Document Title:	**Quantifying the Impact of Intrinsic and Extrinsic Factors to Improve Juvenile Age Estimation**
Author(s):	**Kyra E Stull, Ph.D., D-ABFA**
Document Number:	**306956**
Date Received:	**June 2023**
Award Number:	**2017-DN-BX-0144**

Final Report
National Institute of Justice
Award 2017-DN-BX-0144

Quantifying the Impact of Intrinsic and Extrinsic Factors to Improve Juvenile Age Estimation

PI: Kyra E Stull, PhD, D-ABFA

Table of Contents

Summary of the Project

Major goals and objectives and research questions

The objective of this research was to quantify the impact of intrinsic and extrinsic covariates on juvenile age estimation and assess differences between global and country-specific age estimation models for forensic application using a large diverse international sample of children aged 0 to 15 years.

Two specific aims were pursued throughout this project:

1. Analyze the relationship between skeletal growth and development markers and age in different populations and evaluate the influence of several covariates (sex, population affinity, socio-economic status) on these markers.
2. Explore a global model for age estimation and quantify the impact of country-level SES, inferred by the country's HDI, and inequality of wealth, inferred by the country's Gini coefficient, for all age indicators and covariates.

Research Design, Methods, Analytical and Data Analysis Techniques

Material

The final sample was comprised of 4,902 individuals aged from birth to 22 years (Table 1 and Figure 1). Skeletal and dental age indicators (*i.e.*, epiphyseal fusion, long bone lengths, dental development, and VNC dimensions) were collected from individuals living in socially and geographically diverse countries including Angola, Brazil, Colombia, France, the Netherlands, South Africa, and Taiwan. The international data was combined with the previously collected United States sample collected from the NIJ Award 2015-DN-BX-K409, which is how the sample includes individuals. Data collected from CT scans generated at hospitals and medical examiner's offices included France (n = 577), the Netherlands (n = 215), Taiwan (n = 658), and the United States (n = 1,321). Data from South Africa (n = 1,353), Angola (n = 240), and Brazil (n = 500) were collected from Lodox Statscan images, region-specific conventional x-rays, and dental orthopantomograms, respectively (Table 1). Data from Colombia (n = 43) were collected directly from skeletal remains. All individuals are of known age and sex. The sampling varied at the institution level (Table 2). For example, samples comprised of deceased individuals often presented with unequal distributions across ontogeny and those selected from larger hospital databases are more equally distributed (Figure 2). When all samples are pooled, there is a more uniform distribution (Figure 1).

Individuals were also characterized by their country of origin, or population affinity, and two indicators of socio-economic status/SES: the Human Development Index (HDI), as a measure of socio-economic development and the Gini coefficient, as a measure of social inequality (Jahan 2015). See below for further details on the covariates and Stull and Corron (2021) for further discussion on each of the samples, specifically how they vary in shape and size.

Table 1 Demographics of the samples.

Continent	Country	Sample size (by sex)	Age range (years)	Deceased/Living
Africa	Angola	N = 240 (120 F, 120 M)	0-15 years	Living (x-ray)
	South Africa	N = 1353 (532 F, 821 M)	0-12 years	Living (Lodox)
South America	Colombia	N = 43 (8 F, 35 M)	0-22 years	Deceased (dry bone)
	Brazil	N = 500 (249 F, 251 M)	0-15 years	Living (x-ray)
Europe	France	N = 577 (260 F, 317 M)	0-15 years	Living (CT scan)
	Netherlands	N = 215 (109 F, 106 M)	0-15 years	Living (CT scan)
Asia	Taiwan	N = 658 (268 F, 309 M)	0-16 years	Living (CT scan)
North America	United States (Maryland)	N = 244 (96 F, 148 M)	0-20 years	Deceased (CT scan)
	United States (New Mexico)	N = 1077 (441 F, 636 M)	0-21 years	Deceased (CT scan)

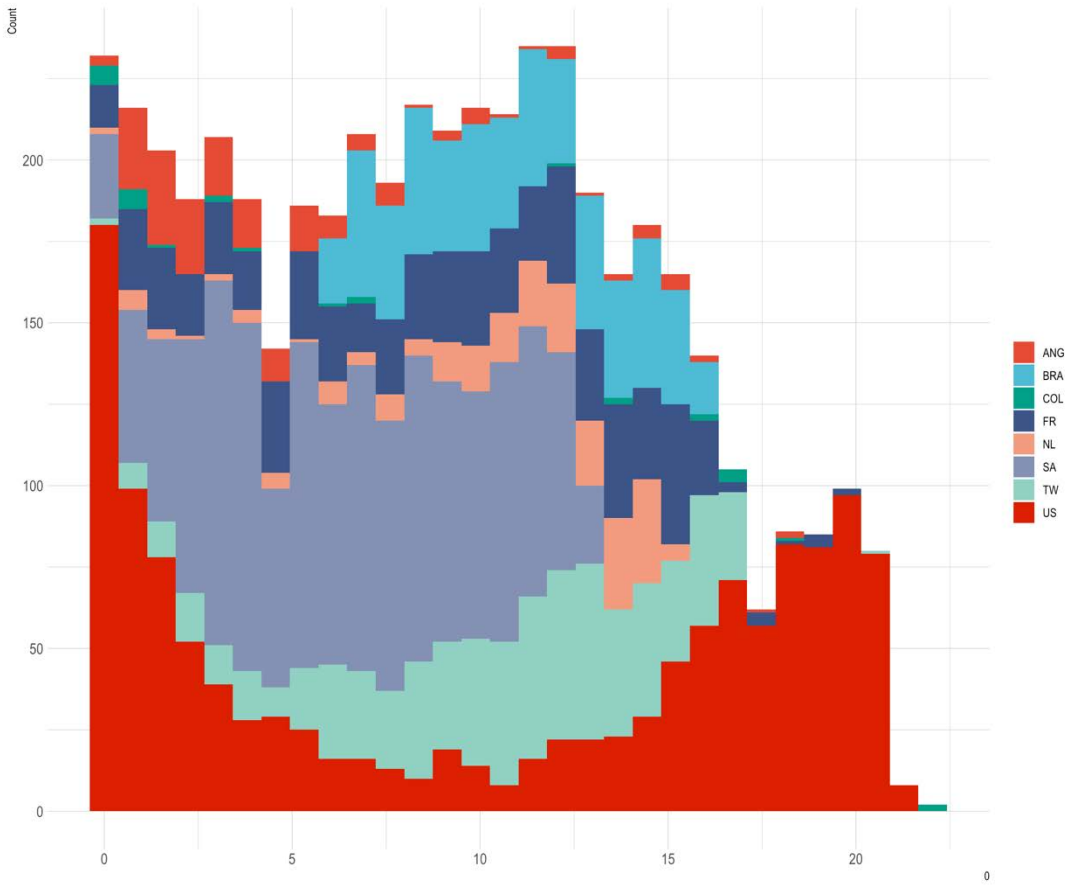

Figure 1. Age distributions and sample size per age.

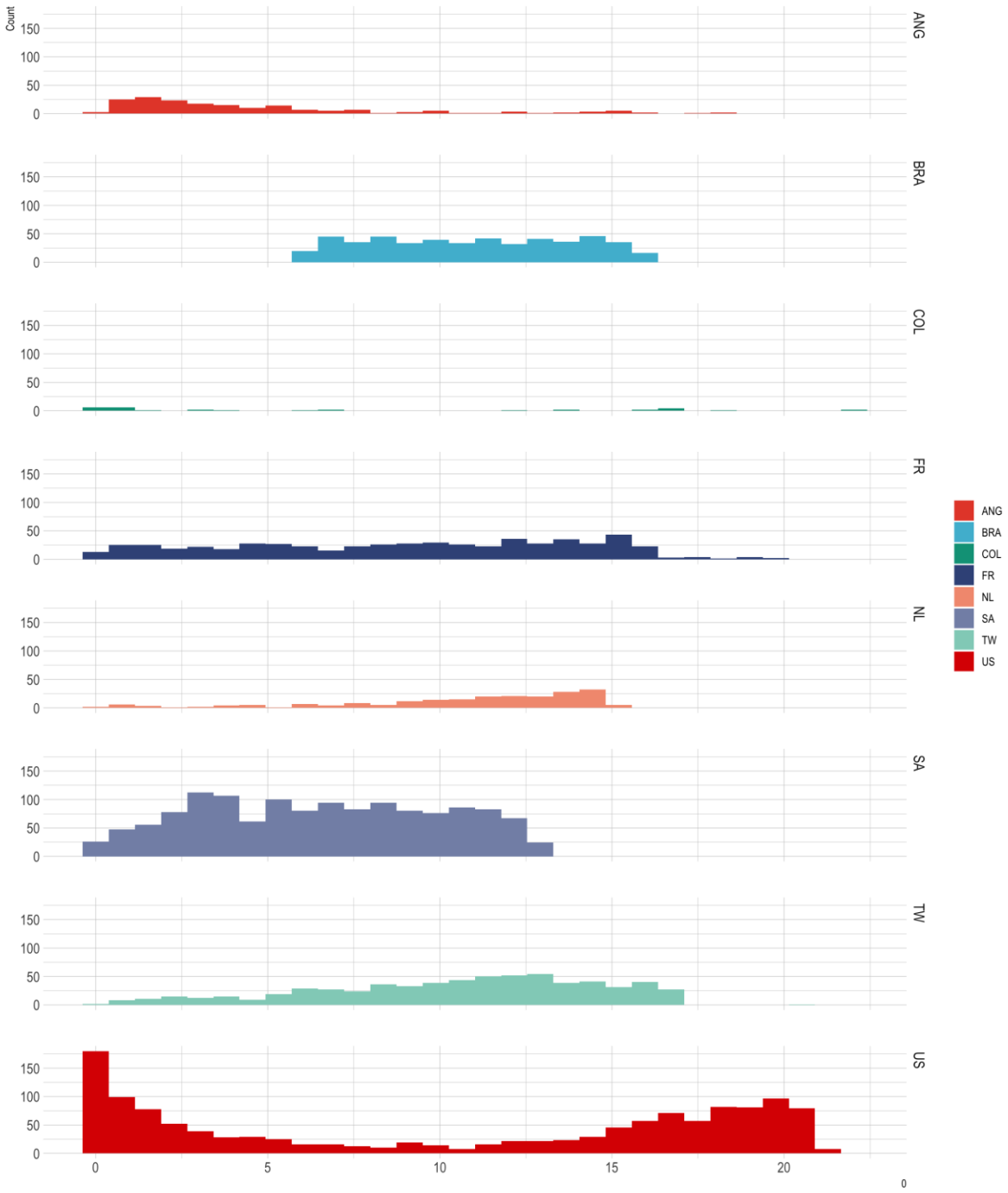

Figure 2. Age distributions and sample size per age.

Data Collection

Numerous modalities were used throughout data collection depending on the collaborating institution and the types of data previously collected and/or available. Specifically, skeletal and dental growth and development indicators were collected on images, dry bones, or virtual bone surfaces. Lodox Statscans were visualized using the DVS software. A Nomad was used to collect dental x-rays from the Colombian sample, which required the MiPacs medical visualization software. The CT scans were visualized using Amira. The virtual surfaces of all six long bones (femur, tibia, fibula, humerus, radius, ulna) and left – unless unavailable and then right – os coxae were segmented and reconstructed three-dimensionally from the CT scans following a standardized segmentation protocol (Stull and Corron 2021a). Segmented surfaces are available for research but do require one to contact the authors for access.

Long Bone Dimensions

Eighteen diaphyseal length and breadth measurements of the six long bones (femur, tibia, fibula, humerus, radius, and ulna) were collected per individual when available (Figure 3). The measurements definitions were those presented by Stull, L'Abbé, & Ousley in 2014, which were based on Fazekas and Kosa (1978) and Moore-Jansen et al. (1994). Long bone data were measured on segmented bone surfaces in the four CT samples (France, the Netherlands, Taiwan, the United States), on Lodox Statscan images in the South African sample (Stull, L'Abbé, and Ousley 2014), and directly on dry bones in the Colombian sample using a digital caliper (0.01mm precision). Left-sided elements were measured by default and their right antimere were measured if the left was unavailable or broken. Length measurements were taken only from unfused long bones. Therefore, the age range for each variable is dependent on its maturation trajectory. Generally, data for the long bones is mostly unavailable after 13 years of age. Full definitions and illustrations of the long bone measurements can be found in Stull and Corron (2021b).

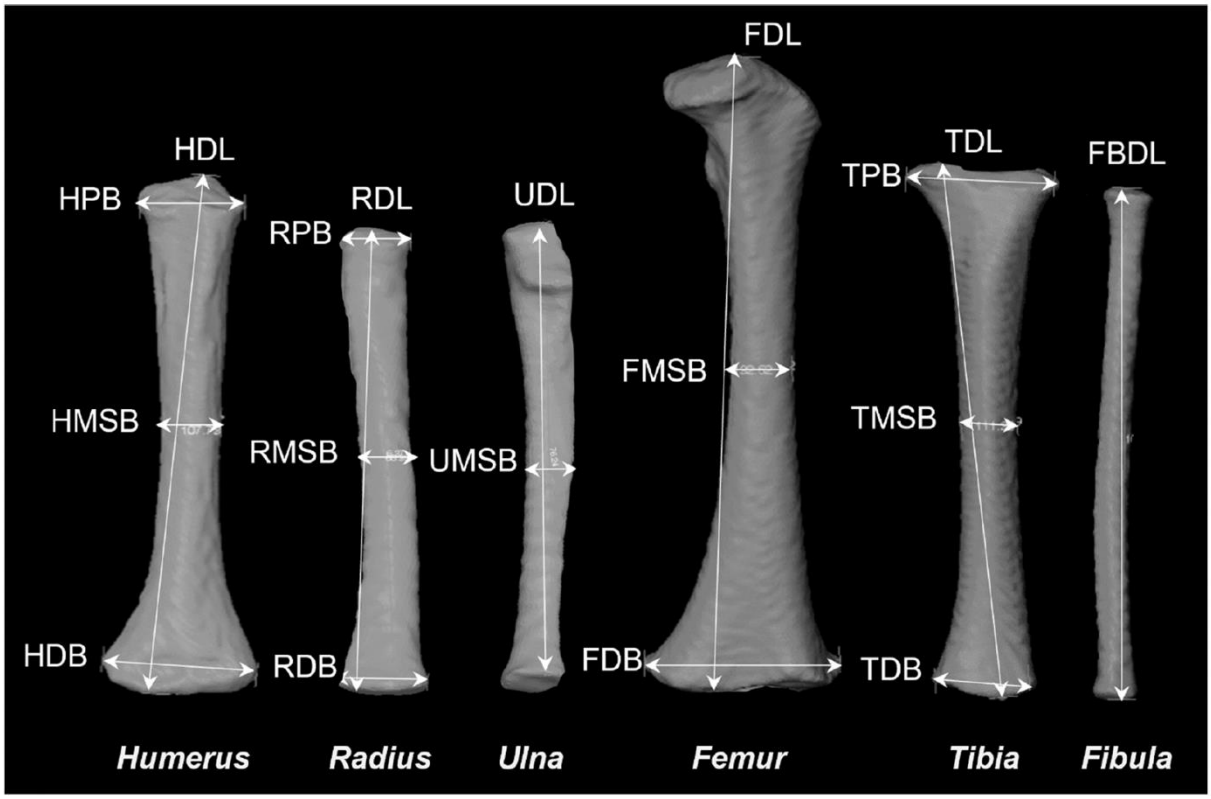

Figure 3. Measurements of the six long bones. Left to right: humerus, radius, ulna. femur, tibia, fibula. Image taken from Stull et al. (2022).

Several publications document the high repeatability of measurements collected on Lodox Statscan and measurements collected on virtual reconstructions of skeletal elements (Colman et al. 2019; Stull et al. 2014; Stull, L'Abbé, and Steiner 2013; Corron, Marchal, Condemi, Chaumoître, et al. 2017; Brough, Bennett, and Morgan 2013; Spake et al. 2020).

Ossification and Epiphyseal Fusion

Epiphyseal fusion stages for proximal and distal long bone epiphyses, the calcaneal tuberosity, the ischiopubic ramus, the ilium and ischium, and ossification of the patella were recorded for both left and right-sided elements. Three different staging systems were employed (Figure 4): a seven-stage system was used for the long bone epiphyses and the calcaneal tuberosity; a three-stage system was used for the pelvic epiphyses; and a binary absent/present was used for the carpals and tarsals, the ossification of the elements of the proximal and distal humerus (e.g., humeral head, lesser tubercle, greater tubercle, capitulum, trochlea, composite epiphyses), and the patella. The ossification of carpals and tarsals were scored as absent (0)/present (1), but then transformed into a component score to indicate the total number of elements present on the left and/or right side. Epiphyseal fusion was scored on CT slices for France, the Netherlands, Taiwan, and the U.S. samples, and on radiographs taken for Angola, Colombia, and South Africa. See Stull and Corron (2021b) for full definitions and descriptions of the staging systems.

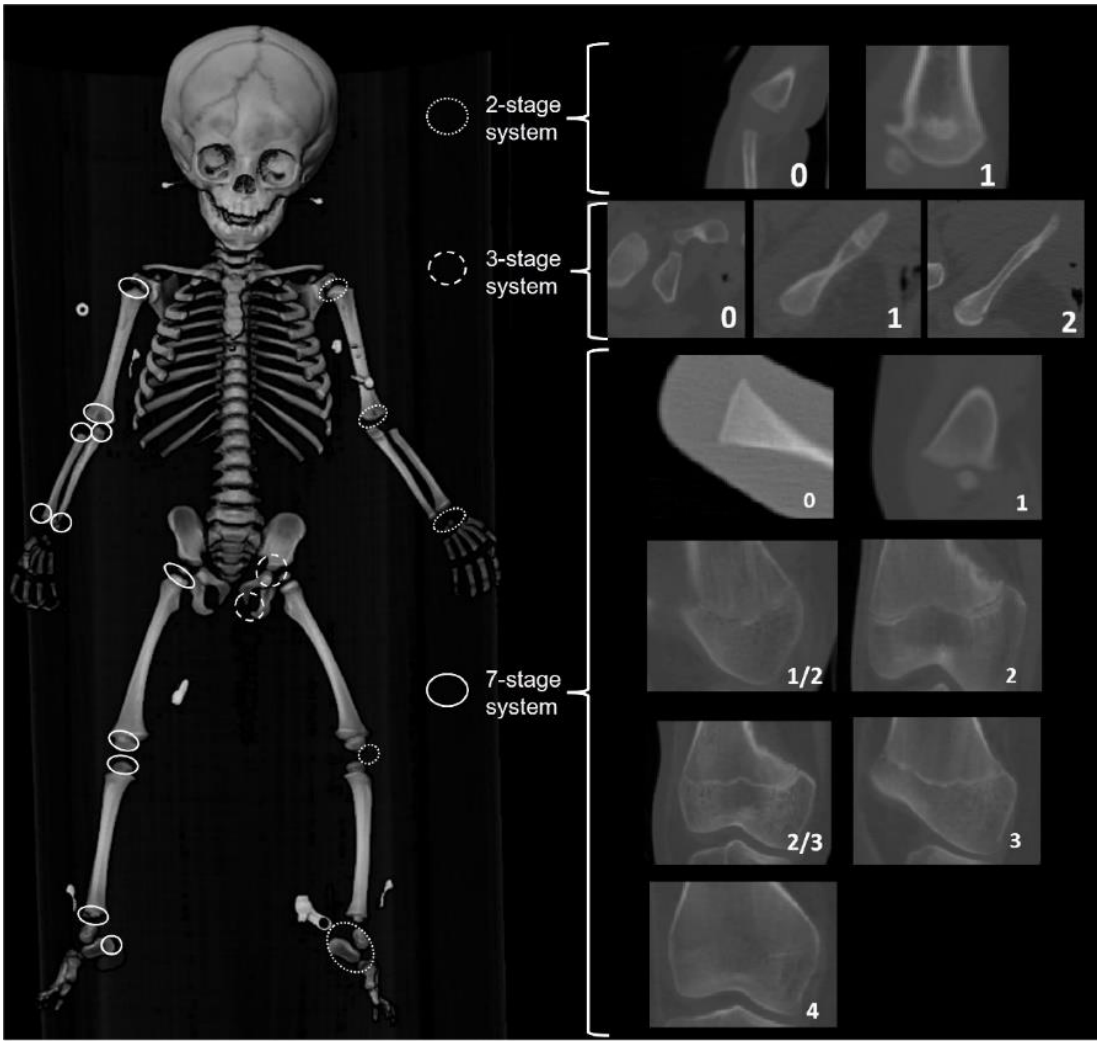

Figure 4. Sites for which ossification and epiphyseal fusion were scored following one of three staging systems developed for CT scans. Image taken from Stull et al.(2022).

Dental Development

Dental development was collected for the 32 permanent teeth following AlQahtani, Hector, and Liversidge (2010) 13-stage revision of the Moorrees, Fanning, and Hunt (1963) mineralization stages for mono-radicular and pluri-radicular teeth. A numerical adaptation rather than the original abbreviations (e.g., Cr ½) of the 13 stages are used. For example, '1' was used rather than 'Ci'. Teeth were scored directly from CT slices, panoramic and conventional radiographs generated at the original institutions, and from radiographs generated on site in Colombia using a portable Nomad Pro2.

Missing data for epiphyseal fusion, dental development, and diaphyseal dimensions is visualized in Figure 5, which highlights the differential growth trajectories of the age indicators.

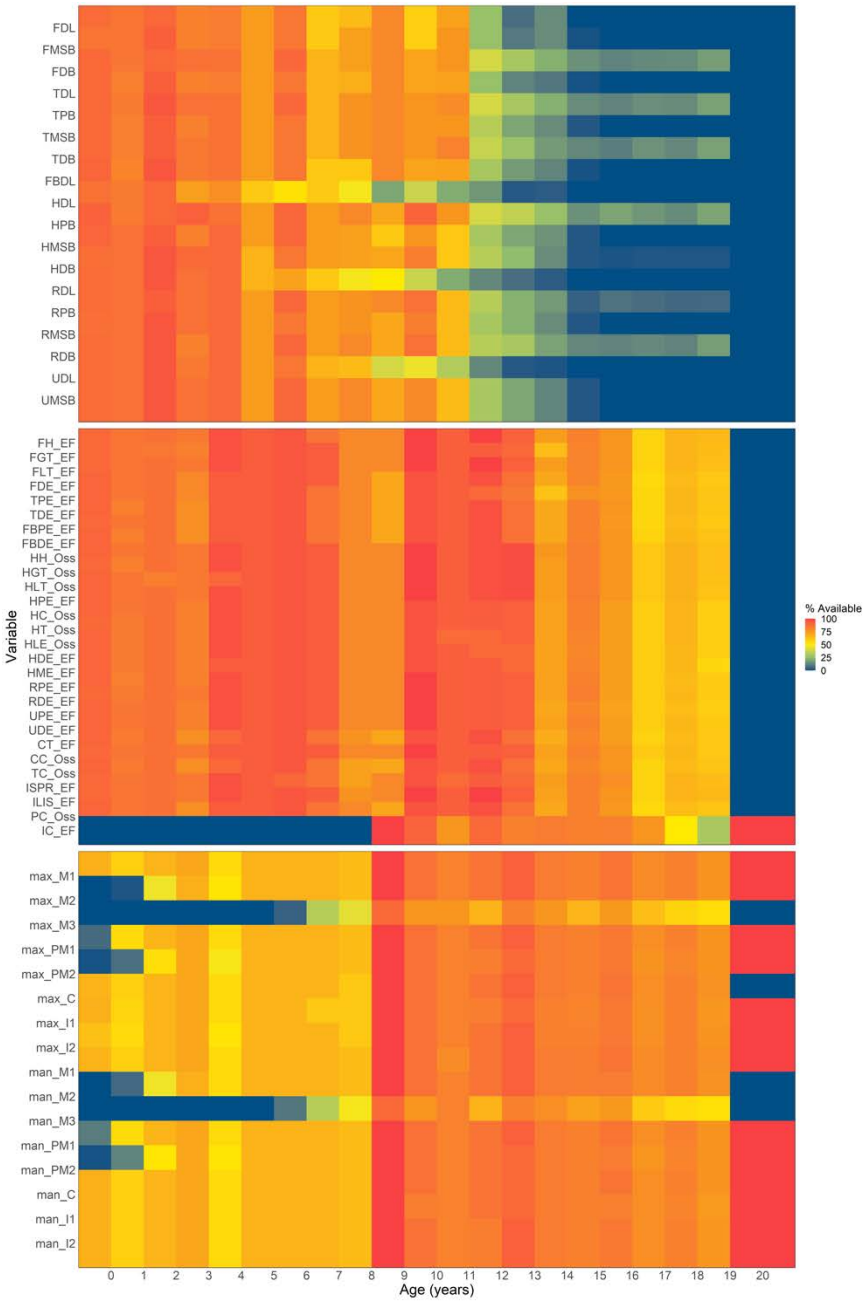

Figure 5. Distribution of recorded and missing data for each variable. Image taken from Stull et al (2022).

Vertebral Neural Canal (VNC) Diameters

Antero-posterior (AP) and transverse (TR) diameters of vertebrae thoracic 10 to lumbar 5 were collected. Data were measured on the virtually reconstructed surfaces from CT scans (Figure 6) and directly on the dry vertebrae in the Colombian sample. This yielded AP and TR VNC diameter measurements for 1404 individuals aged from birth to 22 years. Sample sizes for each country were: Colombia (n = 28), France (n = 484), the Netherlands (n = 23), Taiwan (n = 31), and the United States (n = 838) on which to perform our analyses.

All data has been aggregated to create the Subadult Virtual Anthropology Database (Stull and Corron 2022) and the data is freely available on the Subadult Virtual Anthropology Database Zenodo Community (https://zenodo.org/communities/svad/?page=1&size=20).

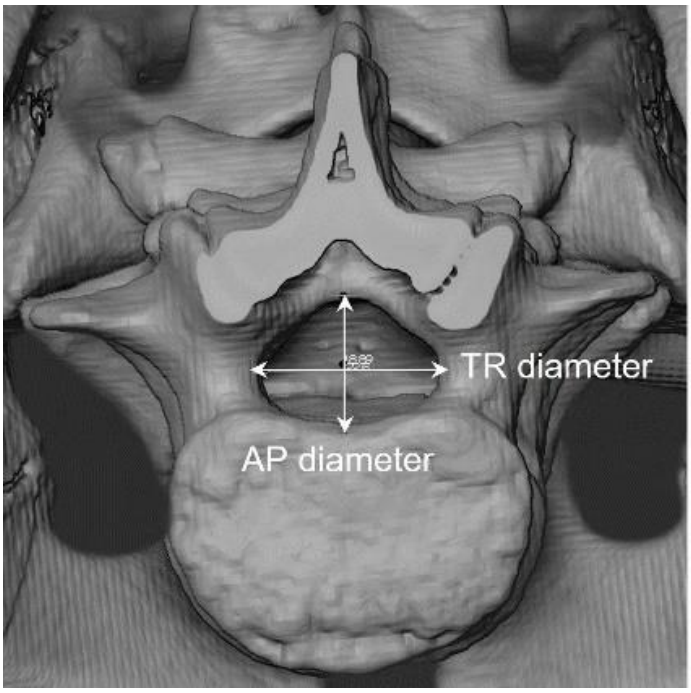

Figure 6. Antero-posterior (AP) and transverse (TR) diameters of the vertebral neural canal measured on a virtually rendered vertebra using the Amira software. Adapted from Corron, Wolfe, and Stull (2023).

Covariates

Two indicators were used to attempt to capture the multifaceted concept of SES: the Human Development Index (HDI) and the Gini coefficient (Table 2). HDI incorporates infant mortality rate, access and quality of education, and economic strength for each country, thus it gives a holistic overview of the country's development, not just from an economic point of view (Steckel 2012; Bogin, Scheffler, and Hermanussen 2017; Jahan 2015). The Gini index is a measure of the level of social inequality of each country (Petrie

and Tang 2008). High inequality results in different economic resources, which in turn leads to different access and quality of food, housing, and healthcare (Sahn and Younger 2005; Bogin, Scheffler, and Hermanussen 2017; Graham 2005; Steckel 2012). The scales for the two parameters are opposite; a high HDI and a low Gini are the ideal values. Additionally, because HDI and the Gini coefficient are independent of one another, they can disagree. For example, a country can have a high HDI (*i.e.*, better overall quality of life), but the country can also have a high Gini coefficient (*i.e.,* a large inequality of wealth). Therefore, they were both included as separate covariates in statistical analyses.

Life History Stages

For analyses that included ontogenetic variation, individuals were grouped into one of four life history stages according to their chronological age for analyses: infancy (birth to 2.99 years), childhood (3 to 7.99 years), juvenile (8 to 11.99 years), and adolescence (12 to 18 years) following Bogin (1999). The life history subsets were done simply to collapse a large amount of data and display trends that are otherwise lost in 20+ chronological ages. Unfortunately, we are unable to use developmental milestones because not individual has all the same available data, leading to a large loss of information (i.e., having an observable m1 to determine life history stage).

Table 2 List of countries and their associated HDI and Gini coefficients. The color of the box reflects if the value; green is used to visualize a higher/better value (the darker the better), red is used to visualize a lower/worse value (the darker the lower), and orange reflects the middle values.

Country	HDI (year)	State-HDI	HDI level	Gini (year) – World bank estimate	Gini level
Angola	0.581 (2017)	0.686 (2018) – Luanda	Medium	42.7 (2008)	Medium
Brazil	0.761 (2018)	• 0.826 (2017) – Sao Paulo • 0.850 (2017) – Brasilia	High	53.3 (2017)	High
Colombia	0.761 (2018)	0.766 (2018) - Antioquia	High	49.7 (2017)	High
France	0.891 (2018)	0.890 (2018) - PACA	Very High	32.7 (2015)	Low/Very low
Netherlands	0.933 (2018)	0.949 (2018) – North Holland	Very High	28.2 (2015)	Low/Very Low
South Africa	0.705 (2018)	0.741 (2018) – Western Cape	High	63.0 (2014)	Very High
Taiwan	0.882 (2014)	NA	Very High	33.8 (2018)	Low/Very Low
United States	0.920 (2018)	• 0.939 (2017) - Maryland • 0.904 (2017) – New Mexico	Very High	41.5 (2016)	Medium

Methods

Intra- and inter-observer errors for continuous (measurements) data were evaluated using Technical Error of Measurement (TEM) and relative TEM (rTEM), Bland-Altman plots, and Intra-Class Correlation Coefficients. Intra- and inter-observer agreement rates for ordinal data (maturation stages) were evaluated using quadratically weighted Cohen's Kappa coefficients. These inter- and intra-observer tests were essential, not only because they ensured reliable data collection across modalities and between observers, but because some of the data had been previously collected by other researchers (NIJ Award 2015-DN-BX-K009; Stull, L'Abbé, and Ousley 2014). All results were published and therefore can be found in Corron et al. (2021). The findings suggest there is no bias in skeletal and dental indicators size that is related to the type of collaborating institution, the context of origin of the individuals (forensic or clinical), or the type of medical image (CT scan or x-ray), and that these indicators could be reliably collected and compared across samples for anthropological research.

Growth and Development

The intrinsic relationship of ontogenetic patterns of skeletal growth and dental development was evaluated on the U.S. individuals, as it was the only sample for which all indicator types could be collected on the same subjects. Structural equation modeling (SEM) was used to assess the relationship between age and all growth indicators. Indicators were analyzed as two latent variables grouping diaphyseal dimensions on the one hand and dental development stages on the other hand (Kline 2016). The measurement models in SEM allow for the extraction of underlying growth factors from multiple measures of diaphyseal dimensions and dental development. SEM is the approach best suited for assessing growth as a whole by building a single model on a latent variable representing several indicators, as opposed to multiple regression which would require building several hierarchical linear models instead and would violate assumptions of homoscedasticity and independence of the predictor variables (Kline 2016; Verleye et al. 2004).

Variables were grouped into three categories to be used as latent variables in the structural equation models (SEM): arm growth, grouping all 10 upper limb measurements (HDL, HPB, HMSB, HDB, RDL, RPB, RMSB, RDB, UDL, and UMSB); leg growth, grouping all 8 lower limb measurements (FDL, FMSB, FDB, TDL, TPB, TMSB, TDB, and FBDL); and dental development, grouping the stages of 14 permanent teeth, excluding the lower and upper third molars. Variables in each latent variable group presented with high intervariable correlation. Measurements were centered and scaled prior to the analyses using their standard deviations. Missing data overlapped for arm and leg growth variables, but not for dental growth variables. Therefore, analyses were conducted for arm and leg variables jointly, but separately for dental variables. SEMs were built to examine growth patterns across ontogeny and within each life history stage. To ensure the variables did not violate any statistical assumptions required by SEM (*i.e.*, normality, homoscedasticity), analyses were run using the square root of age.

Variation in VNC size was evaluated using the antero-posterior (AP) and transverse (TR) diameters of the VNC measured for 1404 contemporary individuals in the SVAD samples from Colombia (N = 28), France (N = 484), the Netherlands (N = 23), Taiwan (N = 31), and the United States (N = 838). Dimensions were plotted against age to visualize their general growth patterns between sexes and countries. Principal component analyses (PCA), MANOVAs, and Partial Least Squares - Discriminant Analyses (PLS-DA) sought to interpret the factors behind VNC variation using the following covariates for each individual in the five samples: population (Colombia, France, the Netherlands, Taiwan, and the United States), sex, HDI level, and Gini coefficient level. Results were also examined according to the individuals' life history stage (Cameron and Bogin 2012; Bogin 1999). This covariate was primarily used to mitigate the wide range of age included in the sample as well as the uneven age distribution. The methodology and results were published in the American Journal of Biological Anthropology (Corron, Wolfe, and Stull 2023)and the data was made freely available (Corron, Wolfe, and Stull 2022).

Age Estimation

Age estimation models were built using the Mixed Cumulative Probit (MCP) algorithm, developed as part of two previous federally-funded grants (Awards NIJ 2015 DN-BX-K409 and NSF BCS-1551913). The MCP is a cumulative probit that combines several features to help address the statistical shortcomings of other age estimation methods, including heteroscedasticity, conditional dependence, mixes of ordinal and continuous variables, missing values, alternative specifications of the mean and noise responses (Stull et al., n.d.). The MCP allows to build both univariate and multivariate models combining any number or type of age indicators. The computational power required to run the MCP algorithm on all 62 indicators and all samples is extensive and comes with long waiting times before obtaining results (*i.e.*, > 4 months). Therefore, only univariate models and a six-variable model based on FDL, RDL, HME-EF, TC_Oss, max_M1, and man_I2 were run to evaluate the performance of global and population-specific models. Following 4 months of processing time, it has been decided that the global models (matching the methodology of Stull et al. [2022]) will be analyzed with a neural network, which is far faster though the inferential capacity of the performance is less.

Participants and other collaborating organizations

Samples originated from ten partner institutions, covering eight countries and four modalities: dry bone (Colombia), computed tomography (CT) scans (France, the Netherlands, Taiwan, the United States), Lodox Statscans (South Africa), and conventional x-rays (Angola and Brazil). Each collaboration provided previously generated medical images resulting in retrospective cross-sectional data used for all analyses (Table 3). The unique parameters per collaborator included: data source (hospital versus medical examiner's office), image modality (CT scan, Lodox Statscan, dry bone, panoramic or conventional radiograph), and type of image (element/region specific versus full body). Three collaborators are included in Table 3, but were not formally a part of this project: the

South African collaborators and the Office of the Chief Medical Examiner, Baltimore, Maryland. Data had been previously collected from these institutions either as part of the NIJ 2015-DN-BX-K409 award or from the PIs other research endeavors, however their data was included in all analyses and therefore included in this table.

Table 3 Context of derivatives per each collaborating institution.

Count	Partner Institution	Context	Modality	Derivatives	Availability
Angola	Departamento de Ciências da Vida, University of Coimbra and private medical cabinets from Luanda, Angola	Private practice	Conventional radiograph	Epiphyseal fusion stages Dental development stages	Derivatives
South Africa	Red Cross War Memorial Children's Hospital, Cape Town Hospital	Hospital	Lodox Statscan	Long bone dimensions	Derivatives
	Forensic Pathology Services, Salt River, Cape Town	Forensic	Lodox Statscan	Long bone dimensions	Derivatives
Colombia	Universidad de Antioquia, Medellin	Forensic	Dry bone (long bone dimensions and epiphyseal fusion) Conventional radiograph (dental development and epiphyseal fusion)	Long bone dimensions VNC diameters Epiphyseal fusion stages Dental development stages	Radiographs and derivatives
Brazil	Universidade de São Paulo (FOUSP)	Dental Practice	Panoramic radiograph	Dental development stages	Panoramic radiographs and derivatives
France	Public hospital services of Marseille (AP-HM)	Hospital	CT scan	Long bone dimensions VNC diameters Epiphyseal fusion stages Dental development stages Segmented bone surfaces	Derivatives
Nether-lands	Amsterdam Medical Center (Hospital)	Hospital	CT scan	Long bone dimensions VNC diameters Epiphyseal fusion stages Dental development stages Segmented bone surfaces	CT scans and derivatives
Taiwan	National Taiwan University Hospital, Taipei City	Hospital	CT scan	Long bone dimensions VNC diameters Epiphyseal fusion stages Dental development stages Segmented bone surfaces	Derivatives
United States*	Office of the Chief Medical Examiner, Baltimore, Maryland	Medico-legal	CT scan	Long bone dimensions Epiphyseal fusion stages	Derivatives
	University of New Mexico Health Sciences Center, Office of the Medical Investigator, Albuquerque, New Mexico	Medico-legal	CT scan	Long bone dimensions VNC diameters Epiphyseal fusion stages Dental development stages Pelvic landmarks Segmented bone surfaces	CT scans** and derivatives

* Age, sex, population affiliation/social race, manner of death and cause of death are also known.
** The University of New Mexico CT scans are directly accessible via the NMDID website and linked to the SVAD derivatives (https://nmdid.unm.edu/).

1) University of New Mexico (UNM) Health Sciences Center, Office of the Medical Investigator (U.S.)

The CT scan images of 1,077 individuals aged between birth and 20 years of individuals who underwent post-mortem full-body CT scans at the UNM OMI were used. CT scanning was conducted using a Phillips Brilliance Big Bore 16-slice multi-detector scanner prior to autopsy, with a 512x512 image matrix, 0.5mm slice thickness and 1mm overlap. Part of the data had been previously collected as part of another NIJ grant (NIJ Award 2015-DN-BX-K009; dental development, diaphyseal dimensions, and epiphyseal fusion). The data was collected from individuals that are also part of the New Mexico Decedent Image Database (NMDID) and freely accessible (https://nmdid.unm.edu/). All skeletal and dental data previously collected as part of that previous NIJ grant were also utilized in this project.

2) Radiology Department of the Public Hospital Services of Marseille (France)

The Radiology Department centralizes all medical images taken on patients from the five public hospitals of Marseille on the Picture Archiving and Communication System (PACS, McKesson Medical Imaging Group, Richmond, BC, Canada). The CT scans were performed with a 64-row multidetector CT scan (Somatom Sensation 64, Siemens®, Erlangen, Germany) with the following scanning parameters: 120 KV, 50-150 mAs, thickness: 0.6 mm. Most scans were obtained after administration of an intravenous contrast media. CT examinations of the pelvis, vertebrae, and some of the long bones of 578 healthy children and adolescents aged 0 to 15 years performed between 2008 and 2018 to detect acute diseases or traumas, or as part of postmortem analysis were acquired.

3) Department of Radiology at the Emma Children's Hospital – Academic Medical Center Amsterdam (the Netherlands)

Anonymized CT images of 218 identified (known age and sex) individuals aged 0 to 15 years from the Netherlands having undergone examinations for various medical reasons were used for this study. Images are stored in a database in the Department of Radiology at the Emma Children's Hospital in Amsterdam and securely transferred to the US for data collection.

4) National Taiwan University Hospital (Taiwan)

CT images of 730 individuals aged 0 to 16 years from Taiwan having undergone examinations for various medical reasons were used for this study. Images were compiled on external hard-drives (one folder per individual) by the on-site partner (Dr An-Di Yim) and sent to the postdoctoral fellow in the U.S., for off-site data collection of the osteometric and qualitative data (epiphyseal fusion and dental development) on reconstructed images of the bones and teeth.

5) Laboratory of Forensic Anthropology, Department of Life Sciences, University of Coimbra (Portugal, with data from Angola)

The x-rays of identified individuals aged 0 to 15 years from a medical practice in Angola used for this study were compiled by the on-site partners (Dr H. Rodrigue and

Pr E. Cunha) and sent to the postdoctoral fellow in the U.S., for off-site data collection of qualitative skeletal data.

6) Department of Anthropology, University of Antioquia, Laboratory of Anthropology and Osteology (Colombia)

The Colombian material included in the present study is composed of the skeletal and dental elements of individuals aged 0 to 15 years dated from the 21st century. Individuals in this sample died between the 1990's and early 2000's and were exhumed from public cemeteries in Medellin and housed at the Universidad de Antioquia. Because the material was skeletonized, a portable x-ray machine (NOMAD Pro2 Handheld Portable Dental X-Ray Aribex) was used to generate radiographic images of the teeth.

7) Universidade de São Paulo (FOUSP) Dental School (Brazil)

The Brazilian material included in this study consists of panoramic dental x-rays of 500 patients (252 males and 248 females) aged 5 to 15 years compiled by our on-site partner, Dr Maria Gabriela Haye Biazevic. Images were sent to the postdoctoral fellow in the U.S., for off-site data collection of dental development stages scored directly on the x-rays.

DICOM images require medical imaging visualization software for data collection. The two-dimensional conventional (Angola, Colombia) and panoramic (Brazil) radiographs are stored in other formats common to digital images (e.g., .JPG, .TIFF formats), thus achieving significant savings in image storage size. Other medical imaging databases composed of .JPG files have been set up and studies have shown that loss of information is minimal and image resolution is preserved after conversions from .dcm files. However, because of the possibility of distortion and/or magnification associated with these three radiographic samples, only ordinal data (*i.e.*, developmental stages), not continuous data (i.e., measurements), were collected on them.

Changes in approach from original design and reason for change, if applicable

Some collaborations were not deemed possible that were included in the original proposal and some new collaborations were added to the project (USP and Taiwan). More collaborations were approved, but the Covid-19 pandemic led to restricted travel and data collection. Two no-cost extensions were approved in hopes the pandemic travel restrictions would be lifted and data collection could proceed.

Accomplishments, Results, and Findings

Reliability, Error, and Agreement of Indicators

Diaphyseal dimensions yielded low intra-observer and inter-observer errors (intra-observer TEM = 0.16, interobserver TEM=0.13 mm) as did vertebral neural canal measurements (intraobserver TEM = 0.10 mm, interobserver TEM=0.19 mm) on virtual surfaces (Stull and Corron 2021b). Agreement rates for epiphyseal fusion (intraobserver Cohen's kappa = 0.92, interobserver Cohen's kappa = 0.86) and dental development (intraobserver Cohen's kappa = 0.939, interobserver Cohen's kappa = 0.965) scored on CT scans were high (Corron et al., 2021). Epiphyseal fusion scored between CT and X-ray (Cohen's kappa = 0.75) and dry bone and X-ray (Cohen's kappa = 0.78) were also high (L. K. Corron et al. 2021). These results confirmed the data acquired from different imaging modalities could be pooled for our analyses.

Univariate Variation in Growth and Development: Sex and Population

Because the age range, sex distribution, and sample sizes were not uniform across samples (20-22 years for Colombia and the U.S., 15-16 years for Angola, Taiwan, France, the Netherlands, 12 years for South Africa) it is difficult to interpret population differences without considering the effect of age truncation and imbalanced age distributions across samples. This is especially true for ordinal data that has developmental stages and chronological ages to consider. For that reason, we rarely explored the data with statistical tests that assume specific data shapes and/or sizes and rather relied on the trends in the visualizations.

Scatterplots revealed diaphyseal lengths had greater consistency among countries compared to the diaphyseal breadths; breadths still displayed overlap among the countries but had more variation (Figure 7). Population differences in growth trajectories are more apparent from childhood onwards (>= 3 years). Sex differences were observed for proximal and distal long bone breadths, with males exhibiting larger breadths than females, usually beginning in childhood (Figure 8). These patterns have also been observed in much younger individuals from Australia and South Africa (Coussens et al. 2002; Kruger, L'Abbe, and Stull 2017; Stull, L'Abbé, and Ousley 2017; Stull et al. 2020). In contrast to proximal and distal breadths, mid-shaft breadth dimensions had more similarities between sexes and among populations in the current samples.

Even with the varied sampling, the dental developmental stages have comparable ranges for chronological ages for all countries (Figure 9). When there were sex differences in dental development, females were more advanced than males (Figure 9), a trend that has been acknowledged in various other studies (Demirjian and Levesque 1980; Moorrees, Fanning, and Hunt 1963; Hunt and Gleiser 1955; AlQahtani, Hector, and Liversidge 2010; Liversidge 2015 Šešelj, Sherwood, and Konigsberg 2019).

Some sexual dimorphism was also apparent in the timing of epiphyseal fusion of the secondary ossification centers corresponding to the proximal and distal long bone epiphyses, the acetabular epiphysis (ILIS_EF), and the calcaneal tuberosity (CT_EF), with

females reaching active fusion stages earlier than males (Figure 10). No sex differences were observed in the timing of the appearance of ossification centers during infancy, childhood, and the juvenile life history stages. Some sex differences between males and females are present in the timing of the appearance of the trochlea, capitulum, and lateral epicondyle, all of which occur around puberty. These patterns follow known trends of sexual dimorphism in the timing and duration of fusion for secondary ossification centers, with females starting fusion earlier than males and males exhibiting a more prolonged maturation phase until early adulthood (Baughan et al. 1980; Cole et al. 2015; Grissom et al. 2018; Tanner 1981).

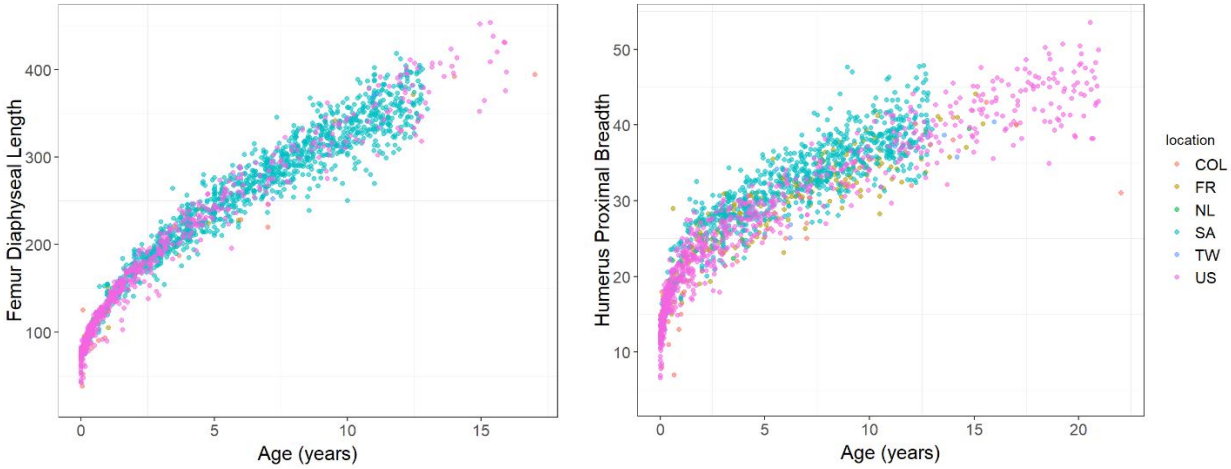

Figure 7. Visualizations of femur diaphyseal length (left) and humerus proximal breadth (right) against age across ontogeny in six samples from the SVAD.

.

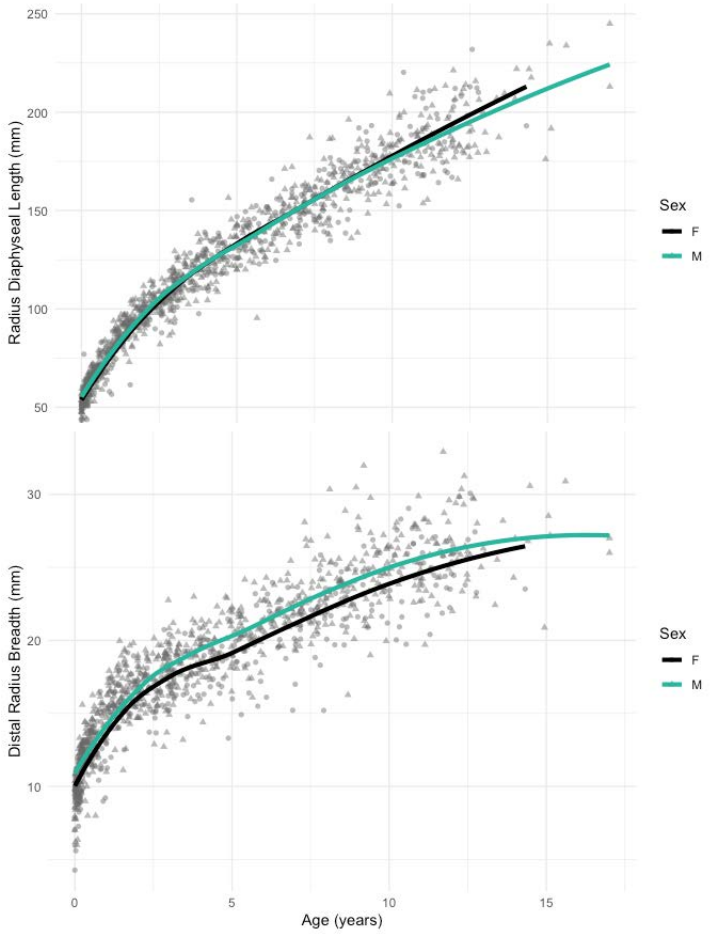

Figure 8. Examples of dental variables exhibiting sex differences in size or timing of development (top) and the generally similar developmental trends when considering all countries (bottom).

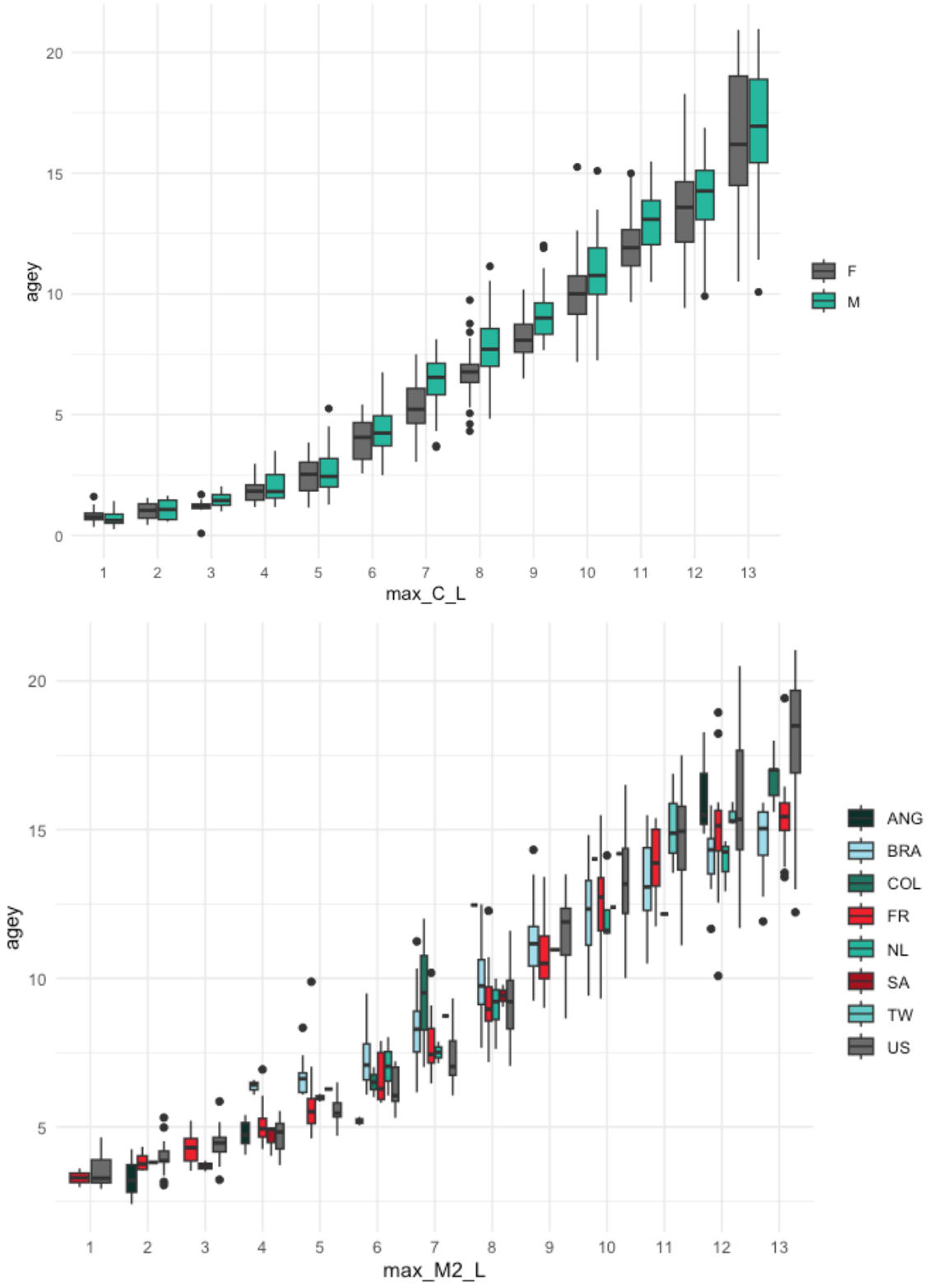

Figure 9. Examples of dental variables exhibiting sex differences in size or timing of development (top, Maxillary Canine) and the generally similar developmental trends when considering all countries (bottom, maxillary second left molar).

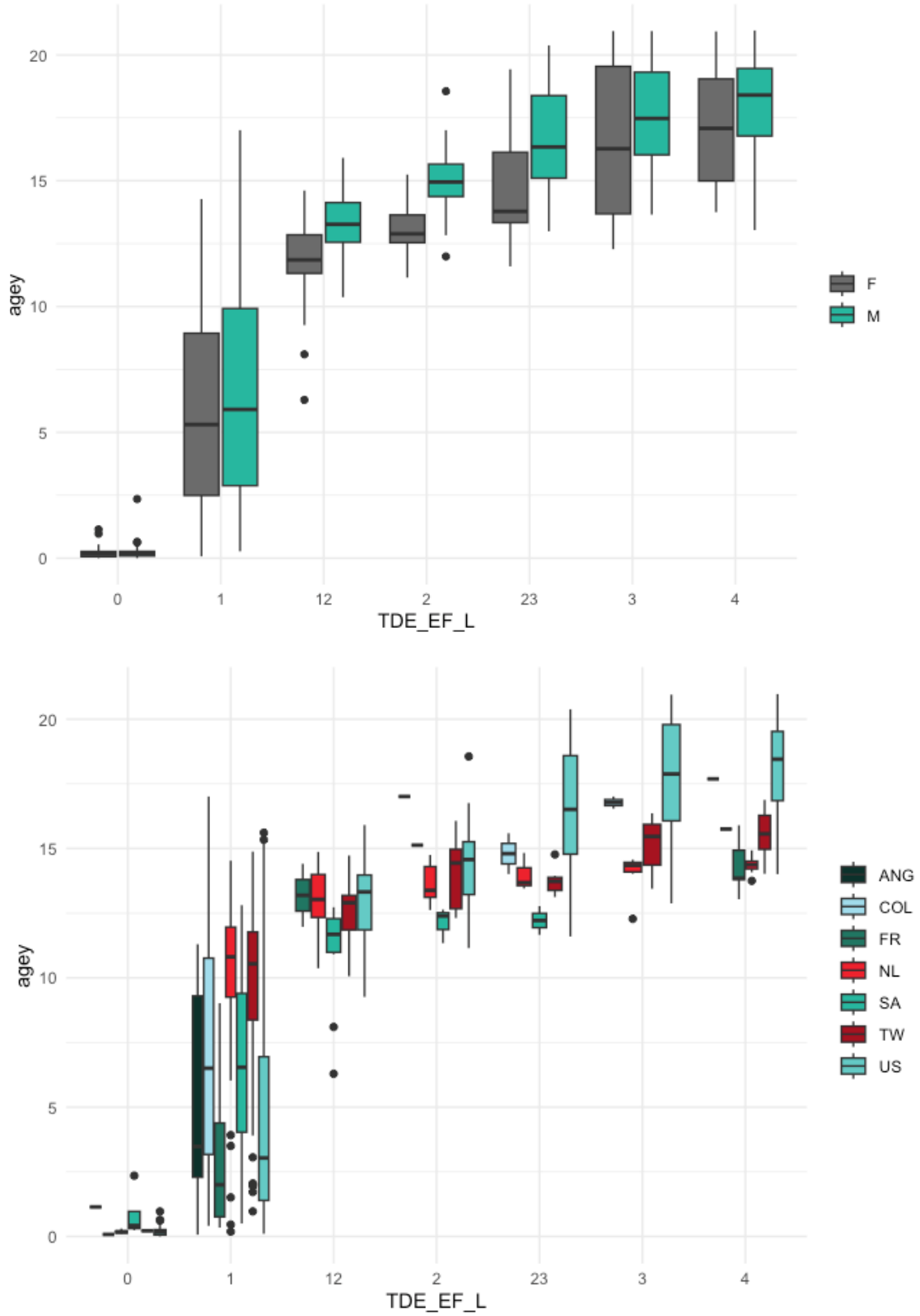

Figure 10. Boxplots of the distribution of epiphyseal fusion stages for the distal tibia epiphysis. Top: Sex differences; Bottom: Population differences.

Structural Equation Modeling

Structural equation modeling (SEM) facilitates the evaluation of patterns, rates, and specific sequences of growth and development of long bones and permanent teeth throughout ontogeny, from birth to 20 years. SEM results show that the variables of the upper and lower limbs loaded well onto the latent growth factors using a linear relationship, while dental development loaded well onto the dental latent growth factor using a quadratic relationship. Factor loadings were all > 0.95 for both types of variables. Results suggest that the relationship between growth variables (arm and leg factors compared to dental factors) may not be the same across different life history stages. Diaphyseal lengths (HDL, RDL, UDL for the arm, and FDL, TDL, FBDL for the leg) are strong growth indicators across all life history stages, while diaphyseal breadths are strong growth indicators during infancy but become less indicative of growth when individuals reach childhood and adolescence. Similarly, dental development does not have the same rate between infancy and childhood or between childhood and adolescence, as illustrated by differences in factor loadings and active variables among life history stages, essentially capturing earlier and later developing permanent teeth. These analyses allow the identification of three different phases in dental development: 1) between birth and one year of age/early infancy with the development of early forming teeth such as the incisors, canines, and first permanent molars, 2) between the age of one year and 13/14 years of age, or late infancy to early adolescence, with the early development of the premolars and molars, and 3) from 13/14 years onward, *i.e*. from mid- to late adolescence when the posterior teeth finalize their development.

Multivariate Variation in Growth and Development: Gini, HDI, and Population

While the trends described herein are quite apparent, it is necessary to acknowledge that the data is very imbalanced for HDI and Gini Categories. The imbalanced classes do substantially impact the training and testing accuracies, with the testing accuracies being lower than the training accuracies, though the overall trends are visually obvious. All classification statistics and visualizations are the testing sample and not the training sample. There are clear separations in the long bones (Figure 11), dental development (Figure 12), and epiphyseal fusion (Figure 13) based on HDI categories, Gini Categories, and Country/Location. FDA analyses showed that classification accuracy based on long bone dimensions, epiphyseal fusion, and dental development was moderate when classifying individuals into HDI groups (83%, 73%, and 88%, respectively), was slightly lower when classifying individuals into Gini groups (96%, 64%, and 76%, respectively), and comparable/slightly lower when classifying individuals into population of origin (83%, 64%, and 75%, respectively). Notably, we would expect the highest classification statistics with HDI, as this only ever had two groups where Gini and Population could have three groups. Not all groups were available to build a model, since the samples were small when there can be no missing data. Therefore, most analyses rely on France, South Africa, the United States, and Brazil.

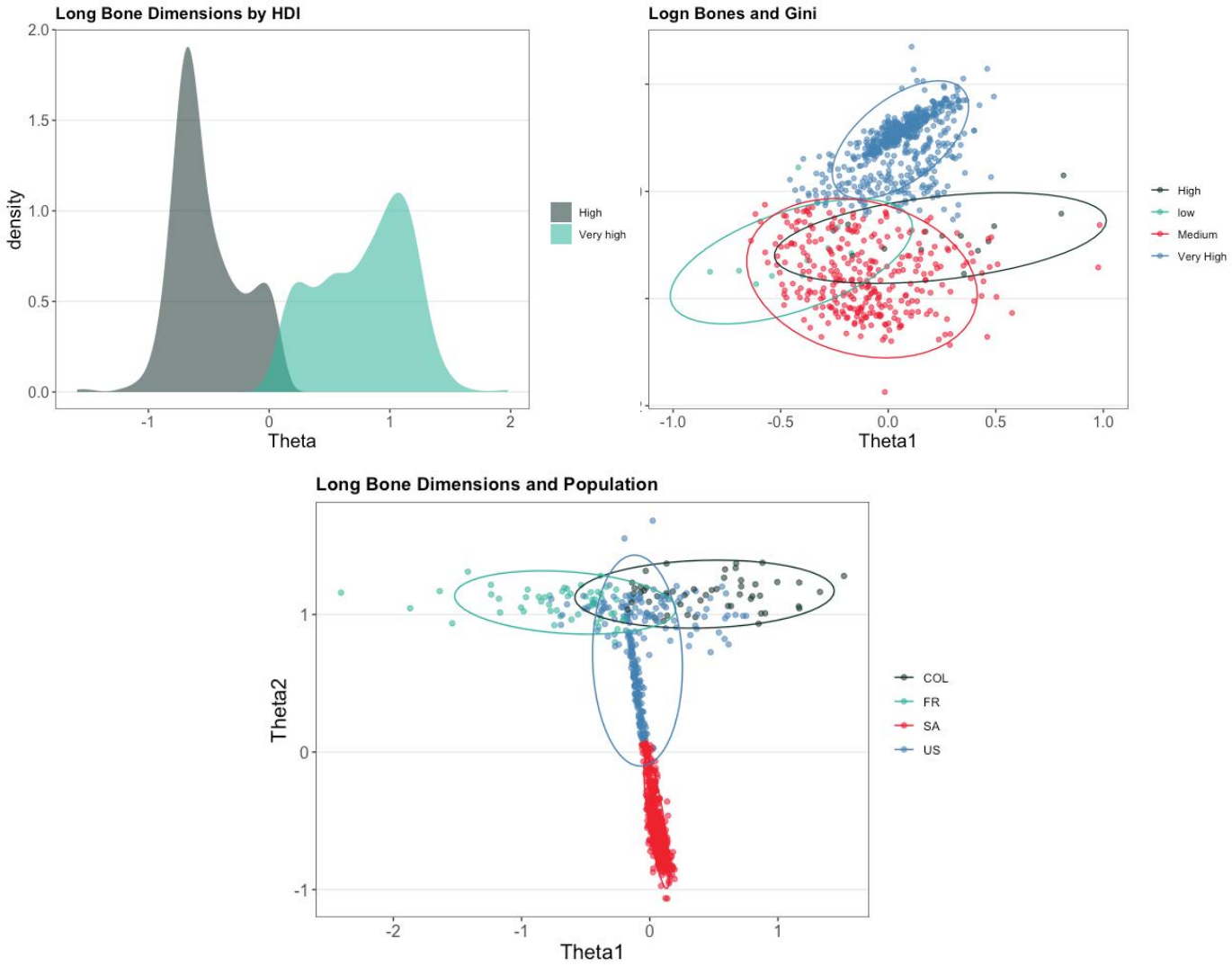

Figure 11. Density plots of FDA classifications into HDI, Gini, and population categories using long bone dimensions.

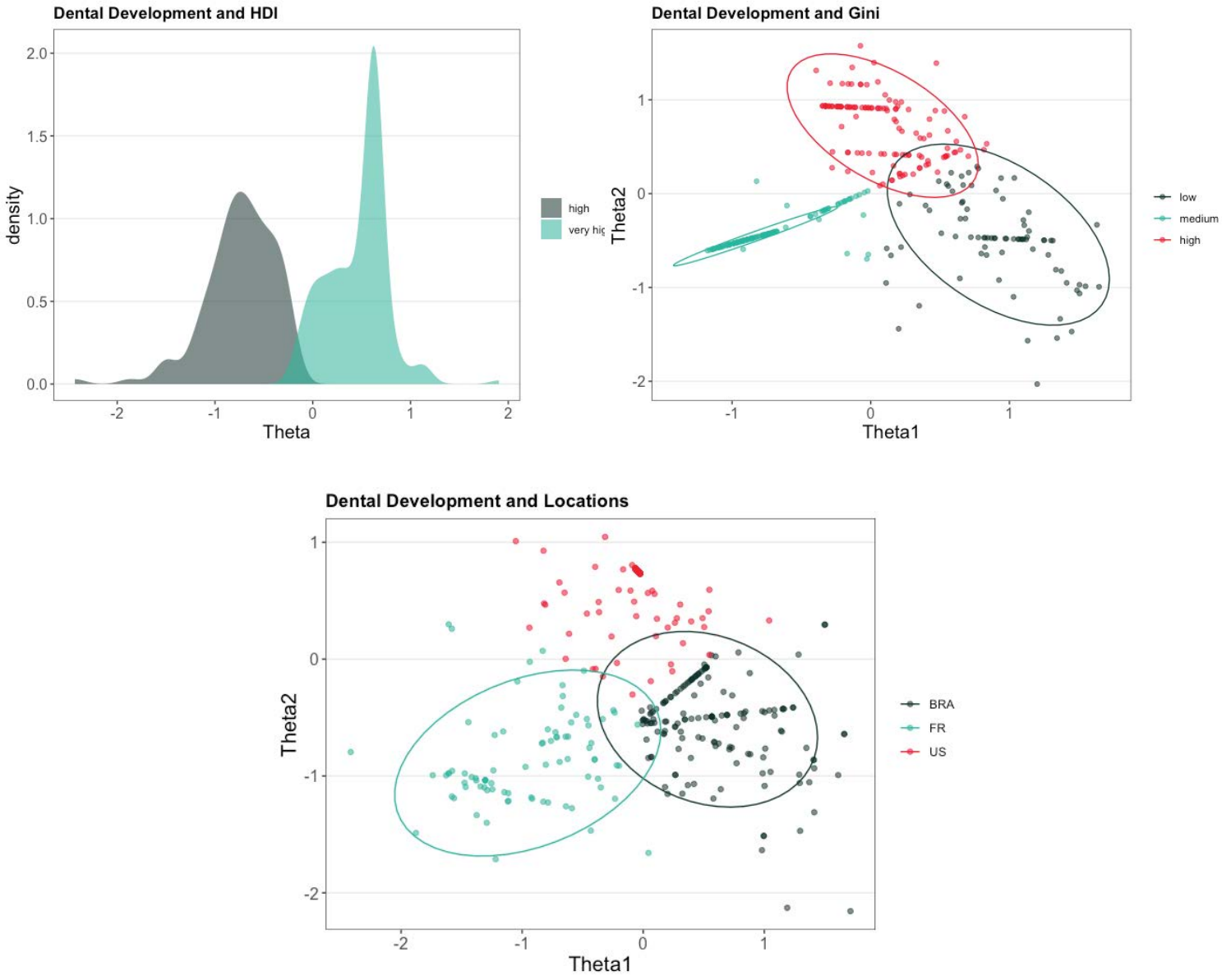

Figure 12. Density plots of FDA classifications into HDI, Gini, and population categories using dental development.

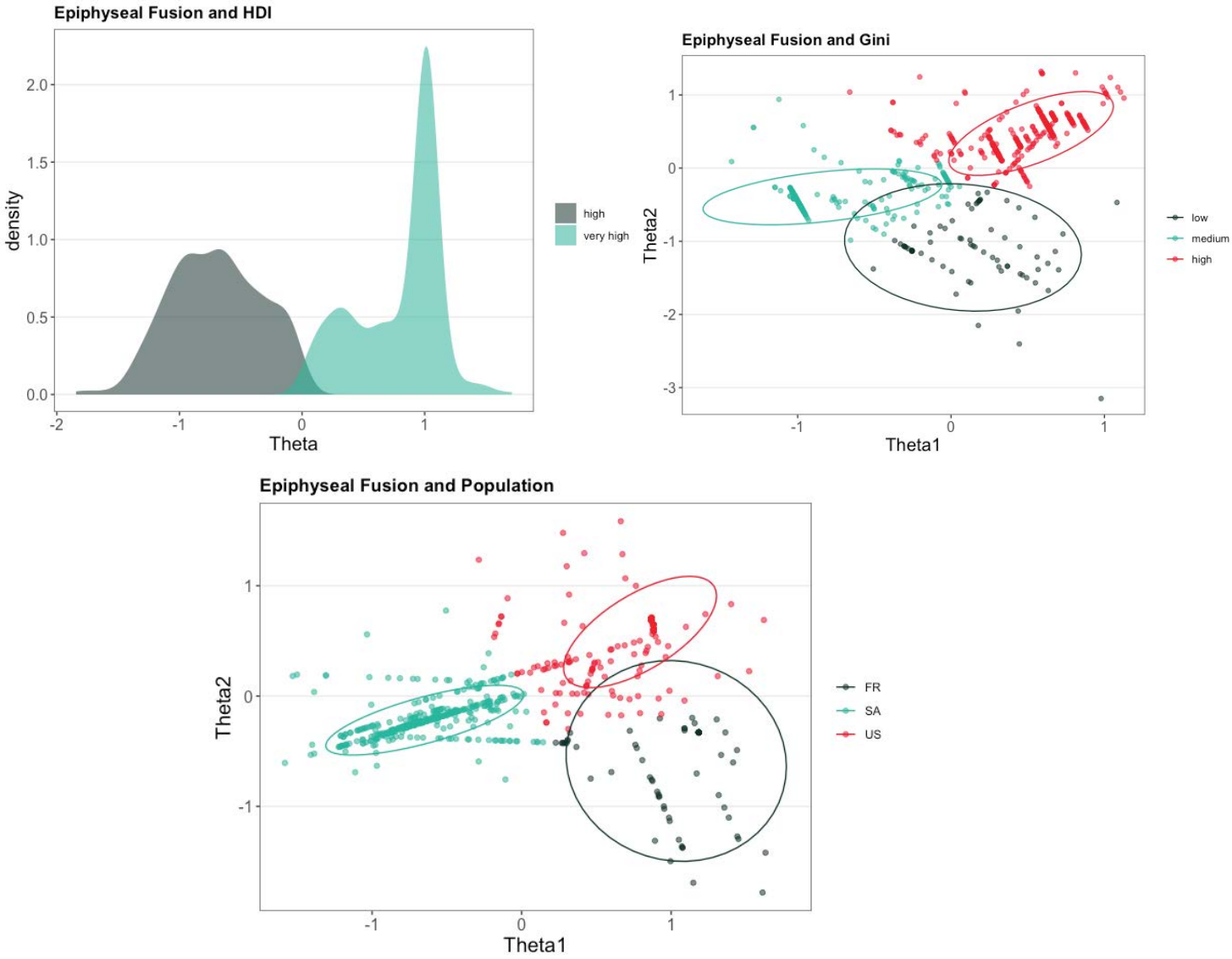

Figure 13. Density plots of FDA classifications into HDI, Gini, and population categories using epiphyseal fusion.

Vertebral Neural Canals

Visualizations of the AP and TR diameters of all vertebrae and age show that Colombian individuals exhibit smaller diameters compared to the four other samples (Figure 14), while individuals from France, the Netherlands, Taiwan, and the U.S. have a comparable range of variation. Unlike raw VNC diameters, VNC ratios (AP/TR diameters) follow comparable patterns and ranges across all samples throughout ontogeny with similar and consistent variation present in each sample (Figure 14).

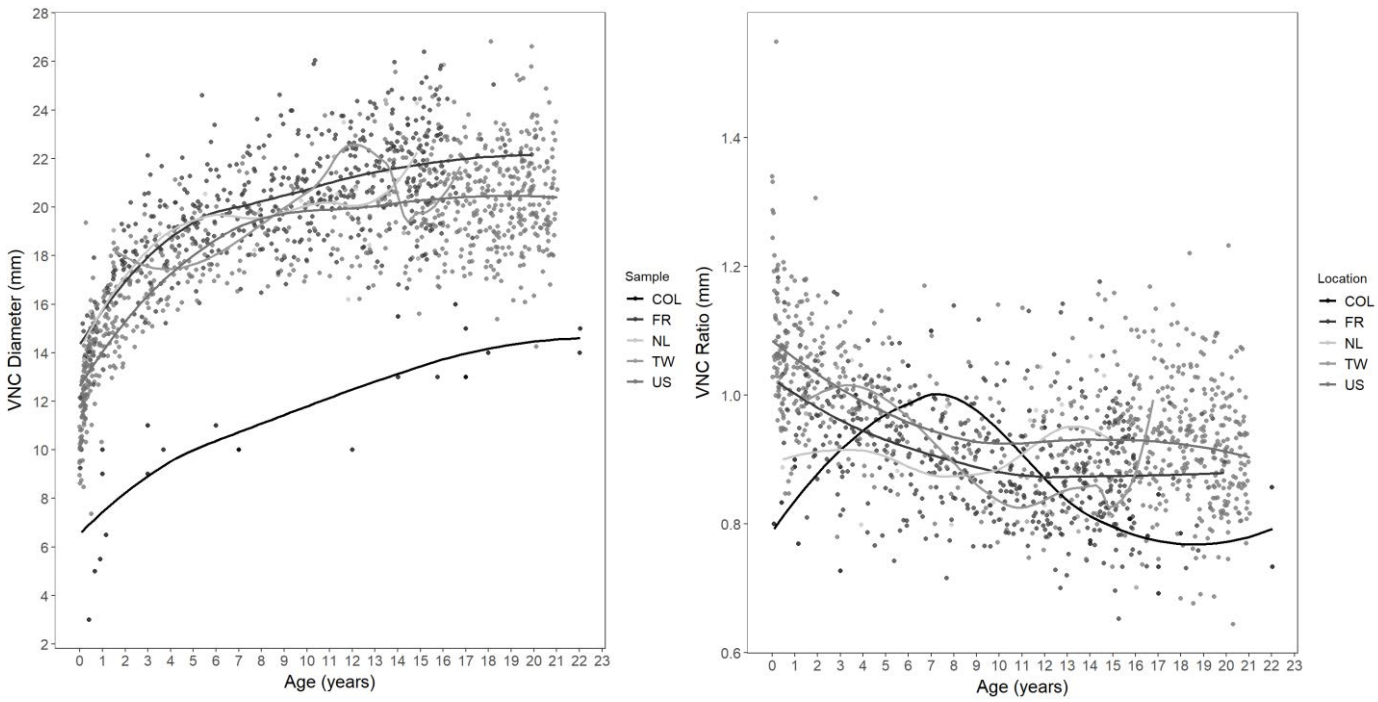

Figure 14. Plots of VNC size (left, Th12 AP diameter) and VNC ratio (right, TH12 AP diameter/TR diameter) against age across ontogeny in five SVAD samples.

The PCA results of the centered and scaled VNC diameters show that PC1 and PC2 contribute to 74.5% and 8.9% of variation, respectively, with TR diameters contributing more to PC1 and lumbar variables contributing more to the loadings of both PC1 and PC2. There are no noticeable differences in variation according to sex; males and females overlap. In terms of life history stage, infants have a distinct cluster compared to children, juveniles, and adolescents, which is to be expected based on the developmental trajectories. There are distinct clusters between the High and Very High HDI, and between the High and Medium/Low Gini categories.

PLS-DA yielded accuracy rates for the test samples of 61.6% and 64.76% when classifying into population of origin (five categories) and Gini (three categories) groups, respectively. Model performance was much higher when classifying individuals into HDI (two categories) groups, with 98.28% accuracy. Results show that PC scores do not differ significantly between the raw VNC data and VNC data scaled using stature. Therefore, these results still highlight that the smaller VNC size in the Colombian sample does not appear to be informed by the inherently smaller stature presented by this population. These findings indicate that quality of life (*i.e.*, HDI) may be a driving factor influencing VNC dimensions. However, the causal relationship between the two, or even correlation between the two, cannot be yet accepted and requires further exploration with more varied reference samples.

Global versus Population-specific Age Estimation Models

Because of the amount of data and models to summarize in this section, we chose to present a smaller number of variables to provide a snapshot of the findings. Six single variable age estimation models are compared in Figures 15-19; models were developed using each available population that had a large enough sample and a global sample that included all eight populations, no matter the sample size. This research design was adopted because not all samples are large enough to develop valid models (i.e., Colombia). We randomly selected two diaphyseal lengths (femur and radius diaphyseal lengths [FDL and RDL, respectively]), two anatomical locations for epiphyseal fusion (medial epicondyle of the humerus [HME-EF] and tarsal count [TC_Oss]), and two teeth (maxillary first molar [max_M1] and mandibular second incisor [man_I2]).

When exploring the performance of the ordinal data, it is obvious that the global sample generally captures most of the variation expressed by each country specific model (Figures 15 – 17). While there is some variation in the age ranges associated with the credible interval (CI), a lot of the variation is most likely linked to the different age ranges associated with each sample. For example, the United States has an age range of birth to 20 years while most international samples only include individuals up to 15 years, and depending on the institution may not include individuals as young as birth (i.e., Brazil). Some of the varying performance is also linked to the data type, which is discussed in detail in (K. E. Stull et al. 2022). The variable age range, data types, and varying sample sizes all affect the performance of the models, and this is especially apparent in Figures 17 and 18. While the global model is not always more precise or less biased, the model appears to be more consistent in its performance (Figures 18 and 19), which is likely because it was trained on larger and more diverse samples from across the entire age range.

Importantly, there are advantages and disadvantages to each modeling approach and there is not a 'one model is best approach'. Instead, it will always be dependent on the context of the forensic case, what data is available, and the confidence associated with the decedent's nationality.

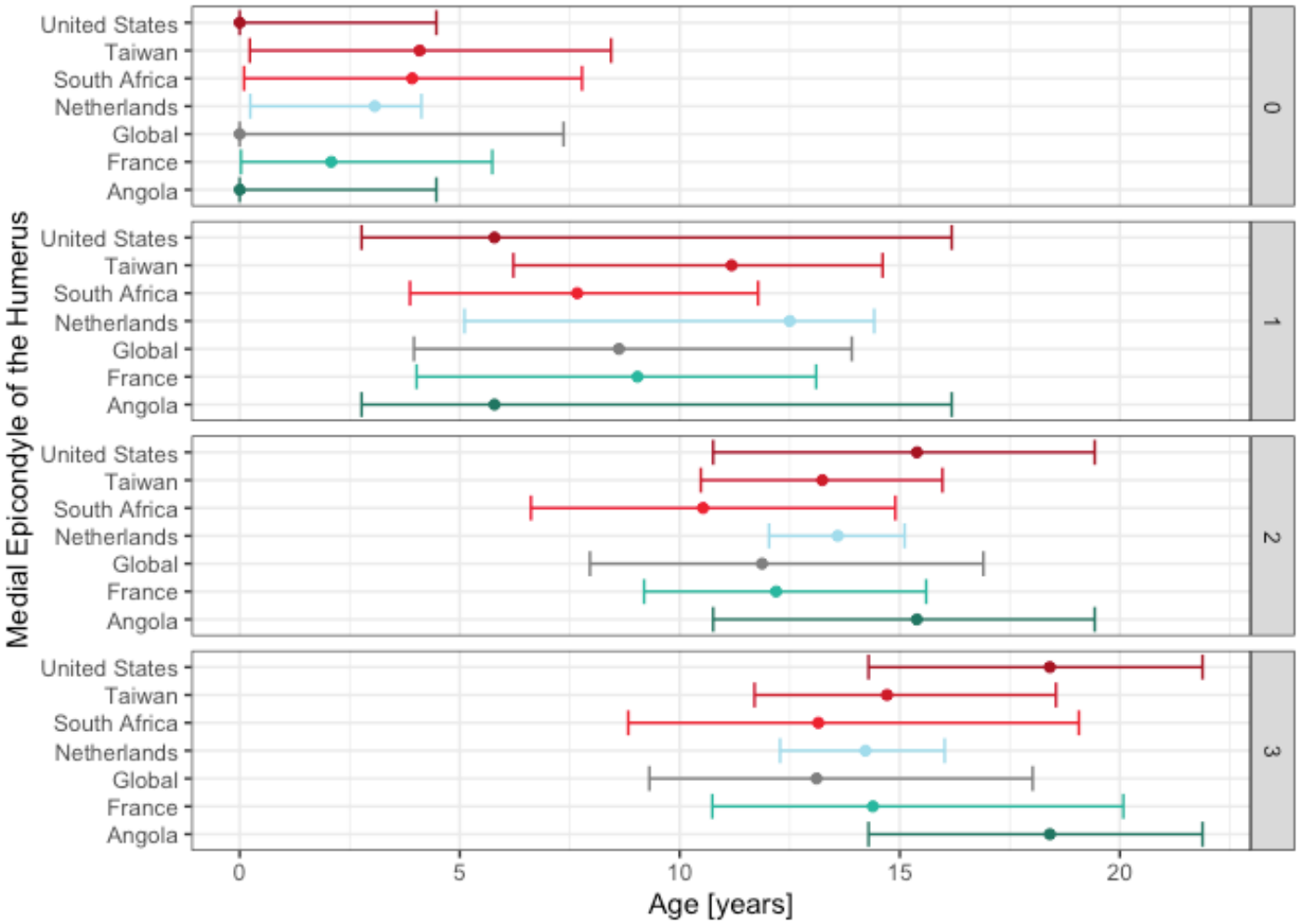

Figure 15. Credible intervals (95%) associated with the six-population specific age estimation models and one global age estimation model using the medial epicondyle of the humerus.

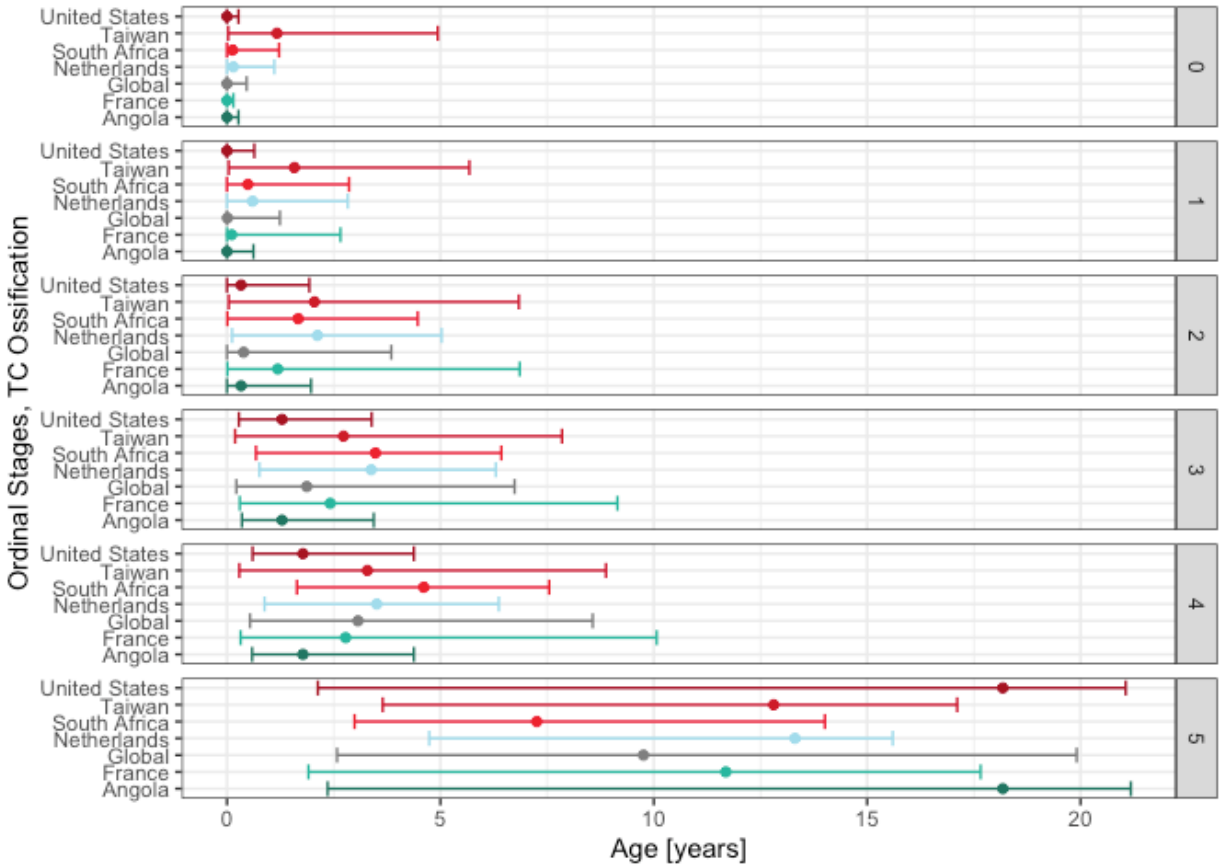

Figure 16. Credible intervals (95%) associated with the six-population specific age estimation models and one global age estimation model using tarsal count.

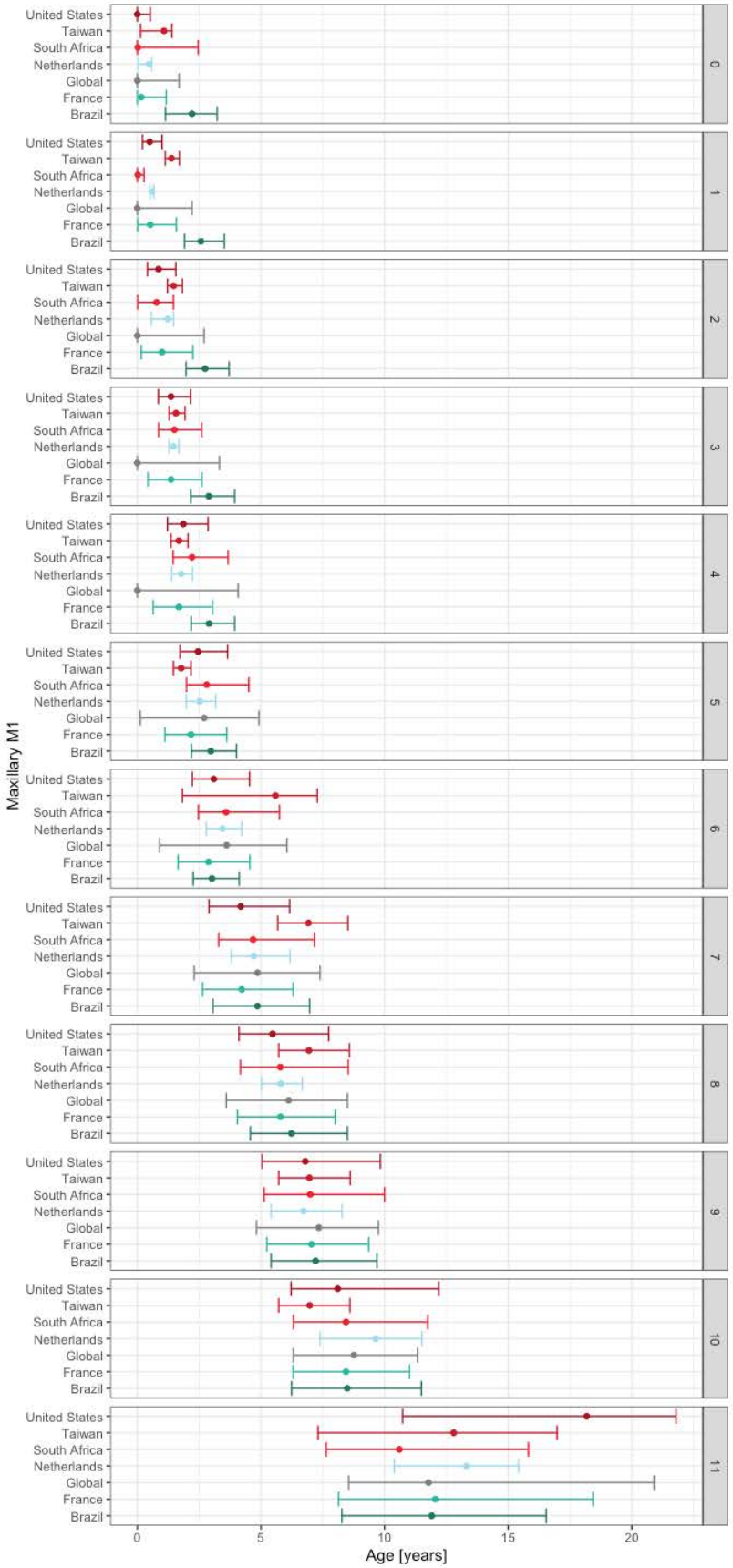

Figure 17. Credible intervals (95%) associated with the six-population specific age estimation models and one global age estimation model using the maxillary first molar.

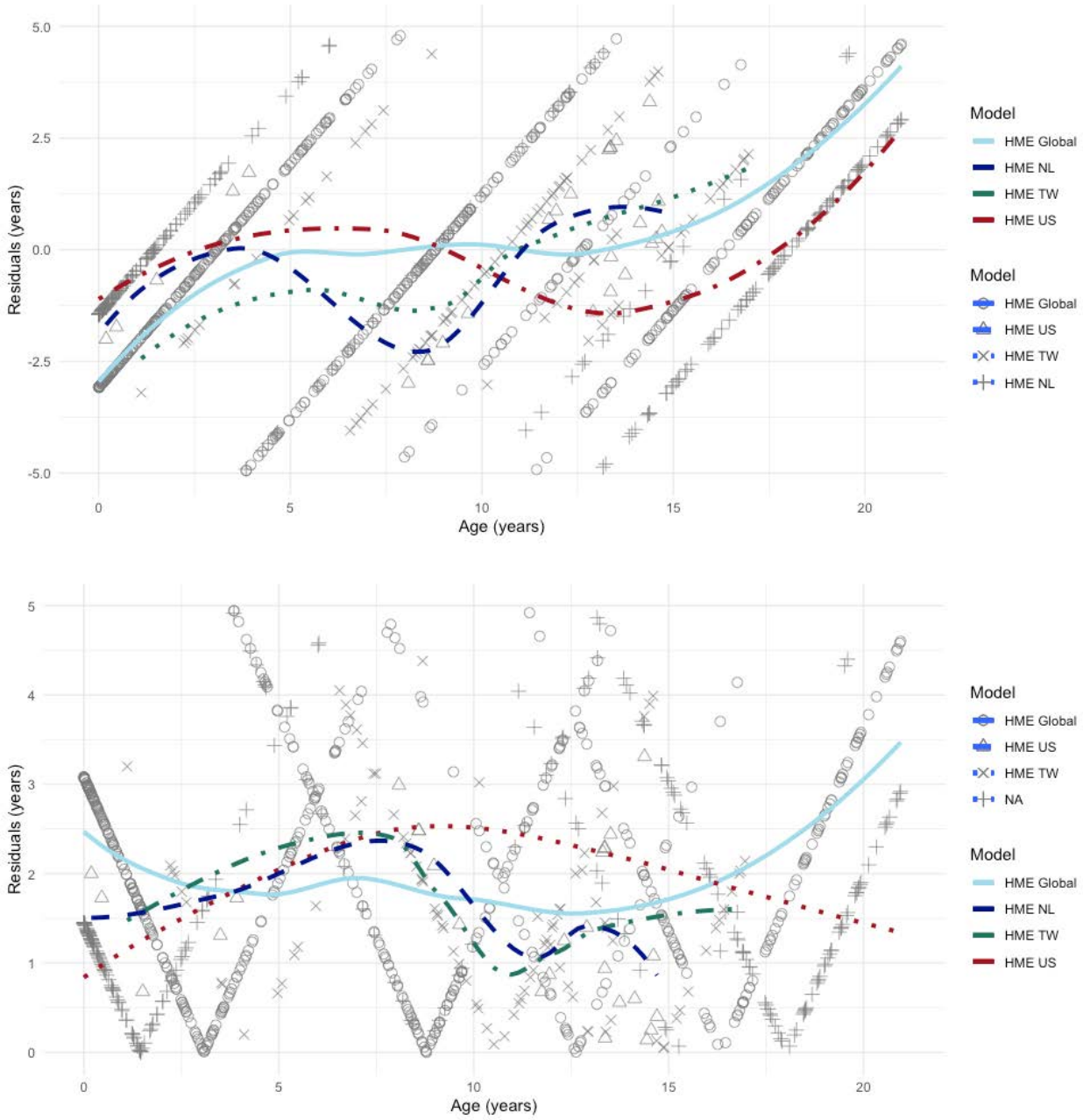

Figure 18. Top: Bias (in years) associated with the medial epicondyle of the humerus age estimation models for the US (HME US), Taiwan (HME TW), the Netherlands (HME NL), and the global model (HME Global). Bottom: Accuracy values (in years) associated with the medial epicondyle of the humerus age estimation models for the US (HME US), Taiwan (HME TW), the Netherlands (HME NL), and the global model (HME Global).

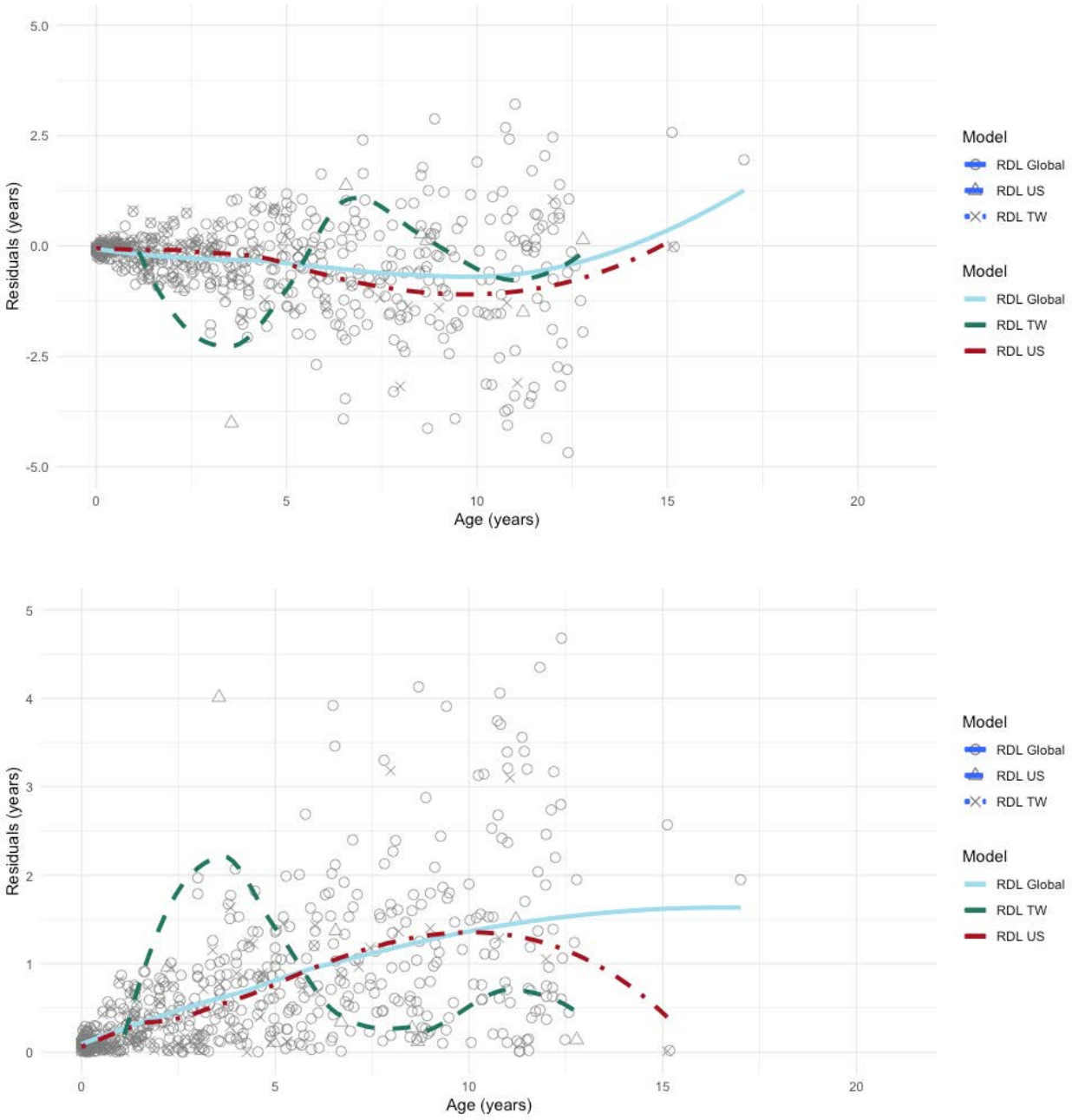

Figure 19. Top: Bias (in years) associated with the radius diaphyseal length age estimation models for the US (RDL US), Taiwan (RDL TW), and the global model (RDL Global). Bottom: Accuracy values (in years) associated with the radius diaphyseal length age estimation models for the US (RDL US), Taiwan (RDL TW), and the global model (RDL Global).

Application of the Global MCP Models

Univariate and multivariate MCP age estimation models based on the U.S. sample are already available through the MCP-S-Age GUI (K. E. Stull et al. 2022). The user can input any number of the 62 age indicators (18 long bone measurements, 28 stages of epiphyseal fusion and ossification, or 16 stages of dental development) and run the corresponding models to obtain a mean point estimate of age and the associated 95% percentiles calculated using the posterior probability of the mean. The MCP is methodologically appropriate for subadult age indicators, however, it does require a substantial amount of processing time when there are large samples and large number of variables in the models. Therefore, we are still waiting on the optimization/completion of all the global multivariate and mixed models. Once all the population-specific and global models are created and validated, they will be published in a peer-review journal and made available in MCP-S-Age for users to choose the type of model(s) and sample(s) they wish to apply to estimate the age of an unknown individual.

Limitations

The current project amassed the largest and most diverse age indicator dataset of individuals between birth and 20 years. However, even with the massive amount of data, there are still limitations to consider in the methodology and interpretations. The idiosyncrasies of each institution determined the type and amount of data. While all the data is helpful, especially when building global models, the sampling impacts the feasibility of developing population-specific models, and the downstream interpretations of the effects of the covariates (HDI, gini index, sex, population, etc.). The more limited the sample, in age range and number of variables per person, the less suitable the data are for a population-specific age estimation model. While not all countries were able to have their own population specific model, all data was included in the global model.

Artifacts

Data Availability

In addition to the new insights on the influence of covariates on ontogeny across populations, and the improvement of subadult age estimation, this project has allowed to amass an unprecedented amount of standardized subadult skeletal and dental reference data and compiled it into a freely accessible repository: the Subadult Virtual Anthropology Database, or SVAD (Stull and Corron 2022). The SVAD hosts derivatives of this research, including data collection protocols, raw datasets, segmented surfaces of long bones and pelves that can be accessed by or provided to fellow researchers and students for their own projects. Each country's complete dataset (.CSV formats) including demographics and age indicators is available for download via the Subadult Virtual Anthropology Database (SVAD) Zenodo Community, a freely accessible online research repository.

SVAD Zenodo community:
https://zenodo.org/communities/svad/?page=1&size=20

SVAD webpage: https://www.unr.edu/anthropology/research-and-facilities/subadult-database

Peer-reviewed Publications

Stull, KE, Chu, EY, Wolfe, CA, Corron, LK, and Price, MW (*in prep*) "Global versus Population Specific Subadult Age Estimation Models." *Forensic Sciences International.*

Corron LK, Yang Y, Stull KE (*in prep*) "Exploring dental development and long bone growth in the United States population." *Annals of Human Biology*

Stull, KE and Corron LK, (*in prep*) "Exploring Differential Growth and Development Patterns Associated with Human Development Index and Gini Coefficient." *Annals of Human Biology*

Corron, Louise K, Christopher A Wolfe, and Kyra E Stull. 2023. "A Multifaceted Exploration of Ontogenetic Variation in Vertebral Neural Canal Size across Contemporary Populations." *American Journal of Biological Anthropology* 180 (2): 328–51. https://doi.org/10.1002/ajpa.24675.

Stull KE, Corron LK (2022) The Subadult Virtual Anthropology Database (SVAD): an accessible repository of contemporary subadult reference data. *Forensic Sciences*. 2(1):20-36.

Corron LK, Stock MK, Cole SJ, Hulse CN, Garvin HM, Klales AR, Stull KE (2021) Standardizing ordinal subadult age indicators: testing for observer agreement and consistency across modalities. *Forensic Science International*. 320

Stull KE, Wolfe CA, Corron LK, Heim K, Hulse CN, Pilloud MA. (2020) A comparison of subadult skeletal and dental development based on living and deceased samples. *American Journal of Physical Anthropology*. 175(1):36-58.

Stock MK, Garvin HM, Corron LK, Hulse CN, Cirillo LE, Klales AR, Colman KL, Stull KE. (2020) The importance of processing procedures and threshold values in CT scan segmentation of skeletal elements: an example using the immature os coxa. *Forensic Science International*. 309

Conference Papers

1. Annual Meeting for the American Academy of Forensic Sciences, Seattle, WA, USA (2018)

Stull, KE and Price, MW. "Exploring the Performance of a Global Subadult Age Estimation Model Using Unsupervised Machine Learning Techniques." 70th Annual Meeting of the American Academy of Forensic Sciences, Seattle, WA, USA.

2. Annual Meeting for the American Academy of Forensic Sciences, Seattle, WA, USA (2018)

Gonzalez, S and Stull, KE. "Comparing Socio-Economic and Population Level Differences and Quantifying their Impact on Subadult Age Estimations." 70th Annual Meeting of the American Academy of Forensic Sciences.

3. Annual Meeting for the Forensic Anthropology Society of Europe (FASE), Marseille, France (2018)

Corron, LK and Stull, KE. "The predictive ability of subadult age indicators according to life history stages: a preliminary study" Annual Meeting for the Forensic Anthropology Society of Europe, Marseille, France

4. Annual Meeting for the American Academy of Forensic Sciences, Baltimore, MD, USA (2019)

Corron, LK, Stull, KE, Price, MW and Yang, Y. "Variations in Skeletal and Dental Growth and Development Patterns and Their Effect on Age Estimation: A Preliminary Study of Five Populations" 71st Annual Meeting for the American Academy of Forensic Sciences, Baltimore, MD, USA.

5. Annual Meeting for the American Association of Physical Anthropology, Cleveland, OH, USA (2019)

Stock, MK, Corron, LK, Cirillo, LE, Garvin, HM, Hulse, CN, Cole, SJ, Stull, KE, and Klales, AR. "Quantifying error in virtual data collection: the impact of MSCT scan segmentation protocol and inter-observer error in 3-D landmark placement on the human subadult pelvis" 88th Annual Meeting for the American Association of Physical Anthropologists, Cleveland, OH, USA

Stull, KE, Corron, LK, Hulse, CN, and Yang, Y. "Exploring the relationship between dental development, population variation, and environment" 88th Annual Meeting for the American Association of Physical Anthropologists, Cleveland, OH, USA

Corron, LK, Stull, KE, and Yang, Y. "Variation of Skeletal Growth and Development Patterns in Populations with Diverse Socio-Economic Backgrounds" 88th Annual Meeting for the American Association of Physical Anthropologists, Cleveland, OH, USA

6. InterForensics 2019 Conference, Sao Paulo, Brazil (2019)

Corron L., Stull K., Hulse C., Wolfe C. "Does variation in skeletal and dental growth and development affect subadult age estimation?" InterForensics 2019 Conference, May 21-24 2019, Sao Paulo, Brazil

Stull K. and Corron L. "KidStats" InterForensics 2019 Conference, May 21-24 2019, Sao Paulo, Brazil

7. Annual Meeting for the American Academy of Forensic Sciences, Anaheim, CA, USA (2020)

"KidStats: improving the subadult biological profile" workshop at the 72nd Annual Meeting of the American Academy of Forensic Sciences, February 17th-23 2020, Anaheim, CA, USA. Chairs: Dr Kyra Stull, Heather Garvin, Alexandra Klales

Stull KE, Price MH, Corron LK. "Subadult Age Estimation Using a Mixed Cumulative Probit and its Application in KidStats". 72nd Annual Meeting of the American Academy of Forensic Sciences, February 17th-23, Anaheim, CA

8. Annual Meeting for the American Academy of Forensic Sciences ONLINE EVENT (2021)

Corron LK, Wolfe CA, Stull KE. "Vertebral Neural Canal (VNC) Dimensions in Contemporary Subadult Samples: Indicators of Stress, Population Variation, or Both?" 73rd Annual Meeting of the American Academy of Forensic Sciences, February 2021.

9. American Association for Anatomy Annual Meeting ONLINE EVENT (2021)

Yim, A-D, Corron LK, Stull KE. "A new tool for medical image processing in biological anthropology and anatomy research". American Association for Anatomy Annual Meeting at Experimental Biology, April 27-30 2021.

10. Annual Meeting for the Forensic Anthropology Society of Europe (FASE), ONLINE EVENT (2021)

Stull KE, Corron LK. "The Subadult Virtual Anthropology Database (SVAD) – An Online Repository of Subadult Reference Material". FASE Advanced Course in Forensic Anthropology "Virtual Anthropology" November 11th, 2021.

Personal Presentations

Stull, KE, Corron, LK, and Price, M (2018) "Exploring variation in growth and development and its impact on subadult age estimation: variables, samples, and models." New Methods In Skeletal Age Estimation For Diverse Populations, August 6th, 2018: Stanford University.

Stull, KE (March 2019) Dental Anthropology Society, American Association of Physical Anthropology; March 2019.

2019 KidStats, Texas State University, San Marcos, April 2019

Subadult Age Estimation, Mercyhurst University Summer Workshop, June 2019

Corron LK (2020) "Using medical images and 3D data in forensic anthropology research on subadult age estimation". TECHNOLOGY EVOLUTIONS AND RESEARCH REVOLUTIONS: A Workshop on the Theory, Practice, and Ethics of New Data, Algorithms, and Tools for Questions of Science and Society. Stanford University, Palo Alto. January 14, 2020

Acknowledgments

All our collaborators who accepted to share their resources and build the SVAD: Dr Maria Gabriela Haye Biazevic, Universidade de São Paulo (FOUSP), Dr Kathia Chaumoitre, AP-HM, Aix-Marseille University, Dr Kerri Colman, University of Amsterdam, Dr Eugenia Cunha, University of Coimbra, Dr Henrique Manuel Vale Fernandes Costa Rodrigues, University of Coimbra, Dr Rick van Rijn, University of Amsterdam, Dr Timisay Monsalve-Vargas, Universidade de Antioquia, Dr An-Di Yim, Truman State University. Additional thanks go to all the colleagues and graduate and undergraduate students who helped collect the derivatives: Drs Michala Stock, Elaine Chu, Laura Cirillo, Stephanie Cole, Cortney Hulse, and Christopher Wolfe.

References

AlQahtani, S.J., M.P. Hector, and H.M. Liversidge. 2010. "Brief Communication: The London Atlas of Human Tooth Development and Eruption." *Am J Phys Anthropol* 142: 481–90.

AlQahtani, SJ., MP. Hector, and HM. Liversidge. 2010. "Brief Communication: The London Atlas of Human Tooth Development and Eruption" 142: 461–90.

Baughan, B, M Brault-Dubuc, A Demirjian, and G. Gagnon. 1980. "Sexual Dimorphism in Body Composition Changes during the Pubertal Period: As Shown by French-Canadian Children." *American Journal of Physical Anthropology* 52: 85–94.

Bogin, B. 1999. *Patterns of Human Growth*. Second Edition. Cambridge University Press.

Bogin, B., C. Scheffler, and M. Hermanussen. 2017. "Global Effects of Income and Income Inequality on Adult Height and Sexual Dimorphism in Height." *American Journal of Human Biology* 29 (2): 1–11.

Brough, AL, J Bennett, and B. Morgan. 2013. "Anthropological Measurement of the Juvenile Clavicle Using Multi-Detector Computed Tomography — Affirming Reliability." *Journal of Forensic Sciences* 58 (4): 946–51.

Cameron, N., and B. Bogin. 2012. *Human Growth and Development*. 2nd edition. Academic Press.

Cole, TJ, EK Rousham, NL Hawley, N Cameron, SA Norris, and JM Pettifor. 2015. "Ethnic and Sex Differences in Skeletal Maturation among the Birth to Twenty Cohort in South Africa." *Archives of Disease in Childhood* 100 (2): 138–43.

Colman, KL., HH. de Boer, JG. Dobbe, NPTJ. Liberton, KE. Stull, M. van Eijnaten, GJ. Streekstra, RJ. Oostra, RR. Van Rijn, and AE. van der Merwe. 2019. "Virtual Forensic Anthropology: The Accuracy of Osteometric Analysis of 3D Bone Models Derived from Clinical Computed Tomography (CT) Scans." *Forensic Science International* 304: 1–10.

Corron, L. K., M. K. Stock, S. J. Cole, C. N. Hulse, H. M. Garvin, A. R. Klales, and K. E. Stull. 2021. "Standardizing Ordinal Subadult Age Indicators: Testing for Observer Agreement and Consistency across Modalities." *Forensic Science International* 320 (March): 110687. https://doi.org/10.1016/j.forsciint.2021.110687.

Corron, L., F. Marchal, S. Condemi, K. Chaumoître, and P. Adalian. 2017. "Evaluating the Consistency, Repeatability and Reproducibility of Osteometric Data on Dry Bone Surfaces, Scanned Dry Bone Surfaces and Scanned Bone Surfaces from Living

Individuals." *Bulletins et Mémoires de La Société d'Anthropologie de Paris* 29 (1–2): 33–53.
Corron, Louise K, Christopher A Wolfe, and Kyra E Stull. 2023. "A Multifaceted Exploration of Ontogenetic Variation in Vertebral Neural Canal Size across Contemporary Populations." *American Journal of Biological Anthropology* 180 (2): 328–51. https://doi.org/10.1002/ajpa.24675.
Corron, Louise, Christopher Wolfe, and Kyra Stull. 2022. "SVAD Vertebral Neural Canal Dimensions [Data Set]." Zenodo. https://doi.org/10.5281/zenodo.6342097.
Coussens, A, T Anson, RM Norris, and M. Henneberg. 2002. "Sexual Dimorphism in the Robusticity of Long Bones of Infants and Young Children." *Anthropological Review* 65: 3–16.
Demirjian, A., and G.-Y. Levesque. 1980. "Sexual Differences in Dental Development and Prediction of Emergence." *Journal of Dental Research* 59 (7): 1110–22. https://doi.org/10.1177/00220345800590070301.
Fazekas, IG, and F. Kosa. 1978. *Forensic Fetal Osteology*. Budapest.
Graham, EA. 2005. "Economic, Racial, and Cultural Influences on the Growth and Maturation of Children." *Pediatrics in Review* 26 (8): 290–94.
Grissom, LE., MP. Harty, GW. Guo, and HH. Kecskemethy. 2018. "Maturation of Pelvic Ossification Centers on Computed Tomography in Normal Children." *Pediatric Radiology* 48 (13): 1902–14.
Hunt, EE, and I. Gleiser. 1955. "The Estimation of Age and Sex of Preadolescent Children from Bones and Teeth." *American Journal of Physical Anthropology* 13: 479–87.
Jahan, S. 2015. "Human Development Report 2015 - Work for Human Development." United Nations Development Programme.
Kline, RB. 2016. *Methodology in the Social Sciences.Principles and Practice of Structural Equation Modeling*. Vol. (4th ed.). Guilford Press.
Kruger, GC, EN L'Abbe, and KE. Stull. 2017. "Sex Estimation from the Long Bones of Modern South Africans." *International Journal of Legal Medicine* 131 (1): 275–85.
Liversidge, HM. 2015. "Controversies in Age Estimation from Developing Teeth." *Annals of Human Biology* 42 (4): 397–406.
Moore-Jansen, PH, and RL. Jantz. 1994. *Data Collection Procedures for Forensic Skeletal Material*. Forensic Anthropology Series.; Report of Investigations (University of Tennessee, Knoxville. Department of Anthropology) 48. Knoxville, Tennesse: Forensic Anthropology Center, Dept. of Anthropology, University of Tennessee.
Moorrees, CFA, EA Fanning, and EE. Hunt. 1963. "Age Variation of Formation Stages for Ten Permanent Teeth." *Journal of Dental Research* 42 (6): 1490–1502.
Petrie, D., and KK. Tang. 2008. "A Rethink on Measuring Health Inequalities Using the Gini Coefficient." Discussion Papers Series 381. School of Economics, University of Queensland, Australia.
Sahn, David E., and Stephen D. Younger. 2005. "Improvements in Children's Health: Does Inequality Matter?" *The Journal of Economic Inequality* 3 (2): 125–43. https://doi.org/10.1007/s10888-005-4494-9.
Spake, L., J. Meyers, S. Blau, H. Cardoso, and N. Lottering. 2020. "A Simple and Software-Independent Protocol for the Measurement of Post-Cranial Bones in Anthropological Contexts Using Thin Slab Maximum Intensity Projection." *Forensic Imaging* 20.

Steckel, Richard H. 2012. “Social and Economic Effects on Growth.” In *Human Growth and Development*, 225–44. Elsevier. https://doi.org/10.1016/B978-0-12-383882-7.00009-X.
Stull, KE, EY Chu, LK Corron, and MH Price. n.d. “Mixed Cumulative Probit: A Novel Algorithm Inspired by Methodological and Practical Shortcomings in Age Estimation.”
Stull, KE, LE Cirillo, SJ Cole, and CN Hulse. 2020. “Subadult Sex Estimation and KidStats.” In *Sex Estimation of the Human Skeleton History, Methods, and Emerging Techniques*, edited by Klales AR, 219–42. Academic Press.
Stull, KE., and LK. Corron. 2021a. “Subadult Virtual Anthropology Database (SVAD) Data Collection Protocol: Amira,” August. https://doi.org/10.5281/ZENODO.5348411.
———. 2021b. “Subadult Virtual Anthropology Database (SVAD) Data Collection Protocol: Epiphyseal Fusion, Diaphyseal Dimensions, Dental Development Stages, Vertebral Neural Canal Dimensions,” August. https://doi.org/10.5281/ZENODO.5348392.
Stull, KE, and LK Corron. 2022. “The Subadult Virtual Anthropology Database (SVAD): An Accessible Repository of Contemporary Subadult Reference Data.” https://doi.org/10.3390/forensicsci2010003.
Stull, KE, EN L’Abbé, and SD. Ousley. 2017. “Subadult Sex Estimation from Diaphyseal Dimensions.” *American Journal of Physical Anthropology* 163 (1): 64–74.
Stull, KE., EN. L’Abbé, and S. Steiner. 2013. “Measuring Distortion of Skeletal Elements in Lodox Statscan-Generated Images.” *Clinical Anatomy* 26: 780–86.
Stull, KE., ML. Tise, Z. Ali, and DR. Fowler. 2014. “Accuracy and Reliability of Measurements Obtained from Computed Tomography 3D Volume Rendered Images.” *Forensic Science International* 238: 133–40.
Stull, Kyra E., Elaine Y. Chu, Louise K. Corron, and Michael H. Price. 2022. “Subadult Age Estimation Using the Mixed Cumulative Probit and a Contemporary United States Population.” *Forensic Sciences* 2 (4): 741–79. https://doi.org/10.3390/forensicsci2040055.
Stull, Kyra E., and Louise K. Corron. 2022. “The Subadult Virtual Anthropology Database (SVAD): An Accessible Repository of Contemporary Subadult Reference Data.” *Forensic Sciences* 2 (1): 20–36. https://doi.org/10.3390/forensicsci2010003.
Stull, Kyra E., Ericka N. L’Abbé, and Stephen D. Ousley. 2014. “Using Multivariate Adaptive Regression Splines to Estimate Subadult Age from Diaphyseal Dimensions.” *Am. J. Phys. Anthropol.* 154 (3): 376–86. https://doi.org/10.1002/ajpa.22522.
Tanner, JM. 1981. “Growth and Maturation during Adolescence.” *Nutrition Reviews* 39 (2): 43–55.
Verleye, G., M-J. Ireton, JC. Carrillo, and RC. Hauspie. 2004. “Latent Variables and Structural Equation Models.” In *Methods in Human Growth Research*, edited by RC. Hauspie, N. Cameron, and L. Molinari, 287–305. Cambridge: Cambridge University Press.