasal bone contour (NBC), posterior zygomatic tubercle (PZT), and zygomatic projection (ZP). The ability to fluidly pan, zoom, and reorient the model in the 3D space relied heavily on the computer's processing capabilities.

Nonmetric traits

Data were collected for a total of 24 nonmetric traits: 16 cranial and 8 mandibular. A complete list of traits and scoring systems can be found in Table 1. Initially, a total of

35 traits were to be scored, but 11 were not consistently reproduced in the 3D model and were removed from the study. Each trait was scored in accordance with the associated trait definitions, character state descriptions, illustrations, and photographs (Rhine 1990; Berg 2008; Hefner 2009; Hefner and Linde 2018; Catalado-Ramirez 2020). Data were incomplete for individuals with damaged or missing bones that impeded trait scoring; individuals missing data for more than three traits were omitted. Data collection was performed in two stages: the primary data collection phase and a secondary rescoring event. During the primary data collection phase, each case was assigned an Excel catalogue number and scored once. The catalogue numbers were later used to randomly select cases for rescoring. The randomly selected sample subset (n=10) was rescored after the initial data collection phase had been completed. The data collected during the rescoring event was used to explore the intraobserver error of the modern sample (see Atkinson and Tallman [2019] for intraobserver error rates for the pre-Contact sample).

Table 1. Nonmetric traits and scoring systems used in the present study.

Trait	Ordinal Scoring System
Anterior nasal spine (ANS)	1 = slight; 2 = intermediate; 3 = marked
Inferior nasal aperture (INA)	1 = smooth transition; 2 = sloping more anteriorly, angular transition; 3 = right angle, no sloping or sill; 4 = weak vertical ridge; 5 = pronounced ridge
Interorbital breadth (IOB)	1 = narrow; 2 = medium; 3 = broad

Malar tubercle form (MTF)	0 = no projection; 1 = trace tubercle; 2 = medium protrusion; 3 = pronounced
Nasal aperture shape (NAS)	1 = teardrop, lateral projection intermediate; 2 = bell shape, greatest lateral project at inferior margin; 3 = bowed, greatest projection at midline
Nasal bone shape (NBS)	1 = no nasal pinching; 2 = superior pinch, minimal lateral bulging; 3 = superior pinch and pronounced lateral bulging; 4 = triangular-shaped
Nasal aperture width (NAW)	1 = narrow; 2 = medium; 3 = broad
Nasal bone contour (NBC)	0 = low, rounded, Quonset-hut; 1 = oval, elongated, rounded lateral walls; 2 = steep walls, broad superior surface, plateau; 3 = steep walls, narrow superior surface; 4 = triangular cross section, no plateau
Nasal overgrowth (NO)	0 = absent; 1 = present
Postbregmatic depression (PBD)	0 = absent; 1 = present
Keeling (KLN)	0 = absent; 1 = slight; 2 = marked
Orbit shape (OS)	1 = rounded; 2 = rectangular; 3 = rhombic
Zygomatic projection (ZP)	1 = retreating; 2 = vertical; 3 = significant projection
Posterior zygomatic tubercle (PZT)	0 = no projection; 1 = weak; 2 = moderate; 3 = marked

Prognathism (PRG)	0 = absent; 1 = moderate; 2 = marked
Dental arcade shape (DAS)	1 = narrow and tapering (parabolic); 2 = wider and smoothly curving (elliptic); 3 = near rectangular (hyperbolic)
Gonial eversion (GE)	0 = absent; 1 = slight; 2 = moderate; 3 = significant flare
Mandibular border shape (MBS)	1 = straight; 2 = undulating (concave curvature); 3 = rocker (convex curvature)
Gonial muscle attachment ridging (GMR)	0 = flat, smooth; 1 = slight; 2 = moderate; 3 = extreme robusticity
Accessory mandibular foramen (AMF)	0 = absent; 1 = present
Chin projection (CP)	0 = retreats; 1 = blunt/vertical; 2 = prominent
Chin shape (CS)	1 = round; 2 = square; 3 = pointed
Mandibular foramen (MF)	0 = absent; 1 = one; 2 = two; 3 = three
Ascending ramus shape (ARS)	1 = pinched; 2 = wide

For the purpose of this research, "nonmetric," or morphoscopic, traits are defined as skeletal attributes that are assessed on an ordinal scale, as opposed to metrical methods. These scoring systems use either ordinal ranges or binomial values to evaluate the form or relative size of these characteristics. Each possible score, or character state,

has an associated description and visual reference. Of the 24 traits observed, 13 have multiple visual references, including illustrations and photographs, available to aid score reliability between observers (Hefner and Linde 2018). Because the established methods of determining trait scores were designed for tangible skeletal remains, and therefore not optimal for digital models, some adjustments were necessary. Hefner and Linde's (2018) directions for scoring an individual's NBC require the use of a contour gauge, which cannot be used on the digital model. To assess NBC, the model was oriented to the Frankfurt plane and the scalpel tool was used to create a vertical slice at the point the contour gauge is placed - approximately 1 cm below nasion (Hefner 2009; Hefner and Linde 2018). The superior surface of the remaining bone was then scored according to the character state descriptions. Similarly, the scalpel tool was utilized while scoring INA, as creating a cross-section of the nasal cavity allowed for better visualization of the inferior border. Additionally, the measurement tool was used to delineate the angle used to assess ZP (Cataldo-Ramirez 2020) and measure the extent of the bony projections observed for MT and PZT (Hefner and Linde 2018).

Bilaterally expressed traits were primarily scored on the left side unless the trait description included instructions to consider both or if the left side was deemed unobservable due to missing or damaged bone. Additional issues that impeded scoring included bone resorption, pathological changes, and the presence of medical appliances. Scores were recorded and input into an Excel spreadsheet. Inconsistencies between the scoring systems used by Atkinson and Tallman (2019) and those standardized in Hefner and Linde (2018) affected two traits (OS and DAS). To address this, OS and DAS were

initially scored using Hefner and Linde's (2018) scoring systems and later converted to those used in Atkinson and Tallman (2019) in a separate Excel spreadsheet using character state descriptions. Figures 1-22 demonstrate the location of each nonmetric trait and how they are replicated in the RadiAnt program.

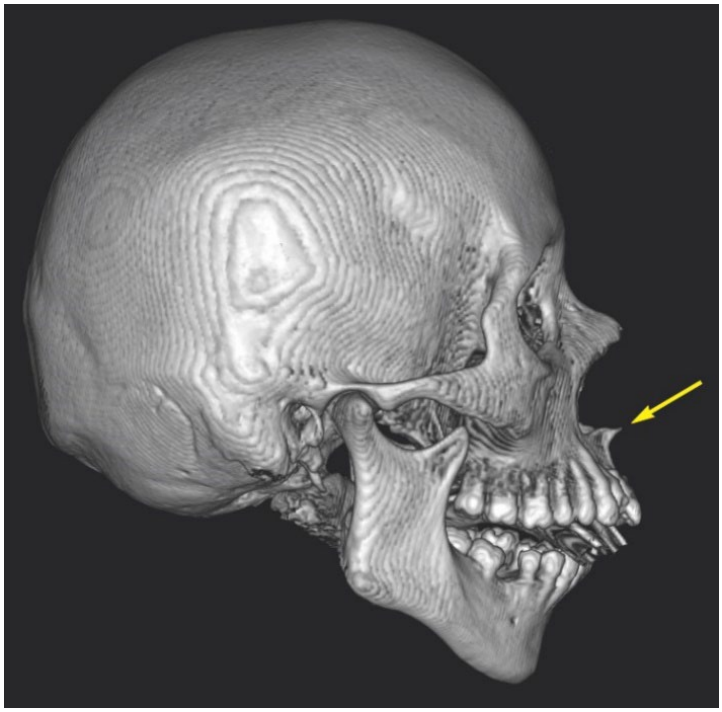


Figure 1. ANS as reproduced in RadiAnt. The yellow arrow indicates anterior nasal spine location.

Figure 2. INA as reproduced in RadiAnt. The red arrows indicate inferior nasal aperture scoring location.

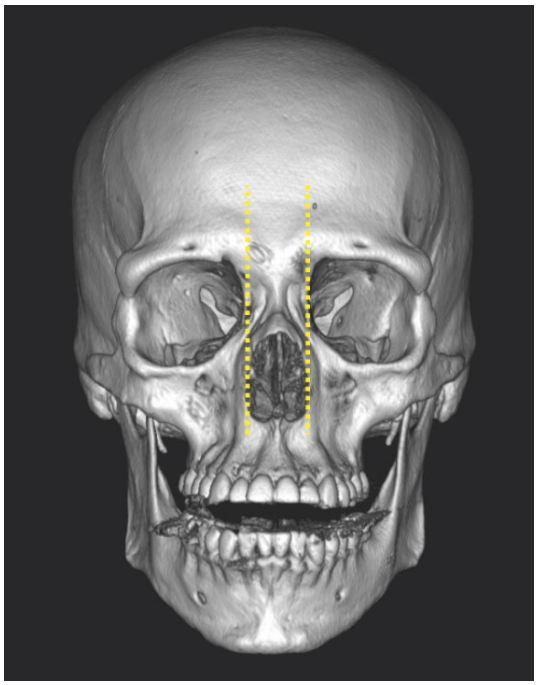


Figure 3. IOB as reproduced in RadiAnt. The yellow dotted lines delineate the interorbital breadth.

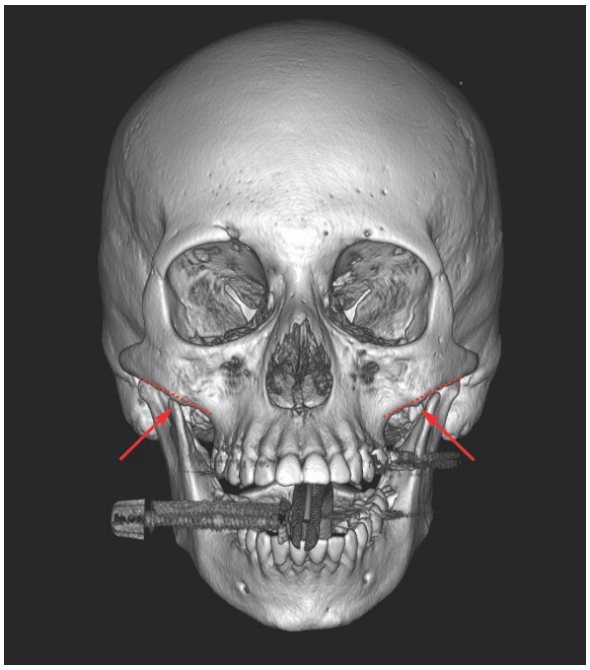


Figure 4. MTF as reproduced in RadiAnt. The red arrows indicate the malar tubercle form scoring location.

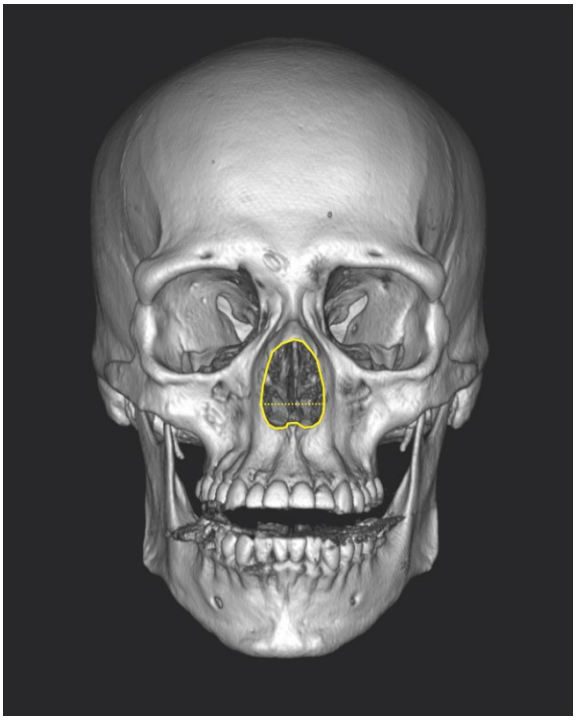


Figure 5. NAS as reproduced in RadiAnt. The nasal aperture is outlined in yellow and the widest point is indicated by the yellow dotted line.

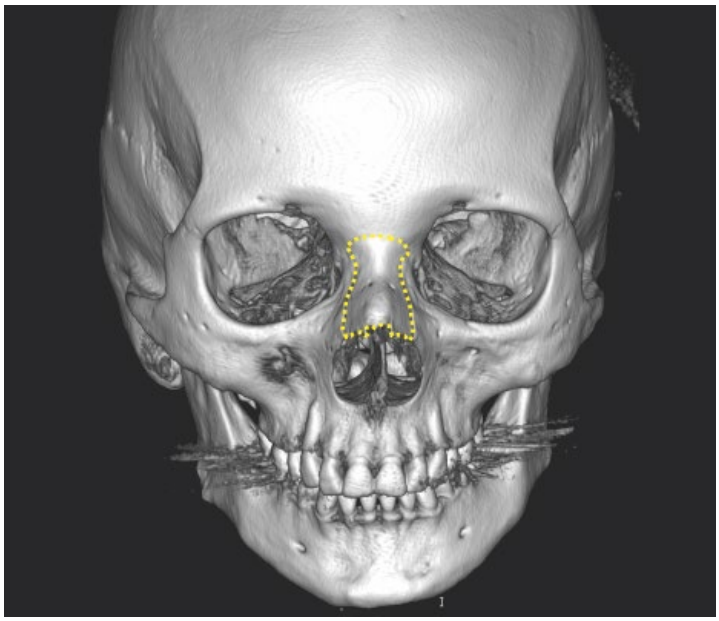


Figure 6. NBS as reproduced in RadiAnt. The nasal bone shape is delineated by the yellow dotted line.

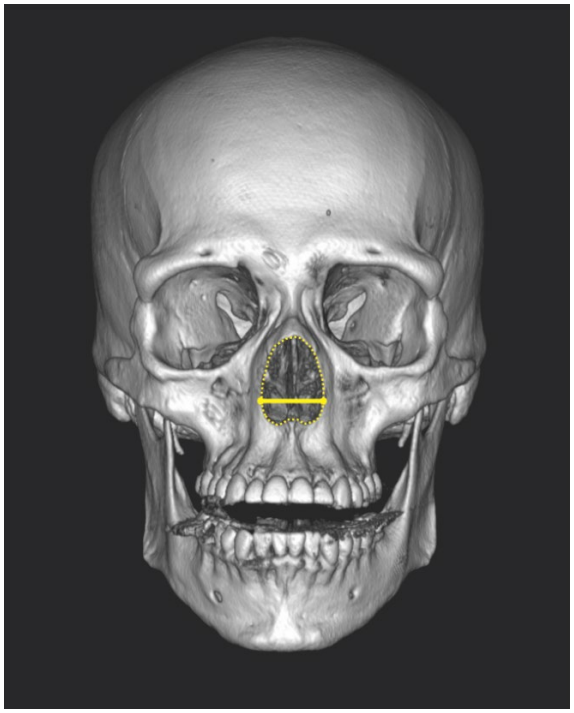


Figure 7. NAW as reproduced in RadiAnt. The nasal aperture is outlined by the yellow dotted line and the widest point is indicated by the solid yellow line.

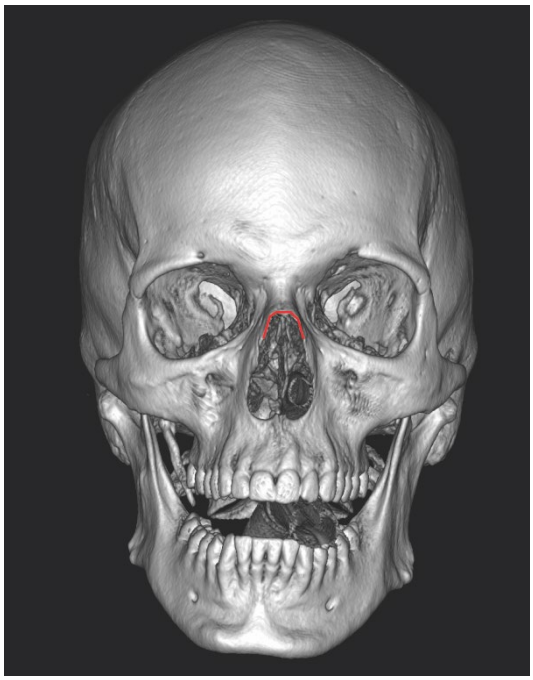


Figure 8. NBC as reproduced in RadiAnt. The nasal bone contour is outlined in red.

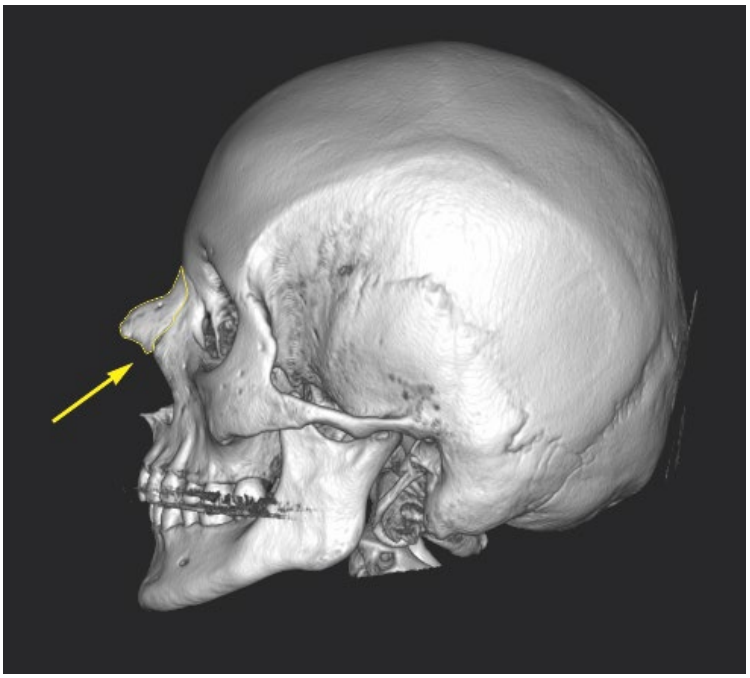


Figure 9. NO as reproduced in RadiAnt. The nasal bone is outlined by the yellow dotted line and the yellow arrow indicates nasal overgrowth scoring location.

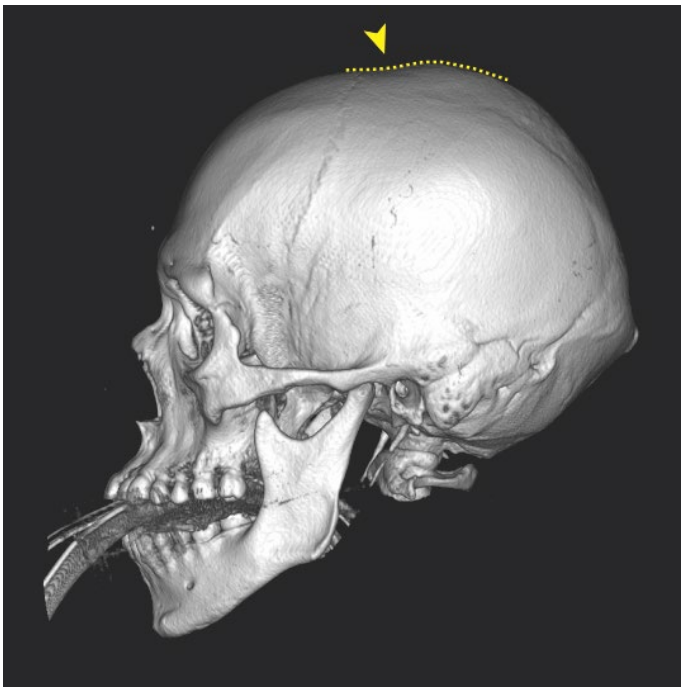


Figure 10. PBD as reproduced in RadiAnt. The postbregmatic depression is indicated by the yellow arrow.

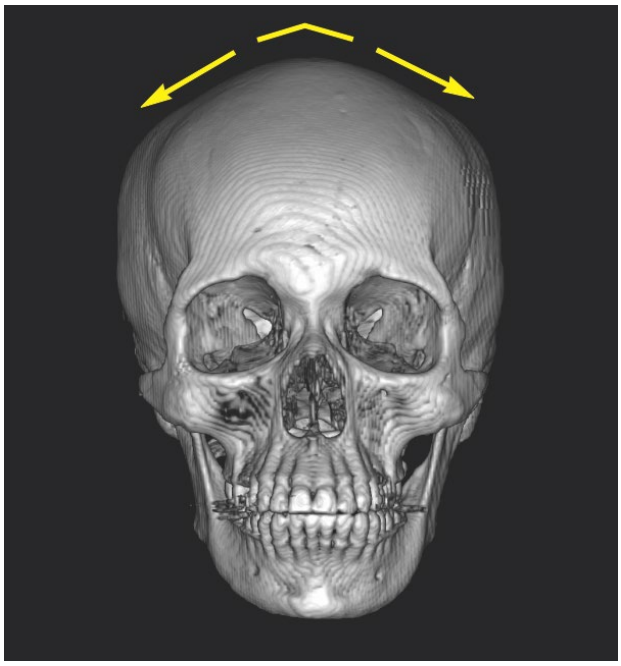


Figure 11. KLN as reproduced in RadiAnt. The shape of keeling is indicated by the yellow arrows.

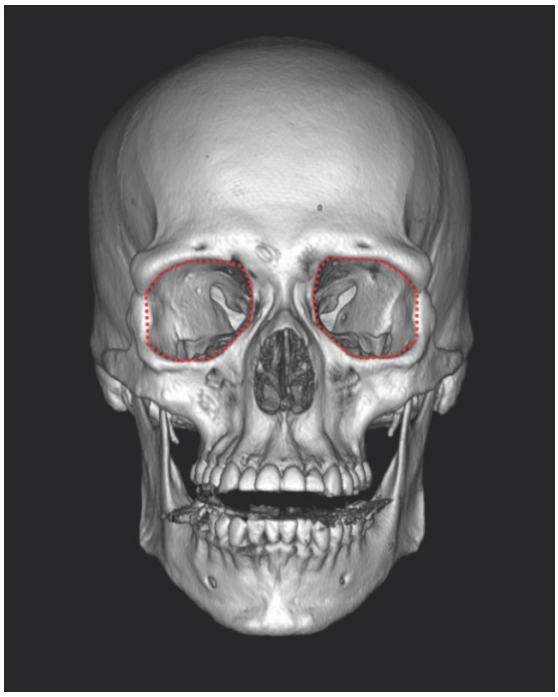


Figure 12. OS as reproduced in RadiAnt. The orbit shape is indicated by the red dotted lines.

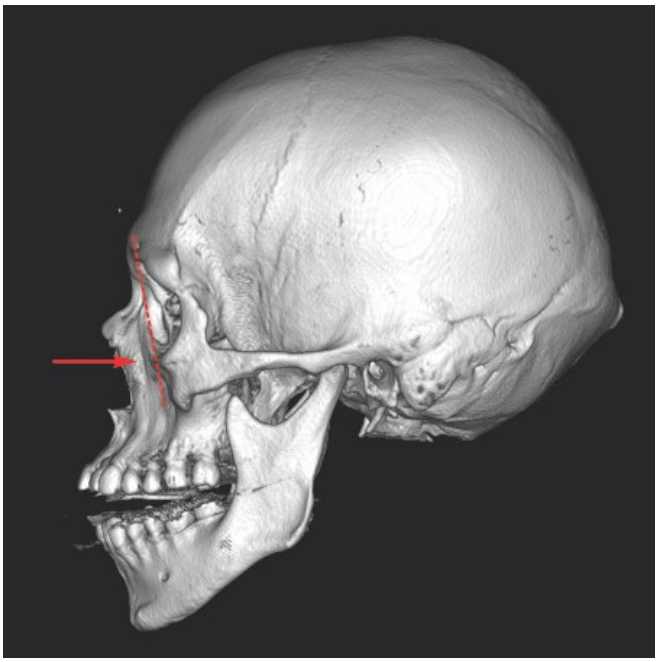


Figure 13. ZP as reproduced in RadiAnt. The zygomatic projection scoring location is indicated by the red arrow.

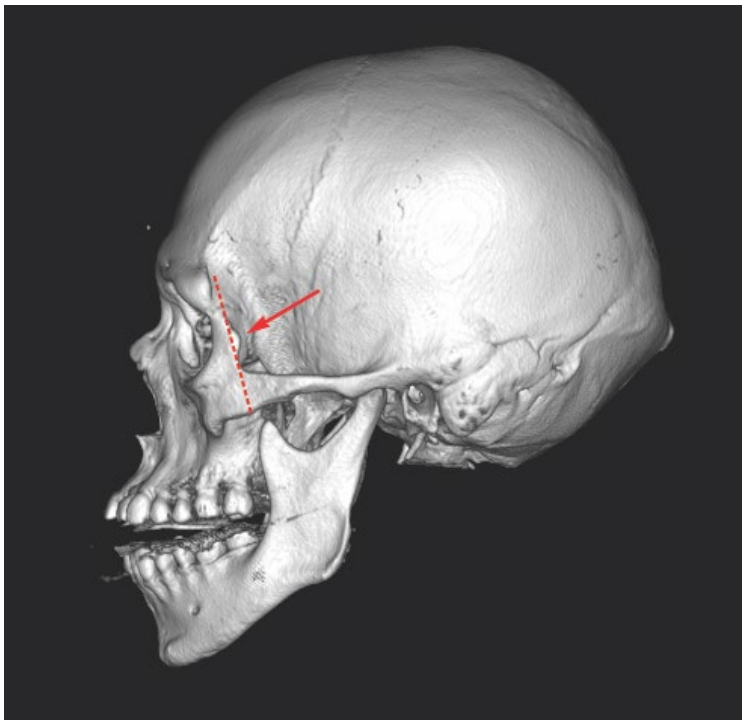


Figure 14. PZT as reproduced in RadiAnt. The posterior zygomatic tubercle is indicated by the red arrow.

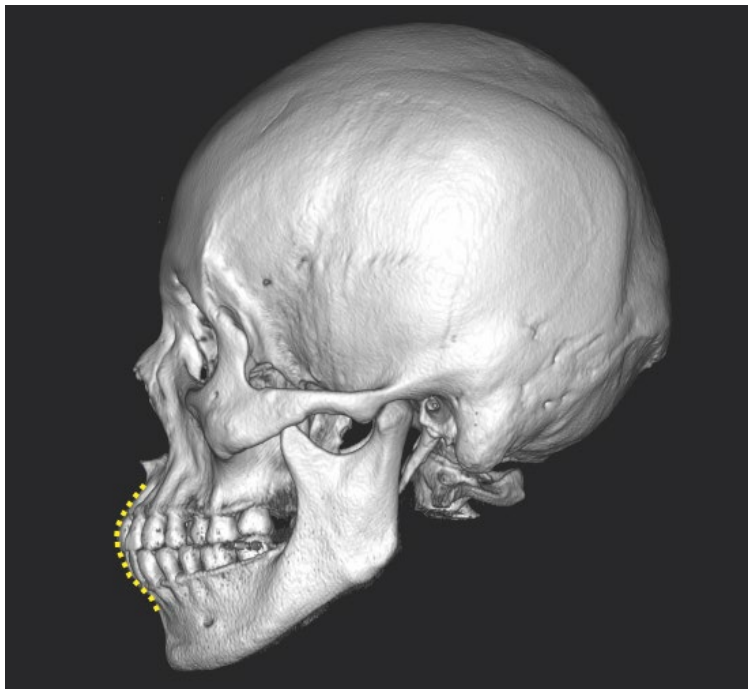


Figure 15. PRG as reproduced in RadiAnt. Prognathism is outlined by the yellow dotted line.

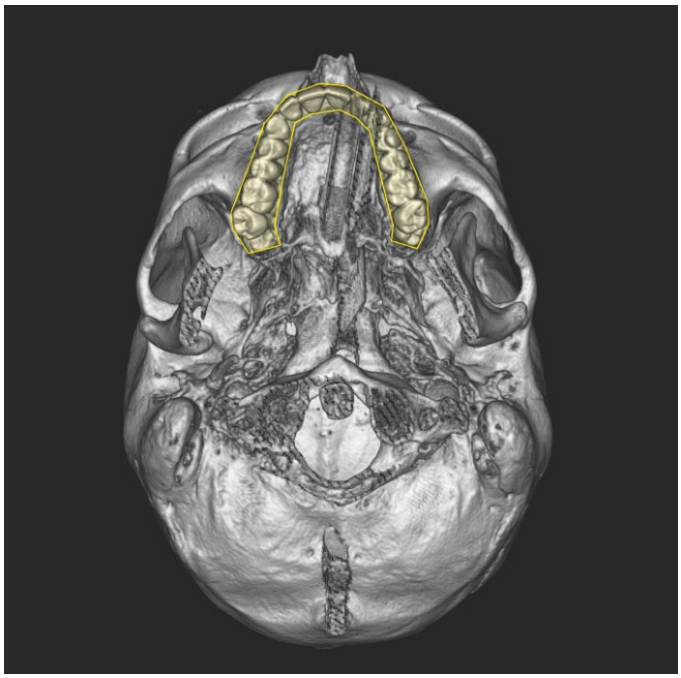


Figure 16. DAS as reproduced in RadiAnt. The dental arcade shape is outlined and highlighted in yellow.

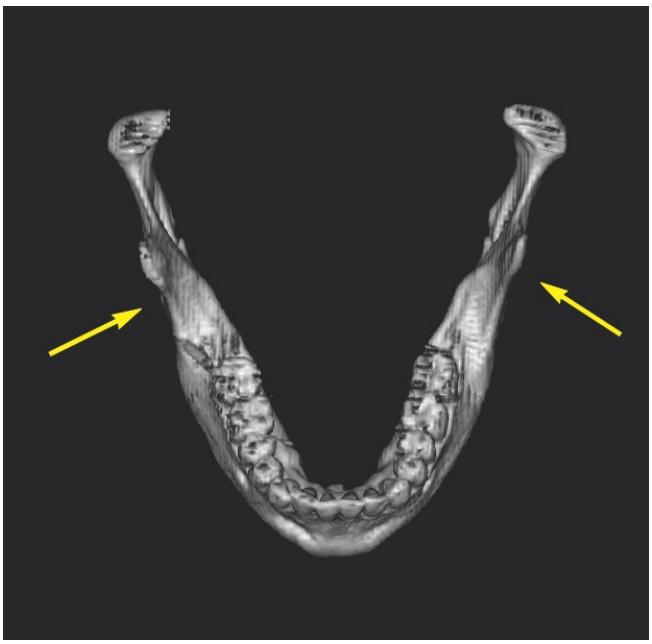


Figure 17. GE as reproduced in RadiAnt. The yellow arrows indicate the scoring location for gonial eversion.

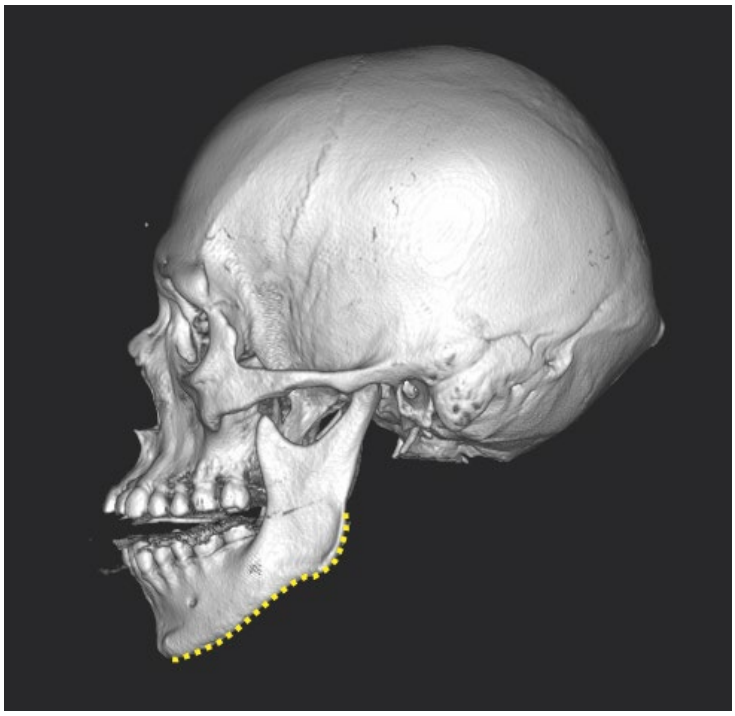


Figure 18. MBS as reproduced in RadiAnt. The mandibular border shape is depicted by the yellow dotted line.

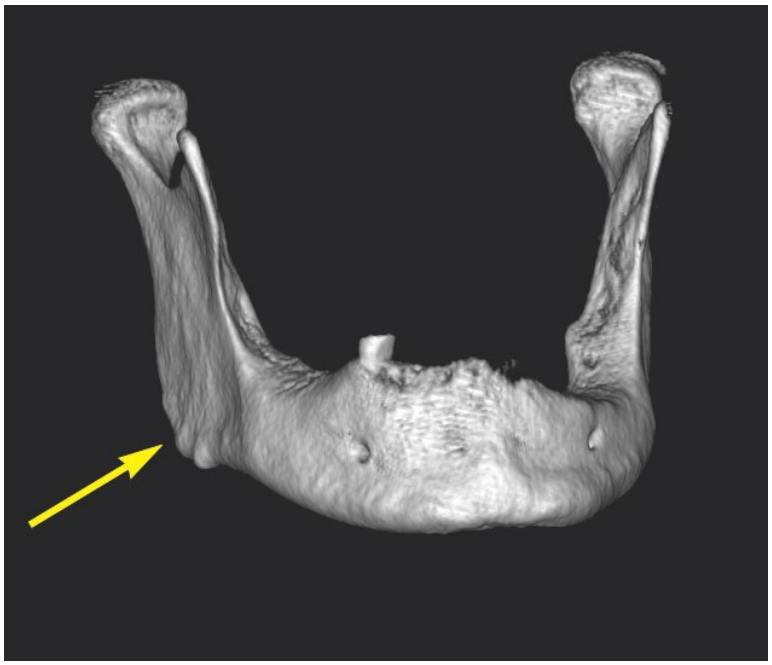


Figure 19. GMR as reproduced in RadiAnt. The arrow indicates the location of gonial attachment ridging.

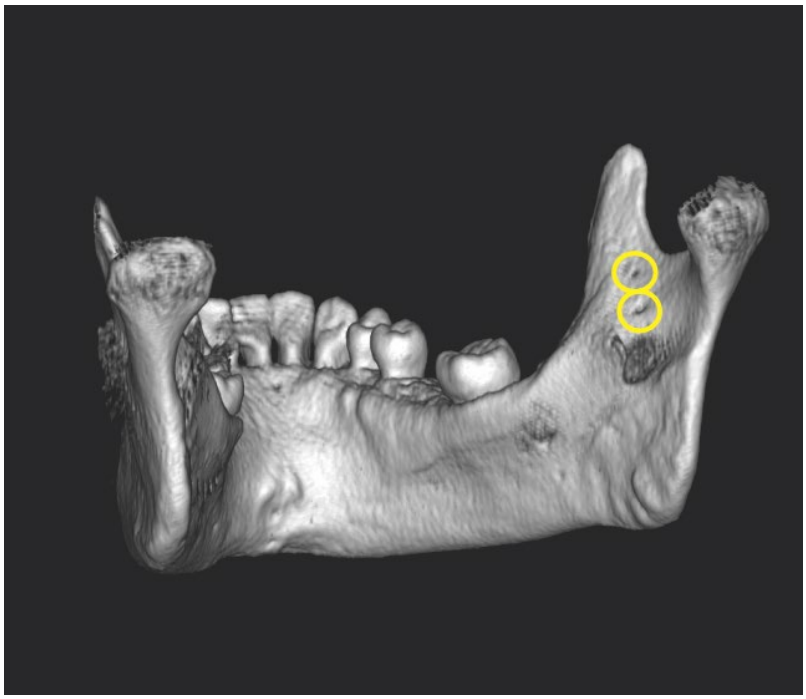


Figure 20. AMF as reproduced in RadiAnt. The accessory mandibular foramina are indicated by the yellow circles.

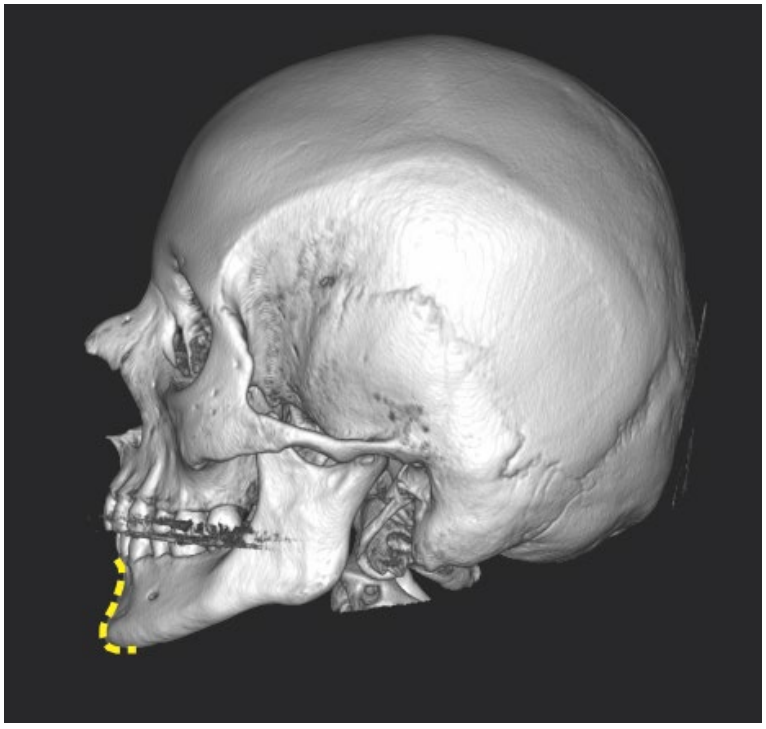


Figure 21. CP as reproduced in RadiAnt. The chin profile is outlined by the yellow dashed line.

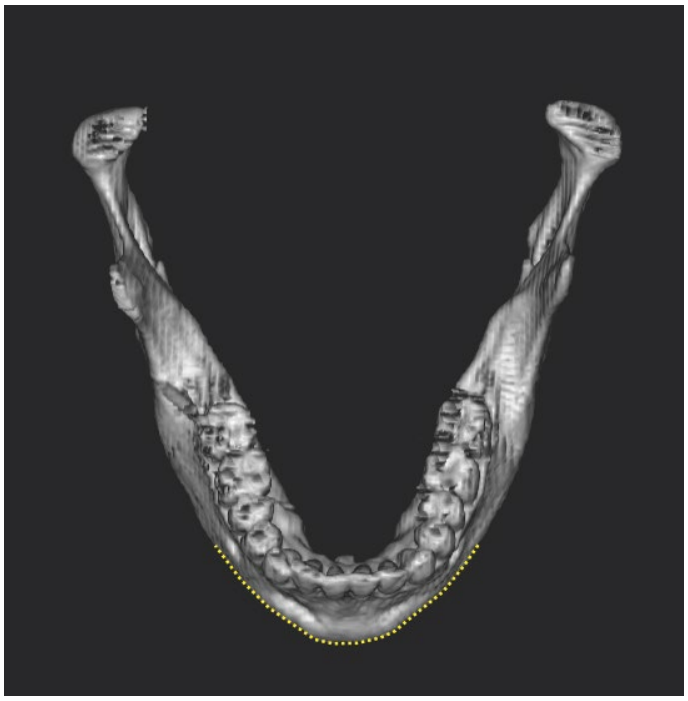


Figure 22. CS as reproduced in RadiAnt. The chin shape is outlined by the yellow dotted line.

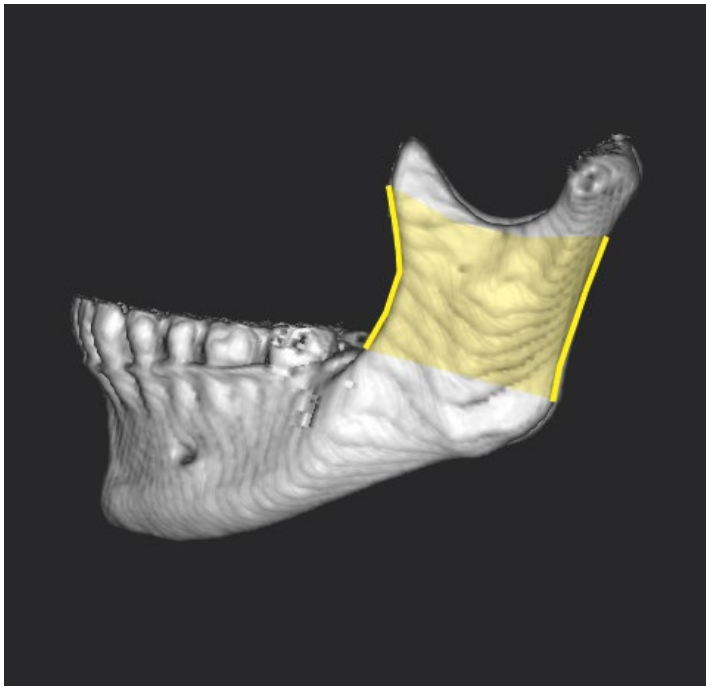


Figure 23. ARS as reproduced in RadiAnt. The ascending ramus shape is outlined in yellow.

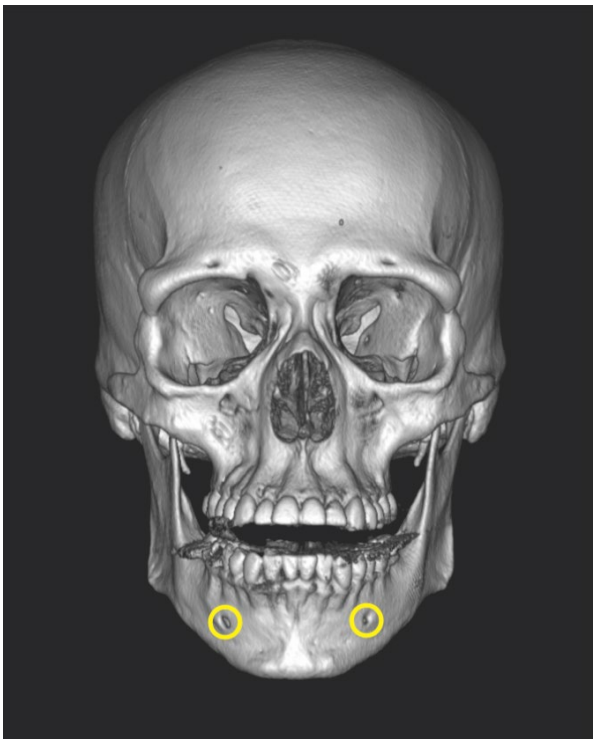


Figure 24. MF as reproduced in RadiAnt. Mandibular foramina are indicated by the yellow circles.

Statistical methods

The statistical analysis of the data was performed using the IBM SPSS Statistics 27 (version 27.0.1) program (IBM Corp. 2020). Because the data gathered was ordinal, it was analyzed using methods that summarize trait frequencies and the differences between the sample groups; the analyses performed included Chi-square analyses, binary logistic regression equations, and Cohen's kappa statistic. Chi-square analyses were used to determine if there were statistically significant differences in trait expression between the two samples. The pre-Contact trait data was used as the expected frequency and the modern data acted as the observed frequency variable. In this statistical analysis, the null

hypothesis states that there is no difference in nonmetric trait expression between the two groups or that there is no association between group membership and trait morphology. An alpha value of 0.05 was used to determine significance; a significance value (*p*-value) less than 0.05 indicated that the trait's presentation was affected by group membership. Missing trait data was excluded pairwise. Outliers were not excluded from the analysis, as the present research sought to explore the extent of variation between the samples.

After determining which traits had statistically significant expression differences between groups, binary logistic regression equations were created to predict the likelihood that an individual would be classified into either the pre-Contact or modern group. Using the variables in the equation, it is possible to infer which trait expressions were most likely to predict membership into each group. Four models were created using the backward Wald function in SPSS; Model 1 used the combined cranial and mandibular traits, Model 2 was developed using seven of Hefner's (2009) macromorphoscopic traits, Model 3 used only cranial traits, and Model 4 included only mandibular traits. A cutoff point of 0.50 was used to determine the individual's group membership, where scores below 0.50 indicated membership in the pre-Contact group (AMNH sample) and scores above 0.50 were indicative of the membership in the modern sample (NMDID sample).

Cohen's kappa statistic (k) was used to explore intraobserver agreement between two scoring events. Landis and Koch (1977) described ranges used to determine the degree of agreement; the ranges are defined as poor agreement (<0.0); slight agreement (0.0-0.2); fair agreement (0.21-0.40); moderate agreement (0.41-0.60); substantial agreement (0.61-0.80); and almost-to-perfect agreement (0.81-1.0) (Landis and Koch

1977). A randomly selected subset representing 10% of the modern sample (n=10) was rescored two weeks after the initial data collection phase was completed.

hefneR ancestry estimation

To further explore how current ancestry estimation analyses classify modern Native Americans, a randomized subset (n=20) of the modern sample was selected, and their scores were input into the Osteomics hefneR program (d'Oliveira Coelho et al. 2020; Osteomics 2023). Osteomics is a freely available web platform that hosts software created to aid in biological profile estimation. The statistical tools were developed and published by David Senhora Navega and João d'Oliveira Coelho; the hefneR algorithm uses a simple Bayesian classifier that uses nonmetric trait scores to estimate ancestry. The algorithm allows the user to input trait scores for the 11 macromorphoscopic traits outlined by Hefner (2009), which is then compared to the population trait data from the same study. Using the entered data, hefneR outputs an individual's probability of congruence with four possible sample groups: African, Asian, European, and Native American. The sample group that the input trait scores are most similar to produces the highest probability output (d'Oliveira Coelho et al. 2020; Osteomics 2020). Three of the traits (SPS, TPS, and ZS) used in this analysis were removed from the study due to their inconsistent reproduction in the 3D model; the ancestry estimations were performed using 8 of the 11 possible traits. Missing data may negatively affect the accuracy of

Bayesian classifier probabilities; to mitigate this effect, only cases with data for the 8 observable traits were analyzed.

CHAPTER FOUR

Results

Frequency distributions and chi-square analyses

Frequency distributions were calculated for each of the 24 observed nonmetric traits and their corresponding character states. The number of individuals observed and the associated trait scores are reflected in Table 2. Chi-square analyses were used to evaluate statistically significant difference in trait expression between the AMNH and NMDID samples. Pearson's chi-square values and significance values are presented in Table 3; an alpha value of 0.05 was used to evaluate if significant change had occurred. Of the 24 traits analyzed, differences were found in the expression of 22 cranial and mandibular traits. Statistically significant change was not observed in two traits, NO and KLN, which are denoted in Table 3 by asterisks.

Table 2. Cranial trait score frequencies (%) for the AMNH and NMDID samples.

			Ordinal Scores					
Trait	Sample	n	0	1	2	3	4	5
ANS	AMNH	149	--	48 (32.2)	78 (52.3)	23 (15.4)	--	--
	NMDID	100	--	12 (12.0)	48 (48.0)	40 (40.0)	--	--
INA	AMNH	150	--	4 (2.7)	40 (26.7)	72 (48.0)	24 (16.0)	10 (6.7)
	NMDID	100	--	9 (9.0)	8 (8.0)	33 (33.0)	38 (38.0)	12 (12.0)
IOB	AMNH	150	--	3 (2.0)	134 (89.3)	13 (8.7)	--	--
	NMDID	100	--	20 (20.0)	61 (61.0)	19 (19.0)	--	--
MTF	AMNH	150	22 (14.7)	75 (50.0)	37 (24.7)	16 (10.7)	--	--
	NMDID	100	34 (34.0)	51 (51.0)	15 (15.0)	0 (0.0)	--	--

NAS	AMNH	150	--	1 (0.7)	138 (92.0)	11 (7.3)	--	--
	NMDID	100	--	70 (70.0)	11 (11.0)	19 (19.0)	--	--
NBS	AMNH	149	--	74 (49.7)	30 (20.1)	45 (30.2)	0 (0.0)	--
	NMDID	90	--	9 (10.0)	47 (52.2)	24 (26.7)	10 (11.1)	--
NAW	AMNH	150	--	3 (2.0)	140 (93.3)	7 (4.7)	--	--
	NMDID	100	--	43 (43.0)	52 (52.0)	5 (5.0)	--	--
NBC	AMNH	145	60 (41.4)	72 (49.7)	7 (4.8)	1 (0.7)	5 (3.4)	--
	NMDID	98	9 (9.2)	33 (33.7)	19 (19.4)	26 (26.5)	11 (11.2)	--
NO	AMNH	136	55 (40.4)	81 (59.6)	--	--	--	--
	NMDID	96	32 (33.3)	64 (66.7)	--	--	--	--
PBD	AMNH	150	150 (100.0)	0 (0.0)	--	--	--	--
	NMDID	99	83 (83.8)	16 (16.2)	--	--	--	--
KLN	AMNH	150	52 (34.7)	77 (51.3)	21 (14.0)	--	--	--
	NMDID	100	37 (37.0)	46 (46.0)	17 (17.0)	--	--	--
OS	AMNH	150	--	13 (8.7)	135 (90.0)	2 (1.3)	--	--
	NMDID	100	--	37 (37.0)	29 (29.0)	34 (34.0)	--	--
ZP	AMNH	150	--	19 (12.7)	16 (10.7)	115 (76.7)	--	--
	NMDID	100	--	60 (60.0)	21 (21.0)	19 (19.0)	--	--
PZT	AMNH	150	16 (10.7)	44 (29.3)	50 (33.3)	40 (26.7)	--	--
	NMDID	100	2 (2.0)	56 (56.0)	34 (34.0)	8 (8.0)	--	--
PRG	AMNH	150	13 (8.7)	131 (87.3)	6 (4.0)	--	--	--
	NMDID	87	21 (24.1)	50 (57.5)	16 (18.4)	--	--	--
DAS	AMNH	150	--	1 (0.7)	149 (99.3)	0 (0.0)	0 (0.0)	--
	NMDID	88	--	41 (46.6)	17 (19.3)	27 (30.7)	3 (3.4)	--
GE	AMNH	149	74 (49.7)	53 (35.6)	21 (14.1)	1 (0.7)	--	--
	NMDID	100	28 (28.0)	37 (37.0)	23 (23.0)	12 (12.0)	--	--
MBS	AMNH	149	--	137 (91.9)	12 (8.1)	0 (0.0)	--	--
	NMDID	99	--	54 (54.5)	42 (42.4)	3 (3.0)	--	--
GMR	AMNH	149	1 (0.7)	40 (26.8)	79 (53.0)	29 (19.5)	--	--
	NMDID	100	6 (6.0)	49 (49.0)	39 (39.0)	6 (6.0)	--	--
AMF	AMNH	148	138 (93.2)	10 (6.8)	--	--	--	--
	NMDID	98	76 (77.6)	22 (22.4)	--	--	--	--
CP	AMNH	149	146 (98.0)	3 (2.0)	0 (0.0)	--	--	--
	NMDID	95	63 (66.3)	31 (32.6)	1 (1.1)	--	--	--

CS	AMNH	149	--	115 (77.2)	1 (0.7)	33 (22.1)	--	--
	NMDID	100	--	45 (45.0)	46 (46.0)	9 (9.0)	--	--
MF	AMNH	149	0 (0.0)	144 (96.6)	5 (3.4)	0 (0.0)	--	--
	NMDID	100	0 (0.0)	0 (0.0)	87 (87.0)	13 (13.0)	--	--
ARS	AMNH	148	--	11 (7.4)	137 (92.6)	--	--	--
	NMDID	100	--	27 (27.0)	73 (73.0)	--	--	--

Table 3. Pearson's chi-square analyses results for traits observed on NMDID sample.

Trait	Value	Significance
ANS	24.642	0.000
INA	32.380	0.000
IOB	32.311	0.000
MTF	23.386	0.000
NAS	174.415	0.000
NBS	60.149	0.000
NAW	68.176	0.000
NBC	76.904	0.000
NO	1.213*	0.271
PBD	25.907	0.000
KLN	0.794*	0.672
OS	102.58	0.000
ZP	84.094	0.000
PZT	27.823	0.000
PRG	27.901	0.000
DAS	97.063	0.000
GE	24.286	0.000
MBS	47.588	0.000
GMR	24.460	0.000
AMF	12.830	0.000
CP	47.391	0.000
CS	80.915	0.000
MF	9.619	0.002
ARS	17.611	0.000

Binary logistic regression

Binary logistic regression formulae are presented in Tables 4-7. Binary logistic regression equations were performed to assess the relationship between the observed traits and group membership regarding the AMNH and NMDID samples. Models were created using the combined cranial and mandibular traits, the six observed traits from Hefner (2009), and cranial and mandibular traits, separately. For all developed models, the AMNH sample was coded as 0, the NMDID sample was coded as 1, and a cut-off value of 0.5 was used to classify observances. An alpha value of 0.05 was used to evaluate significance. Because the chi-square analyses found no significant difference in expression between the AMNH and NMDID samples, NO and KLN were excluded from the analyses.

When the 22 combined cranial and mandibular traits were entered into the binary logistic regression equation (Model 1, Table 4), four traits contributed significantly to predicting group membership and the model correctly classified 95.1% of the sample. Using the seven observed of Hefner's (2009) 11 macromorphoscopic traits (Model 2, Table 5), three traits contributed significantly to the equation and correctly classified 88.4% of the sample. In a model created using only the cranial traits (Model 3, Table 6), six of 12 traits contributed significantly to the equation and correctly classified 93.5% of the sample. Using only the mandibular traits (Model 4, Table 7), four of ten traits contributed significantly to predicting group membership and correctly classified 88.3% of the sample.

Table 4. Binary logistic regression equations using combined cranial and mandibular traits (Model 1).

	Coefficients - Model 1					% Correct		
Group	NAW	NBC	DAS	CP	CON	AMNH	NMDID	Total
AMNH vs. NMDID	-3.752	1.597	-3.831	3.801	7.189	96.4	92.3	95.1

Table 5. Binary logistic regression equations using seven of 11 Hefner (2009) traits (Model 2).

	Coefficients - Model 2				% Correct		
Group	MTF	NAW	NBC	CON	AMNH	NMDID	Total
AMNH vs. NMDID	-1.075	-2.535	0.898	1.732	93.1	81.4	88.4

Table 6. Binary logistic regression equations using only cranial traits (Model 3).

	Coefficients - Model 3							% Correct		
Group	MTF	NAS	NBS	NAW	NBC	ZP	CON	AMNH	NMDID	Total
AMNH vs. NMDID	-1.554	-2.58	1.276	-4.653	0.963	-1.233	7.183	95.8	89.7	93.5

Table 7. Binary logistic regression equations using only mandibular traits (Model 4).

	Coefficients - Model 4					% Correct		
Group	DAS	GE	GMR	CP	CON	AMNH	NMDID	Total
AMNH vs. NMDID	-3.299	1.231	-2.011	3.485	4.431	94.6	76.0	88.3

hefneR ancestry probabilities

The online ancestry classification tool, hefneR, was used to produce probable ancestry classifications for each individual in both the AMNH (n=150) and NMDID (n=100) samples. The number of individuals predicted to most probably be associated with each ancestral group (African, American Indian, Asian, and European) and the accompanying percentages are presented in Table 8. Neither sample resulted in high rates of American Indian classification, with the AMNH sample producing rates (14%) only

slightly higher than the NMDID sample (11%). The majority of the AMNH sample resulted in probable African (45.33%) ancestry while most of the NMDID sample was predicted most likely be European (65.0%); therefore, the null hypothesis is rejected, and the hefneR ancestry estimation algorithm does not reliably predict American Indian ancestry in pre-Contact or modern Indigenous individuals.

Table 8. Ancestry estimation totals (%) of the AMNH and NMDID samples using hefneR.

Sample	n	Ancestry Groups			
		African	American Indian	Asian	European
AMNH	150	68 (45.3)	21 (14.0)	40 (26.7)	21 (14.0)
NMDID	100	16 (16.0)	11 (11.0)	8 (8.0)	65 (65.0)

Intraobserver error analysis

Intraobserver error rates for the observed traits were assessed on 10% of the NMDID sample by the author (n=10). Table 9 presents the k-value and corresponding level of agreement, as described by Landis and Koch (1977). The level of agreement ranges between fair to perfect agreement, with most of traits receiving k-values falling within the moderate to substantial categories. Three traits (MTF, ZP, and CS) produced the lowest levels of agreement, with k-values corresponding with fair agreement, while five (ANS, NBS, NO, PBD, and MF) produced k-values falling into the almost perfect to perfect range.

Table 9. Intraobserver error rates and levels of agreement following Landis and Koch (1977).

Trait	Cohen's k	Agreement
ANS	0.831	Almost Perfect
INA	0.692	Substantial
IOB	0.756	Substantial
MTF	0.400	Fair
NAS	0.737	Substantial
NBS	0.815	Almost Perfect
NAW	0.600	Moderate
NBC	0.733	Substantial
NO	1.000	Perfect
PBD	1.000	Perfect
KLN	0.524	Moderate
OS	0.630	Substantial
ZP	0.318	Fair
PZT	0.615	Substantial
PRG	0.600	Moderate
DAS	0.781	Substantial
GE	0.559	Moderate
MBS	0.615	Substantial
GMR	0.516	Moderate
AMF	0.545	Moderate
CP	0.524	Moderate
CS	0.348	Fair
MF	1.000	Perfect
ARS	0.737	Substantial

CHAPTER FIVE

Discussion

This chapter contextualizes and interprets the data collected, their implications, and the issues encountered during data collection and how they were addressed. The purpose of this study was to explore whether change has occurred in cranial and mandibular trait expression between two temporally distinct Indigenous groups from the same geographic region. Documenting and understanding the extent of potential change in nonmetric trait expression is imperative because it may affect the accuracy of ancestry estimation methods used by forensic anthropologists. Though the use of ancestry estimation methodologies, and specifically morphology-based ancestry estimation, have come into question in recent years (DiGangi and Bethard 2020; DiGangi and Bethard 2021), the techniques continue to be taught in forensic anthropology programs. Because we understand that human populations are not immutable, the techniques used to individuate decedents must be reviewed to ensure their continued relevance and assess accuracy.

Nonmetric methods of ancestry estimation have largely focused on African and European descended populations until recent years (Hurst 2012; Spradley 2013; Tallman 2016; Atkinson and Tallman 2020) but there continues to be a paucity of research regarding North American Indigenous populations. The nonmetric ancestry estimation methods currently employed by forensic anthropologists rely on trait data derived from archaeological samples for Indigenous populations (Rhine 1990, Gill 1998; Hefner 2009; Hefner and Linde 2018). Based on statistical analysis of data collected on 24 cranial and

mandibular nonmetric traits, the results of this research indicate that data from archaeological samples do not result in accurate ancestry estimation for modern Indigenous decedents. For the purpose of this study, the AMNH and NMDID datasets act as proxies for pre-Contact (e.g., ancient) and post-Contact (e.g., modern) era Indigenous peoples of the American Southwest, broadly speaking. While the samples should not be taken to fully encompass the diversity of past and present Indigenous groups that have lived in the American Southwest, these samples can be used to explore differences in trait expression between two temporal periods of the broad Native American group.

Trait frequencies

The 24 assessed nonmetric traits were separated by their location on the skull, cranial or mandibular, for a total of 16 cranial traits and eight mandibular traits. The results of the chi-square analyses indicate that statistically significant change has occurred in 22 (91.7%) of the 24 observed nonmetric traits. Only two cranial traits, NO and KLN, were not shown to have undergone change in morphological expression.

It has long been acknowledged that the Indigenous peoples of North America do not display skeletal homogeneity, as previous anthropologists have been documenting and attempting to parse out variability in nonmetric trait expression between Indigenous populations for decades (Ossenberg 1976; Rhine 1990; Oettle et al. 2017; Cataldo-Ramirez et al. 2019). Microevolutionary forces, such as genetic drift, are a potential cause for trait frequency differences between the AMNH and NMDID samples. Research

indicates that the Indigenous populations of the American Southwest did not live in static, isolated groups. Throughout history, these groups have interacted for various reasons, including trade, migration, and warfare; interactions which affected each population's composition. Evidence of these dynamic group relationships can be seen in material culture, studies of linguistic development, and genetic investigations into pre-Contact admixture (Malhi et al. 2003). Another factor to consider when discussing the Indigenous peoples of the Americas is the rapid decline in population size since the Contact era (Leibmann et al. 2016; Jordan et al. 2019). Though the field of forensic anthropology has not definitively established how much heritability factors into nonmetric trait expression within a population, it is reasonable to assume that significant changes in population composition would affect trait expression frequencies.

Another potential explanation for the disparity between the two samples is that nonmetric trait expression has been affected by large-scale secular change. Jantz and Meadows Jantz's (2000) metrical study of five craniofacial variables in white and Black American individuals indicated that secular change affected the gross craniofacial morphology within a sample spanning 139 years. While this change was mostly witnessed in the cranial vault, it also affected the viscerocranium and, subsequently, the traits examined in this study. The bony structures of the midface were noted to have become higher and narrower overall (Jantz and Meadows Jantz 2000). The first study to document secular change in nonmetric traits, Kilroy et al. (2020), demonstrated statistically significant change in a population's nonmetric trait frequencies has occurred

in as little time as half a century. The study analyzed 23 nonmetric cranial and mandibular traits, 17 of which were also examined in the current study.

Further studies on secular change have indicated that environmental factors, such as shifts in diet and nutrition may contribute to these changes in craniofacial skeletal morphology. For instance, maternal and early childhood nutrition may influence the growth of facial bones, as this portion of the cranium displays plasticity through development. Similarly, external factors regarding diet, such as masticatory stress, have also been found to influence facial and mandibular morphology (Martin and Danforth 2009; Paschetta et al. 2010; Freidline et al. 2015). Kilroy et al.'s (2020) study supports the idea that major sociocultural shifts can affect skeletal morphology, noting that some of the most significant changes in nonmetric trait expression occurred around the turn of the twentieth century, during an era of technological advancement and industrialization. If we consider the magnitude of time that separates the AMNH and NMDID samples, which can be roughly estimated to range between 3,400 to 550 years, it is likely that such secular change has occurred within the Indigenous populations of the American Southwest and that these changes have affected the craniofacial morphology, altering the efficacy of the nonmetric anthropological assessment. The lack of literature on modern Indigenous populations suggests that the extant ancestry estimation methods have not yet addressed the potential impact of secular change or genetic drift on nonmetric traits and continue to rely on archaeological samples as the basis of ancestry estimation in modern Indigenous populations. This is true for both metric and nonmetric ancestry estimation methods, as illustrated by hefneR's lack of a modern indigenous sample and FORDISC's

minute (n=20) modern indigenous cases (Ousley and Jantz 2005; Hefner 2009). There is currently a paucity of literature exploring how potential disparity between comparative sample groups and modern populations affects missing or unknown persons identification efforts.

Binary logistic regression

Binary logistic regression equations were used to determine which traits influence group membership and assess the magnitude of their effect. The binary logistic regression equations presented in the current study are comprised of models created with the combined cranial and mandibular traits (Model 1), the seven observed of Hefner's (2009) 11 macromorphoscopic traits (Model 2), cranial traits (Model 3), and mandibular traits (Model 4). The two traits found to not be statistically significant in distinguishing between the two samples, NO and KLN, were excluded from the binary logistic regression equations.

Model 1 performed the best, correctly predicting 96.4% of the AMNH sample and 92.3% of the NMDID sample, culminating into 95.1% overall. Of the 22 combined cranial and mandibular traits, four traits (NAW, NBC, DAS, and CP) were determined to be statistically significant to predicting group membership. The binary logistic regression equation found that a NAW score of 1 (*narrow*) was most closely associated with membership in the NMDID group and a score of 3 (*broad*) was most likely to predict AMNH group membership, with the likelihood of NMDID decreasing by a factor of

0.023 as the score increased incrementally. The model found that an NBC score of 0 (*low, rounded,* "*Quonset-hut*"; Hefner and Linde 2018) was most likely to predict AMNH group membership and a score of 4 (*triangular cross section, no plateau*; Hefner and Linde 2018) was most closely related to NMDID group membership, with the probability of NMDID group membership increasing by a factor of 4.940 as the score increased by one. A DAS score of 1 (*parabolic*; Hefner and Linde 2018) was found to be most closely associated with membership in the NMDID group, with a score of 3 (*hyperbolic*; Hefner and Linde 2018) being most likely to predict AMNH group membership; the likelihood that an individual would belong to the NMDID group decreased by a factor of 0.022 as the score increased incrementally. A CP score of 0 (*retreats*; Hellman 1928; Turner 1989; Rhine 1990; Atkinson and Tallman 2020) was found to be most closely associated with ANMH group membership and a score of 2 (*prominent*; Hellman 1928; Turner 1989; Rhine 1990; Atkinson and Tallman 2020) being most likely to predict NMDID membership; the probability of NMDID membership increased by a factor of 44.731 as the score increased by one.

Model 2 correctly classified 93.1% of the AMNH sample, 81.4% of the NMDID sample, and 88.4% overall. Using seven of the 11 macromorphoscopic traits outlined in Hefner (2009), three traits (MTF, NAW, and NBC) were found to significantly contribute to group membership predictions. In Model 2, a MTF score of 0 (*no projection*; Hefner and Linde 2018) was most closely associated with NMDID group membership while a score of 3 (*pronounced*; Hefner and Linde 2018) was most likely to predict membership in the AMNH sample; the probability that and individual would belong in the NMDID

group decreased by a factor of 0.341 as the score increased by an increment of one. As in Model 1, Model 2 found that a NAW score of 1 (*narrow*; Hefner and Linde 2018) was most closely associated with NMDID group membership and score of 3 (*broad*; Hefner and Linde 2018) was most likely to predict membership in the AMNH group, with the likelihood of NMDID group membership decreasing by a factor of 0.079 as the score increased by an increment of one. In Model 2, the binary logistic equations also found that an NBC score of 0 (*low, rounded,* "*Quonset-hut*"; Hefner and Linde 2018) was most likely to be associated with AMNH group membership, while a score of 4 (*triangular cross section, no plateau*; Hefner and Linde 2018) was most likely to predict NMDID membership; in Model 2 the probability of NMDID group membership increased by a factor of 2.455 as the score increased by an increment of one.

Model 3 demonstrated correct group classification in 95.8% of the AMNH sample, 89.7% in the NMDID sample, and 93.5% overall. Of the 12 cranial traits analyzed, six were determined to significantly contribute to the binary logistic equations. As in the previous model, an MTF score of 0 (*no projection*; Hefner and Linde 2018) was most likely to predict NMDID group membership while a score of 3 (*pronounced*; Hefner and Linde 2018) was most likely to indicate membership in the AMNH sample; the probability that and individual would belong in the NMDID group decreased by a factor of 0.211 as the score increased by an increment of one. A NAS score of 1 (*teardrop, lateral projection intermediate*; Hefner and Linde 2018) was found to be most closely linked to NMDID group membership and a score of 3 (*bowed, greatest projection at midline*; Hefner and Linde 2018) was most closely related to membership in the AMNH

sample, with an individual's probability of membership in the NMDID group decreasing by a factor of 0.076 as the assigned score increased incrementally. The model found that a NBS score of 1 (*no nasal pinching*; Hefner and Linde 2018) was most likely to predict AMNH group membership and a score of 4 (*triangular-shaped*; Hefner and Linde 2018) was most closely associated with membership in the NMDID sample, while an individual's potential of NMDID group membership increased by a factor of 3.581 as the score increased by one. As in previous models, Model 3 found that a NAW score of 1 (*narrow*; Hefner and Linde 2018) was mostly closely associated with NMDID group membership while a score of 3 (*broad*; Hefner and Linde 2018) was most likely to predict membership in the AMNH sample; the likelihood of an individual's membership in the NMDID group decreased by a factor of 0.01 as the score increased incrementally. Model 3 also found NBC scores of 0 (*low, rounded, "Quonset-hut"*; Hefner and Linde 2018) to be most closely linked to AMNH group membership and scores of 4 (*triangular cross section, no plateau*; Hefner and Linde 2018) most likely to predict NMDID group membership, with the probability of an individual's membership in the NMDID group increasing by a factor of 2.619 as the score increased by one. The model found that a ZP score of 1 (*retreating*; Rhine 1990) was most likely to indicate membership in the NMDID group, while a score of 3 (*significant projection*; Rhine 1990) was most closely associated with membership in the AMNH sample; the probability that an individual would belong in the NMDID group decreased by a factor of 0.291 as the score increased by an increment of one.

Model 4 had the lowest levels of accurate group membership prediction, classifying 94.6% of the AMNH sample correctly, 76.0% of the NMDID, and 88.3% overall. Five mandibular traits (DAS, GE, GMR, CP, and ARS) were determined to significantly contribute to the group membership predictions. It found that as the DAS score increased by an increment of one, the likelihood of the individual's membership in the NMDID group decreased by a factor of 0.37. This indicates that a score of 1 (*parabolic*; Hefner and Linde 2018) was most closely associated with membership in the NMDID sample, while a DAS score of 3 (*hyperbolic*; Hefner and Linde 2018) was most closely associated with membership in the AMNH sample. A GE score of 0 (*absent*; Hellman 1928; Turner 1989; Rhine 1990; Atkinson and Tallman 2020) was most likely to predict membership in the AMNH sample, with the likelihood of the individual belonging in the NMDID sample increasing by a factor of 3.425 as the score increased by an increment of one. A score of 3 (*significant flare*; Hellman 1928; Turner 1989; Rhine 1990; Turner 1990; Atkinson and Tallman 2020) was most closely associated with membership in the NMDID sample. As the GMR score increased by an increment of one, the likelihood of the individual's membership in the NMDID group decreased by a factor of 0.134. A GMR score of 0 (*flat*/*smooth*; Hellman 1928; Turner 1989; Atkinson and Tallman 2020) is most closely associated with membership in the NMDID sample, while a score of 3 (*extreme robusticity;* Hellman 1928; Turner 1989; Atkinson and Tallman 2020) is most closely associated with AMNH sample membership. This indicates that the modern Indigenous sample trends toward less robust muscle attachment ridging of the gonial angle. Model 4 found that an individual's membership in the NMDID sample

increased by a factor of 32.638 as the assigned score increased by an increment of one. A CP score of 0 (*retreats*; Hellman 1928; Turner 1989; Rhine 1990; Atkinson and Tallman 2020) was most likely to predict membership in the AMNH sample, while a score of 2 (*prominent*; Hellman 1928; Turner 1989; Rhine 1990; Atkinson and Tallman 2020) was most closely related to membership in the NMDID sample. An ARS score of 1 (*pinched*; Rhine 1990; Turner 1990; Berg 2008) was most closely associated with membership in the NMDID sample, while a score of 2 (*wide*; Rhine 1990; Turner 1990; Berg 2008) was most likely to predict membership in the AMNH sample. As ARS was assessed with a binary scoring system, there was a factor of 0.227 in an individual's likelihood of membership between the two samples.

Models 1-4 reveal shifting trends in nonmetric trait morphology between the pre-Contact and modern Indigenous samples. Broadly speaking, the presentation of the midface morphology has become higher and narrower while mandibular traits have become more gracile. Gonial eversion (GE) was the sole exception; pronounced gonial eversion was found to be most likely to predict membership in the modern group. These findings support the conclusions drawn by Jantz and Meadows Jantz (2000), particularly the change demonstrated in nasal aperture width (NAW), nasal aperture shape, nasal bone contour (NBC), and zygomatic projection (ZP). The modern sample's trend toward triangular nasal bone shape (NBS) may also be connected to this shift in facial morphology. The change in presentation of the malar tubercle (MTF) and the ascending ramus shape (ARS) are consistent with variation described by Kilroy et al. (2020), which also demonstrated a pattern of increasingly gracile morphology as their sample became

more contemporary. In contrast to the current research, Kilroy et al. (2020) found that gonial eversion (designated as gonial angle flare) had decreased in robusticity over time. Martin and Danforth's (2009) metrical study of secular change in the mandible also found a decrease in bigonial breadth. While bigonial breadth can loosely be used to discuss gonial eversion, the measurement assesses the overall bigonial diameter and does not explicitly define how much the gonial morphology influences the breadth.

It is important to emphasize that the shifts in nonmetric trait expression described above are only general trends observed between the two samples studied and should not be used as a typological trait list for modern North American Indigenous populations. While these observations are consistent with findings from previous secular change studies (Jantz and Meadows Jantz 2000; Martin and Danforth 2009; Kilroy et al. 2020), the results of this study cannot be used to make broad statements about pan-Indigenous morphological variation. The conclusions drawn here only function to demonstrate the presence of statistically significant differences between a pre-Contact comparative sample and a modern population. As such, pre-Contact Indigenous individuals should not be used to develop biological profile methods for modern U.S. populations.

Ancestry estimation: hefneR

To further examine how current ancestry estimation methods classify Indigenous individuals, the hefneR classification tool was used to assess both the AMNH and NMDID samples. The hefneR tool uses a simple Bayesian classification analysis and data

from Hefner (2009) to predict membership into one of four samples: African, American Indian, Asian, and European. The NMDID sample data were entered into the hefneR tool but, because three traits used in the hefneR algorithm had been eliminated from the study, SPS, TPZ, and ZS were entered as "missing." The algorithm predicted probable African ancestry for 16 (16%) individuals, 11 (11%) as American Indian, 8 (8%) as Asian, and 65 (65%) as European. The low level of classification accuracy with the NMDID sample was not unexpected, as hefneR's American Indian sample was composed of archaeologically derived populations (Hefner 2009), and rather highlights the necessity of reassessing these antiquated methods which are still being used by forensic anthropologists.

Of the 150 individuals in the AMNH sample, 68 (45.3%) were classified as African, 21 (14%) as American Indian, 40 (26.7%) as Asian, and 21 (14%) as European. While Hefner's (2009) Indigenous sample and the ANMH sample are temporally and spatially dissimilar, the number of individuals whose ancestry was accurately predicted is exceedingly low. The algorithm classified the majority of the sample's individuals to be of African descent. Even if we consider the shared population history between Asian and Indigenous populations and group them together, the combined American Indian and Asian results fall short of the African predictions at 61 (40.7%). The currently accepted discourse regarding the peopling of the Americas does not support significant interaction between pre-Contact Indigenous peoples and European or African groups until after the start of colonialism, and as such, the AMNH sample should be able to act as a proxy for Indigenous peoples of the North American southwest prior to admixture with these

groups. Even if we attempt to explain the low level of American Indian classification in the NMDID sample as admixture, it is unexpected that American Indian classification would have the lowest rates in both samples. The low levels of accurate ancestry estimations produced suggest that the hefneR algorithm does not reliably predict American Indian ancestry in pre-Contact or modern Indigenous samples, an issue that likely stems from the comparative sample.

The current study demonstrates that using the Hefner (2009) American Indian sample as the comparative basis for Indigenous ancestry estimation is fundamentally flawed. While the sample is somewhat temporally and spatially diverse, it does not wholly encompass the extent of variation present in North American Indigenous populations. Of the 268 individuals included in the sample, 166 (61.9%) individuals were derived from archaeological sites radiocarbon dated to between 1000-200 BP (950-1750 CE), while the remaining 102 (38.1%) individuals were recovered from sites dated 1550 CE or later (Hefner 2009). The sample's individuals were derived from sites across the United States, including Alaska, California, Florida, New Mexico, and South Dakota. When compared to the AMNH sample, there is a relatively short window of temporal overlap but only 47 (17.5%) individuals were recovered from the American Southwest. While it is evident that hefneR's American Indian sample is not reflective of the individuals represented in the AMNH sample, it also fails to include Indigenous populations from large geographic regions (i.e., the Northeastern United States). While these gaps in representation may be due to lack of accessible samples, it must be stressed that the comparative sample currently used to predict group membership does not reliably

predict American Indian ancestry. The use of this tool may be contributing to misclassification of unknown individuals and is potentially hindering identification efforts. This also underscores the fact that Indigenous groups were (and are) incredibly diverse in terms of genetics, population histories, and sociocultural practices and that skeletal assemblages/datasets with superficial smatterings of pre-Contact individuals from across the continent are not resolute or robust enough to classify some pre-Contact (or any modern) Indigenous individuals. Thus, the field should move away from ideas of one broad Indigenous/Native American group, similar to other advancements with problematic and overly broad "Hispanic" and "Asian" groups (Birkby et al. 2008; Hurst 2012; Tallman 2016; Atkinson and Tallman 2020).

Intraobserver error

As outlined in the previous chapter, Cohen's kappa was used to explore intraobserver error between the two scoring events. Of the 24 observed traits, five had *almost perfect to perfect* levels of agreement, nine had *substantial* agreement, seven had *moderate* levels of agreement, and three were categorized as having *fair* levels of agreement following Landis and Koch (1977). This indicates that most of the scoring systems developed on dry bone are reliable in their capacity to produce consistent scores on digitized individuals. While the majority of traits achieved acceptable levels of agreement, the three traits (CS, MTF, and ZP) with the highest rates of error between

scoring events required further examination. There are a few possible explanations for this conflict between scoring events.

One factor that likely influenced error rates was the observer's experience level. My experience with scoring cranial and mandibular nonmetric traits is limited to coursework-related study on dry bone specimens and did not previously extend to digital specimens. All familiarity with digital specimens is attributed to the scoring processes of the current study. Similarly, nearly all experience with CT scans, three dimensional renderings, and the programs necessary to create them has been gained from use during the current study. It is possible that intraobserver error rates were exacerbated by this initial unfamiliarity and subsequent familiarization between the first and second scoring events. The scoring systems for MTF and ZP evaluate robusticity, with character states that become more pronounced as the ordinal score increases. For MTF and ZP, all inconsistencies between scoring events reflect the observer lowering an individual's score by one. By the second scoring event, the observer's increased familiarity with the extent of the variation present in the observed traits and its digital reproduction likely contributed to the adjustment of scores from extreme to moderate presentation.

An additional factor that likely contributed to increased error rates was the methodologies chosen. Of the multiple existing methods used to assess ZP, the procedure described by Rhine (1990) was used in this study to maintain consistency with the ZP scoring technique employed by Atkinson and Tallman (2020) to gather the AMNH sample data. While Rhine's (1990) method continues to be referenced in nonmetric trait literature, a study by Cataldo-Ramirez et al. (2020) evaluated the utility of two ZP

scoring methods and found that only one reliability differentiated Indigenous peoples from a pooled sample of white and Black individuals from the United States. The study compared the scoring techniques described by Rhine (1990) and Bass (1995), both of which have previously been shown to result in interobserver reliability issues (L'Abbe et al. 2011). Cataldo-Ramirez et al. (2020) study supported previous conclusions that both methods have interobserver reliability issues, with Rhine's (1990) method resulting in lower levels of agreement between observers. Any interobserver error resulting from the method itself may have been intensified by implementation on a digital model.

As described in Rhine (1990), the ZP scoring method relies on the observer's ability to place the cranium in *norma lateralis* and visually assess the angle created between the Frankfurt plane and an imaginary line that runs through the middle of both the upper and lower margins of the orbit. This angle dictates the ZP score; an angle of more than 90 degrees is considered *projecting*, 90 degrees is *vertical*, and less than 90 degrees is *receding/retreating*. This process is further complicated when performed in the RadiAnt program; while RadiAnt does provide tools that allow the user to draw lines on the 3D model, the program does not anchor drawn lines and they disappear as the model is manipulated in the 3D space. To place a line through the middle of the upper and lower orbit margins, the observer must estimate the placement while the model is in *norma lateralis*. Estimating the placement of the vertical component of Rhine's (1990) angle introduces an additional avenue for error. Furthering the potential for error, the program lacks a means to measure the created angle, which compels the observer to estimate the score.

When compared to previous cranial and mandibular nonmetric trait research, the current study produced slightly higher levels of intraobserver error. Hefner (2009) demonstrated *moderate* levels of intraobserver agreement in two traits, ANS and SPNS (ANS k = 0.422; SPNS k = 0.468), citing the difficulties inherent to scoring morphological traits. Tallman (2016) reported their lowest levels of agreement in two traits, location of posterior ramus inversion (LPRI) (k = 0.158) and SF (k = 0.373), classified as *slight* and *fair* agreement, respectively. Kilroy et al. (2020) achieved *substantial* to *almost perfect* agreement in all 23 traits observed. Atkinson and Tallman (2020) reported their lowest level of agreement, designated as *fair*, in INA (k = 0.375); subsequently choosing to remove INA from the analyses. An important distinction between the aforementioned studies and the current research is the matrix on which these observations were made. The previous research (Hefner 2009; Tallman 2016; Kilroy et al. 2020; Atkinson and Tallman 2020) performed their research on human osteological specimens, which they were likely familiar with. The slightly higher levels of intraobserver error present in the current study are likely explained by a learning curve present when applying techniques developed for physical material to digital samples.

These potential explanations for the discrepancies between the first and second scoring events fall in line with the avenues of intraobserver error outlined by Kamnikar et al. (2018). In their study, the authors identified three patterns related to disagreement which may occur during macromorphoscopic trait analysis: the observer's experience level, the introduction of novel or recently refined technologies, and error resulting from the methodology, itself. The observer's experience level directly influences nonmetric

trait scoring, as those with more experience have been exposed to a greater range of human morphological variation and are more likely to confidently assign trait scores. Less experienced observers are unlikely to have been exposed to extreme trait values, making them more likely to misclassify moderate presentations as extreme.

Using new or recently updated technologies can influence error rates as mistakes are more likely during the adjustment period inherent to learning unfamiliar programs or becoming used to modifications to pre-existing procedures. In their study, Kamnikar et al. (2018) referenced updates made to the Macromorphoscopic Module in Osteoware which included shifting extant scoring system values by an increment of one. This slight alteration in data collection procedure has the potential to cause imprecise data collection, resulting in flawed analyses and inaccurate conclusions if left uncorrected. Error inherent to the method, such as unclear character state definitions or implementation procedure, can make it difficult for an observer to make accurate, replicable datasets. Any ambiguity in procedure is likely to result in inconsistent observer interpretation.

While the study anticipated meaningful levels of error, most of the observed traits had negligible levels of intraobserver error and were deemed suitable for continued use in Osteoware's Macromorphoscopic Module Database. The authors made several recommendations intended to mitigate these sources of error, highlighting the importance of training observers on nonmetric trait scoring prior to data collection, and integrating standardizing technologies into methodological approaches. Kamnikar et al. (2018) counsel researchers to be conscientious of ongoing modification to extant techniques and

suggested further intraobserver error testing on nonmetric trait methodologies to ensure their reliability.

Extant scoring methods and 3D models

The methods used to assess nonmetric traits have not yet been optimized for use on 3D models. Several techniques are aided by physical palpation of bony idiosyncrasies or require the implementation of tools, neither of which have been replicated by software capabilities. Two of the traits used in this study, INA and NBC, are scored with such methods. To address this, alternative approaches were developed to assess these traits. While Hefner and Linde (2018) may not explicitly instruct anthropologists to palpate the INA, tactile exploration can aid in the scoring process. This is especially true when evaluating individuals whose INA morphology does not visually indicate a specific score. To better assess the INA, RadiAnt's built-in scalpel tool was used to remove the right half of the model skull, to include the vomer insertion. Removing this portion of the model provides an unobstructed view of the slope of the nasal aperture floor and any potential guttering or silling, all of which are factors detailed in the character state descriptions used to assign an INA score (Figure 25).

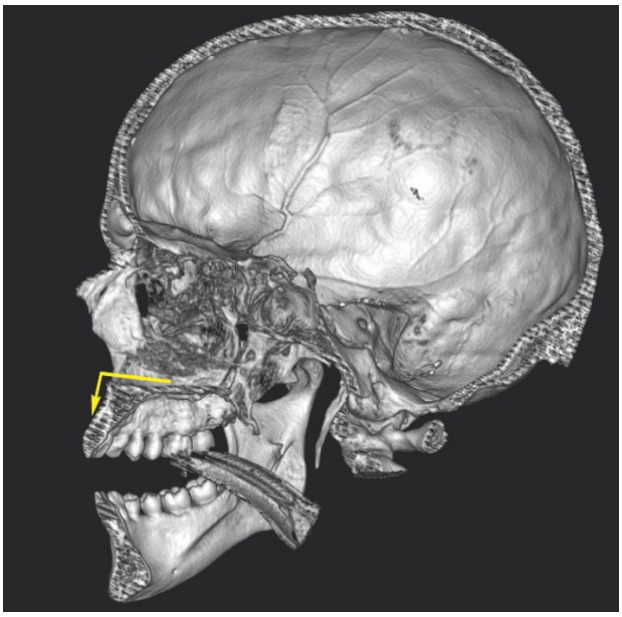


Figure 25. Rendering with left portion of the skull removed to better assess INA. INA scoring location highlighted by yellow arrow.

According to the methodology outlined in Hefner and Ousley (2018), NBC observation is reliant on the use of a contour gauge. Because there is no contour gauge equivalent in the RadiAnt program, the technique needed to be adjusted to the program's capabilities. Using the measuring tool, a point located one inch below the nasofrontal suture was found and the scalpel tool was used to remove the portion of the nasal bones distal below this point. This creates a cross-section of the nasal bones at the location that the contour gauge is conventionally placed. Using the measurement tool, the NBC can be delineated, and any plateaus can be quantified.

The scoring system for ARP, as outlined by Berg (2008) utilizes an angle measurement to establish character state. This measurement reflects the difference between the gonial angle and 90 degrees, falling into one of three categories: *straight* (0-

10 degrees), *medium* (11-20 degrees), and *slanted* (>21 degrees). Due to this method's reliance on angle measurement, this trait is better suited to metrical analysis (Berg 2008) and assessment using a mandibulometer. Though it is possible to score ARP morphoscopically, concerns about score reliability and the RadiAnt program's lack of angle measuring tools led to the trait's removal from the study.

Traits removed from observation

During the initial phases of data collection, 11 of 35 traits included in Atkinson and Tallman's (2020) dataset were excluded from the study due to inaccessibility, poor resolution, or inconsistent reproduction in the RadiAnt program. Traits concerned with sutural presentation were particularly susceptible to reproduction issues. While the larger sutures of the cranial vault were often visible on the models, CT scan stacks taken on the same individual often produced conflicting trait presentations or were not detailed enough to accurately score. This issue negatively affected visualization of SF and WB, necessitating their removal from the study. The smaller, thinner sutures of the viscerocranium, NS and SNS, were therefore eliminated due to similar reproduction difficulties. Additionally, areas of thin bone were often incompletely captured in the models, leaving traits such as ZMS and TPS absent or too unreliable for observation (Figure 26).

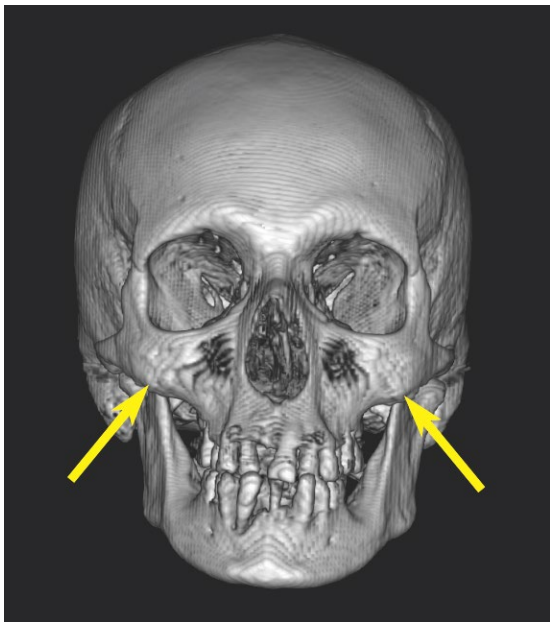


Figure 26. A case in which the ZMS was not reproduced in the rendering. The yellow arrows indicate where the zygomaticomaxillary suture should be.

PRI was removed from the study due to poor image resolution, replicability issues, and artifact-related distortion. While PRI was often visible notwithstanding artifact-related obstructions, the character state presented differed between CT scan stacks, making scores unreliable. Because accurate scoring of LPRI is reliant on clearly defined origination and termination points of PRI, LPRI was also removed from the study. Similar issues with trait reproduction differing between CT scan stacks resulted in inconsistent MB scoring, causing MB to also be excluded.

Another issue which interfered with trait observation was the presence of medical and dental appliances in the scans. Evidence of medical interventions present at the time of death were not always removed before images of a decedent were taken. These instruments, including endotracheal and nasogastric tubes (Figures 27), were captured

and reproduced in the 3D models. The position of these objects often obscured the traits being observed and required painstaking removal. Although removing these inclusions was often possible, the process could be difficult; while the program's scalpel tool allows the user to remove portions of the rendering, it does not allow the user to isolate objects in the 3D space. Removing a section of the model removes all portions of the rendering within the highlighted area, including bone. Dental appliances, such as partial or full dentures and amalgam fillings, also impeded the observation of maxillary and mandibular traits, such as MT. Amalgam fillings created starburst-like artifacts in both the scans and the rendered models, which obscured or distorted large portions of the mandible (Figure 28).

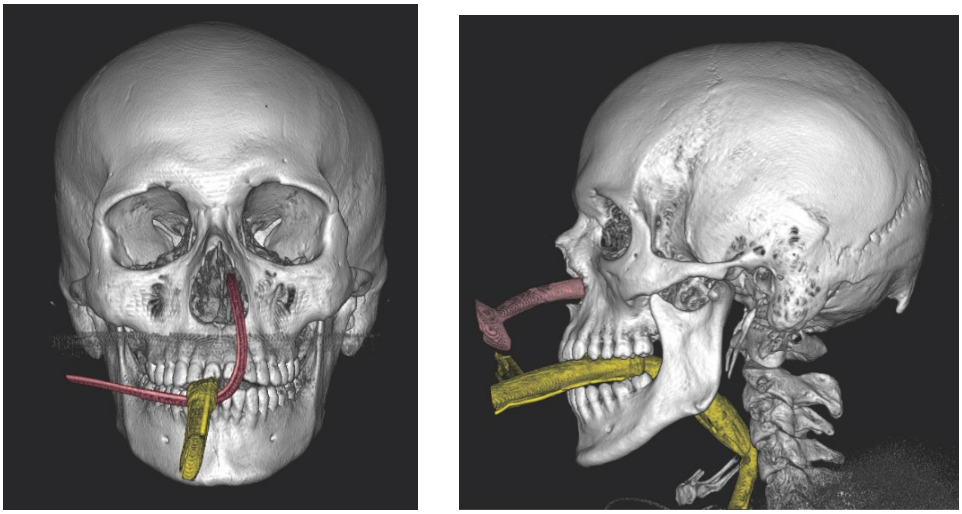


Figure 27. Medical instruments included in models. Nasogastric tube highlighted in red. Endotracheal tubing highlighted in yellow.

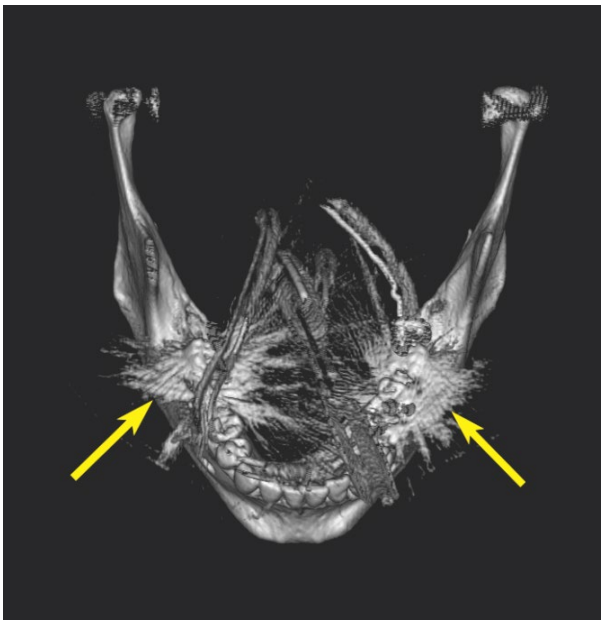


Figure 28. Mandible reproduction obscured by medical tubing and "starburst" effect. "Starburst" effect indicated by yellow arrows.

The Missing and Murdered Indigenous Women and Girls crisis in North America

The results produced by this study are particularly troubling due high levels of violence, trafficking, and homicide experienced by Indigenous peoples in the United States and Canada. Research has found that indigenous women are 4.5 times more likely to experience violence than the rest of the general population (Saramo 2016; Ficklin et al. 2021.) The National Institute for Justice provided insight into violence committed against Indigenous women and girls in the United States, reporting that 56.1% experienced sexual violence in their lifetime, 84.3% experienced violence in their lifetime, and 97.0% who had experienced one or more acts of violence reported that the violence had been committed by someone of another "race" (Rosay 2016). Canada's General Social Survey on Victimization has also noted a trend of increasing violence

against Indigenous. The 2017 survey found that Indigenous people, who only represent approximately 4.9% of the Canadian population, made up 24% of all homicide victims in the country. In the same year, the rate of homicides committed against Indigenous women saw an increase of 32% and those against Indigenous men increased by 2%. Despite these elevated rates of violence and homicides, Indigenous peoples are underrepresented in federal missing persons databases and tend to receive less media attention (Salam 2019; Ficklin et al. 2021.)

Due to underreporting, lack of media attention, or exclusion of indigenous cases from national databases, Indigenous communities have turned to social media to amplify the stories of their missing and murdered persons (Ficklin et al. 2021). Storytelling cultural practices, which is prevalent in Indigenous communities, can help the listener recognize the value of the missing or murdered person and prevents them from being forgotten as a nameless statistic. Storytellers are also in a unique position to shift the narrative away from harmful conceptions of indigeneity and, instead, focus on the importance of the person discussed (Ficklin et al. 2022). Ficklin et al. (2022) notes that indigenous storytelling practices do not always translate well to medias valued by western perspectives, such as academic papers or statistics. Additionally, while the victimization of Indigenous women can be, in part, linked to mainstream victimization such as substance abuse, childhood trauma, and mental health issues, the study posits that the higher rates of violence committed against Indigenous women can also be linked to the colonial and Western conceptions of gender roles, the sexualization of Indigenous women in popular culture, and the otherization of Indigenous groups (Ficklin et al. 2022).

Unable to rely on mainstream avenues, Indigenous peoples have taken it upon themselves to shed light on these issues. (Ficklin et al. 2022; *New York Times* 2019).

To address the lack of centralized information concerning Missing and Murdered Indigenous Peoples, the Missing and Murdered Peoples Database was created (Ficklin et al. 2023; Sovereign Bodies Institute 2023). The database, formerly known as the Missing and Murdered Women and Girls Database, was founded by Annita Lucchesi, an Indigenous survivor of intimate partner violence, sexual assault, and trafficking. (Sovereign Bodies Institute 2023). The database collects data about the victims, the perpetrators, the violence/case, police activity or legal processes, and location information. The database has collected information on cases dating as far back as the early 1900s (Sovereign Bodies Institute 2023). The website also provides access to relevant resources such as webinars and research performed on the topic (Sovereign Bodies Institute 2023). The MMIP Database is no longer the only database in the United States that collects and documents information specifically focused on violence committed against Indigenous peoples; the Missing and Murdered Women, Girls, and Two Spirit Database focuses specifically on Alaskan Native groups (Data for Indigenous Justice 2023). Recently, the United States congress has passed legislation, called the "Not Invisible Act of 2020," designed to address the crisis of violence committed against Indigenous peoples (Bureau of Indian Affairs 2023). This led to the establishment of the Missing and Murdered Unit, a joint effort between the Department of the Interior and the Department of Justice.

Future research

Given the results of the previously discussed analyses, further investigation into the efficacy of nonmetric ancestry estimation methods is necessary if the field of forensic anthropology continues to teach and employ them. While these analyses were performed on two small samples, the demonstrated variation in nonmetric trait expression between the two groups and hefneR's low level of accurate ancestry prediction illustrates the fault in using historic populations as a comparative sample for modern decedents. As with all methods used to develop the biological profile, the value in ancestry estimation techniques relies heavily on their ability to provide information with an acceptable level of accuracy. As discussion on whether forensic anthropologists should continue to use ancestry estimation as part of the biological profile persists, the value of updating extant techniques can also be called into question.

The inaccurate ancestry estimations produced by hefneR is likely due to reliance on a sample that does not reflect the variability present in North American Indigenous populations. To mitigate this issue, efforts to incorporate trait data from modern populations may help to increase the algorithm's accuracy. Further, the problematic nature of using the four heavily referenced geographically derived ancestral groups as the only possibilities for population affinity should also be reexamined. Given what recent research on secular change and the current study have illustrated about the ever-changing nature of human variation, it is unlikely that modern humans will continue to reliably be affiliated with such dated, simplistic population limitations. If the samples are not modernized and trait data is not continuously updated, this method will become less

reliable as time goes on. It is the author's opinion that, if reliable ancestry estimation methods cannot be achieved, the practice should be eliminated from the forensic anthropological toolkit.

As technology advances, the field of forensic anthropology has the potential to become more accurate, more accessible, and less invasive. As demonstrated by this study, scanning and digitizing specimens allows researchers to study samples without necessitating physical handling of the remains. This could be particularly useful for specimens whose condition has become compromised or are inaccessible due to factors such as reburial, physical distance, or pandemic associated lockdowns. While there were issues encountered during the process of this study, standardization of scanning methods and the development of a program geared toward forensic anthropological needs could mitigate many of those problems.

To improve scanning methods for use by forensic anthropologists, procedures would need to be developed and disseminated to those performing the imaging. Images optimized for forensic anthropologists would not include medical appliances or metal zippers and bag closures, as they can be too close to the body to be removed or may create artifacts that distort the images. Additionally, the decedent would need to be placed into a standardized position, as close to anatomical position as possible. If it is not possible, ensuring that limbs are not positioned over the head or torso and that the head is not resting on the shoulder or chest would be ideal. Standardizing the position of the decedent would ensure that all areas of the body are accessible and are positioned in a manner that the observer is accustomed to. Areas used for multiple analytical methods,

such as the skull and pelvis, would ideally be scanned in with the highest resolution possible. Any renderings produced must be both as morphologically and metrically accurate to the specimen as possible, as anthropologists use a combination of nonmetric and metric techniques to draw their conclusions.

To perform methods used by forensic anthropologists on digital specimens, a DICOM viewing program or another software capable of rendering three dimensional images would be necessary. Such a program would need the capabilities to separate fleshed specimens into layers that can be toggled on and off, provide a tool to remove portions of the rendering, and provide measurement and angle tools. It would also be helpful to include "snap to" positions commonly used by forensic anthropologists, such as anterior view, posterior view, and Frankfurt plane. To produce images for inclusion in reports, the program would also need to allow the user to toggle a scale on and off, layer arrows or circles to highlight areas of interest, and isolate individual bones. Many of these features already exist in extant programs, such as RadiAnt and Blender, but examination of digital specimens in one program would be ideal for standardization and reliability amongst future forensic anthropologists.

CHAPTER SIX

Conclusions

The present study sought to explore the variation in cranial and mandibular nonmetric trait expression between pre-Contact and modern Southwest indigenous populations and assess how those changes effect the efficacy of ancestry estimation techniques currently employed by forensic anthropologists. The results of the chi-square analyses demonstrate that statistically significant change occurred in 22 of the 24 observed traits, with NO and KLN being the only exceptions. Models created to explore the relationship between the nonmetric trait expression and group membership found that the modern sample was characterized by higher, more narrow presentation in craniofacial morphology and features of the mandible trend toward more gracile presentations with the exception of gonial eversion. These findings support conclusions drawn from previous studies of secular change in cranial and mandibular morphology, which indicate that the gross morphology of the skull has changed over time (Jantz and Meadows Jantz 2000; Martin and Danforth 2009; Kilroy et al. 2020).

An investigation into the effectiveness of an extant ancestry estimation tool, hefneR, found that the algorithm produced low levels of accurate ancestry prediction for both the pre-Contact and modern indigenous samples. It predicted that the majority of the pre-Contact sample was most likely to be of African ancestry while most of the modern sample was determined to be most likely to be of European descent. Predictions of American Indian ancestry yielded the lowest rates from both samples, correctly classifying 14% of the AMNH sample and 11% of the NMDID sample. While these low

rates of accurate ancestry estimation may be partially explained by nonmetric trait discrepancies caused by secular change demonstrated here, it is likely that the American Indian sample the algorithm is based on does not truly reflect the extent of biological variation present in North American indigenous populations and cannot be reliably used as a comparative basis for assessing population affinity in either pre-Contact or modern indigenous samples.

The present research illustrates the necessity of continually updating comparative sample databases to ensure they accurately reflect modern human biological variation. While hefneR may produce reliable ancestry estimations for other groups, this study has demonstrated that it does not result in reliable population affinity predictions for the two assessed indigenous samples. The use of inaccurate approaches in forensic casework could negatively affect identification efforts through the dissemination of incorrect ancestry estimates. This is especially relevant to the ongoing issue of Missing and Murdered Indigenous Women and Girls, as misclassify ancestry estimates may be hindering identification and contributing to their continued missing status; as of June 2023, there are more than 800 missing persons categorized as American Indian/Alaskan Native active on NamUS (NamUS 2023). Law enforcement officials rely on the proficiency of subject matter experts, such as forensic anthropologists, to accurately inform them about the scientific aspects of their cases. Additionally, all scientific methods used in forensic casework must meet reliability thresholds put forth by the Daubert standard (Daubert v. Merrell Dow Pharmaceuticals, Inc), which require all admissible scientific testimony be based on methods that have been tested for validity,

peer-reviewed, are accepted by the scientific community, have documented error rates, and have established procedures to guide their implementation. Failure to address the effect of secular change, not only in ancestry estimates, but also in all facets of the biological profile, will likely negatively impact the reliability of forensic assessments.

The recent discourse regarding the future of ancestry estimation questions the validity of the practice, citing lack of heritability data for many of the nonmetric traits used in forensic anthropological methods and condemns the typological nature of past research (DiGangi and Bethard 2021a; DiGangi and Bethard 2021b). Some practitioners recommend that ancestry estimation is eliminated from the biological profile while others argue that forensic anthropologists acknowledge the influence of racial biases on scientific research and currently strive to address previous failings (Stull et al. 2021). Contemporary biological anthropologists understand that a race is a social construct; while statistical analyses may produce a population affinity estimate that has been associated with a social race category, it is acknowledged they are not the same and the "assigned" social race may not correspond with how an individual self-identifies. The disparity between biological population affinity and perceived social race can be problematic, as social race cannot be extracted from skeletal material but is integral to an individual's lived experience, including their death.

The elimination of ancestry estimation practices does pose a few problems. While DNA analysis has become the gold standard for identifying unknown decedents, DNA testing can be a costly endeavor and may not always be feasible in smaller jurisdictions.

The process is destructive and acquiring a viable DNA sample is not always possible. Furthermore, identification of an unknown individual through DNA testing requires a comparative sample. If no comparative sample is available, any DNA extracted is stored until one is acquired. Though the biological profile is unlikely to be the sole factor in the identification of unknown individuals, the information it provides can be used to narrow down search results. Additionally, some methods used to develop estimates for other facets of the biological profile (e.g., stature) are population specific and are performed after an ancestry estimate has been established. If the field continues to use nonmetric trait expression to assess population affinity, more work needs to be done to address the concerns raised about their heritability and how microevolutionary forces affect group membership probabilities.

Further research should seek to continue to explore the reliability of extant approaches used to develop the biological profile and determine the effect of secular change on these analyses. Efforts to modernize samples used for comparative analyses should continue to be made for all major population groups, to ensure accurate representation of modern human biological variation. Furthermore, nonmetric and metrical approaches should be optimized for digital specimens. Forensic anthropological techniques will require adaptation as technology continues to advance.

The reproduction issues encountered during this study could be mitigated in future research through the standardization of CT scan procedure and the development or adaptation of 3D rendering software for use by forensic anthropologists. The ability to

create precise 3D renderings of skeletal samples has many potential benefits, including increased accessibility to skeletal samples which would eliminate obstacles such as geographic distance or financial limitations. It would also decrease the need for continuous handling of physical remains which would otherwise inevitably lead to their degradation. Digitization would also create an avenue for preservation of skeletal data in deteriorating samples, as that information eventually becomes lost when osseous material becomes friable to handle or specimens are damaged. Additionally, 3D models created from images taken during autopsy could contribute to the creation of modern skeletal samples; these samples would not need to be subjected to the maceration process nor would they require physical curation.

BIBLIOGRAPHY

Albanese, J., & Saunders, S. R. (2006). Is It Possible to Escape Racial Typology in Forensic Identification? In A. Schmitt, E. Cunha, & J. Pinheiro (Eds.), Forensic Anthropology and Medicine: Complementary Sciences From Recovery to Cause of Death (pp. 281–316). Humana Press. https://doi.org/10.1007/978-1-59745-099-7_12

Angel, J. L. (1976). Colonial to modern skeletal change in the U.S.A. American Journal of Physical Anthropology, 45(3), 723–735. https://doi.org/10.1002/ajpa.1330450339

Angel, J. L. (1982). A new measure of growth efficiency: Skull base height. American Journal of Physical Anthropology, 58(3), 297–305. https://doi.org/10.1002/ajpa.1330580308

Atkinson, M. L., & Tallman, S. D. (2020). Nonmetric Cranial Trait Variation and Ancestry Estimation in Asian and Asian-Derived Groups. Journal of Forensic Sciences, 65(3), 692–706. https://doi.org/10.1111/1556-4029.14234

Ayers HG, Jantz RL, Moore-Jansen PH (1990) Giles and Elliot race discriminant functions revisited: a test using recent forensic cases. In: Gill GW, Rhine S (eds) Skeletal attribution of race. Maxwell Museum of Anthropology, Albuquerque, pp 65–71

Bass WM. Human osteology: a laboratory and field manual, 4th edn. Columbia, MO: Missouri Archaeological Society, 1995.

Bethard, J. D., & DiGangi, E. A. (2020). Letter to the editor—Moving beyond a Lost Cause: Forensic anthropology and ancestry estimates in the United States. Journal of Forensic Sciences, 65(5), 1791–1792. https://doi.org/10.1111/1556-4029.14513

Berg, G. E. (2008). Biological Affinity and Sex Determination Using Morphometric and Morphoscopic Variables from the Human Mandible. 192.

Bertoglio, Corradin, S., Cappella, A., Mazzarelli, D., Biehler-Gomez, L., Messina, C., Pozzi, G., Sconfienza, L. M., Sardanelli, F., Sforza, C., De Angelis, D., & Cattaneo, C. (2020). Pitfalls of Computed Tomography 3D Reconstruction Models in Cranial Nonmetric Analysis. Journal of Forensic Sciences, 65(6), 2098–2107. https://doi.org/10.1111/1556-4029.14535

Berry, A. C. & Berry, R. J. (1967). Epigenetic variation in the human cranium. Journal of Anatomy, 101(Pt 2), 361–379.

Birkby, W. H., Fenton, T. W., & Anderson, B. E. (2008). Identifying Southwest Hispanics Using Nonmetric Traits and the Cultural Profile. Journal of Forensic Sciences, 53(1), 29–33. https://doi.org/10.1111/j.1556-4029.2007.00611.x

Blumenbach, Bendyshe, T., Marx, K. F. H., Flourens, P., Wagner, R., Hunter, J., Bendyshe, Thomas, Marx, K. F. H. (Karl Friedrich Heinrich), Flourens, P. (Pierre), Wagner, Rudolph, & Hunter, John. (1865). The anthropological treatises of Johann Friedrich Blumenbach ... Published for the Anthropological Society, by Longman, Green, Longman, Roberts, & Green.

Brooks, S. & Suchey, J. M. (1990) Skeletal age determination based on the os pubis: A comparison of the Acsadi-Nemeskeri and Suchey-Brooks methods. Human Evolution 5:227-238.

Brues. (1958). Identification of Skeletal Remains. Journal of Criminal Law, Criminology & Police Science, 48(5), 551–563. https://doi.org/10.2307/1139795

Brues A. 1990. The Once and Future Diagnosis of Race. In: Gill G, and Rhine S, editors. Skeletal Attribution of Race. Albuquerque NM: University of New Mexico: Maxwell Museum of Anthropological Papers No. 4.

Bureau of Indian Affairs. (2023). Missing and murdered indigenous people. Missing and Murdered Indigenous People | Indian Affairs. https://www.bia.gov/service/mmu

Byrnes, J. F., Kenyhercz, M. W., & Berg, G. E. (2017). Examining Interobserver Reliability of Metric and Morphoscopic Characteristics of the Mandible. Journal of Forensic Sciences, 62(4), 981–985. https://doi.org/10.1111/1556-4029.13349

Cataldo-Ramirez, C., Garvin, H. M., & Cabo, L. (2020). A Quantitative Assessment of Zygomatic Projection for Ancestry Estimation. Journal of Forensic Sciences, 65(2), 580–590. https://doi.org/10.1111/1556-4029.14188

Chambellan V (1883) Étude anatomique et anthropologique sur les os Wormiens. thèse de médecine, Paris, p 66

Cunha, E., & Ubelaker, D. H. (2020). Evaluation of ancestry from human skeletal remains: A concise review. Forensic Sciences Research, 5(2), 89–97. https://doi.org/10.1080/20961790.2019.1697060

Decker, Davy-Jow, S. L., Ford, J. M., & Hilbelink, D. R. (2011). Virtual Determination of Sex: Metric and Nonmetric Traits of the Adult Pelvis from 3D Computed Tomography Models. Journal of Forensic Sciences, 56(5), 1107–1114. https://doi.org/10.1111/j.1556-4029.2011.01803.x

DiGangi, E. A., & Bethard, J. D. (2021). Uncloaking a Lost Cause: Decolonizing ancestry estimation in the United States. American Journal of Physical Anthropology, 175(2), 422–436. https://doi.org/10.1002/ajpa.24212

d'Oliveira Coelho, J., Curate, F., & Navega, D. (2020). Chapter 5.1 - Osteomics: Decision support systems for forensic anthropologists. In Z. Obertová, A. Stewart, & C. Cattaneo (Eds.), Statistics and Probability in Forensic Anthropology (pp. 259–273). Academic Press. https://doi.org/10.1016/B978-0-12-815764-0.00005-8

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

Dwight. (1890). The Closure of the Cranial Sutures as a Sign of Age. The New England Journal of Medicine, 122(17), 389–392. https://doi.org/10.1056/NEJM189004241221701

Edgar, H. J. H. (2013). Estimation of Ancestry Using Dental Morphological Characteristics. Journal of Forensic Sciences, 58(s1), S3–S8. https://doi.org/10.1111/j.1556-4029.2012.02295.x

Edgar, H. j. h., & Lease, L. r. (2007). Correlations between deciduous and permanent tooth morphology in a European American sample. American Journal of Physical Anthropology, 133(1), 726–734. https://doi.org/10.1002/ajpa.20564

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

Ficklin, Tehee, M., Killgore, R. M., Isaacs, D., Mack, S., & Ellington, T. (2022). Fighting for our sisters: Community advocacy and action for missing and murdered Indigenous women and girls. Journal of Social Issues, 78(1), 53–78. https://doi.org/10.1111/josi.12478

Freidline, S. E., Gunz, P., & Hublin, J.-J. (2015). Ontogenetic and static allometry in the human face: Contrasting Khoisan and Inuit. American Journal of Physical Anthropology, 158(1), 116–131. https://doi.org/10.1002/ajpa.22759 Fuentes, A. (2021). Biological anthropology's critical engagement with genomics, evolution, race/racism, and ourselves: Opportunities and challenges to making a difference in the academy and the world. American Journal of Physical Anthropology, 175(2), 326–338. https://doi.org/10.1002/ajpa.24162

Garcia de Leon Valenzuela, M.J.. (2016). Three-dimensional image technology in forensic anthropology: assessing the validity of biological profiles derived from CT-3D images of the skeleton. M.S. thesis, Boston University. https://open.bu.edu/handle/2144/15358

Giles E, and Elliot O. 1962. Race Identification from Cranial Measurements. Journal of Forensic Sciences 7:147-157.

Gill, G. W. (1995). Challenge on the Frontier: Discerning American Indians from Whites Osteologically. Journal of Forensic Sciences, 40(5), 15384J. https://doi.org/10.1520/JFS15384J

Gill, GW (1998) Craniofacial Criteria in the Skeletal Attribution of Race. In, Forensic Osteology, 2nd edition, ed. Kathleen Reichs. Springfield, Illinois: Charles C. Thomas.

Gill, G. W., Hughes, S. S., Bennett, S. M., & Miles Gilbert, B. (1988). Racial Identification from the Midfacial Skeleton with Special Reference to American Indians and Whites. Journal of Forensic Sciences, 33(1), 12440J. https://doi.org/10.1520/JFS12440J

Go, & Hefner, J. T. (2020). Morphoscopic ancestry estimates in Filipino crania using multivariate probit regression models. American Journal of Physical Anthropology, 172(3), 386–401. https://doi.org/10.1002/ajpa.24008

Gould. (1981). The mismeasure of man (1st ed.). Norton.

Government of Canada, D. of J. (2019). Indigenous overrepresentation in the criminal justice system—JustFacts. https://www.justice.gc.ca/eng/rp-pr/jr/jf-pf/2019/may01.html

Hamm, E. (2019). Christian Thought, Race, Blumenbach, and Historicizing. Zygon®, 54(1), 237–245. https://doi.org/10.1111/zygo.12495

Hanihara, T. (2008). Morphological variation of major human populations based on nonmetric dental traits. American Journal of Physical Anthropology, 136(2), 169–182. https://doi.org/10.1002/ajpa.20792

Hawes, M. B., Slakoff, D. C., & Anguelov, N. (2023). Understanding the Missing and Murdered Indigenous Women Crisis: An Analysis of the NamUs Database. Criminal Justice Policy Review, 34(2), 184–207. https://doi.org/10.1177/08874034221098909

HefneR. (n.d.). Retrieved November 18, 2020, from http://osteomics.com/hefneR/

Hefner. (2007). The statistical determination of ancestry using cranial nonmetric traits. Ph.D. dissertation, University of Florida. ProQuest Dissertations Publishing.

Hefner, J. T. (2009). Cranial Nonmetric Variation and Estimating Ancestry. Journal of Forensic Sciences, 54(5), 985–995. https://doi.org/10.1111/j.1556-4029.2009.01118.x

Hefner, J. T. (2018). The macromorphoscopic databank. American Journal of Physical Anthropology, 166(4), 994–1004. https://doi.org/10.1002/ajpa.23492

Hefner, J.T., Linde, K.C. 2018. Atlas of human cranial macromorphoscopic traits. San Diego, CA: Elsevier, Academic Press.

Hefner, J. T., & Ousley, S. D. (2014). Statistical Classification Methods for Estimating Ancestry Using Morphoscopic Traits. Journal of Forensic Sciences, 59(4), 883–890. https://doi.org/10.1111/1556-4029.12421

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

Hellman M. Racial characters in human dentition: part I: a racial distribution of the Dryopithecus pattern and its modifications in the lower molar teeth of man. Proceedings of the American Philosophical Society 1928;67(2):157–74.

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

Hooton, EA (1930) The Indians of Pecos Pueblo. Yale University Press, New Haven.

Howells WW. 1973. Cranial Variation in Man. A Study by Multivariate Analysis of Patterns of Differences Among Recent Human Populations. Papers of the Peabody Museum of Archeology and Ethnology, vol. 67, pp. 259. Cambridge, Mass.: Peabody Museum.

Howells WW. 1989. Skull Shapes and the Map. Craniometric Analyses in the Dispersion of Modern Homo. Papers of the Peabody Museum of Archaeology and Ethnology, vol. 79, pp. 189. Cambridge, Mass.: Peabody Museum.

Howells WW. 1995. Who's Who in Skulls. Ethnic Identification of Crania from Measurements. Papers of the Peabody Museum of Archaeology and Ethnology, vol. 82, pp. 108. Cambridge, Mass.: Peabody Museum.

Hrdlička, A. (1898). Anthropological investigations on one thousand white and colored children of both sexes, the inmates of the New York Juvenile Asylum, with additional notes on one hundred colored children of the New York colored orphan asylum. New York : s.n. http://archive.org/details/anthropologicali00hrdl

Hrdlička, A. (1917). The genesis of the American Indian ... https://catalog.hathitrust.org/Record/100560988

Hrdlička. (1920). Shovel-shaped teeth. American Journal of Physical Anthropology, 3(4), 429–465. https://doi.org/10.1002/ajpa.1330030403

Hrdlička. (1925). The old Americans, by Aleš Hrdlička. The Williams & Wilkins Company, 1925.

Hubbe, & Neves, W. (2007). On the Misclassification of Human Crania. Current Anthropology, 48(2), 285–288. https://doi.org/10.1086/512985

Hurst, C. V. (2012). Morphoscopic Trait Expressions Used to Identify Southwest Hispanics. Journal of Forensic Sciences, 57(4), 859–865. https://doi.org/10.1111/j.1556-4029.2012.02080.x

Jantz, L. M., & Jantz, R. l. (1999). Secular change in long bone length and proportion in the United States, 1800–1970. American Journal of Physical Anthropology, 110(1), 57–67. https://doi.org/10.1002/(SICI)1096-8644(199909)110:1<57::AID-AJPA5>3.0.CO;2-1

Jantz, R. l., & Meadows Jantz, L. (2000). Secular change in craniofacial morphology. American Journal of Human Biology, 12(3), 327–338. https://doi.org/10.1002/(SICI)1520-6300(200005/06)12:3<327::AID-AJHB3>3.0.CO;2-1

Jantz R, and Ousley S. 2005. FORDISC 3.0: Personal Computer Forensic Discriminant Functions. Knoxville: Departments of Anthropology, University of Tennessee.

Jordan, K. L., Rishishwar, L., & Conley, A. B. (2019). Native American admixture recapitulates population-specific migration and settlement of the continental United States. PLoS Genetics, 15(9), e1008225–e1008225. https://doi.org/10.1371/journal.pgen.1008225

Kamnikar, Plemons, A. M., & Hefner, J. T. (2018). Intraobserver Error in Macromorphoscopic Trait Data. Journal of Forensic Sciences, 63(2), 361–370. https://doi.org/10.1111/1556-4029.13564

Kennedy, K. A. R. (1995). But Professor, Why Teach Race Identification if Races Don't Exist? Journal of Forensic Sciences, 40(5), 15386J. https://doi.org/10.1520/JFS15386J

Kilroy, G. S., Tallman, S. D., & DiGangi, E. A. (2020). Secular change in morphological cranial and mandibular trait frequencies in European Americans born 1824–1987. American Journal of Physical Anthropology, 173(3), 589–605. https://doi.org/10.1002/ajpa.24115

L'Abbé, Van Rooyen, C., Nawrocki, S., & Becker, P. (2011). An evaluation of non-metric cranial traits used to estimate ancestry in a South African sample. Forensic Science International, 209(1), 195.e1–195.e7. https://doi.org/10.1016/j.forsciint.2011.04.002

Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. Biometrics, 33(1), 159–174. https://doi.org/10.2307/2529310

Liebmann, M. J., Farella, J., Roos, C. I., Stack, A., Martini, S., & Swetnam, T. W. (2016). Native American depopulation, reforestation, and fire regimes in the Southwest

United States, 1492–1900 CE. Proceedings of the National Academy of Sciences, 113(6), E696–E704. https://doi.org/10.1073/pnas.1521744113

Linnaeus, Carl (1758). Systema naturae per regna tria naturae :secundum classes, ordines, genera, species, cum characteribus, differentiis, synonymis, locis (in Latin) (10th ed.). Stockholm: Laurentius Salvius.

Little, Michael A, & Kennedy, K. A. R. (2010). Histories of American physical anthropology in the twentieth century. Lexington Books.

Maier, C., Craig, A., & Adams, D. M. (2021). Language use in ancestry research and estimation. Journal of Forensic Sciences, 66(1), 11–24. https://doi.org/10.1111/1556-4029.14601

Malhi, R. S., Mortensen, H. M., Eshleman, J. A., Kemp, B. M., Lorenz, J. G., Kaestle, F. A., Johnson, J. R., Gorodezky, C., & Smith, D. G. (2003). Native American mtDNA prehistory in the American Southwest. American Journal of Physical Anthropology, 120(2), 108–124. https://doi.org/10.1002/ajpa.10138

Martin, D. C., & Danforth, M. E. (2009). An analysis of secular change in the human mandible over the last century. American Journal of Human Biology, 21(5), 704–706. https://doi.org/10.1002/ajhb.20866

Medixant. RadiAnt DICOM Viewer [Software]. Version 2021.1. Jun 27, 2021. URL: https://www.radiantviewer.com

Missing and Murdered Indigenous People Crisis | Indian Affairs. (n.d.). Retrieved August 8, 2022, from https://www.bia.gov/service/mmu/missing-and-murdered-indigenous-people-crisis

Monchalin, L., Marques, O., Reasons, C., & Arora, P. (2019). Homicide and Indigenous peoples in North America: A structural analysis. Aggression and Violent Behavior, 46, 212–218. https://doi.org/10.1016/j.avb.2019.01.011

Morton SG. 1839. Crania Americana. Philadelphia: J. Dobson

Murad, T. A., & Murad, T. D. (2000). The Postmortem Fate of Pat Gregory: A Disinterred Native American. Journal of Forensic Sciences, 45(2), 14713J. https://doi.org/10.1520/JFS14713J

NamUs | Home. (n.d.). https://namus.nij.ojp.gov/

New Mexico Decedent Image Database—Welcome Home. (n.d.). Retrieved October 26, 2020, from https://nmdid.unm.edu/

Oettlé, A. C., Demeter, F. P., & L'abbé, E. N. (2017). Ancestral Variations in the Shape and Size of the Zygoma. The Anatomical Record, 300(1), 196–208. https://doi.org/10.1002/ar.23469

Ossenberg, N. S. (1976). Within and between race distances in population studies based on discrete traits of the human skull. American Journal of Physical Anthropology, 45(3), 701–715. https://doi.org/10.1002/ajpa.1330450337

Ousley, S., & Jantz, R. (2013). Fordisc 3. Rechtsmedizin, 23(2), 97–99. https://doi.org/10.1007/s00194-013-0874-9

Paschetta, C., de Azevedo, S., Castillo, L., Martínez-Abadías, N., Hernández, M., Lieberman, D. E., & González-José, R. (2010). The influence of masticatory loading on craniofacial morphology: A test case across technological transitions in the Ohio valley. American Journal of Physical Anthropology, 141(2), 297–314. https://doi.org/10.1002/ajpa.21151

Phenice TW. A newly developed method of sexing the os pubis. American Journal of Physical Anthropology 1969; 30: 297– 302.

Pilloud, M. A., Adams, D. M., & Hefner, J. T. (2019). Observer error and its impact on ancestry estimation using dental morphology. International Journal of Legal Medicine, 133(3), 949–962. https://doi.org/10.1007/s00414-018-1985-3

Pilloud, M. A., Hefner, J. T., Hanihara, T., & Hayashi, A. (2014). The Use of Tooth Crown Measurements in the Assessment of Ancestry. Journal of Forensic Sciences, 59(6), 1493–1501. https://doi.org/10.1111/1556-4029.12540

Pilloud, M. A., Maier, C., Scott, G. R., & Edgar, H. J. H. (2018). Molar crenulation trait definition and variation in modern human populations. HOMO, 69(3), 77–85. https://doi.org/10.1016/j.jchb.2018.06.001

Rhine S. 1990. Nonmetric skull racing. In: Gill G and Rhine S, editors. Skeletal attribution of race: methods for forensic anthropology. Albuquerque: University of New Mexico. p. 9-20.

Rosay, A. B. (2016, June 1). Violence Against American Indian and Alaska Native Women and Men. National Institute of Justice. https://nij.ojp.gov/topics/articles/violence-against-american-indian-and-alaska-native-women-and-men

Ross, A. H., & Pilloud, M. (2021). The need to incorporate human variation and evolutionary theory in forensic anthropology: A call for reform. American Journal of Physical Anthropology, 176(4), 672–683. https://doi.org/10.1002/ajpa.24384

Ross, A. H., & Williams, S. E. (2021). Ancestry Studies in Forensic Anthropology: Back on the Frontier of Racism. Biology, 10(7), Article 7. https://doi.org/10.3390/biology10070602

Salam, M. (2019). Native American Women Are Facing a Crisis. The New York Times. https://www.nytimes.com/2019/04/12/us/native-american-women-violence.html

Saramo, S. (2016). Unsettling Spaces: Grassroots Responses to Canada's Missing and Murdered Indigenous Women During the Harper Government Years. Comparative American Studies An International Journal, 14(3–4), 204–220. https://doi.org/10.1080/14775700.2016.1267311

Sauer, N. J. (1992). Forensic anthropology and the concept of race: If races don't exist, why are forensic anthropologists so good at identifying them? Social Science & Medicine, 34(2), 107–111. https://doi.org/10.1016/0277-9536(92)90086-6

Scott, R. G., & Turner, C. G. (1997). The Anthropology of Modern Human Teeth: Dental Morphology and Its Variation in Recent Human Populations 11. New York: Cambridge University Press, 382 pp.

Scott, G. R., & Piiloud, M. A., Navega, D., d'Oliveira Coelho, J., Cunha, E., &Irish, J. D. (2018). rASUDAS: A new web-based application for estimating ancestry from tooth morphology. Forensic Anthropology, 1, 18-31.

Slakoff, D. C., & Fradella, H. F. (2019). Media Messages Surrounding Missing Women and Girls: The "Missing White Woman Syndrome" and Other Factors that Influence Newsworthiness. Criminology, Criminal Justice, Law & Society, 20(3), 80–102.

Snow, M. H., Durand, K. R., & Smith, D. G. (2010). Ancestral Puebloan mtDNA in context of the greater southwest. Journal of Archaeological Science, 37(7), 1635–1645. https://doi.org/10.1016/j.jas.2010.01.024

Sommers, Z. (2016). Missing White Woman Syndrome: An Empirical Analysis of Race and Gender Disparities in Online News Coverage of Missing Persons. The Journal of Criminal Law and Criminology (1973-), 106(2), 275–314.

Sovereign Bodies Institute. (2022). MMIP database. https://www.sovereign-bodies.org/mmiw-database

Spiros, & Hefner, J. T. (2020). Ancestry Estimation Using Cranial and Postcranial Macromorphoscopic Traits. Journal of Forensic Sciences, 65(3), 921–929. https://doi.org/10.1111/1556-4029.14231

Spradley, M. K. (2016). Metric Methods for the Biological Profile in Forensic Anthropology: Sex, Ancestry, and Stature. Academic Forensic Pathology, 6(3), 391–399. https://doi.org/10.23907/2016.040

Spradley, Stull, & Hefner. (2016). Craniofacial Secular Change in Recent Mexican Migrants. Human Biology, 88(1), 15. https://doi.org/10.13110/humanbiology.88.1.0015

Stull, K. E., Bartelink, E. J., Klales, A. R., Berg, G. E., Kenyhercz, M. W., L'Abbé, E. N., Go, M. C., McCormick, K., & Mariscal, C. (2021). Commentary on: Bethard JD, DiGangi EA. Letter to the Editor—Moving beyond a lost cause: Forensic anthropology and ancestry estimates in the United States. J Forensic Sci. 2020;65(5):1791–2. doi: 10.1111/1556-4029.14513. Journal of Forensic Sciences, 66(1), 417–420. https://doi.org/10.1111/1556-4029.14616

Tallman, S. D. (2016). The Evaluation and Refinement of Nonmetric Sex and Ancestry Assessment Methods in Modern Japanese and Thai Individuals.

Thomas, R. M., Parks, C. L., & Richard, A. H. (2017). Accuracy Rates of Ancestry Estimation by Forensic Anthropologists Using Identified Forensic Cases. Journal of Forensic Sciences, 62(4), 971–974. https://doi.org/10.1111/1556-4029.13361

Tise, M. L., Kimmerle, E. H., & Spradley, M. K. (2014). Craniometric Variation of Diverse Populations in Florida: Identification Challenges Within a Border State. Annals of Anthropological Practice, 38(1), 111–123. https://doi.org/10.1111/napa.12046

Tsai, J., Cerdeña, J. P., Khazanchi, R., Lindo, E., Marcelin, J. R., Rajagopalan, A., Sandoval, R. S., Westby, A., & Gravlee, C. C. (2020). There is no 'African American Physiology': The fallacy of racial essentialism. Journal of Internal Medicine, 288(3), 368–370. https://doi.org/10.1111/joim.13153

Turner CG II. Teeth and prehistory in Asia. Sci Am 1989;260(2):88–96. 37. Turner CG II. Major features of Sundadonty and Sinodonty, including suggestions about East Asian microevolution, population history and late Pleistocene relationships with Australian Aboriginals. American Journal of Physical Anthropology 1990;82(3):295–317.

Turner C. G., Nichol, C. R., Scott, G. R. (1991) Scoring procedures for key morphological traits of the permanent dentition: the Arizona State University Dental Anthropology System. In: Kelley, M. A., Larsen, C. S. (eds) Advances in dental anthropology. Wiley-Liss, New York, pp 13-31.

United States Supreme Court. 1993. Daubert v. Merrell Dow Pharmaceuticals, Inc. 509 U.S. 579.

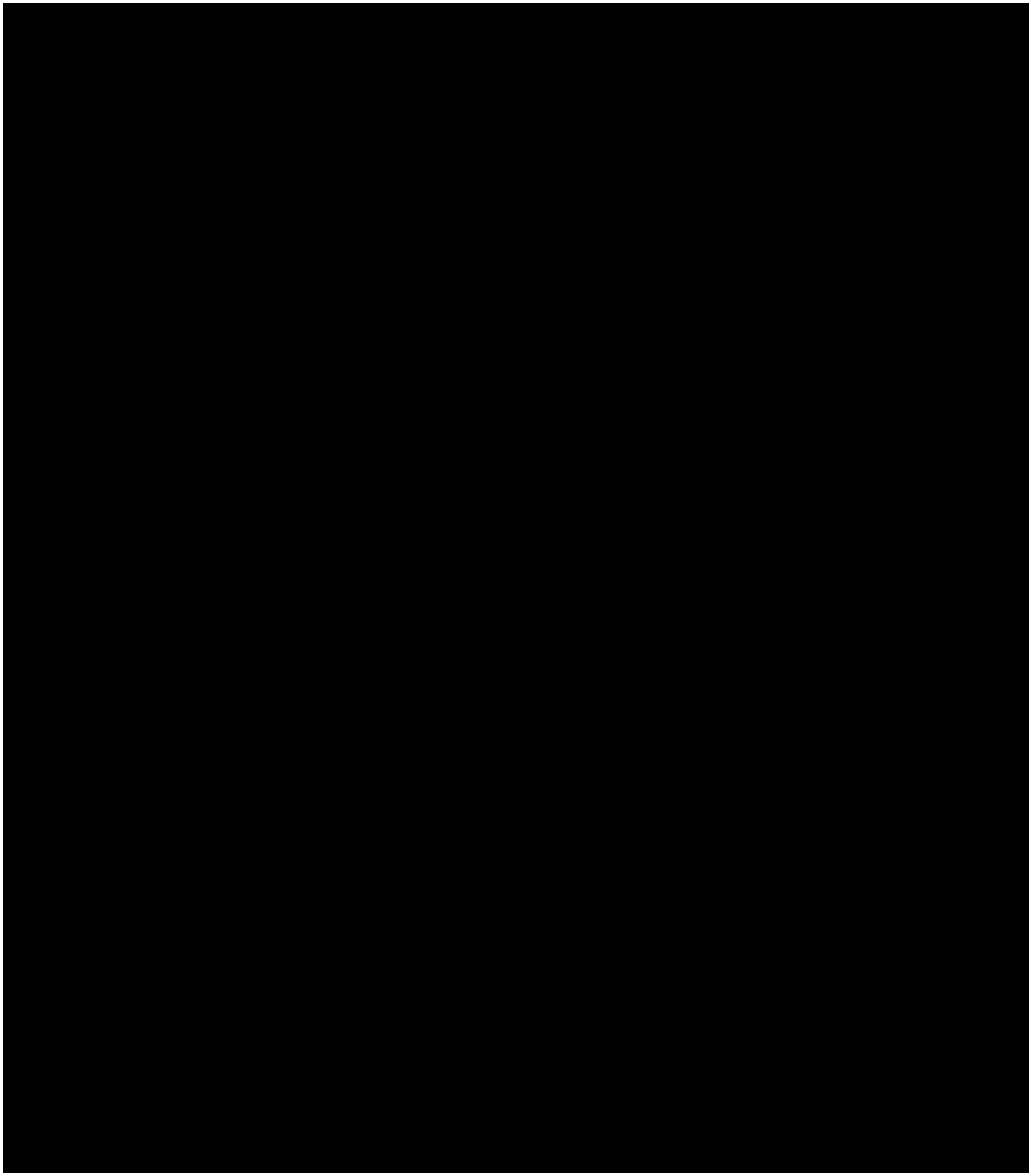
Weisberg. (2014). Willful ignorance. Wiley. https://doi.org/10.1002/9781118839539

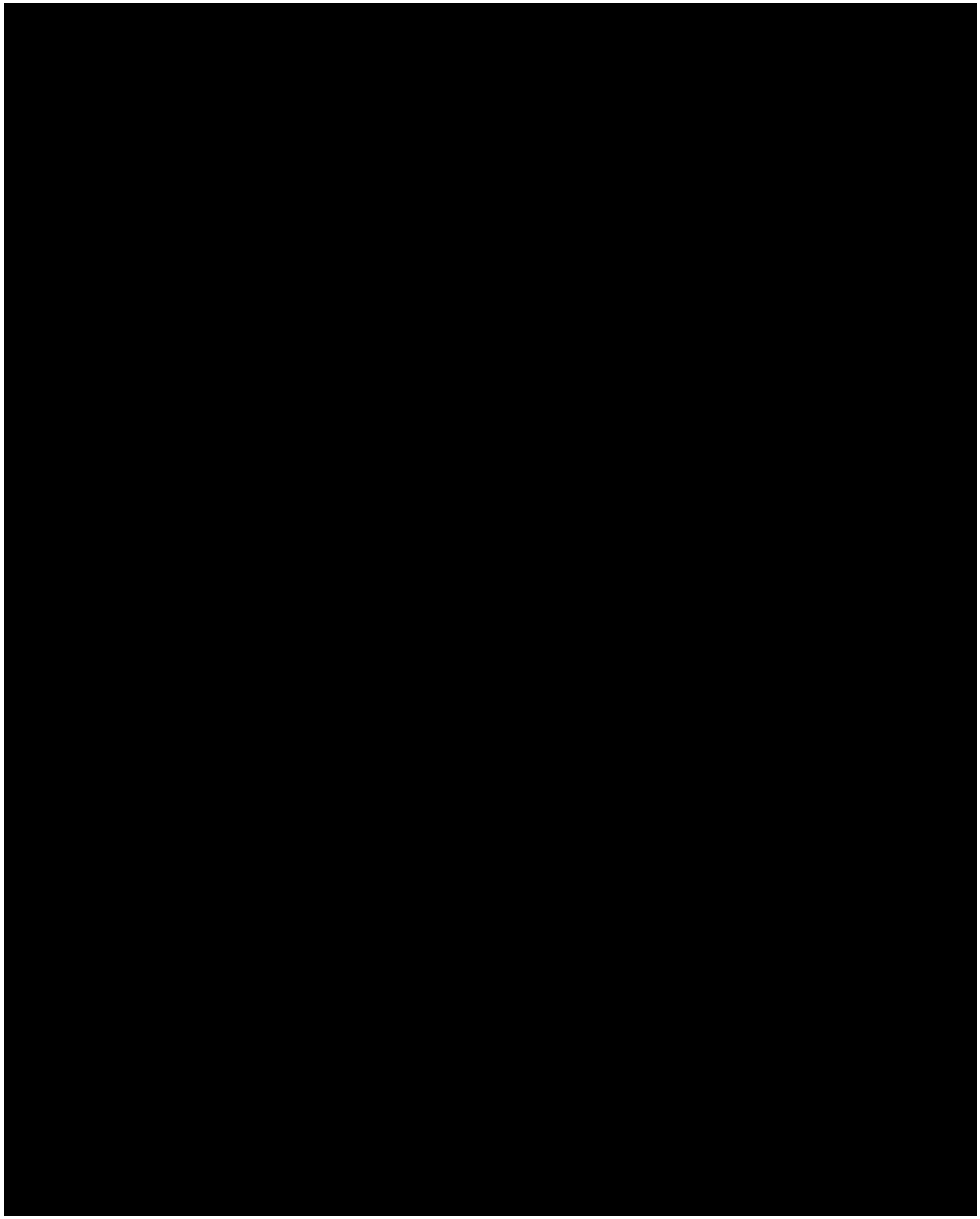
Wescott. (2005). Population Variation in Femur Subtrochanteric Shape. Journal of Forensic Sciences, 50(2), JFS2004281–8. https://doi.org/10.1520/JFS2004281

Winburn, A. P., & Clemmons, C. M. J. (2021). Objectivity is a myth that harms the practice and diversity of forensic science. Forensic Science International: Synergy, 3, 100196. https://doi.org/10.1016/j.fsisyn.2021.100196

Williams, Belcher, R., & Armelagos, G. (2005). Forensic Misclassification of Ancient Nubian Crania: Implications for Assumptions about Human Variation. Current Anthropology, 46(2), 340–346. https://doi.org/10.1086/428792

CURRICULUM VITAE





References

Dr. Matthew Laird, RPA
The James River Institute for Archaeology
223 McLaws Circle, Suite 1
Williamsburg, VA 23185
Phone: (757) 880-3302

Jacob Tumelaire, RPA, Principal Investigator
Independent Archaeological Consulting, llc
34 Dover Point Road, Suite 300
Dover, NH 03820
Phone: (603) 430-2970
Email: jtumelaire@iac-llc.net

Jesse Cofelice, RPA, Project Manager/Principal Investigator
Independent Archaeological Consulting, llc
34 Dover Point Road, Suite 300
Dover, NH 03820
Phone: (603) 430-2970
Email: jcofelice@iac-llc.net

Dr. Thomas Crist, Professor
Department of Physical Therapy
Utica College
1600 Burrstone Road
Utica, NY 13502
Phone: (315) 792-3390
Email: tcrist@utica.edu