University of Nevada, Reno

Analyzing the Impact of Population Level Differences and Socioeconomic Status on Subadult Age Estimation

A thesis submitted in partial fulfillment
of the requirements for the degree of

Bachelor of Arts in Anthropology and the Honors Program

by

Samantha M. Gonzalez

Dr. Kyra Stull, Thesis Advisor

December, 2018

UNIVERSITY OF NEVADA RENO

THE HONORS PROGRAM

We recommend that the thesis
prepared under our supervision by

SAMANTHA M GONZALEZ

entitled

ANALYZING THE IMPACT OF POPULATION LEVEL DIFFERENCES AND SOCIOECONOMIC STATUS ON SUBADULT AGE ESTIMATION

be accepted in partial fulfillment of the
requirements for the degree of

BACHELORS OF ARTS

Kyra E. Stull, PhD, Thesis Advisor

Tamara Valentine, Ph.D., Director, Honors Program

December, 2018

Abstract

Age estimation is usually the sole contribution to the biological profile of a skeletally immature individual. Forensic anthropologists use dental and skeletal development to estimate age. It is regularly stated that there needs to be population-specific methods of age estimation that encompass population level and socioeconomic status differences (Esan, Yengopal, & Schepartz, 2017). While both population differences and SES is recognized as influencing growth, it is not known at what age the population differences or SES differences become apparent or influential to age estimations. The purpose of this study is to compare the skeletal and dental formations of South African and United States individuals aged 0-6 as a means to determine if differences in diaphyseal length and molar formation exist between the populations and if so, to quantify the magnitude. Furthermore, the study will create population specific and global models and test the accuracy and precision of MARS models.

Acknowledgements

Special thanks must be given to Kyra Stull, who has been a trusted mentor for two years and a wonderful thesis mentor for one year. You've helped me see things in myself I hadn't thought to look for before, and that, above all the instructional teaching, was the most valuable thing you've given me.

My friends and family have always been my support group, and Cyn and Mom deserve special shout-outs. Thanks to Mom for picking up the phone even at 3 in the morning, and thanks to Cyn for *everything*.

My writing consultant, Melissa, has been a wonderful ray of sarcastic sunshine that brightened every Tuesday morning. Many thanks to a wonderful consultant who helped me stay on track during a STRESSFUL time.

Table of Contents

Introduction

When analyzing unidentified human remains, a forensic anthropologist will estimate the biological profile to narrow the list of possible individuals and facilitate identification. The adult biological profile includes biological sex, age, stature, ancestry, pathology, trauma, and identifying/unique characteristic traits, such as medical implantations (Márquez-Grant, 2015). In contrast to adult biological profiles, subadult biological profiles generally consist of age alone. Since human growth is relatively predictable and stature and age have a strong correlation, stature can be estimated in place of age. However, application of adult stature methods on subadults consistently underestimate the height of subadults (Brits, Manger, & Bidmos, 2017; Cardoso, 2009a). Sex estimations have been problematic in the past on subadults, due to the low level of sexual dimorphism found in subadults compared to adults (Cardoso, 2008b; Klales & Burns, 2017; Stull, L'Abbé, & Ousley, 2014).

Human growth and development has been a source of scientific interest as early as the 18th century. One of the first seminal publications on human growth and development was the 1835 publication by Adolphe Quetelet, *A Treatise on Man and the Development of his Faculties*. Quetelet, a prominent scientist of his time, introduced statistical methods to the study of human growth and development, and was the first scientist to introduce the bell curve to represent normality (Quetelet, 1841). Concurrently, Count du Montbeillard measured his son every six months, from birth to 18 years and presented one of the first studies of longitudinal growth (Quetelet, 1841). The impact of this introduction of human growth research by Quetelet cannot be underestimated. He gave researchers the foundation to establish a scientific field of study. A century later came another seminal work in the field of human growth and development, Todd's atlas of skeletal maturation. The book was published in 1937 after a decade of studying radiographs of over 1,000 skeletally immature individuals (Todd, 1937). The goal of the study was to create a useful and useable method to estimate age of a living individual. The method is extremely useful for living individuals, but the study itself was influential by introducing methodological ways to study growth and development for biological anthropology. Other studies of human growth, such as Quetlet (1841), influenced many other fields to incorporate human growth and development studies, such as economics, nutrition, and biological anthropology. The large gap in time between the works of Quetlet and Todd can be explained by the difference in goals of biological anthropology and human development studies at the time. At the time, biological anthropology was far more concerned with abnormal human growth while human development studies were concerned with tracking normal patterns of human growth.

Population level differences can be seen in daily life by looking at hair color, skin color, height differences, and many other traits. These can be shown biologically through various traits ranging from skin color to morphological differences in the human skeleton. Most sources use population specific methods for age, sex, or stature estimation (Colman et al., 2018; Jeong & Jantz, 2016; Singal & Sharma, 2017), but very few pose a global, or universal, model due to the socioeconomic status and population differences.

This project will assist the field of forensic anthropology by confirming the applicability of a global model for subadults under 6 years. After this has been established, further research to increase the precision of global models for subadults can occur. These global models can be used to estimate the age of a subadult when the ancestry is unknown. This will be especially useful in the modern age with the amount of war refugees emigrating from the violence of their area. This paper attempts to quantify the level of socioeconomic status and population level differences and asses the validity of a global model.

Literature Review

Age estimation is one of the tools of a biological anthropologist. Estimating the age at death of an individual can give information about the circumstances around the death of an individual, inform a possible identification, and it can also be used to establish the age of a living individual whose age is in question. Age estimation comes in many different methods, from epiphyseal[1] presence and fusion, to the shape of the pubic symphysis (Shirley et al. 2013, Stull et al. 2014). Age estimation of adults differs from subadults based on the concept of whether the indicator is developing or degenerating over time. While these two ideas are considerably general, almost all methods to estimate age look at either patterns of growth or degradation of a specific indicator (H. F. V. Cardoso, Spake, & Liversidge, 2016; H. F. V Cardoso, 2008; Márquez-Grant, 2015; Passalacqua, 2013; Ríos & Cardoso, 2009; Sarajlić & Gradaščević, 2012; Shirley, Fazlollah, & Tersigni-Tarrant, 2013; Stull, L'Abbé, & Ousley, 2014b).

There are two broad life periods, namely growth and maintenance/breakdown. When creating methods to estimate the age of subadults, researchers look at patterns of growth in skeletal and dental elements. When estimating age for adults, researchers look at the patterns of degradation in skeletal elements. For instance the face of the pubic symphysis is not an accurate age indicator until the individual has reached skeletal maturity, because the level of degradation and presence of new morphology, such as lipping, does not occur until the maintenance/degeneration phase of life (Brooks &

[1] An epiphysis is the spongy end of a long bone (femur, humerus, tibia, etc.) that ossifies on its own from the shaft of the diaphysis and fuses at various known times in an individual's life.

Suchey, 1990; Sarajlić & Gradaščević, 2012). There are some age indicators for adults that do not measure degeneration over time, rather they measure generation of a sign of degradation over time, such as osteoarthritis. Adult age indicators range from the closure of cranial sutures to the morphology of the surface of joints such as the pubic symphysis and hip joint (Corron et al., 2018; Shirley et al., 2013).

One age estimator used on the adult that can also be used to estimate the age of a subadult is the medial clavicular epiphysis. The medial clavicular epiphysis is on the end of the clavicle (the collar bone) that meets the sternum (the breastbone). This epiphysis on the sternal end of the clavicle is an ossification site in the human body and fuses last, between 22 and 30 years of age. This method estimates the age of an individual by scoring the fusion site according to the stages defined in Shirley-Jantz (2010), with age ranges that accompany each stage. The Shirley-Jantz method has high precision and accuracy estimating the age of older adolescents and young adults, individuals 16 to 30 (Langley-Shirley & Jantz, 2010; Shirley et al., 2013). This method of age estimation illustrates the blurred line between subadult and adult age estimation. Adult age estimations were the primary focus of past biological anthropologists, and therefore had a distinct influence on how subadult age estimation was studied methodologically.

Subadult Age Estimation

Subadult age indicators range from the length of bones, such as the femur and clavicle, to the ossification of secondary ossification sites, such as the hand and wrist (H. F. V. Cardoso, 2008a, 2009b; H. F. V. Cardoso & Ríos, 2011; H. F. V. Cardoso, Spake, & Humphrey, 2017; H. F. V. Cardoso et al., 2016; Corron et al., 2018; Franklin, 2010).

The most widely used indicators for subadult age estimation with high accuracy are dental formation of molars, diaphyseal lengths, and epiphyseal fusion (Franklin, 2010; Márquez-Grant, 2015). Dental formation of molars is the growth process of molars with specific interest in length and root/cusp formation. This is measured on x-rays with various detailed scoring methods that are widely used (Alqahtani, Hector, & Liversidge, 2014; H. F. V. Cardoso, 2007, 2009b; Conceição & Cardoso, 2011; Heuzé & Cardoso, 2008). Dental eruption, on the other hand, is measured as the eruption of a tooth through the alveolar bone (Franklin, 2010; Márquez-Grant, 2015), and is extremely variable and is therefore not used to estimate age. Diaphyseal length is the length of the shafts of long bones, such as the femur, tibia, humerus, and radius (H. F. V. Cardoso, Abrantes & Humphrey, 2014; H. F. V Cardoso, 2008; Stull, L'Abbé & Ousley, 2014). It is used to estimate age because of the strong correlation between stature and age. Epiphyseal fusion is the process of secondary ossification centers, such as the ends of the radius and ulna, fusing at known times throughout childhood and adolescence (Márquez-Grant, 2015; Shirley et al., 2013). This project will use dental formation and diaphyseal length to estimate subadult ages across two populations.

Socioeconomic Status and Population Level Differences

Populations are defined as groups of people with similar socioeconomic status, genetic backgrounds, similar ancestries, and similar economic insults[2] due to the socioeconomic status of the region (Krieger, 2012). Socioeconomic status can be

[2] Insults: in the context of growth and development, an insult is the result of a negative stressor, such as poor nutrition or disease.

measured through methods known as the Human Development Index (HDI) and the Gini Coefficient (Programme, 2008; UNDP, 2018). The HDI of a region is calculated through life expectancy at birth, expected years of schooling, and gross net income per family (Jahan, 2015). Another economic statistic that captures SES differences is the Gini Coefficient. Specifically, the Gini coefficient is used to measure the inequality of a region through examination of the distribution of wealth of the area (Petrie & Tang, 2008). These factors inform the socio-economic status of populations.

Historically, population specific models have been favored to estimate the age of all individuals, subadult or adult (Jeong & Jantz, 2016; Colman et al., 2018). While population specific models are believed to be more precise than global models, they are problematic for an individual of an unknown population. Population specific models have been specifically created for one population and tested on individuals from the same population. For example, a researcher will create a model using a sample of French individuals and test it on a different sample of French individuals to evaluate the precision and accuracy of the model. On the contrary, global, or universal, models are trained using individuals from all or many populations, and tested on individuals from any population within the training sample. For example, a researcher will create a model using a sample of Czech, Hungarian, French, and Danish individuals. Next, the researcher will test the model on a set of individuals encompassed within the global model (*i.e.*, a different group of Hungarian, Czech, Danish, or French individuals). The application of a global model has generally been seen to be less precise in estimation of age, with assumptions of larger prediction intervals and lower accuracies, however publications have shown global models to not only be successful but more precise because of the larger sample size (Stull & Price, 2017). In an attempt to look more closely at global models, some researchers have proposed exploring subadults in specific age ranges to more closely evaluate the onset of differences in populations or SES groups rather than subsuming the entire time period of growth into one analysis (Gonzalez & Stull, 2018). According to Gonzalez and Stull (2018), the economic insults that arise from low socioeconomic status do not accumulate until after 6 years of age. Similarly, the stabilization of the size of infants within a womb to prevent harm to the mother during birth (*i.e.*, stabilizing selection), is one of the potential reasons the level of population differences seen in adults is not apparent in individuals less than 6 years of age. To clarify, the authors conclude that the growth of an individual is not substantially impacted by the environment until a certain age, and before this point a universal model can be used to estimate the age of an individual.

Human growth is affected by many factors, such as the environment, ancestry, and socioeconomic status. Socioeconomic status (SES) is social standing, or class standing, measured by one's education, income, and occupation (H. F. V. Cardoso, 2007; Conceição & Cardoso, 2011; Howe et al., 2012; McCrory et al., 2017). In populations of different SES levels, the heights of subadults differed; those in higher SES communities are taller than those in lower SES communities (de Onis & Branca, 2016; Howe et al., 2012; McCrory et al., 2017). In contrast, very little variation is found in birth weight across racial and regional boundaries, and the differences found were more likely as a result of reduced average gestational time found across regions (de Onis & Branca, 2016; Howe et al., 2012; McCrory et al., 2017). Since birth weight is non-indicative of childhood well-being and economic status, height is supported by a number of regional studies to be a valid indicator of biological stress (Conceição & Cardoso, 2011; de Onis & Branca, 2016; McCrory et al., 2017).

The goal of this research is to evaluate the difference in diaphyseal lengths and molar formations of individuals from the United States and individuals from South Africa under the age of 6 years. Diaphyseal lengths of the femur, tibia, humerus, and ulna will be measured digitally on computed tomography (CT) scans or Lodox Statscan radiographic images. First mandibular and first maxillary molars will be scored according to AlQahtani and colleagues (2010) within the same digital scans. Multivariate adaptive regression splines (MARS) will be employed to create population specific and global models to estimate the age of an individual using the above traits. A MARS model is beneficial for age estimation of subadults due to the non-linear pattern of human growth. Additionally, MARS models have fluctuating prediction intervals, giving the model the ability to adapt to increased variation with age, or heteroscedasticity[3], which is also a trademark of growth. Additionally, MARS models allow for both continuous and ordinal data to be compared, which allows for comparison of long bone lengths (continuous) and molar formation (ordinal).

The expected results for this project are that 1) individuals from South Africa and the United States will not have statistically significant differences in diaphyseal lengths, 2) molar formation will not be statistically significantly different between individuals from the United States and South Africa, and 3) global models can be created with high precision for individuals under 6 years even when the sample is from geographically and

[3] Scedasticity: the amount of variation in data. Data with the same amount of variation over time is homoscedastic and data with different amounts of variation over time is heteroscedastic

socioeconomically diverse groups, such as the United States and South Africa. This research will impact the field of forensic anthropology by confirming the applicability of a global model for subadults under 6 years. After this has been established, further research to increase the precision of global models for subadults can occur.

Methods

Sample Collection

A total of 553 subadults were used in this project, which consisted of 333 individuals from the United States and 220 individuals from South Africa. Computed Tomography (CT) scans of individuals from the United States were collected at the University of New Mexico Health Sciences Center and Office of the Medical Investigator. Radiographs of individuals from South Africa were collected at Red Cross Children's Hospital in Cape Town. All individuals chosen for this project were between birth and 6 years of age (K. Stull, personal communication, February 2, 2018). All individuals chosen for this project were not failure to thrive individuals, and outliers as defined by outlier tests (individuals more than three standard deviations away from the mean), and I examined all "outliers" as defined by the tests and removed the extreme outlier individuals.

These two populations are excellent candidates for a comparison and validity test of global models due to a number of significant factors, including mortality[4] of the individual, socioeconomic status, and ancestry. The South African population is a sample of living children, while the institutions in the United States drew from forensic samples. Furthermore, the major ancestry of the United States is Western European, while the major ancestry of South Africa is Sub-Saharan African.

The South African sample was collected using Lodox Statscan, a device created specifically for the diamond mining industry in the area. It is a fast-acquisition radiograph

---

[4] Whether a person is living or dead

device that can generate images with minimal distortion, allowing for measurement of diaphyseal lengths using the accompanying software (Stull et al., 2014a; Stull, L'Abbé, & Steiner, 2013). Molar stages of formation were collected according to a modified scoring system based on AlQahtani and colleagues (2012). The system uses 13 different points of identification and evaluation of molar formation (see Table 2) (AlQahtani, Hector, & Liversidge, 2010).

Table 1

Age Distributions of Samples.

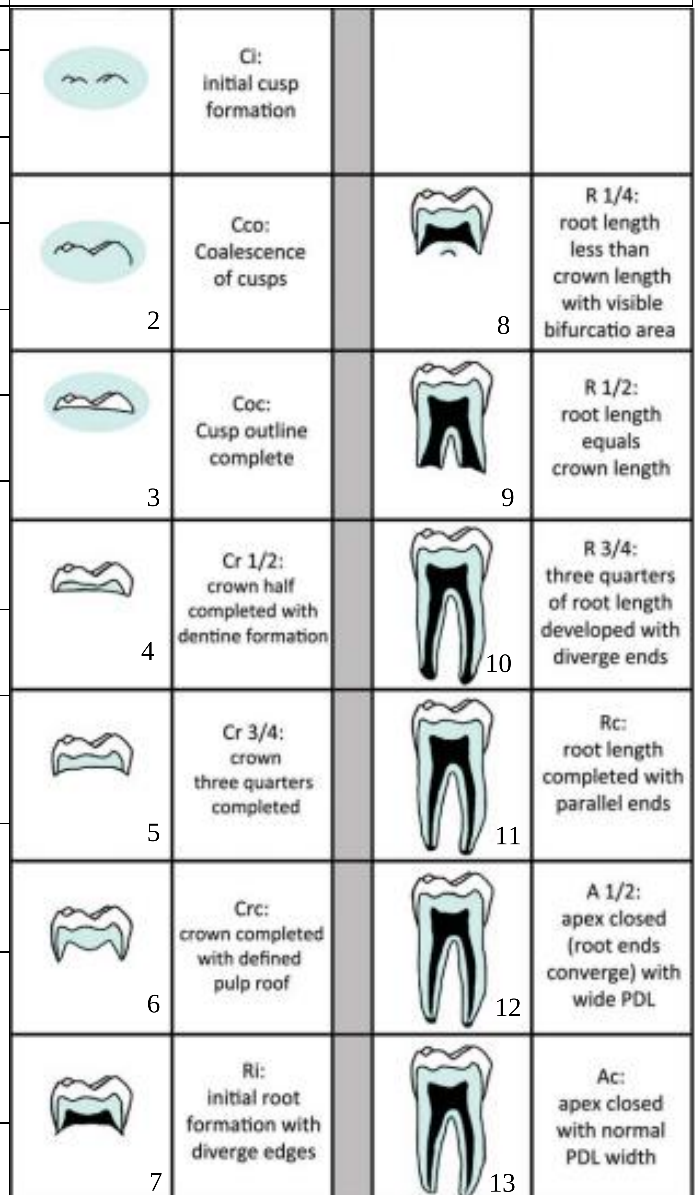
Age	South African	United States
0.0-0.5	1	135
0.5-1.0	3	40
1.0-1.5	10	29
1.5-2.0	19	27
2.0-2.5	18	14
2.5-3.0	39	18
3.0-3.5	19	15
3.5-4.0	21	11
4.0-4.5	23	11
4.5-5.0	20	13
5.0-5.5	24	12
5.5-6.0	23	8
Total	220	333

Table 2

Molar Development Stages

Stage Number	**AlQahtani Explanation of Stage**
1	initial cusp formation
2	coalescence of cusps
3	cusp outline complete
4	crown half completed with dentine formation
5	crown three quarters completed [1]
6	crown completed with defined pulp roof
7	initial root formation with diverge edges
8	root length less than crown length with visible bifurcation area
9	root length equals crown length
10	R ¾: three quarters of root length developed with diverge ends
11	Rc: root length completed with parallel ends
12	A ½: apex closed (root ends converge) with wide periodontal ligament space
13	Ac: apex closed with normal periodontal ligament space width

Original AlQahtani Stages



Note. Molar development classification system for scoring molar formation. Figures on right (AlQahtani et al., 2010)

MARS

MARS is a regression analysis that does not require linearity or homoscedasticity for its data, which is the reason it is used for this project, as human growth is known to be nonlinear and heteroscedastic (Corron, Marchal, Condemi, Chaumoître, & Adalian, 2017; Friedman, 1997; Milborrow, 2018; Stull et al., 2014a). Linear regressions are a form of statistical analysis that fit the best equation to explain the data (Figure 1). By fitting small sections of linear regressions to the data, MARS creates a "hinged" non-linear model to fit the data more precisely than a standard linear regression (Figure 1). A hinge is the location at which the regression equation changes.

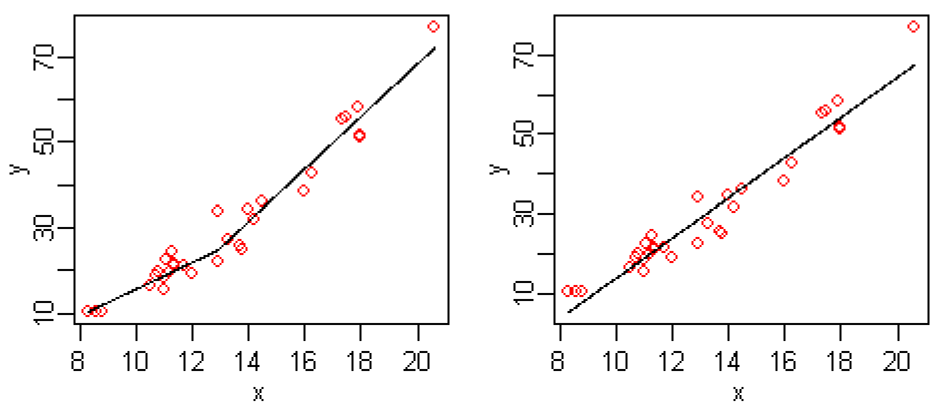


Figure 1. MARS vs linear regressions. This image shows the difference between a MARS model (left) and a simple linear regression (right) using the same data.

MARS aims to find the relationship between variables in a sample by subdividing it into regions and estimating the regression equation for each region. MARS outputs are a hinge function that combines the most accurate regression equations for each region (Corron et al., 2017; Friedman, 1991; Milborrow, 2018; Stull et al., 2014a). Hinge functions combine multiple equations, and the correct equation is chosen for each datum based on which equation will yield a positive result. This results in one equation that hinges the most accurate regression equations for certain age cohorts, rather than one regression equation for all ages (Figure 1). MARS models are unique in that the accuracy of a model can be controlled, affecting the precision of the model. Accuracy is the closeness of a predicted value to the true value, precision is when a model repeatedly predicts values in a small area (Supplementary Figure 1). For the ideal age estimation method, the model would be accurate and precise, but in reality, accuracy and precision often have a negative relationship. For this project, I try to create the most precise age estimation models for a set accuracy: 95%.

MARS models have been proposed for use in age estimation in recent studies by forensic anthropologists (Corron et al., 2017; Stull et al., 2014a). In an article by Corron and colleagues, MARS was used to estimate the age of individuals with measurements of the ilium[5] (Corron et al., 2017). Corron and colleagues addressed the general lack of methodological consistency in subadult age estimation. They proposed a new method, using MARS models of unidimensional and bidimensional ilium measurements to estimate age. Dr. Corron collected a sample of 244 patients, from birth to twelve years that had a CT scan for any medical reason (excluding those with pathologies known to affect skeletal development or growth) at a hospital in Marseilles, France (Corron et al., 2017). A total of 176 individuals fit the requirements, and the CT scans were measured for length, width, and module (shape) of the ilium. Multivariate MARS models had

---

[5] Ilium: one of the three bones comprising the hipbone

increased precision compared to univariate or bivariate models, as presented in previous works on subadult age estimation (Corron et al., 2017; Stull et al., 2014a). All models had very strong cross-validated R-squared values, and overall the ranges for prediction intervals were smaller than previous models such as linear or polynomial regressions, indicating its validity for subadult age estimation models (Corron et al., 2017).

Statistical Analysis

In order to test the level of population differences between the United States (US) and South Africa (SA), I performed Welch two sample t tests on diaphyseal length and Kruskal-Wallis tests on mandibular and maxillary molar development between both populations.

Null Hypothesis: There will be differences in the diaphyseal lengths or molar formation between US and SA populations.

Alternative Hypothesis: There will be no differences in the diaphyseal lengths or molar formation between US and SA populations.

I performed a simple correlation test to establish the correlation of diaphyseal length to age by population. In order to evaluate the difference between skeletal and dental development, I created correlation graphs of average molar formation stage by average diaphyseal length for the South African and United States populations.

Model Creation and Application

Univariate and multivariate models were created using the following age indicators: development of maxillary first molars (MxM1), development of mandibular first molars (MnM1), femur diaphyseal length (FDL), tibia diaphyseal length (TDL), humerus diaphyseal length (HDL), and radius diaphyseal length (RDL). Each age indicator was used to create a univariate model and three multivariate models were created using different combinations of the age indicators. The first multivariate model used FDL, TDL, HDL, and RDL, and will be referred to as the long bone model. The second multivariate model used MxM1 and MnM1, and will be referred to as the molars model. The final multivariate model used FDL, TDL, HDL, RDL, MxM1, and MnM1. Each univariate and multivariate model was created as a population specific model for the United States, a population specific model for South Africa, and a global model that pooled both populations.

All models were created with a prediction interval of 95%, meaning that 95% of variation in the sample is captured through the upper and prediction intervals. Therefore, with two populations of differing ancestry and socioeconomic status, there should be overlap between the upper and lower boundaries of each populations. Applying an individual of one population to the other population would lead to an incorrect age estimation if the population differences are big enough. To apply South African individuals to the United States model would be a perfect worst-case scenario by testing the age estimation of an individual using a population specific model for which that individual is not a part of. If a model is still able to correctly estimate the age of an individual outside of the population and socioeconomic status it was trained on, a global model should be much closer to the realm of possibility than previously thought. To apply the models to the other population, I added the dataset of individuals as the new data predicted by the model. By cross-applying the models in this way, an extreme validity test is performed where the percentage of individuals that was correctly estimated by age has a positive relationship with the validity of the model for global populations.

Results

Welch two sample t tests resulted in all long bones and molar formations having a p value of ≤0.05 (Table 3), which leads me to reject the null hypothesis and accept the alternative hypothesis that diaphyseal lengths and molar formations do not show a statistically significant difference between South African and United States populations.

Table 3

Testing Population Level Differences

Age indicator	Population	
	p value	Adjusted *p* value
MxM1	<2.2 e-16	--
MnM1	<2.2 e-16	--
FDL	<2.2 e-16	1.760 e-15
TDL	<2.2 e-16	1.760 e-15
HDL	<2.2 e-16	1.760 e-15
RDL	<2.2 e-16	1.760 e-15

Note. FDL (femur diaphyseal length), TDL (tibia diaphyseal length), HDL (humerus diaphyseal length), and RDL (radius diaphyseal length) *p* values found using Welch Two Sample T Test. MxM1 (development of upper first molar) and MnM1 (development of lower first molar) *p* values found using Kruskal-Wallis tests. No adjustment was needed for MxM1 and MnM1 *p* values.

Table 4

Correlation of age and each age indicator by population

Age indicator	South Africa	United States
FDL	0.92	0.87
TDL	0.90	0.87
HDL	0.90	0.87
RDL	0.89	0.87
MnM1	0.91	0.90
MxM1	0.90	0.90

Note. Pearson’s correlation values for each age indicator listed and age presented above, where 0 indicates a weak correlation and 1 indicates a strong correlation.

The Pearson's correlation test resulted in femur diaphyseal length having the greatest correlation to age for the South African population, with formation of lower and upper first molars following accordingly (Table 4). For the United States population, lower and upper first molar scored equally in correlation for age and femur diaphyseal length followed closely (Table 4). The comparison of average molar formation and average diaphyseal length showed that diaphyseal averages are higher for South Africans starting in early infancy, but the pattern shifts as childhood begins and US individuals have greater diaphyseal lengths (Figures 2 and 3).

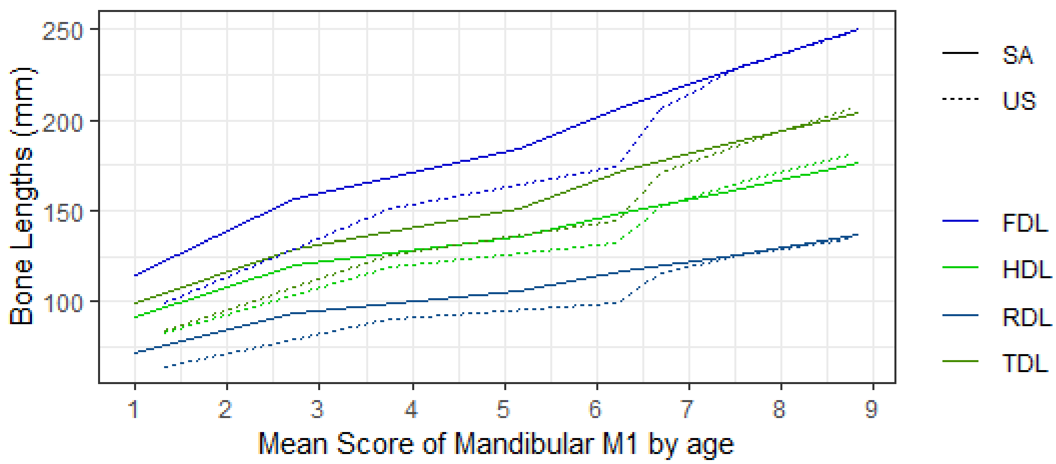


Figure 2. Correlation graph of MxM1 and diaphyseal length. The bone length of the four long bones used in this project were plotted against the average MnM1 stage.

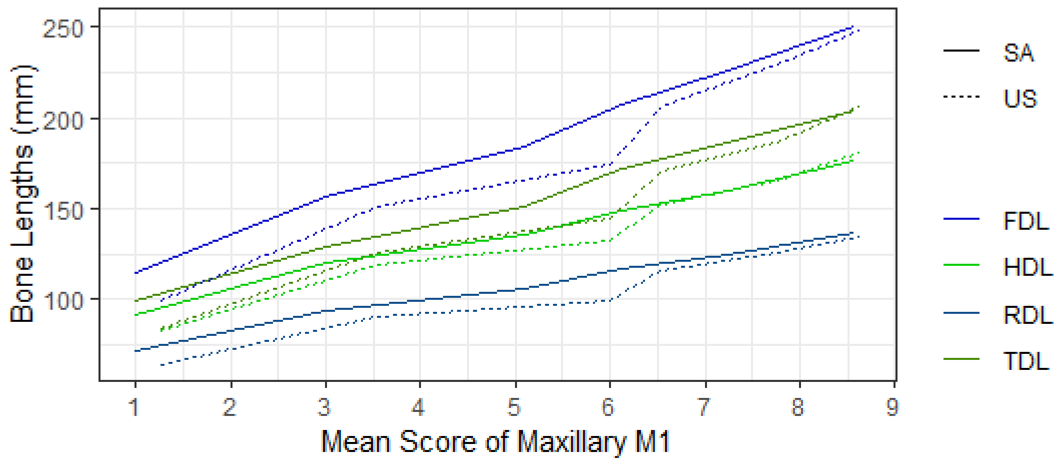


Figure 3. Correlation graph of MnM1 and diaphyseal length. The bone length of the four long bones used in this project were plotted against the average MnM1 stage.

When comparing bone lengths by population for each age cohort (*i.e.*, 1 year olds pooled, 2 year olds pooled, 3 year olds pooled, etc), it is apparent that mean bone length for US individuals overtakes South African individuals between 3 and 4 years of age (See figures 4, 5, and 6).

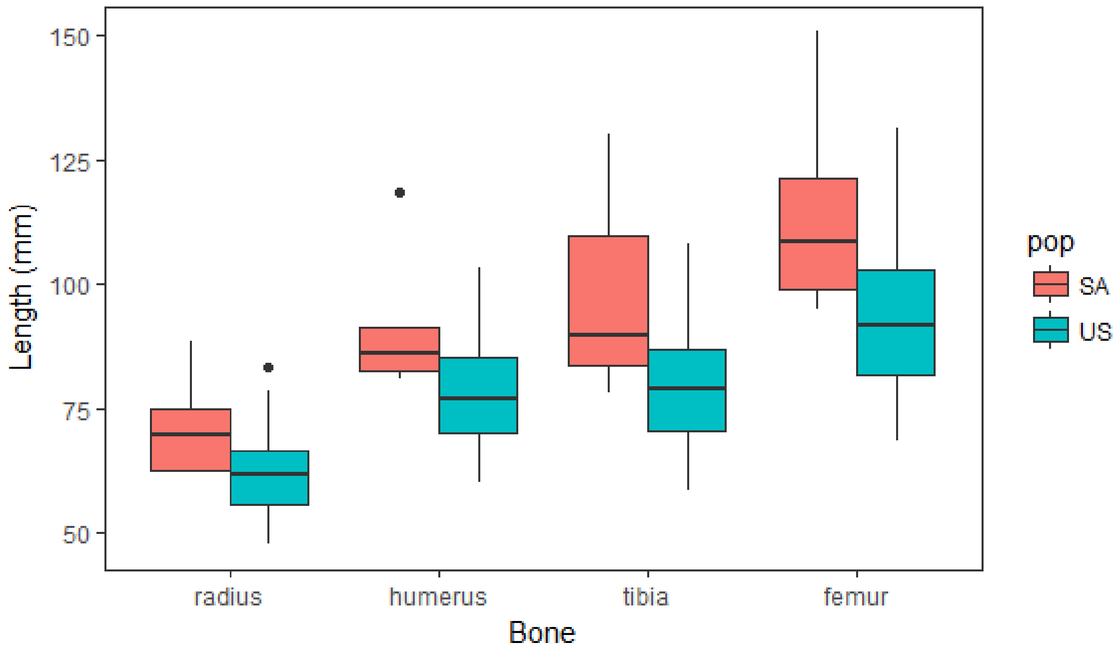


Figure 4. Box and whisker plot of long bone length by population (Ages 0-0.99). The box shows the middle 50% of individuals, the dark line through the box shows the median of bone lengths, and the lines coming off each side of the box represent the outer 50% of individuals. This shows that prior to 1 year old, South Africans have longer long bone lengths than individuals from the United States.

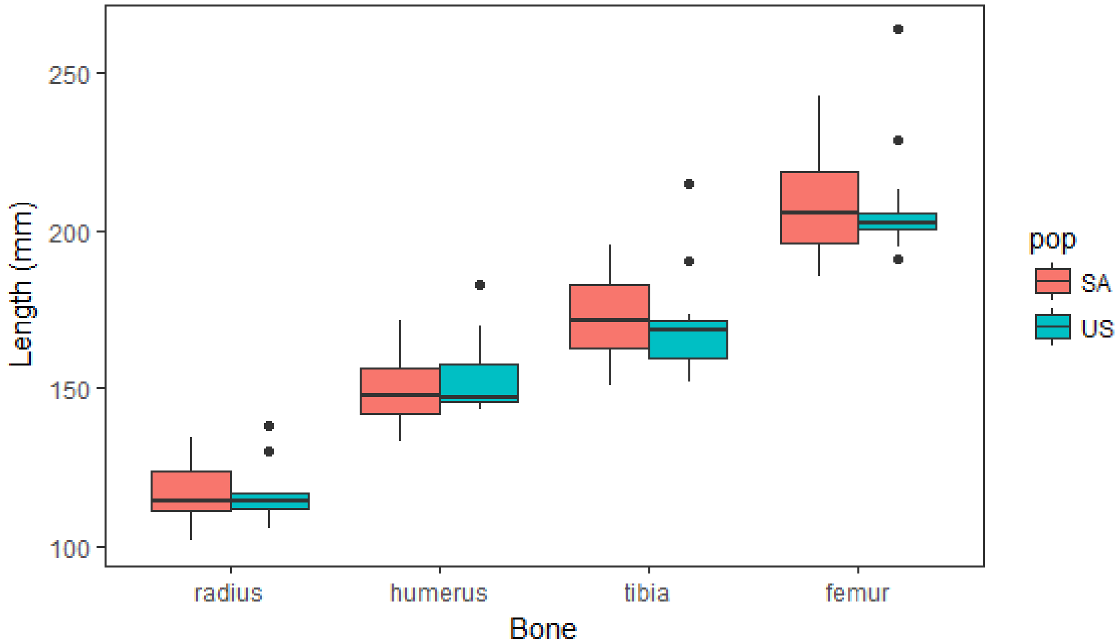


Figure 5. Box and whisker plot of long bone lengths by population (Ages 3-3.99). This graph shows that individuals from the United States are near similar to individuals from South Africa at 3 years old.

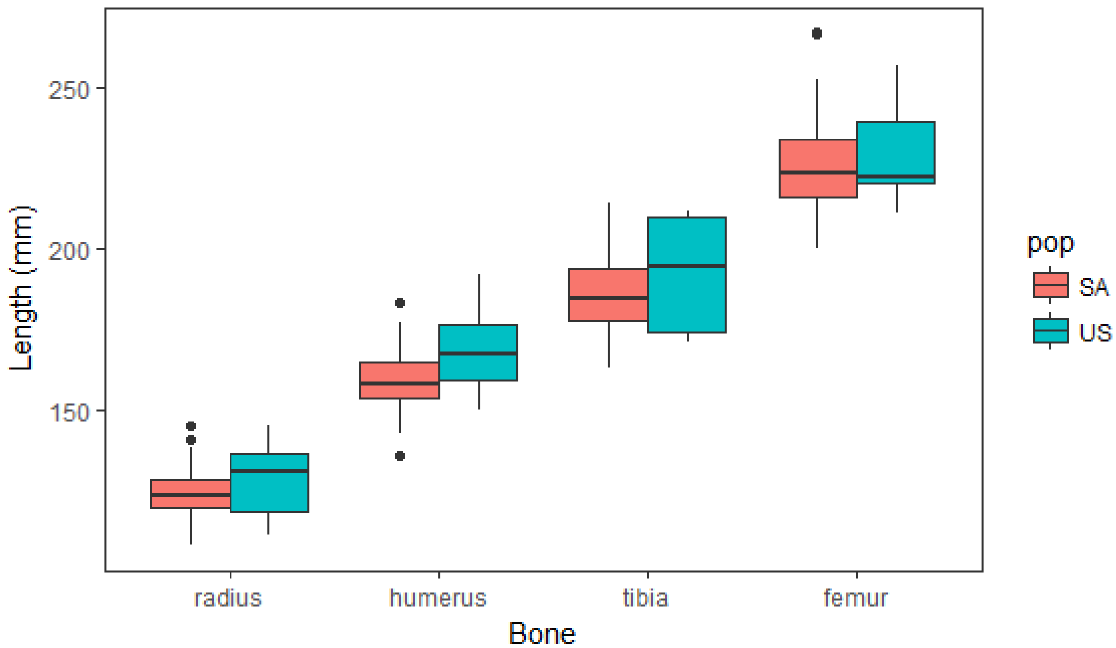


Figure 6. Box and whisker plot of long bone lengths by population (Ages 4.00-4.99). This graph shows that between ages 4 and 5 individuals from the United States have surpassed individuals from South Africa in diaphyseal length in all long bones except femur diaphyseal length.

Model Accuracy and Precision

For the United States sample, the most precise univariate model was FDL, with TDL following closely behind (Table 5), and for the South African population, MnM1 was the most precise univariate model, with FDL following closely behind (Table 5). Multivariate models proved to be more precise than univariate models, with smaller prediction intervals overall (Table 5). All models had greater precision for younger individuals with decreasing precision as age increased (Figures 7 and 8). Global models were more precise in one model only, and only more precise than the United States model, not the South African model (Table 4). All models are presented with prediction interval (PI) minimum and maximums, indicating the size of the smallest and largest prediction interval ranges. The Cross-Validated R Squared (CVR^2) value is reflective of the accuracy if used on a data set outside of the training material, where 0 is low and 1 is high. Plots of each univariate model can be found in Supplementary Figures, multivariate models were not plotted, as only one variable would be plotted at a time.

Table 5

Prediction Intervals of all MARS Models

	South Africa			United States			Global Model		
	PI Min	PI Max	CVR^2	PI Min	PI Max	CVR^2	PI Min	PI Max	CVR^2
MxM1	1.294	2.770	0.795	0.3491	5.2975	0.882	1.164	3.832	0.877
MnM1	1.943	2.129	0.825	0.9038	6.0012	0.932	1.479	3.332	0.874
FDL	1.081	2.430	0.869	0.2539	2.5389	0.959	0.3148	3.1132	0.945
TDL	1.138	2.895	0.825	0.4320	3.3999	0.964	0.2869	3.7508	0.926
HDL	1.369	2.850	0.805	0.2590	3.7347	0.961	0.2342	3.4387	0.923
RDL	1.812	2.755	0.808	0.4004	4.2710	0.954	0.3079	4.0254	0.919
All	1.594	1.776	0.633	1.434	1.873	0.901	0.4599	2.2605	0.971
LB	1.349	2.457	0.83	1.439	1.892	0.907	0.8376	3.2612	0.91
Molars	1.520	2.236	0.838	3.422	4.593	0.814	0.4666	2.4357	0.967

Note. Prediction intervals and Cross-Validated R Squared values for each MARS model. LB: Long bones (HDL, RDL, FDL, TDL). A prediction interval is the range of the age estimation that forensic anthropologists would give when using this model. Smaller prediction intervals indicate higher precision.

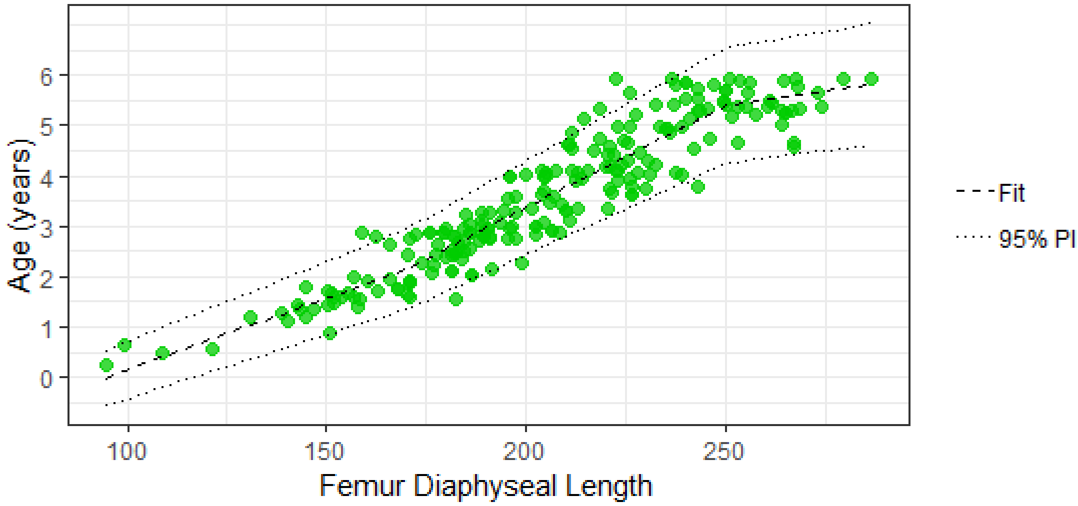


Figure 7. South African FDL Model. The graph shows a MARS model with the individuals it was trained on. Each dot is a point of data used to create the MARS model, and the MARS model is shown as lines of the fit (hinge equation) and the upper and lower 95% prediction intervals.

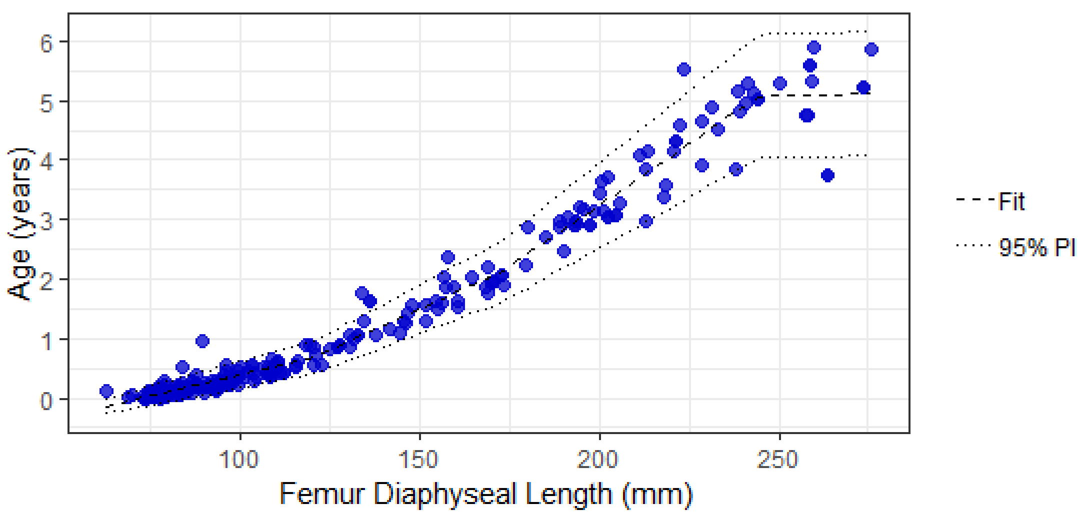


Figure 8. United States HDL MARS Model. This graph shows the United States humerus diaphyseal length model along with all individuals it was trained on.

Cross-Model Application

Some of the cross-applied models proved useable for estimating the age of a subadult if using the incorrect population's model, for example, running the US population data through the model created for the SA population (Table 6). The results had an overall successful age estimation, with TDL and MxM1 as the most accurate indicators of age for both populations (Table 6). While the cross-applied models were not as accurate as a model estimating the age of an individual within its training population, they showed that a population level difference is small enough to consider the application of a global model for subadult age estimation.

Table 6

Cross Applying the Models

	US -> SA			SA -> US		
Indicator	n	Correct	% (Accuracy)	n	Correct	% (Accuracy)
Max M1	254	187	73.6	197	167	84.8
Man M1	209	155	74.2	206	159	77.2
FDL	248	190	76.6	219	204	91.3
TDL	249	241	96.8	211	197	93.4
HDL	250	224	89.6	220	198	90.0
RDL	252	245	97.2	220	211	95.9
All	169	131	77.5	187	168	89.8

Note. Cross-Applied Validity test of South African and United States models, where n is the number of samples from the given population applied to the other population's model.

Adjusting for Age

When looking at the distribution of ages present in both samples (Figure 7), it is apparent that the United States sample is disproportionately younger, while the South African sample has very few individuals under 1 year old (Figure 9). This difference in the distribution of ages can cause a large difference in model precision and accuracy on other samples without similar distributions. To see the effect this skew was having on the cross-application, I adjusted the United States sample to only include individuals over 6 months old to remove a large group of individuals whose age cohort was not represented in the South African models. Two models are presented below, one that was not adjusted (Figure 10), and one that was adjusted so that age is greater than 6 months (Figure 11). Note on Figure 10, that the cluster of individuals from or near birth do not fall within the South African model, as the South African model does not have many individuals under 1 year old, and the youngest individual in the South African sample is 0.24 years, compared to 0.00 years for the United States sample. Because of the effect this had on accuracy, I recreated the experiment for all cross applied models (Table 7).

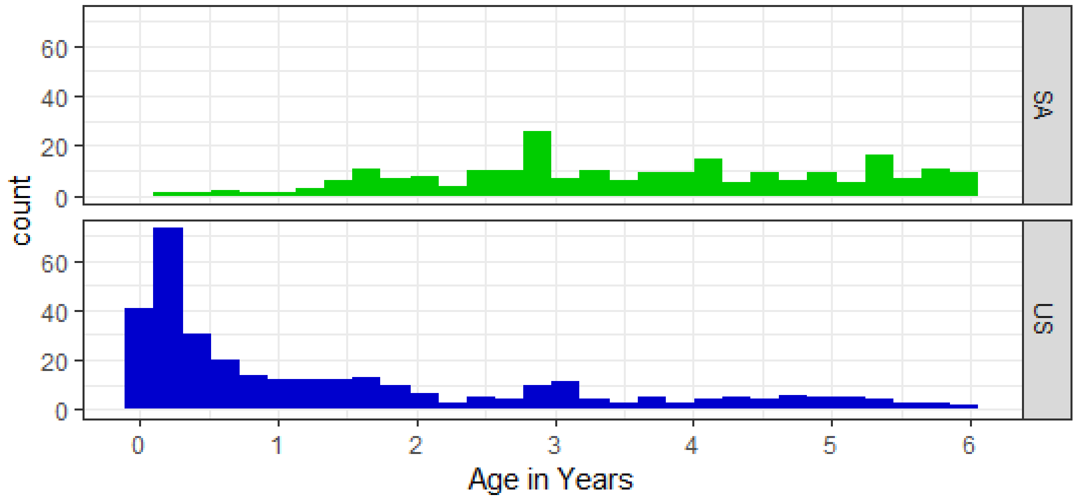


Figure 7. Histogram of ages by population.

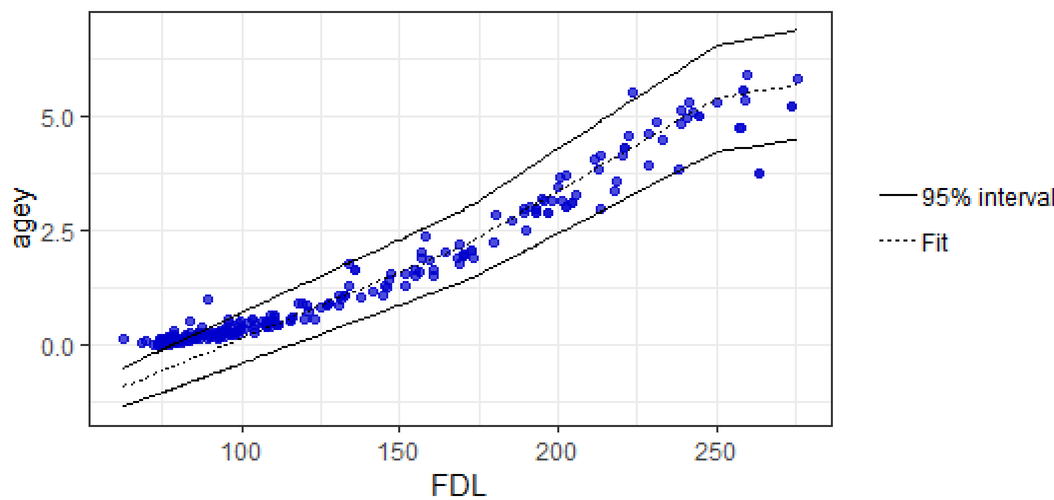


Figure 8. Cross applied South African FDL model to United States individuals, no adjustment for age

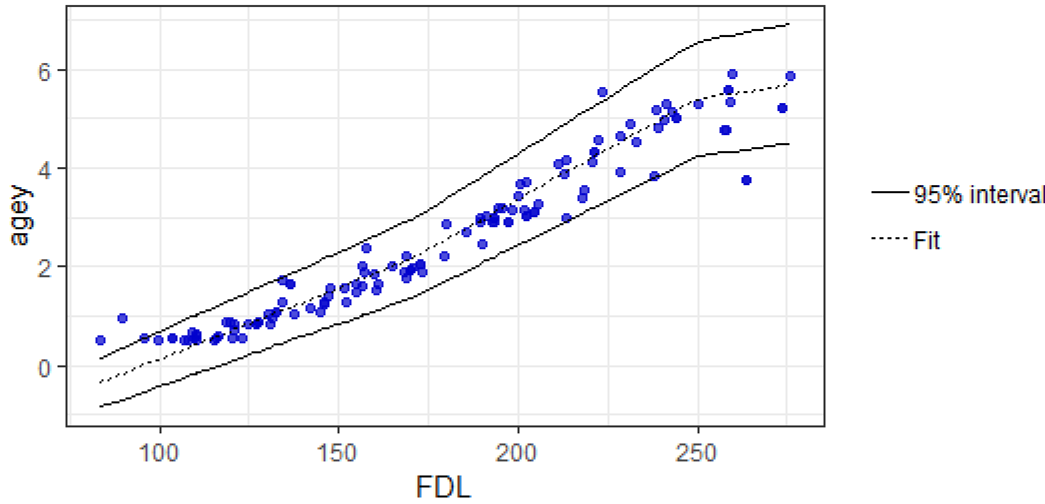


Figure 9. Cross applied South African model to United States population, adjusted so that age is greater than 0.5 years.

Table 7

Cross Applying the Models with Adjusted Ages

	US -> SA			SA -> US		
Age Indicator	n	Correctly estimated	% (Accuracy)	n	Correctly Estimated	% (Accuracy)
Max M1	151	104	68.9	197	167	84.8
Man M1	207	155	75.0	206	159	77.2
FDL	124	119	96.0	219	204	91.3
TDL	126	122	96.8	211	197	93.4
HDL	126	122	96.8	220	198	90.0
RDL	128	124	96.9	220	211	95.9
All	87	80	92.0	187	168	89.8

Note. Cross-Applied Validity test of South African and United States models with an adjusted age range for the US population, where n is the number of samples from the given population applied to the other population's model.

Table 8

Global Model Precision of Birth to 4 Years vs Birth to 6 Years

	Birth – 4 years			Birth – 6 years		
Indicator	PI Min	PI Max	CVR^2	PI Min	PI Max	CVR^2
MxM1	1.211	2.762	0.690	1.164	3.832	0.877
MnM1	2.052	2.634	0.630	1.479	3.332	0.874
FDL	1.278	1.493	0.835	0.3148	3.1132	0.945
TDL	1.370	1.694	0.810	0.2869	3.7508	0.926
HDL	1.216	2.225	0.778	0.2342	3.4387	0.923
RDL	1.540	1.959	0.763	0.3079	4.0254	0.919
All	1.247	1.262	0.893	0.4599	2.2605	0.971
LB	1.161	1.549	0.865	0.8376	3.2612	0.91
Molars	0.8901	2.5833	0.761	0.4666	2.4357	0.967

Note. Both global models presented, the first was trained on US individuals under 4 years old South African individuals under 4 years old. The second was the original global model, based on all US individuals under 6 years and all South African individuals under 6 years.

Discussion

While many authors have cited the need for population specific methods for age estimation, population level differences were not present between South Africa and the United States, meaning that a global model should be successful. The lack of a statistically significant difference for diaphyseal length and molar formation reveals that population level differences and economic insults are not substantial in subadults before the age of six. The present study identified the shift between US and South African diaphyseal lengths between 3 and 4 years of age, indicating that economic insults and population level differences are present after 4 years old, but it is still unknown at what age economic insults and population level differences accumulate enough to cause need for population specific models. I pose that global models are more accurate for subadults under 4 years old, but are less precise after time and population specific models are more precise after 3 years old than global models are.

Model Accuracy

Population specific models proved to be more precise in general when estimating the age of a subadult however, global models were not so imprecise to be unusable for an individual of unknown ancestry. Prediction intervals for global models were slightly larger than those of population specific models, but the smallest prediction intervals of global models were significantly smaller than the smallest prediction intervals of population specific models. Cross application of models was better than chance, though the accuracy of cross-applied models would not hold up to scientific speculation. MARS models proved to be extremely useful in creating age estimation models for the samples, and yielded precise and accurate models to estimate age of any group of individuals due to the nonlinear and heteroscedastic form of the analysis. Models for age estimation based on molar formation were far less precise and accurate than originally expected in this project. Dental formation has been repeatedly thought to be the most accurate and precise age estimator across all populations, so its position as the least accurate cross-applied model was confounding. However, it cannot be ignored that the samples have different age distributions, which could be causing the inefficacy of cross-model application. This is supported by the increase in accuracy of cross-model application after adjusting the samples to have a more similar age distribution (Table 7).

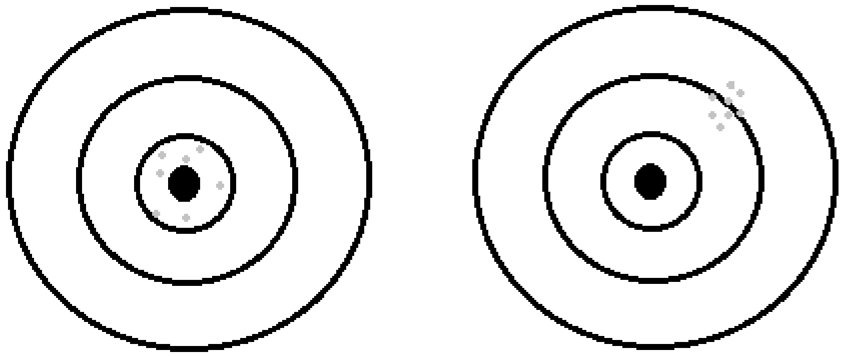
Limitations of the Current Study

The current study focused only on two populations, South Africa and the United States, out of the many that exist across the globe. Incorporating many populations of differing ancestries and socioeconomic status levels would greatly improve the applicability of a global model, which could be achieved through use of modern skeletal collections across the world or compilation of subadult data from other researchers. Furthermore, the present samples were differing in age distribution, which could have had an effect on the accuracy of the cross-applied models. Likewise, the sample sizes were relatively small, with South Africa only having a total of 137 individuals. An increase in the amount of individuals for this research will only benefit the field of forensic anthropology by allowing us to craft more precise and representative samples

Implications for Future Research

While this study was useful for establishing that there were no statistically significant difference in diaphyseal lengths and molar formation of subadults from South Africa and subadults from the United States, it is not enough to bring global models the validity needed for future use by researchers and forensic anthropologists. Many more populations and variables must be tested before global models can be validated for all age estimation methods, such as hand-wrist radiographs, vertebral fusion, and many others. Future research would benefit from investigating the specific nature of each age indicator used to estimate age for subadults and its relationship with SES and population level differences to establish which age indicators would be valid for a global model and which age indicators would be invalid for global models. After this research, it would be beneficial to explore the population level differences and SES differences across populations and establish either ability or inability to use global models across all populations or if broader categories such would be more applicable to subadult age estimation.

Supplementary Figures



Supplementary Figure 1. Accuracy vs Precision. Accuracy is the closeness to the goal, precision is consistency in test values. In this example, the left is accurate and consistently reaches close to the goal. The right side is precise and may not be close to the goal, but is repeatedly estimated in the same area.

Bibliography

Alqahtani, S. J., Hector, M. P., & Liversidge, H. M. (2014). Accuracy of dental age estimation charts: Schour and Massler, Ubelaker and the London Atlas. *American Journal of Physical Anthropology 154*(1), 70-78. https://doi.org/10.1002/ajpa.22473

AlQahtani, S. J., Hector, M. P., & Liversidge, H. M. (2010). Brief communication: The London atlas of human tooth development and eruption. *American Journal of Physical Anthropology 142*(2), 481-490. https://doi.org/10.1002/ajpa.21258

Brits, D., Manger, P. R., & Bidmos, M. A. (2017). Assessing the use of the anatomical method for the estimation of sub-adult stature in Black South Africans. *Forensic Science International 283*, 221.e1-221.e9. https://doi.org/10.1016/j.forsciint.2017.11.024

Brooks, S., & Suchey, J. M. (1990). Skeletal age determination based on the os pubis: A comparison of the Acsádi-Nemeskéri and Suchey-Brooks methods. *Human Evolution 5*(3), 227-238. https://doi.org/10.1007/BF02437238

Cardoso, H. F. V. (2007). Environmental effects on skeletal versus dental development: Using a documented subadult skeletal sample to test a basic assumption in human osteological research. *American Journal of Physical Anthropology*, *132*(2), 223–233. https://doi.org/10.1002/ajpa.20482

Cardoso, H. F. V. (2008a). Epiphyseal union at the innominate and lower limb in a modern Portuguese skeletal sample, and age estimation in adolescent and young adult male and female skeletons. *American Journal of Physical Anthropology 135*(2), 161-170. https://doi.org/10.1002/ajpa.20717

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

Cardoso, H. F. V. (2009a). A test of three methods for estimating stature from immature skeletal remains using long bone lengths. *Journal of Forensic Sciences, 54*(1), 13-19. https://doi.org/10.1111/j.1556-4029.2008.00916.x

Cardoso, H. F. V. (2009b). Accuracy of developing tooth length as an estimate of age in human skeletal remains: The permanent dentition. *American Journal of Forensic Medicine and Pathology 30*(2), 127-133. https://doi.org/10.1097/PAF.0b013e3181879d07

Cardoso, H. F. V., Abrantes, J., & Humphrey, L. T. (2014). Age estimation of immature human skeletal remains from the diaphyseal length of the long bones in the postnatal period. *International Journal of Legal Medicine 128*(5), 809-824. https://doi.org/10.1007/s00414-013-0925-5

Cardoso, H. F. V., & Ríos, L. (2011). Age estimation from stages of epiphyseal union in the presacral vertebrae. *American Journal of Physical Anthropology, 144*(2), 238-247. https://doi.org/10.1002/ajpa.21394

Cardoso, H. F. V., Spake, L., & Humphrey, L. T. (2017). Age estimation of immature human skeletal remains from the dimensions of the girdle bones in the postnatal period. *American Journal of Physical Anthropology 163*(4), 772-783. https://doi.org/10.1002/ajpa.23248

Cardoso, H. F. V., Spake, L., & Liversidge, H. M. (2016). A Reappraisal of Developing

Permanent Tooth Length as an Estimate of Age in Human Immature Skeletal Remains. *Journal of Forensic Sciences*, *61*(5), 1180–1189. https://doi.org/10.1111/1556-4029.13120

Cardoso, H. F. V. (2008). Age estimation of adolescent and young adult male and female skeletons II, epiphyseal union at the upper limb and scapular girdle in a modern Portuguese skeletal sample. *American Journal of Physical Anthropology 137*(1), 97-105. https://doi.org/10.1002/ajpa.20850

Colman, K. L., Janssen, M. C. L., Stull, K. E., van Rijn, R. R., Oostra, R. J., de Boer, H. H., & van der Merwe, A. E. (2018). Dutch population specific sex estimation formulae using the proximal femur. *Forensic Science International 286*, 268.e1-268.e8. https://doi.org/10.1016/j.forsciint.2017.12.029

Conceição, E. L. N., & Cardoso, H. F. V. (2011). Environmental effects on skeletal versus dental development II: Further testing of a basic assumption in human osteological research. *American Journal of Physical Anthropology*, *144*(3), 463-470. https://doi.org/10.1002/ajpa.21433

Corron, L., Marchal, F., Condemi, S., Chaumoître, K., & Adalian, P. (2017). A new approach of juvenile age estimation using measurements of the ilium and Multivariate Adaptive Regression Splines (MARS) Models for better age prediction. *Journal of Forensic Sciences*, *62*(1), 18–29. https://doi.org/10.1111/1556-4029.13224

Corron, L., Marchal, F., Condemi, S., Telmon, N., Chaumoitre, K., & Adalian, P. (2018). Integrating growth variability of the ilium, fifth lumbar vertebra, and clavicle with Multivariate Adaptive Regression Splines models for subadult age estimation.

Journal of Forensic Sciences. https://doi.org/10.1111/1556-4029.13831

de Onis, M., & Branca, F. (2016). Childhood stunting: A global perspective. *Maternal and Child Nutrition*, *12*, 12-16. https://doi.org/10.1111/mcn.12231

Esan, T. A., Yengopal, V., & Schepartz, L. A. (2017). The Demirjian versus the Willems method for dental age estimation in different populations: A meta-analysis of published studies. *PLoS ONE*, *12*(11), e0186682. https://doi.org/10.1371/journal.pone.0186682

Franklin, D. (2010). Forensic age estimation in human skeletal remains: Current concepts and future directions. *Legal Medicine*, *12*(1), 1–7. https://doi.org/10.1016/j.legalmed.2009.09.001

Friedman, J. H. (1991). Multivariate Adaptive Regression Splines. *The Annals of Statistics*, *19*(1), 1–67. https://doi.org/10.1214/aos/1176347963

Heuzé, Y., & Cardoso, H. F. V. (2008). Testing the quality of nonadult Bayesian dental age assessment methods to juvenile skeletal remains: The Lisbon collection children and secular trend effects. *American Journal of Physical Anthropology*, *135*(3), 275-283. https://doi.org/10.1002/ajpa.20741

Howe, L. D., Tilling, K., Galobardes, B., Smith, G. D., Gunnell, D., & Lawlor, D. A. (2012). Socioeconomic differences in childhood growth trajectories: At what age do height inequalities emerge? *Journal of Epidemiology and Community Health*, *66*(2), 143-148. https://doi.org/10.1136/jech.2010.113068

Jeong, Y., & Jantz, L. M. (2016). Developing Korean-specific equations of stature estimation. *Forensic Science International*, *260*, 105.e1-105.e11. https://doi.org/10.1016/j.forsciint.2015.12.048

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

Kreiger, N. (2012). Who and what is a “population”? *The Milbank Quarterly*, *90*(4), 634–681. https://doi.org/10.1111/j.1468-0009.2012.00678.x

Langley-Shirley, N., & Jantz, R. L. (2010). A bayesian approach to age estimation in modern Americans from the clavicle. *Journal of Forensic Sciences*, *55*(3), 571-583. https://doi.org/10.1111/j.1556-4029.2010.01089.x

Márquez-Grant, N. (2015). An overview of age estimation in forensic anthropology: perspectives and practical considerations. *Annals of Human Biology*, *42*(4), 308–322. https://doi.org/10.3109/03014460.2015.1048288

McCrory, C., O’Leary, N., Fraga, S., Ribeiro, A. I., Barros, H., Kartiosuo, N., … Zins, M. (2017). Socioeconomic differences in children’s growth trajectories from infancy to early adulthood: Evidence from four European countries. *Journal of Epidemiology and Community Health*, *71*(10), 981-989. https://doi.org/10.1136/jech-2016-208556

Milborrow, S. (2018). Notes on the earth package. Retrieved from http://www.milbo.org/doc/earth-notes.pdf

Passalacqua, N. V. (2013). Subadult age-at-death estimation from the human calcaneus. *International Journal of Osteoarchaeology*, *23*(4), 471-474. https://doi.org/10.1002/oa.1255

Petrie, D., & Tang, K. K. (2008). A Rethink on Measuring Health Inequalities Using the Gini Coefficient. *Discussion Papers Series*.

Programme, U. N. 2006 D. (2008). Gini coefficient. *Stat*. https://doi.org/10.4135/9781412963930.n237

Quetelet, A., & Knox, R. (1841). *A treatise of man and the development of his functions*. Edinburgh, Scotland.

Ríos, L., & Cardoso, H. F. V. (2009). Age estimation from stages of union of the vertebral epiphyses of the ribs. *American Journal of Physical Anthropology*, *140*(2), 265-274. https://doi.org/10.1002/ajpa.21065

Sarajlić, N., & Gradaščević, A. (2012). Morphological characteristics of pubic symphysis for age estimation of exhumed persons. *Bosnian Journal of Basic Medical Sciences*, *12*(1), 51. https://doi.org/10.17305/bjbms.2012.2534

Shirley, N., Fazlollah, A., & Tersigni-Tarrant, M. (2013). Age estimation methods. In *Forensic Anthropology: An Introduction* (pp. 161–179).

Singal, S., & Sharma, S. (2017). Age estimation by position of mental foramen in Haryana population: A radiographic study. *Journal of Advanced Meidcal and Dental Sciences*, *5*(12), 68–73. https://doi.org/https://dx.doi.org/10.21276/jamdsr.2017.5.12.20

Stull, K. E., L'Abbé, E. N., & Ousley, S. D. (2014a). Using multivariate adaptive regression splines to estimate subadult age from diaphyseal dimensions. *American Journal of Physical Anthropology*, *154*(3), 376–386. https://doi.org/10.1002/ajpa.22522

Stull, K. E., L'Abbé, E. N., & Ousley, S. D. (2014b). Using multivariate adaptive regression splines to estimate subadult age from diaphyseal dimensions. *American Journal of Physical Anthropology*, *154*(3), 376–386.

https://doi.org/10.1002/ajpa.22522

Stull, K. E., L'Abbé, E. N., & Steiner, S. (2013). Measuring distortion of skeletal elements in Lodox Statscan-generated images. *Clinical Anatomy*, *26*(6), 780-786. https://doi.org/10.1002/ca.22225

Todd, T. W. (1937). *Atlas of Skeletal Maturation*. St. Louis, MO: The C.V. Mosby Company. Retrieved from https://unr.summon.serialssolutions.com/#!/search?bookMark=ePnHCXMw42JgAfZbU1kZOIzBZzuBbvvgMDQxNwVdtMPJIONYAmwVKuSnKRRnA8tVYANTIRd0giXYGTwMLCVFpcBiMcTNNcTZQxd6HYBuormRialucmoysLJLTEkzMExLTLNISzIwtAAKpVgamCanmBgmG5ulJpknGVkkpQBTVmKKiQXoRi3QgU6pSSnAlAq6GB1i

UNDP. (2018). Human Development Data (1990-2015) | Human Development Reports. https://doi.org/10.1007/s00338-005-0033-1