A comparison of subadult skeletal and dental development based on living and deceased samples

Kyra E. Stull[1,2,*], Christopher A. Wolfe[1], Louise K. Corron[1], Kelly Heim3, Cortney N. Hulse[1] and Marin A. Pilloud[1]

[1]Department of Anthropology, University of Nevada, Reno, Nevada
[2]Department of Anatomy, Faculty of Health Sciences, University of Pretoria, Pretoria, South Africa
[3]3 St Petersburg, Florida

*Correspondence to: Kyra E. Stull, Department of Anthropology, University of Nevada, Reno, 1664 N Virginia St MS 0096, Reno, NV 89557.
Email: kstull@unr.edu

Funding information

Directorate for Biological Sciences, Grant/Award Number: BCS-13 1551913; National Institute of Justice, Grant/Award Numbers: 2015-DN-BX-K409, 2017-DN-BX-K409

ABSTRACT

Objectives: A fundamental assumption in biological anthropology is that living individuals will present with different growth than non-survivors of the same population. The aim is to address the question of whether growth and development data of non-survivors are reflective of the biological consequences of selective mortality and/or stress.

Materials and Methods: The study compares dental development and skeletal growth collected from radiographic images of contemporary samples of living and deceased individuals from the United States (birth to 20 years) and South Africa (birth to 12 years). Further evaluation of deceased individuals is used to explore differential patterns among manners of death (MOD).

Results: Results do not show any significant differences in skeletal growth or dental development between living and deceased individuals. However, in the South African deceased sample the youngest individuals exhibited substantially smaller diaphyseal lengths than the living sample, but by 2 years of age the differences were negligible. In the US sample, neither significant nor substantial differences were found in dental development or diaphyseal length according to MOD and age (>2 years of age), though some long bones in individuals <2 years of age did show significant differences. No significant differences were noted in diaphyseal length according to MOD and age in the SA sample.

Discussion: The current findings refute the idea that contemporary deceased and living individuals would present with differential growth and development patterns through all of ontogeny as well as the assumptions linking short stature, poor environments, and MOD.

KEYWORDS: ontogeny, population variation, reference samples, sampling bias, selective mortality

A central question to the study of subadult skeletal material is whether non-survivors are representative of the surviving sample of children who make-up later age cohorts (i.e., the survivors), or whether there is differential selection of those who have died and entered the skeletal record (i.e., selective mortality) (DeWitte & Stojanowski, 2015; Hoppa, 2000; J. Wood et al., 1992). The non-random entry into the death record, or selective mortality, relates to the fact that the deceased do not represent all individuals within a particular age cohort. Because the deceased have failed to survive, their pattern of growth and development may not reflect that of the true population and they may exhibit delayed linear growth and development (i.e., bones that are small for age). Saunders and Hoppa (1993) concluded nearly three decades ago that there were minimal differences in skeletal linear growth between survivors and non-survivors in both archeological and contemporary samples, and that the methodological biases (e.g., cross-sectional data, measurement error), could contribute more to these observable differences than actual biological or health factors. Yet other researchers have claimed that contemporary non-survivors displayed differential growth patterns in comparison to survivors using various types of anthropometric data (e.g., Chen et al., 1980; McDonald et al., 2013; Olofin et al., 2013; Smedman et al., 1987; Spake & Cardoso, 2019a). Subsequently, there is a continued presumption that deceased individuals from the skeletal record or medicolegal context (i.e., non-survivors) have differential growth patterns compared to the living population (i.e., survivors). Implications of this have led researchers to postulate that methods and reference samples must be appropriate to the “target sample” on which they are applied. In other words, to evaluate components of the biological profile (i.e., age, sex, stature) on a deceased individual, the method chosen must be based on a deceased reference sample. Recent suggestions have been made (e.g., Spake & Cardoso, 2018) to move beyond the binary non-survivors and survivors and into context-specific samples, such that “forensically-relevant” samples should be exclusively used in a forensic context. The nature of such a distinction, the lack of consensus on the subject, and the implications in terms of subadult methods beg the question: are there statistically and practically significant growth and development differences between deceased and living subadults and between context-specific samples?

The current research proposes to evaluate the growth and development of subadults based on samples of living and deceased individuals within the same, or similar, populations. This project will explore sample differences in four main ways: 1) a comparison of diaphyseal lengths between deceased and living subadults, 2) a comparison of dental development between deceased and living subadults, 3) a comparison of diaphyseal lengths and dental development of deceased subadults based on manner of death (MOD), and 4) a comparison of diaphyseal lengths in relation to the forensic setting (i.e., differences between individuals analyzed by a forensic pathologist and a forensic anthropologist). The overarching goal is to inform anthropologists on appropriate reference samples for forensic or bioarchaeological subadult methods and to identify any skeletal and dental differences between living and deceased individuals or between individuals with different contexts of death. Furthermore, results can be extrapolated to how subadults are analyzed and their biological data interpreted within an archeological context. The current study has significant implications for bioarchaeologists and forensic anthropologists and provides information on the variability present in subadult assemblages of both deceased and living individuals.

1 THE INHERENT COMPLEXITIES OF REFERENCE SAMPLES

It is critical to evaluate a reference sample when developing and validating methods in biological anthropology. Use of skeletal reference material that is diverse enough to capture human variation increases the accuracy rate for all parameters of the biological profile (i.e., age, sex, ancestry, stature). However, the demographics of donated skeletal collections, which are regularly used to develop methods in forensic anthropology and bioarchaeology, do not always mimic forensic or bioarchaeological samples or even the true population (Komar & Grivas, 2008). Obtaining representative samples is a constant challenge, especially when accounting for variation because of secular change, transnational migration, and large-scale global demographic shifts as well as the broad objectives of forensic anthropology, which results in shifting target populations. Unfortunately, the accumulation of contemporary skeletal reference material is slower than these other phenomena and is still not sufficient to address uncertainties in both forensic and bioarchaeological work. Therefore, sampling biases inherent to modern skeletal reference material may mean they do not accurately represent the “target populations” of forensic anthropologists or bioarchaeologists (Albanese, 2003; Kimmerle, 2014; Kimmerle et al., 2008; Spradley, 2014). While it is acknowledged that donated skeletal collections do not necessarily reflect the demographics of the target samples, it is typically only the age distribution of the reference sample that is discussed, particularly in regard to age mimicry and how a biased reference sample may lead to inaccurate results in regression-based age estimation methods (Bocquet-Appel & Masset, 1982; Boldsen et al., 2002; Getz, 2020; Konigsberg & Frankenberg, 1994).

Most methods aimed to estimate biological parameters for the deceased are in fact based on deceased samples, but not all deceased samples are comprised of the same type of individuals. For example, the methods developed by Fazekas and Kósa (1978) were based on human fetuses between the third and tenth lunar month who were subject to autopsy, therefore, representing individuals who died in utero and were failure to thrive non-survivors. The age estimation methods developed by McKern and Stewart (1957) were also based on non-survivors: a military population of males who died during military efforts in Korea. However, although these individuals are non-survivors, they died from a violent death and could very well have been in perfect health. A major exception to deceased individuals comprising reference samples involve age estimation in subadults and forensic age estimation of the living. Charts and atlases of dental development and long bone lengths have been based on radiographs of living individuals (e.g., (Demirjian et al., 1973; Heim, 2018; Maresh, 1970; Moorrees et al., 1963) and applied or adapted for use in forensic age estimation of both living and deceased individuals. Specific to subadult skeletal material, the main concern is always the fact that a child in a skeletal collection represents a non-survivor, and therefore, may not be representative of the healthy, living subadult population (Johnston, 1962).

The primary reason living subadults are more commonly used to develop methods is because there is a scarcity of subadults compared to adults in skeletal collections. Furthermore, the age distribution is usually positively skewed, which impacts the validity of the developed methods. Anthropologists can increase sample sizes of modern individuals by working with conventional radiography images (x-ray), computed tomography scans (CT), and magnetic resonance images generated in either medical examiner's offices or hospital settings (e.g., Berry & Edgar, 2020; Garvin & Stock, 2016; Stock et al., 2016). By including these imaging modalities, researchers can also obtain a more uniform age distribution, reducing the risk of bias and effects of age mimicry (Bocquet-Appel & Masset, 1982; Boldsen et al., 2002). Of note here though, depending on the type of institution, images composing these samples can be taken from either living or deceased individuals, thereby representing both the living (and sometimes healthy) population and the deceased (non-survivors, and unhealthy) population. The appropriateness of using either type of subadult reference source (i.e., deceased or living) or combining them to increase sample size for the creation of anthropological methods remains uncertain. These examples demonstrate that methods have been based on both survivors and/or non-survivors, largely because of sample size, availability, and context of study. The appropriateness of each in application is not only dependent on the target population, but also if the goal of the research is inferential or predictive.

While modern medical records of deceased children could very well represent individuals of any socioeconomic status (SES), it has been argued that a deceased child is more likely to have low SES (Spake & Cardoso, 2018), as mortality rates are closely and inversely linked to SES level (Bairagi & Chowdhury, 1994; Braudt et al., 2019; Kim & Saada, 2013; Spencer, 2004). The higher mortality rates and biological effects of low SES can be linked to numerous factors, such as parental education level, level of access to medical care, family income-to-needs ratio, and living quality (housing), among others. Within the deceased sample, it has been suggested that children with different MODs may represent different sectors of the population and will, therefore, also exhibit different developmental trajectories (Spake & Cardoso, 2018). Yet research comparing stature between individuals with a MOD of homicide and accident in three different countries found no consistent differences. In conclusion, Spake and Cardoso (2018) suggest that the composition of reference samples be thoroughly scrutinized and to specifically incorporate child abuse and neglect cases and data from medical examiners may be preferable to data from hospital sources for methods applied in forensic anthropology.

Biological differences are assumed to exist between individuals from medicolegal and hospital databases. However, the composition of clinical samples is as equally complex as the composition of deceased samples. Samples of living children based on data obtained from hospitals and clinical settings could yield images of children with both acute and chronic conditions, individuals from all SES levels, and individuals that may enter the medicolegal record with any type of MOD. Chronic conditions could result in growth and development being delayed and/or stunted, thereby making these children potentially more similar to lower SES children than “normal” or higher SES children. In addition, in countries where healthcare is accessible, independent of income, the clinical sample could include individuals from all SES levels, illustrating a range of variation in growth and development, including stature.

The relationship among subadults, MOD, and forensic anthropological casework in particular has ramifications for the process behind building reference samples and for thinking about the target population(s). According to the Center for Disease Control's (CDC) National Vital Statistics 2016 report, in the United States, 4,114 children between the ages of birth and 14 years died that year from accidental causes, 443 from suicide, and 901 from documented homicide. Relative to the population, this is approximately 30 accidental deaths for every 100,000 recorded deaths and 7 childhood homicides for every 100,000 deaths (Xu et al., 2018). Throughout all of ontogeny, accidental deaths are more common than homicides (Cunningham et al., 2018). While forensic pathologists who work within medical examiner's offices are statistically more likely to examine subadult individuals who died of accidental deaths than homicides (He et al., 2015), according to the Forensic Databank (FDB), forensic anthropologists are more likely to examine subadult individuals who died of homicide. While the work of both the forensic pathologist and the forensic anthropologist is within a medico-legal context, the two are tasked to work on very different cases. Typically, the forensic anthropologist is asked to examine remains with some manner of decomposition, while the forensic pathologist works on every case that comes through their office, regardless of the body's condition. While circumstantial, it could be argued that a homicide is more likely to result in human remains being recovered some time after death, related to some manner of concealment of the body and resulting in more advanced decomposition, whereas remains from accidental deaths are likely to be recovered immediately and less likely to have advanced levels of decomposition. Further, forensic anthropologists might be asked to assist in cases without decomposition when trauma is complex or there is question about the cause of skeletal trauma *(*i.e., accident vs. homicide). The slight nuances can explain the discordant levels of MODs by different medicolegal practitioners and demonstrates the medicolegal context is not homogenous across disciplines. While the above describes the target population for the medical examiner's office, the role of the forensic anthropologist also extends into age estimation of the living and victim identification in natural and mass disasters. Therefore, the target population, MOD, and appropriate reference sample are always shifting.

2 LEVELS OF VARIATION IN SKELETAL AND DENTAL INDICATORS OF GROWTH AND DEVELOPMENT

Diaphyseal dimensions, commonly used as subadult age indicators, are also used as non-specific indicators of stress and population health for both subadults and adults (e.g., public health, economics, anthropology, human biology) (e.g., Cardoso et al., 2017; Conceição & Cardoso, 2011; Dhavale et al., 2017; Mummert et al., 2011; Pinhasi et al., 2014; Sohn, 2016). Subsequently, children from a lower SES are usually also assumed to be shorter for their age and lighter in weight when compared to children of the same age and population, but from a higher SES, even though it is recognized that growth-related measurements remain non-specific indicators of health and are sensitive to many factors. However, it is difficult to determine if, at one particular point, a child exhibiting smaller skeletal dimensions experienced prolonged insults or if that individual is genetically or epigenetically predisposed to smaller stature. Just as importantly, we are unable to measure an individual's capacity for catch-up growth, and therefore, it is unknown if that child could later experience such growth to attain their genetic potential (Cameron, 2012; Kuzawa & Bragg, 2012; Martorell & Zongrone, 2012). Stunted growth, which is strongly correlated to long bone lengths, is also linked to mortality. In a deceased sample, a higher mortality bias is linked to a taller average stature whereas a lower mortality bias is linked to a shorter average stature because in the latter context, only the frailest, or the most vulnerable (who could also be the shortest), die (Wood et al., 1992). As stated earlier, research by Saunders and Hoppa (1993) demonstrated that subadult non-survivors were consistently shorter than their surviving peers, but the maximum difference in femoral length was no greater than 3 mm. This difference is comparable to measurement error and smaller than the normal variation in long bone lengths per age (e.g., Maresh, 1970). Dental development has been found to be less susceptible to environmental insults than diaphyseal dimensions (H. Cardoso, 2007a; S. Saunders, 2008; Smith, 1991). In addition, mortality bias has not been recognized in dental development or eruption (Holman et al., 2004; Spake & Cardoso, 2019a). Therefore, a deceased child would theoretically be more likely to present with normal dental development, but exhibit a shorter stature for age compared to a living child of the same age.

3 MATERIALS AND METHODS

The samples used to explore differences between deceased and living subadults and within deceased subadult samples originate from the United States (US) for dental development and South Africa (SA) and the US for skeletal growth. The majority of comparisons are country specific, such that there is only one instance where the SA sample is compared to the US sample. Thus, the population is controlled in an effort to remove a source of variation. Both living samples from the US and SA come from a clinical/hospital setting and the deceased samples come from a medico-legal setting, which is either a medical examiner (ME) or forensic anthropology (FA) database. Therefore, regardless of the continent, the general context between the comparative samples is similar, allowing for valid interpretations from the comparisons.

TABLE 1. Age distribution for the deceased samples in the United States, which includes individuals from UNM, OCMEM, Patricia, and the FDB. Patricia and UNM/OCMEM samples were used for dental comparisons, while the FDB and UNM/OCMEM samples were used for long bone comparisons

Age (years)	**Sample**	**Number of individuals**		**Age (years)**	**Sample**	**Number of individuals**	
		Diaphyseal	***Dental***			***Diaphyseal***	***Dental***
0	UNM	205	-	8	UNM	14	9
	OCMEM	50	-		OCMEM	2	-
	Patricia	-	-		Patricia	-	7
	FDB	7	-		FDB	-	-
1	UNM	77	-	9	UNM	21	9
	OCMEM	23	-		OCMEM	4	-
	Patricia	-	-		Patricia	-	5
	FDB	3	-		FDB	-	-
2	UNM	42	-	10	UNM	10	6
	OCMEM	17	-		OCMEM	-	-
	Patricia	-	-		Patricia	-	-
	FDB	3	-		FDB	2	-
3	UNM	30	-	11	UNM	22	18
	OCMEM	9	-		OCMEM	1	-
	Patricia	-	-		Patricia	-	-
	FDB	-	-		FDB	4	-
4	UNM	33	-	12	UNM	24	14
	OCMEM	7	-		OCMEM	2	-
	Patricia	-	-		Patricia	-	-
	FDB	-	-		FDB	3	-
5	UNM	21	14	13	UNM	-	24
	OCMEM	8	-		OCMEM	-	-
	Patricia	-	23		Patricia	-	-
	FDB	-	-		FDB	-	-
6	UNM	12	6	14	UNM	-	27
	OCMEM	3	-		OCMEM	-	-
	Patricia	-	14		Patricia	-	-
	FDB	-	-		FDB	-	-
7	UNM	19	16	15	UNM	-	32
	OCMEM	1	-		OCMEM	-	-
	Patricia	-	15		Patricia	-	-
	FDB	-	-		FDB	-	-
				Totals		679	239

As discussed earlier, an inherent problem with subadult samples is the disparity in age distributions (Tables 1 and 2; Figures 1 and 2). In general, the mortality pattern seen in modern children is bimodal (Figure 1); there is a high mortality rate for infants, followed by a reduction in mortality in early to mid-childhood, and then an increase in mortality in the teenage years (Ousley, 2013). In comparison, a living sample coming from the hospital/clinical setting shows a more uniform age distribution with rather equal representation across the age groups except for infancy, where smaller sample sizes are typically found (see Figure 2). The discrepancy between each of these distributions impacts certain statistical tests that follow, and while some interpretations are limited because of the lack of statistical power, the general pattern is informative to the specific questions.

TABLE 2. Age distribution for the living sample in the United States, which includes individuals from UTSA

Age (years)	Number of individuals	Age (years)	Number of individuals
5	12	11	61
6	24	12	43
7	72	13	32
8	94	14	31
9	66	15	17
10	60		

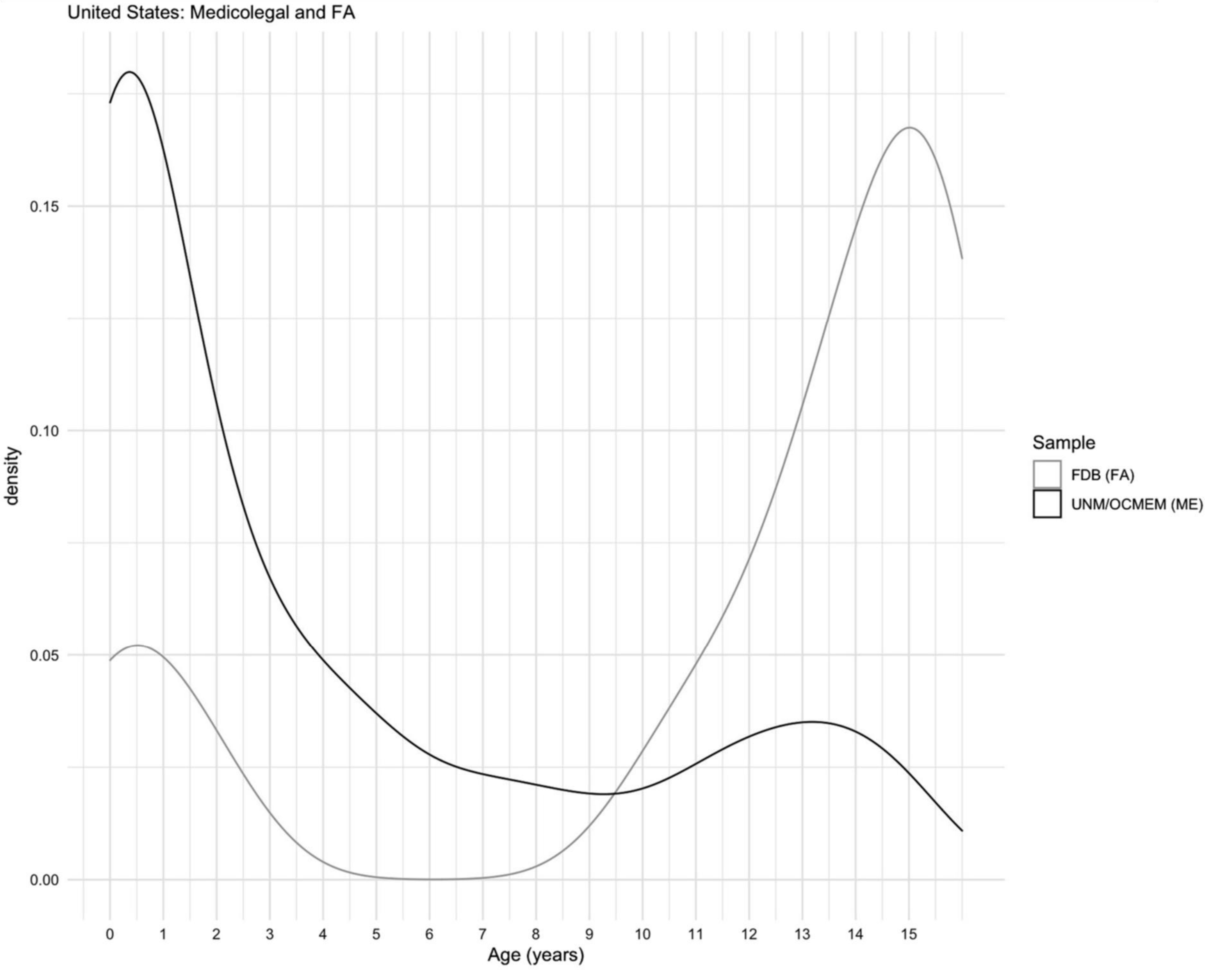

FIGURE 1. Age distribution of the deceased individuals in the United States, which is comprised of a sample from UNM and OCMEM (ME) in black and a sample from the FDB (FA) in gray

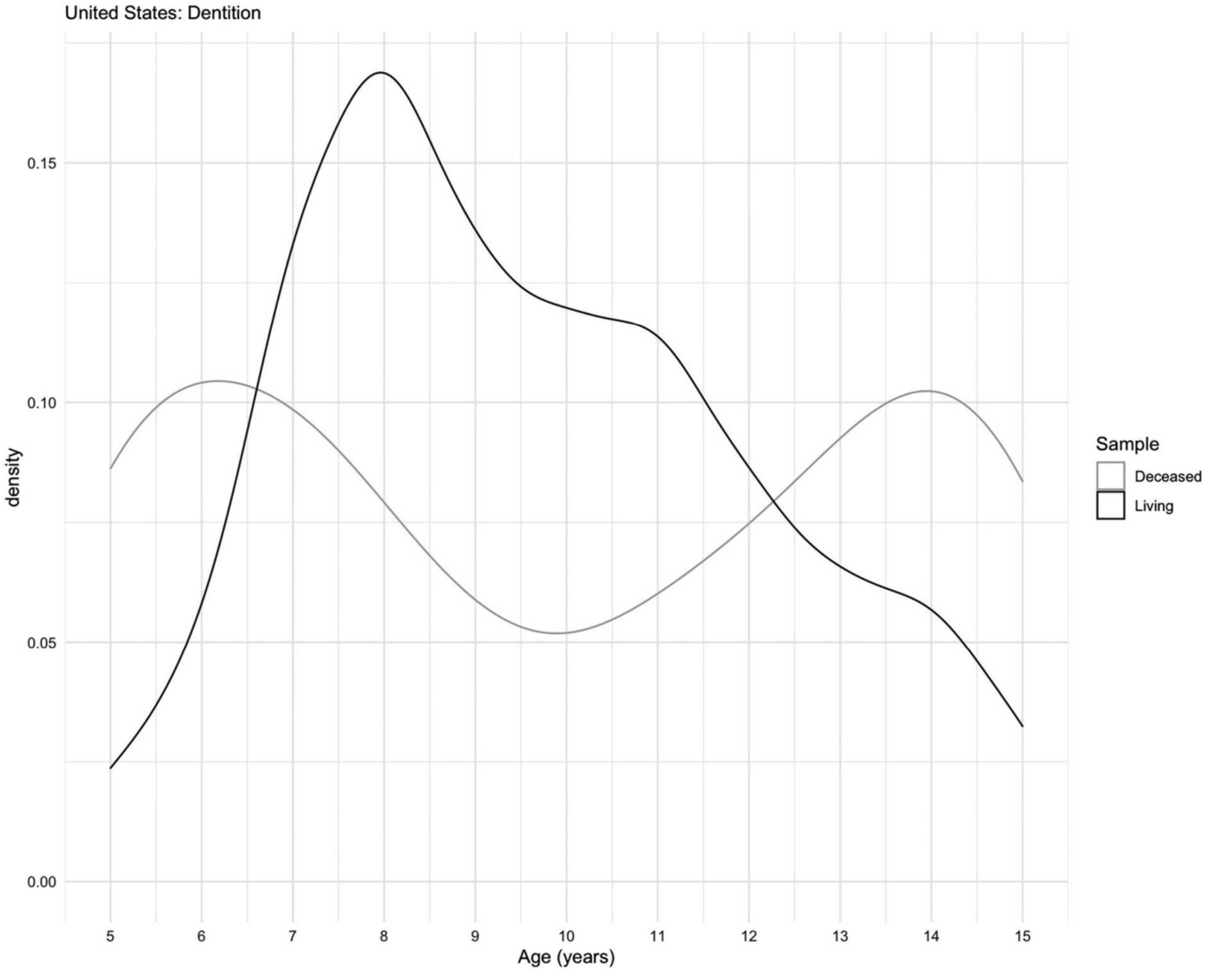

FIGURE 2. Histogram of the samples (Deceased = UNM/Patricia; Living = UTSA) used for the dental comparisons in the US

4 COMPARATIVE DATA

4.1 US: Deceased, medico-legal

The deceased ME sample from the US is comprised of individuals from: 1) University of New Mexico (UNM) Health Sciences Center, Office of the Medical Investigator, 2) the Office of the Chief Medical Examiner, State of Maryland (OCMEM), and from 3) Patricia (Pediatric Radiology Interactive Atlas (http://math.mercyhurst.edu/~sousley/databases/radiographic_database) (see Table 1 and Figure 1).

The sample from UNM (*n* = 613) comprises the bulk of the deceased sample and contains both skeletal and dental information (NIJ Award 2015-DN-BX-K409 and NSF BCS-1551913). Diaphyseal lengths of individuals ranging in age from birth to 12 years were collected from bone surfaces reconstructed from CT scans with the Amira™ software. Information regarding the CT and segmentation processes can be found in Stock et al. (2020) and Stock et al. (2016). All dental development scores were assessed by scrolling through the individual CT slices and not through a reconstructed isosurface or volume rendering. The age range for individuals in the dental analyses is truncated between 5 and 15 years to correspond with the study design of the living sample (see below).

The OCMEM data (n = 127) comprise a smaller portion of the deceased data and only contribute to the long bone length analyses. The OCMEM CT images were reconstructed in a contiguous fashion, similar to the UNM images, but using the GE Advanced CT Workstation (AW-2) (Version: aws-2.0-5.5).

Patricia is a database of digital radiographs with demographic data compiled from numerous ME's and coroner's offices across the US that are freely accessible for research (Ousley, 2013). The database was queried for all individuals with high-quality radiographs that included the dentition (n = 57). The Patricia data are included to supplement the UNM dental sample for the ages of 5 to 9 years, as some ages had especially small sample sizes. There were no differences in dental developmental trajectories between UNM and Patricia samples, and therefore, the samples were pooled for the analyses (n = 239) (Table 1, Figure 2).

4.2 US: Deceased, forensic anthropological

Deceased subadult data were also compiled from the FDB, which is a large repository of skeletal and demographic data submitted by forensic anthropologists (Jantz & Moore-Jansen, 1988). The individuals either stem from forensic anthropological casework or skeletal collections in the US (Jantz & Moore-Jansen, 1988), and therefore, the individuals may or may not have come from a ME setting. All data were collected by the submitting anthropologist and maintained at the University of Tennessee, Knoxville (Jantz & Moore-Jansen, 1988; NIJ 85-IJ-CX-0021).

The subadult sample from the FDB (n = 103) was included to evaluate the skeletal differences between a “typical” forensic anthropologist sample compared to a forensic pathologist sample (UNM/OCMEM). As discussed earlier, the MOD for cases evaluated by a forensic anthropologist may differ from those typically seen by a forensic pathologist. Diaphyseal measurements were used from all available individuals (see Table 1 and Figure 1).

4.3 US: Living

Dental development data were collected from individuals who received orthodontic treatment at the University of Texas Health Science Center at the San Antonio School of Dentistry (UTHSCSA) between 2005 and 2017. The sample (n = 512) is comprised of individuals between the ages of 5 and 15 years (Table 2 and Figure 2). Data were collected from routinely generated orthopantomograms; the images were queried and deidentified by UTSA faculty prior to data collection (Heim, 2018).

4.4 SA: Deceased and living

Individuals from the South African Long Bone Database (SALB) were included to compare long bone growth between deceased and living individuals (Stull et al., 2014,b,c). Diaphyseal measurements were collected from Lodox Statscan (Lodox Systems [Ltd], Sandton, SA) radiographic images of children between birth and 12 years of age from Cape Town, SA (DVS 2.8.8.2, Lodox Systems, SA). The living sample (n = 423) comes from the Red Cross War Memorial Children's Hospital (RXH), while the deceased sample (n = 200) comes from the Salt River Forensic Pathology Laboratory (SR). The age distributions can be seen in Table 3 and Figure 3.

TABLE 3. Age distribution of the South African deceased (SR) and living (RXH) samples

Age (years)	Sample size	
	Deceased (SR)	**Living (RXH)**
0	45	4
1	23	16
2	25	37
3	22	39
4	16	32
5	9	26
6	3	50
7	5	53
8	3	49
9	4	23
10	3	24
11	3	34
12	4	14

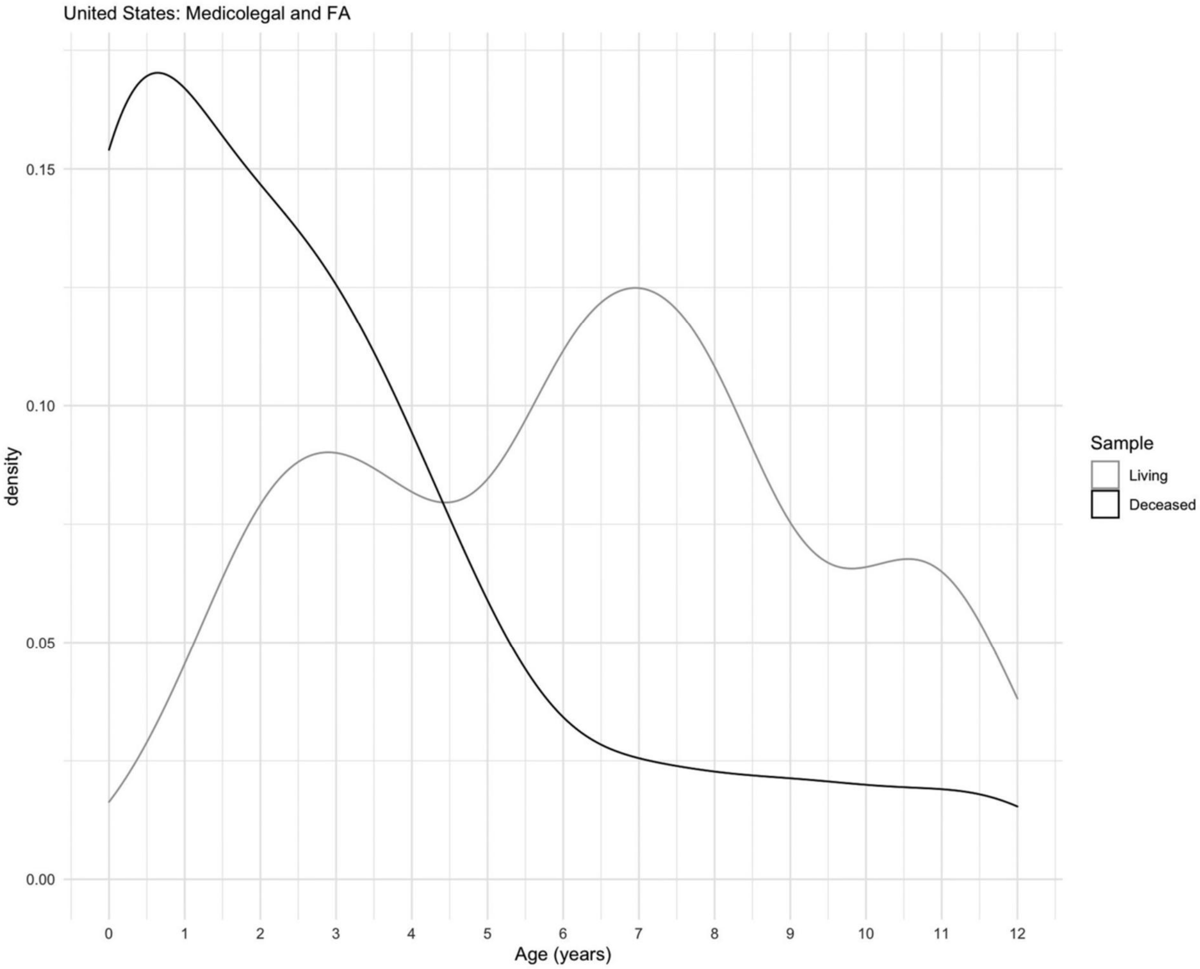

FIGURE 3. Age distributions of the two South African samples (Deceased = SR; Living = RXH) used for the long bone analyses

5 DATA COLLECTION

Maximum diaphyseal length was collected on the femur, tibia, fibula, humerus, radius, and ulna. The definitions presented by Stull, L'Abbé, and Ousley (2014) were followed, which were modified from Fazekas and Kósa (1978). Left-sided elements were measured and if unavailable, their right-sided counterparts were measured. All US long bone data were measured directly on reconstructed bone surfaces, while the SA long bone data were measured using DVS, the Lodox imaging software (see Stull, L'abbé, & Ousley, 2014, 2014). Previous publications document the high repeatability and accuracy of Lodox measurements (Stull et al., 2013; Stull, L'abbé, & Ousley, 2014, 2014) as well as high precision, accuracy, and repeatability in measurements collected on CT scans or virtual reconstructions of skeletal elements (Colman et al., 2017; Colman et al., 2019; Corron et al., 2017; Stull, Tise, et al., 2014).

Dental development was scored on mandibular and maxillary permanent teeth of individuals from UNM and Patricia using the staging system developed by AlQahtani et al. (2010), which is a modified version of the Moorrees et al. (1963) permanent dentition developmental stages. The UTHSCSA dentition was initially scored in the Moorrees et al. (1963) 14-stage system and was later recoded to match the AlQahtani et al. (2010) modification. Specifically, AlQahtani et al. (2010) do not include the “initial cleft formation” code (Cl_i, Stage 8 in Moorrees et al., 1963), so therefore, this stage was collapsed into the “initial root formation” stage (R_i, Stage 7 in AlQahtani et al., 2010 and Moorrees et al., 1963). Teeth from the left dental arcade were used for each comparison unless unavailable or unobservable; in which case the right antimere was substituted. Because dental data from two deceased samples were pooled for the analyses, inter-observer agreement was quantified using a weighted Cohen's kappa. The results suggest high rater agreement with kappa values ranging between 0.74 and 1.0. The developmental scores with the greatest agreement rates (kappa = 1.0) were the mandibular central incisor and the maxillary first molar; the tooth with the lowest agreement rate (kappa = 0.74) was the maxillary lateral incisor. The majority (75%) of kappa values ranged between 0.82 and 0.94 indicating the majority of teeth had excellent agreement (Landis & Koch, 1977).

6 STATISTICAL ANALYSES

All statistical analyses were completed in the R programming environment (R Core Team, 2019). Analyses were separated according to the major objectives of the research.

1. *Comparing skeletal data of deceased and living individuals*: *South Africa*

The first set of skeletal analyses compares long bone lengths of the deceased (SR) and living (RXH) individuals from the SALB. Diaphyseal lengths and age, by context, were visualized to elucidate overall trends between the samples, which were accentuated by loess lines. Diaphyseal lengths for the deceased and living subsets per age were analyzed using a Kolmogorov–Smirnov (KS) test, a nonparametric test that evaluates the cumulative distributions of samples, and therefore, tests for differences in median, variances, and distributions (Lehmann, 2006). The p-values had a Holm's adjustment to account for the number of statistical tests, which directly increases the likelihood of a Type I error (Holm, 1979). Note, the p.adjust function provides an interface for the multiplicity adjustments, but maintains the 0.05 level for ease of interpretation (R Core Team, 2019).

- 2 *Comparing dental data of deceased and living individuals*: *United States*

Variation in dental development between the deceased and living samples from the US was visualized using boxplots of dental developmental stage for age and separated by context. Dental development stages per age were analyzed using a KS test. Similar to the first objective, p-values had a Holm's adjustment to account for the number of statistical tests (Holm, 1979).

- 3 *Comparing skeletal and dental data of deceased individuals between the ME and FA contexts and by MOD*

The last set of analyses specifically address skeletal and dental developmental variation in deceased samples to answer the third and fourth research objectives. An analysis of covariance (ANCOVA) was used to elucidate the effect of MOD (i.e., accident, homicide, or natural) and age (predictor variables) on each long bone length or dental development of each tooth (response variable). The interaction between age and MOD on long bone length and dental development stage was specifically of interest rather than the main effects in each model. Long bones from the US and SA samples and dental development stages for each tooth from the US sample were used for this analysis. Bivariate scatterplots of long bone lengths and age were created, and loess lines were used to reveal the trends per MOD using the SA and US long bone length data. Boxplots of dental development stage and age were used to expose any differential patterns between MOD using the US deceased sample. Visualizations were also made to reveal the similarities/dissimilarities between ME and FA cases in the US. Because of the unequal variances and sizes of the samples or unavailable individuals for comparison, the interpretations were based solely on visual comparisons and loess lines for this final objective.

7 RESULTS

1. *Comparing skeletal data of deceased and living individuals*: *South Africa*

The adjusted p-values from the KS tests indicated that no age groups presented with statistically significant differences ($p > 0.05$) between the deceased and living SA samples for all long bones. While there were no statistically significant differences, there were some larger than expected mean differences between the living and deceased samples in the first 2 years of life (i.e., <2 years) (Table 4). The mean long bone lengths for the RXH (living) sample were consistently larger than the SR (deceased) sample in the first 2 years of life, though the differences decreased as age increased. In the bivariate scatterplots, the loess lines representative of each sample track very closely to one another with minimal differences beyond 2 years of age (Figure 4). The positive skew in the density plots on the x-axes emphasizes the higher number of deceased individuals in the youngest ages. In contrast, the living sample has the smallest number of individuals in the youngest ages and a more uniform distribution between approximately 2 years to 12 years.

- 2 *Comparing dental data of deceased and living individuals*: *United States*

TABLE 4. Mean bone lengths per age and sample (RXH = living sample, SR = deceased sample) from birth to 2 years for the South African sample

Age (years)	**Collection**	**Mean lengths (mm)**					
		Femur	**Tibia**	**Fibula**	**Humerus**	**Radius**	**Ulna**
0	RXH	115.34	101.16	98.46	95.40	73.00	84.22
	SR	103.56	87.02	82.61	79.77	65.73	74.49
1	RXH	159.54	131.96	129.90	122.9	98.59	110.80
	SR	146.12	120.58	116.14	111.36	87.52	99.68
2	RXH	184.20	152.31	150.48	135.08	104.87	117.06
	SR	183.93	150.38	147.45	133.49	103.52	115.74

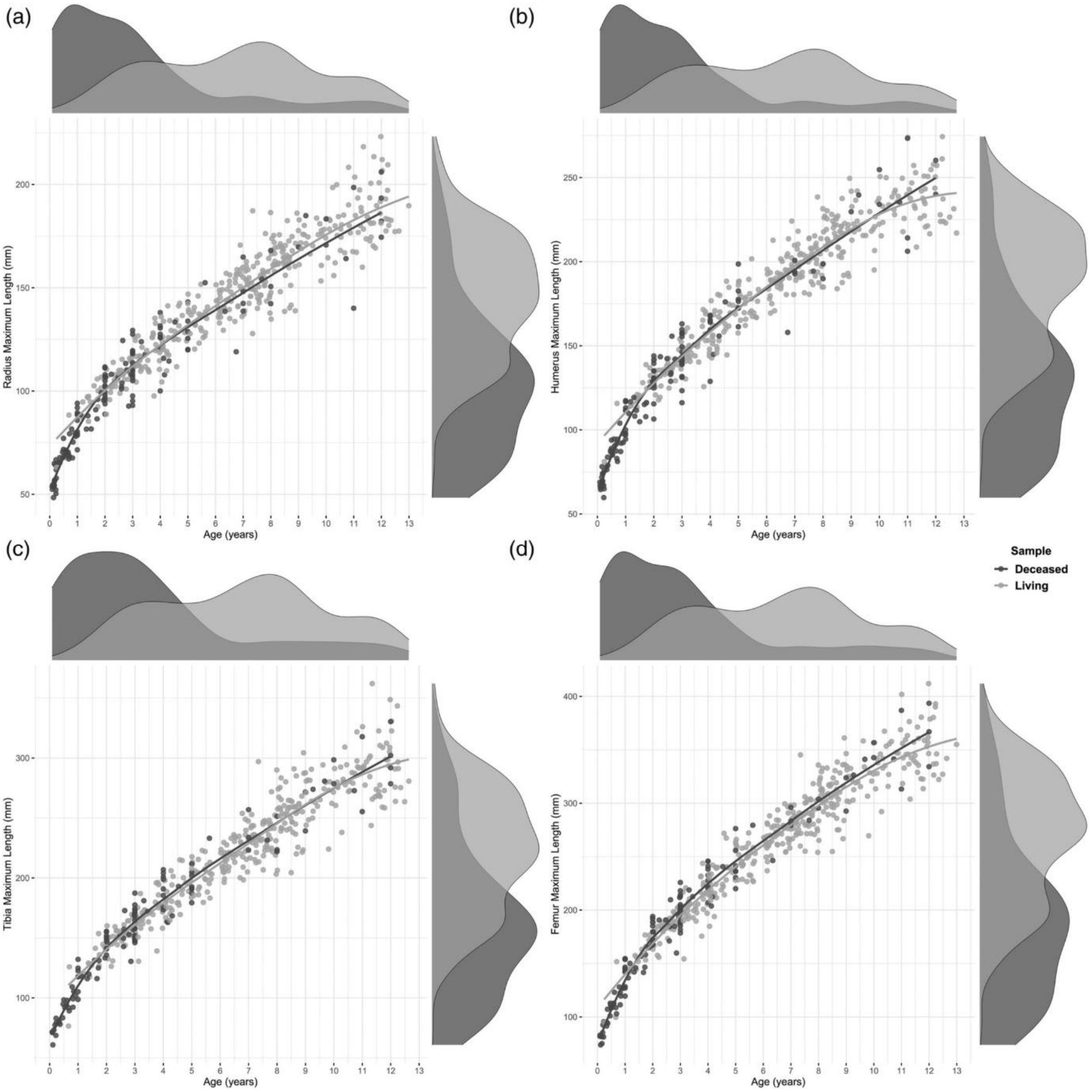

FIGURE 4. - Scatterplots of radius (a), humerus (b), tibia (c), and femur (d) diaphyseal lengths according to age in the RXH (living) and SR (deceased) samples from South Africa. The density plots above the scatterplot shows the age distributions of each sample and the density plots to the right of the scatterplot shows the distribution of the measurement for each sample

The KS tests were conducted on single year age cohorts of the US samples for all maxillary and mandibular dentition and the vast majority of comparisons were not significant (Figure 5, see Supplemental Material for all teeth). However, there were three instances of statistical significance with adjusted p-values <0.05, namely the development scores of the lower central incisor among 6 and 7-year-olds, and those of the mandibular lateral incisor for 11-year-olds. In these three comparisons, there were less than 4 individuals total for the comparisons; therefore, the differences are likely spurious because of limited sample size and ultimately questionable when such few individuals are being compared.

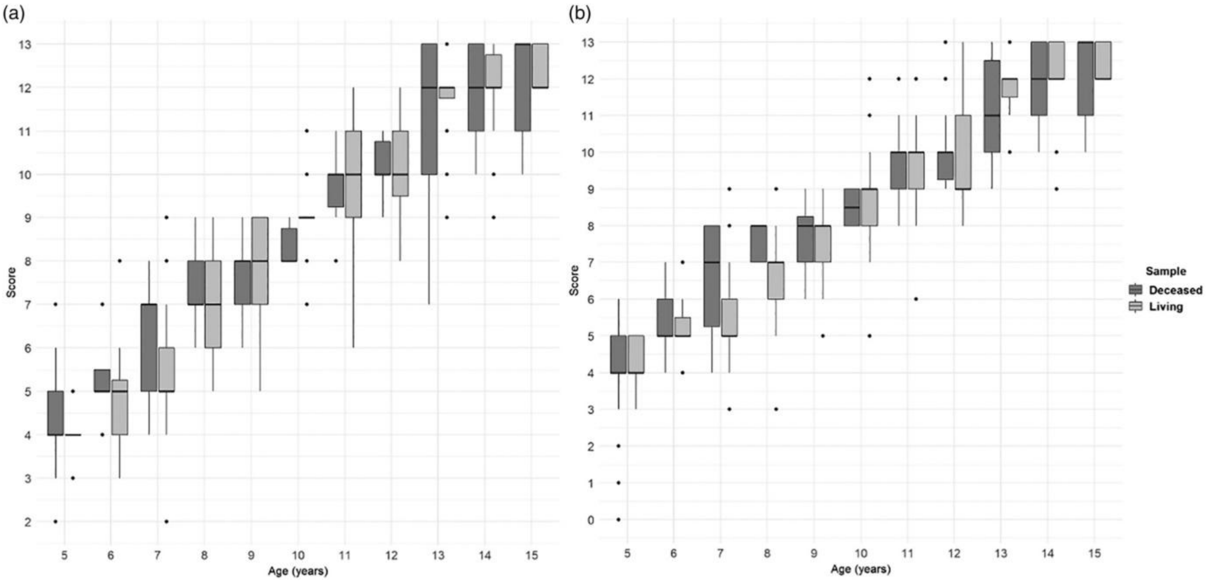

FIGURE 5. - Boxplots of the dental development scores of mandibular (a) and maxillary (b) second permanent molars in the deceased sample (dark gray) and the living sample (light gray). The second mandibular molar is visualized because it yielded the widest range of developmental scores for the current sample

3(a) Comparing skeletal data of deceased individuals by MOD: *South Africa and United States*

Further exploration of the long bones of the SA and US deceased samples by MOD was accomplished via visual comparisons and statistical analyses. In the US sample, when all ages were included, there were no significant differences in the interaction between MOD and age for the femur, tibia, fibula, and radius; there were significant interactions between MOD and age for the humerus and ulna. No significant differences in the interaction between MOD and age for any long bone existed when individuals 2 years and older were included from the US sample. For the SA long bone data, there were no significant differences in the interaction between MOD and age. Because the assumptions were violated in some of the subsets, the visual comparisons were used to corroborate the results. Visually, there appears to be little to no difference in diaphyseal lengths among MODs (Figures 6 and 7). Some sample sizes per age and long bone were small in the SA sample, and as age increases the number of individuals decreases in the US sample. Both situations suggest conservative interpretations of the loess lines, as loess lines require dense datasets, and therefore, are quite susceptible to outliers. Yet, even with the small sample sizes per age and bone groups, there are not obvious size differences per MOD.

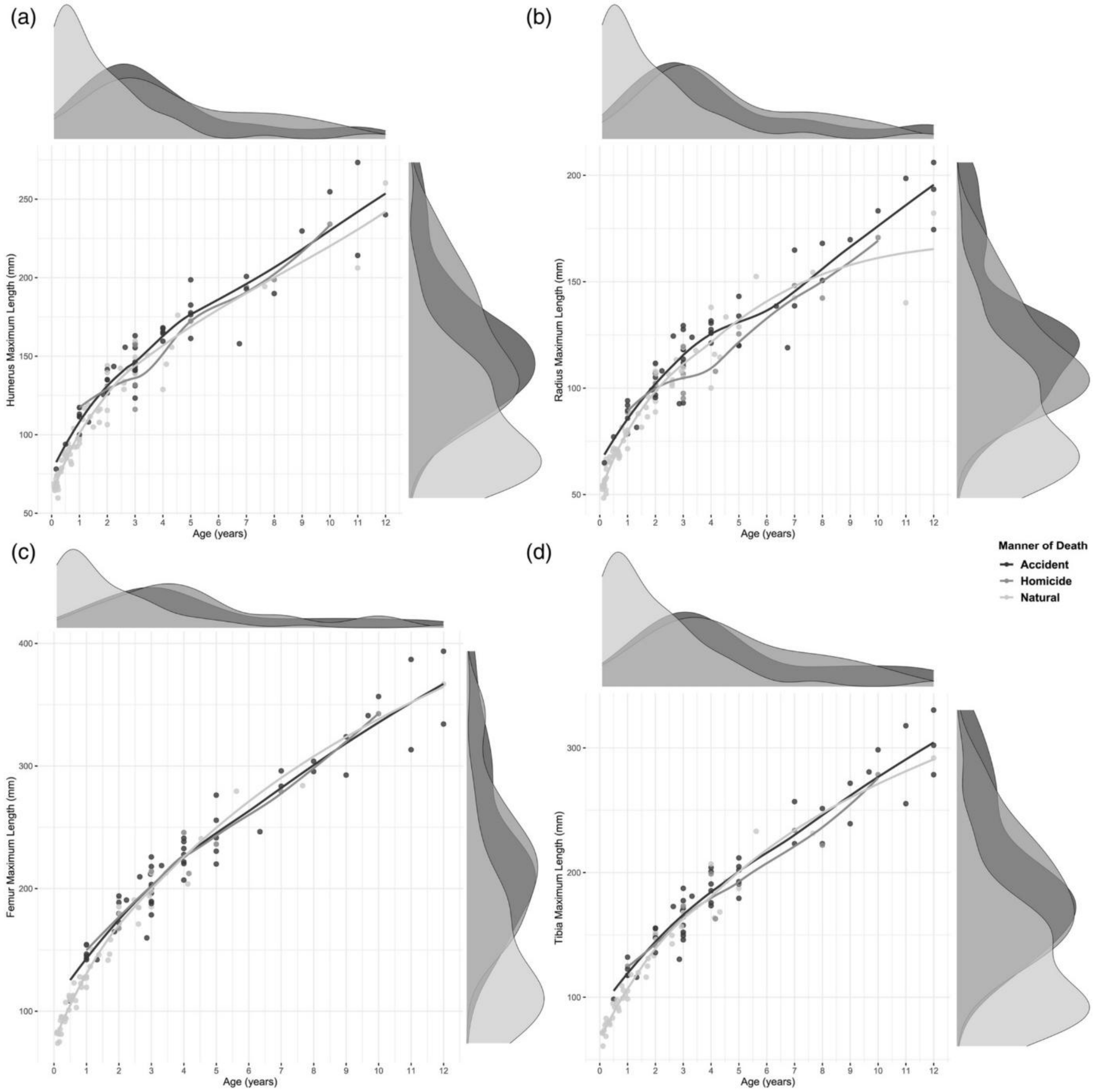

FIGURE 6. Scatterplots of humerus (a), radius (b), femur (c), and tibia (d) diaphyseal lengths according to age and manner of death (MOD) in the South African deceased sample (SR): accidental (dark gray), homicide (medium gray), and natural (light gray)

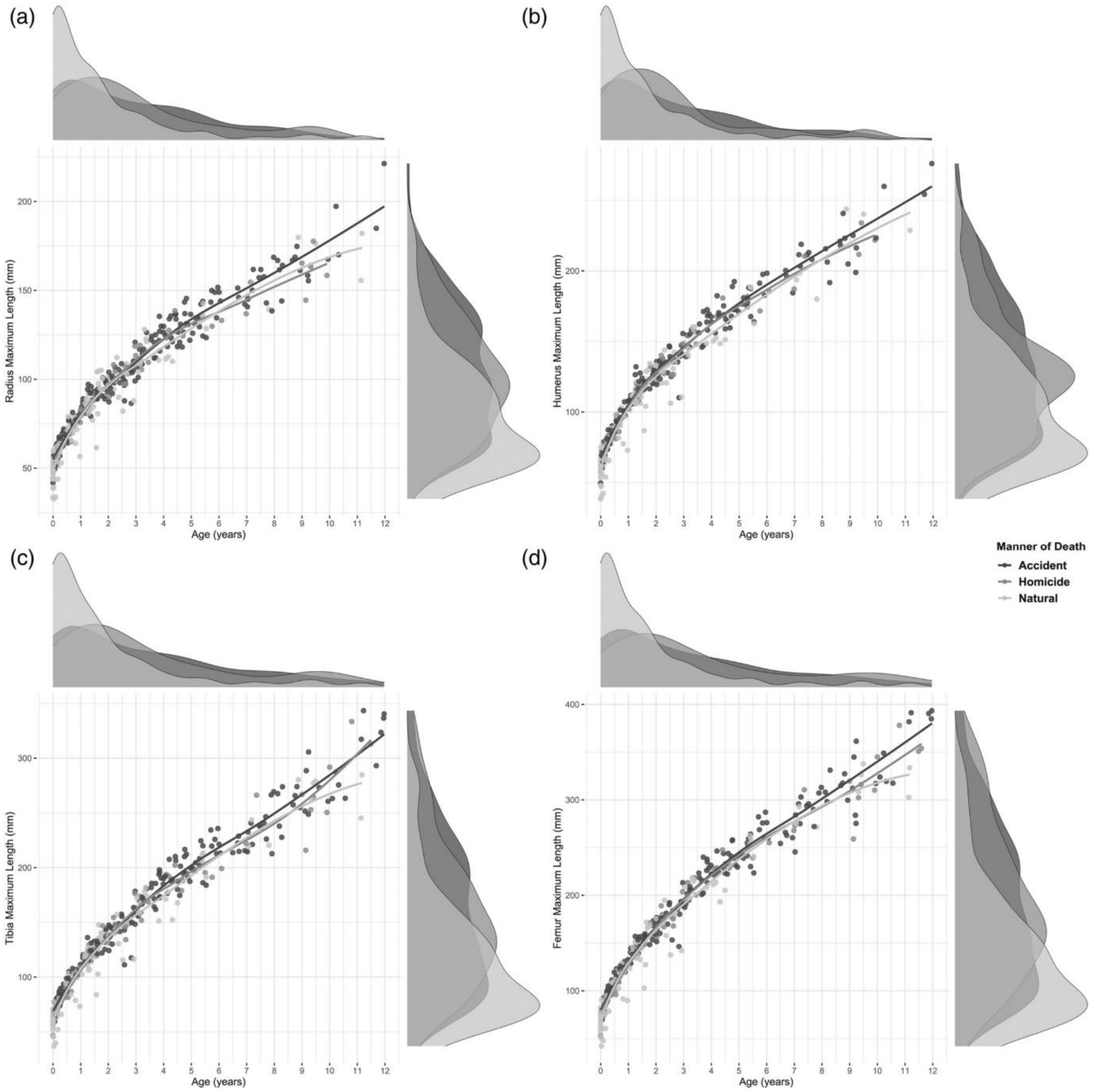

FIGURE 7. – Scatterplots of humerus (a), radius (b), femur (c), and tibia (d) diaphyseal lengths according to age and manner of death (MOD) in the US deceased sample: accidental (dark gray), homicide (medium gray), and natural (light gray)

There are markedly more individuals who are less than 1 year old with a natural MOD (*n* = 41) compared to an accident (*n* = 2) or a homicide (*n* = 0) MOD for the SA sample. However, the average size of individuals less than a year old with a natural MOD is not substantially different from the individuals with an accident or homicide MOD. For individuals in the US that are younger than 1 year, there is a slightly different pattern, but similarly the smallest sample size is homicides (*n* = 17), while there are large numbers of individuals with natural (*n* = 89) and accident (*n* = 70) MODs. In all long bones and both samples, there is a larger number of natural deaths in the youngest age groups, and as age increases, the likelihood of a natural death decreases. Figure 8 displays the remarkable reduction in natural deaths as age increases for the SA sample. The density plots on both the x and y axes in Figures 6 and 7 also reveal the greater similarity in distributions for accident and homicide MODs compared to natural MODs for all long bones.

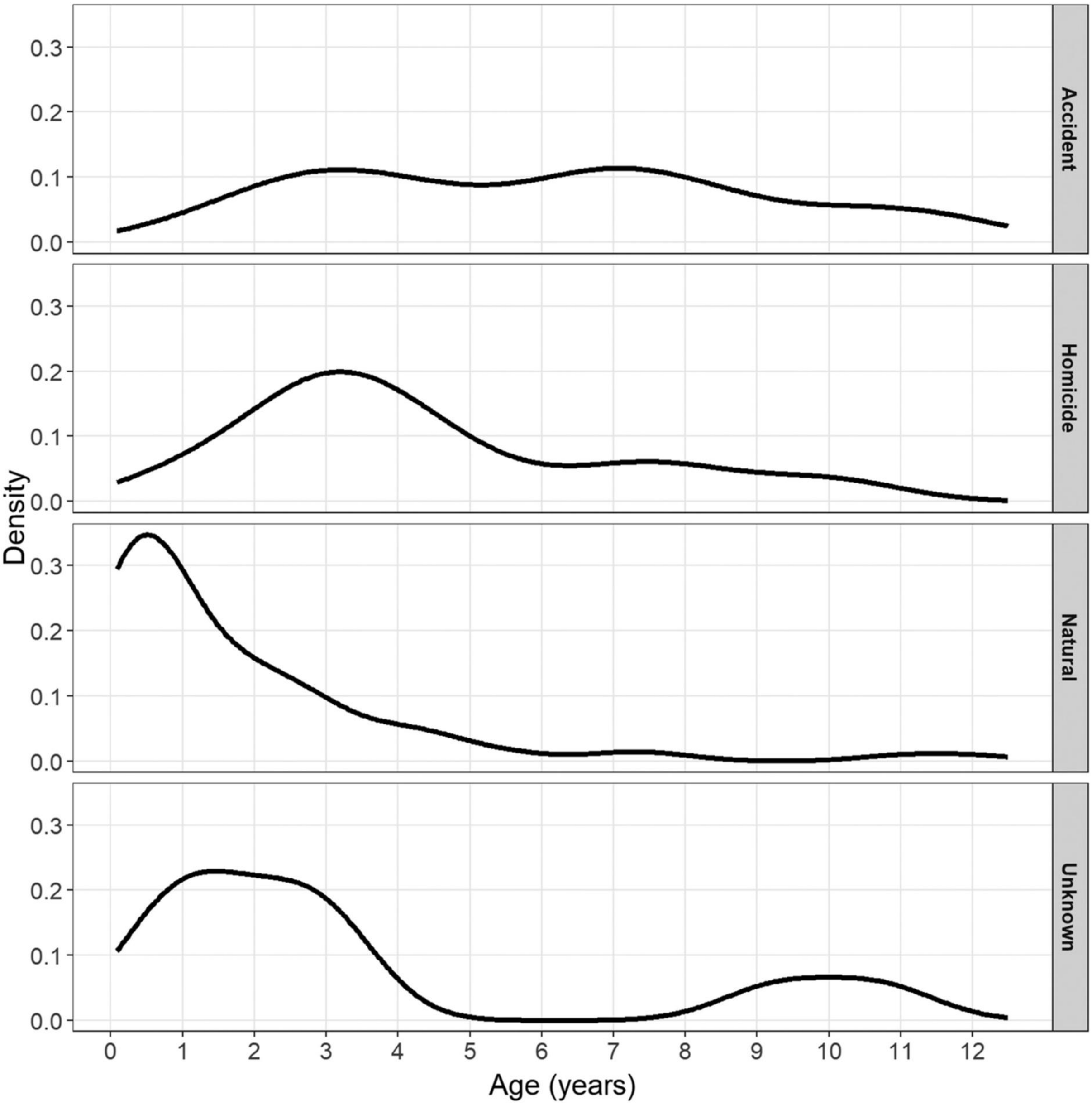

FIGURE 8. Density plots visualizing the frequency of individuals per each MOD and age in the South African deceased sample

3(b) Comparing dental data of deceased individuals by MOD: *United States*

The ANCOVA exploring the interaction effect of MOD and age on the dental development per tooth in the US data revealed no significant differences. Similar to the long bone data, visualizations were used to corroborate the results because the assumptions were violated in some of subsets. Visualizations of dental development according to MOD do not reveal any substantial trends in dental development (Figure 9, see Supplemental Material for all teeth). Boxplots display that, even when there is a different median age per dental development stage, the age ranges for the same stage is comparable and/or overlapping. Admittedly, there are some categories with small sample sizes or even without any data available.

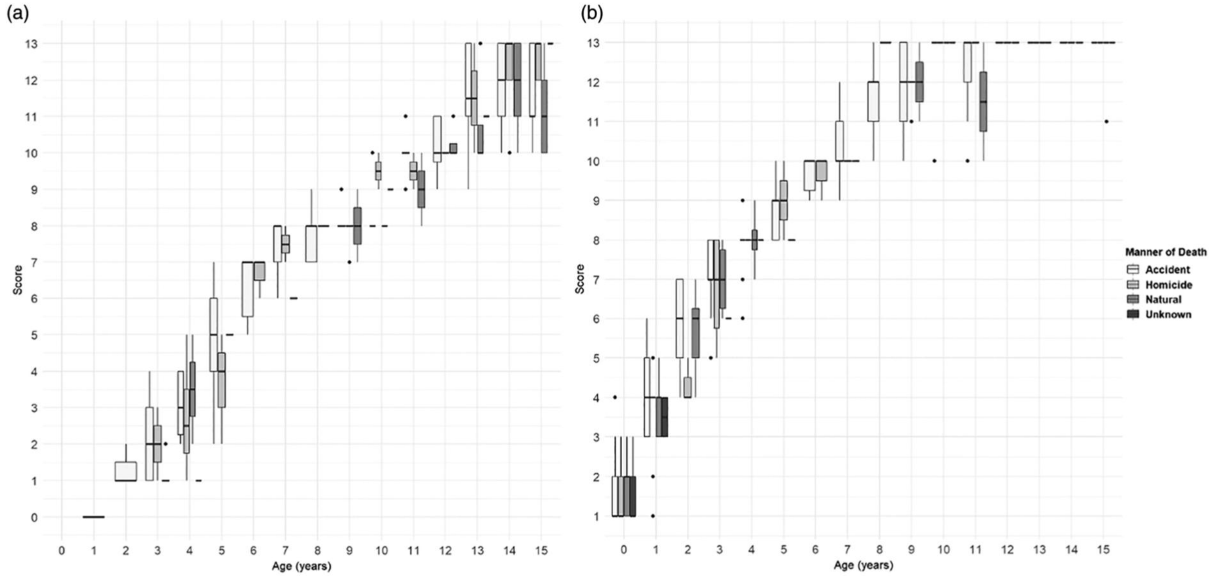

FIGURE 9. Boxplots of dental development stages of the permanent second mandibular (a) and maxillary (b) molars in the US sample separated by MOD

3(c) Comparing skeletal data of deceased individuals between the ME and FA contexts: *United States*

Visualizations of long bone lengths and age reveal similar ranges in size between the FA (FDB sample) and ME contexts (US deceased sample) (Figure 10). The FA sample had few individuals less than 3 years of age, no individuals between 3 and 9 years of age, and the largest sample size was for individuals above 10 years of age. In contrast, the ME (UNM/OCMEM) sample had the smallest sample for the oldest individuals and the largest sample for the younger individuals (Table 1). In both samples, the fewest numbers of individuals are in mid-childhood.

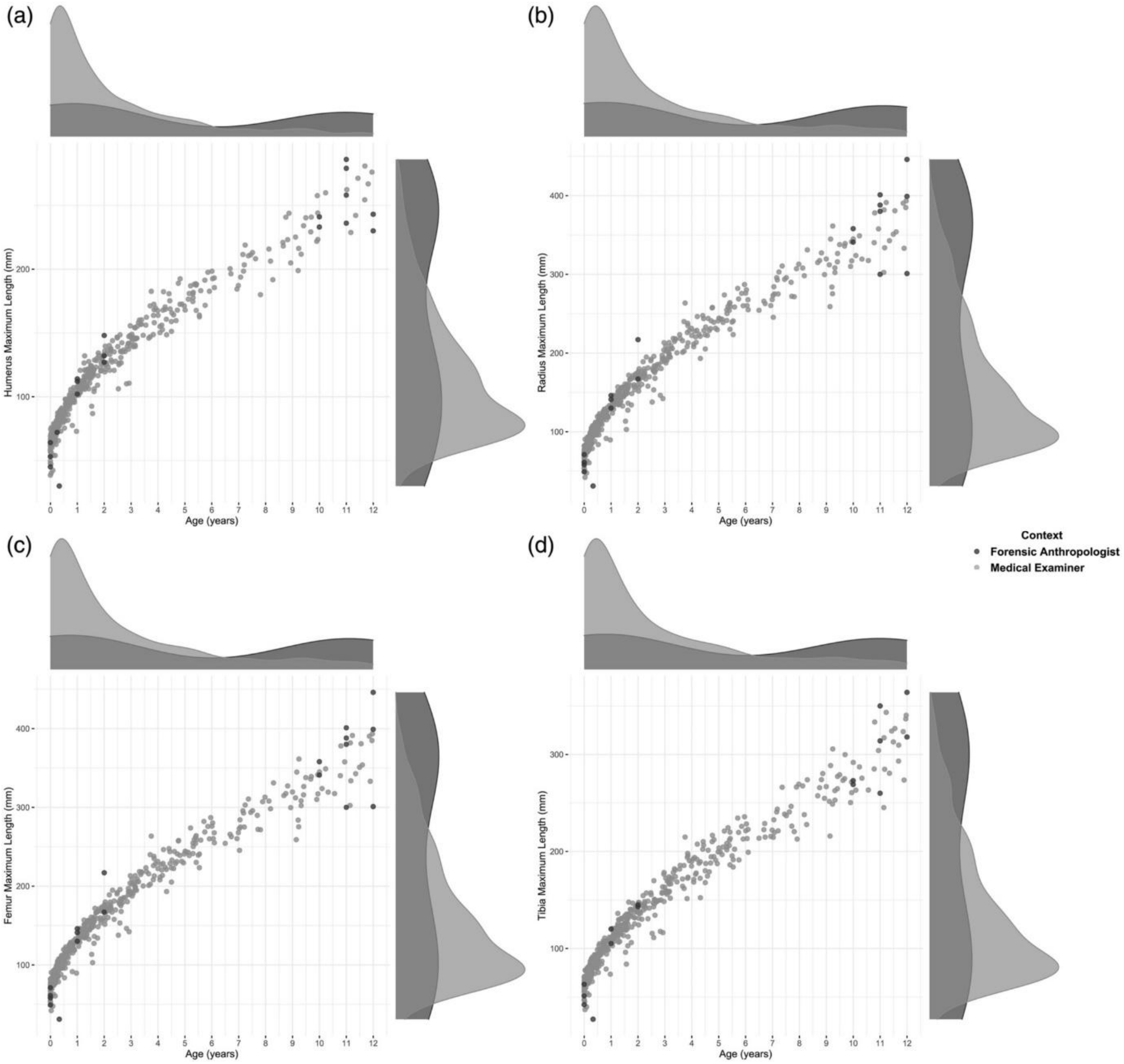

FIGURE 10. Scatterplots of humerus (a), radius (b), femur (c), and tibia (d) diaphyseal lengths according to age and the origin of the US samples: FA (FDB) (dark gray) and ME's office (UNM/OCMEM) (light gray). The density plots above show the age distributions for each origin (same color code as the points/lines)

Because it has been suggested by some authors that forensic anthropological cases of subadults are more likely to be a victim of child abuse/neglect or malnourished based on published case studies (e.g., Spake & Cardoso, 2018), we delved deeper into the available metadata and identified nine victims of child abuse in our sample of 740 subadults from the UNM/OCMEM database, which equates to 1.2% of the sample. Zero children had malnourishment reported as a cause of death or a contributing cause of death in the entire database. Visualizations are provided that highlight the long bone lengths of the victims of child abuse to demonstrate their relationship to the remaining sample. For all long bones, the victims of child abuse were consistently within the variation of the entire sample (Figure 11).

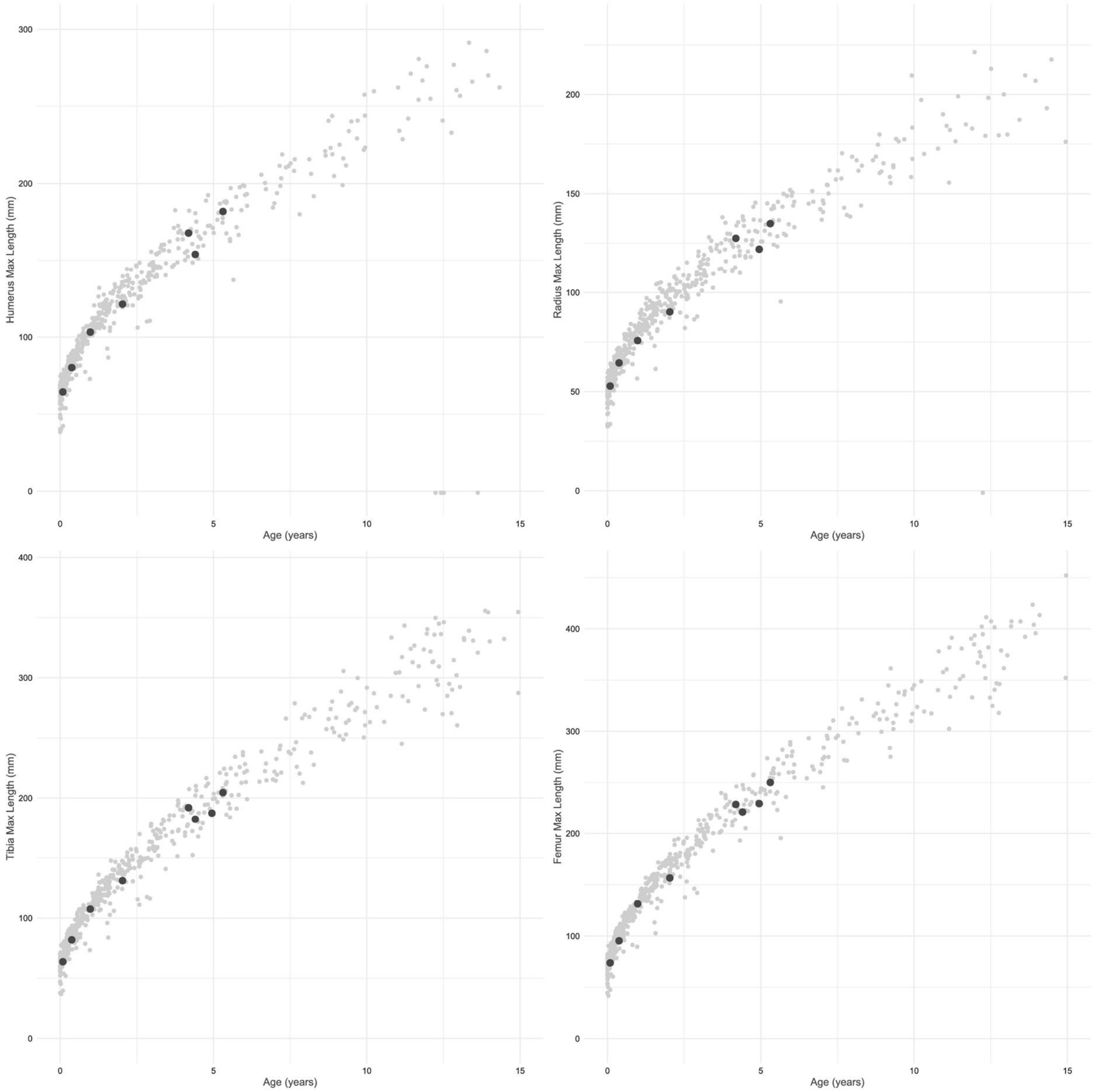

FIGURE 11. Bivariate scatterplots of long bone length and age. Individuals with a contributing cause of death of "victim of child abuse" were highlighted (larger black circles) to illustrate their comparative sizes against the entire deceased sample from the US (UNM/OCMEM)

8 DISCUSSION

8.1 Diaphyseal lengths and dental development: Deceased versus living

Our results showed no statistical differences between the long bones of deceased individuals and living individuals for all age groups. The findings contradict the expected findings of the biological anthropological community, which is there should be differences in group phenotypes because of differential exposures to environmental stressors, with individuals exposed to higher levels of stress potentially exhibiting stunting or stunted linear growth, which correlates to higher mortality risks (Buikstra & Cook, 1980; DeWitte & Wood, 2008; Goodman et al., 1988; McDonald et al., 2013; Prendergast & Humphrey, 2014; Spake & Cardoso, 2019b; Steckel, 2005). However, and of practical significance, is the size differences between the SA living and deceased individuals less than 2 years of age. While the KS tests did not find statistically significant differences in their cumulative distributions, there is a progression from large differences in mean lengths of individuals under 2 years of age to minimal differences in mean lengths of 2 year-olds; this is present in all long bones (see Table 4). The results for this age group may be an artifact of sampling bias, concomitant with the differential hospitalization trends between adults and children, and/or infant mortality rates. For example, only four individuals in the living sample are compared to 45 individuals in the deceased sample between birth and 0.99 years. Because the living sample is associated with an emergency room environment, there could be a bias to older children rather than younger children (Merrill & Owens, 2007; Witt et al., 2016). Furthermore, there is research to suggest that childhood exposure to radiation from imaging modalities may be linked to certain adult cancers (e.g., Don, 2004; Preston-Martin & White, 1990). Therefore, even when very young children are admitted to an emergency room, it may be unlikely that there is any imaging, which makes it difficult to obtain a large sample of radiographs of living children at very young ages. If a larger sample of living children under the age of 2 years could be obtained, comparisons could confirm or refute the size differences among these ages.

As with diaphyseal lengths, no statistically significant differences were found in dental development between the US living and deceased samples. This is not unexpected, as dental development is considered to be less influenced by environment, and it could be hypothesized that there would be minimal differences between the living and deceased samples (Cardoso, 2007b). Our results did not elucidate differential trends in dental development trends between deceased and living individuals.

8.2 Diaphyseal lengths, dental development, MOD, and context

Further exploration of the SA deceased sample clearly revealed a steep decrease in number of individuals who died of a natural death after 2 years of age (Figure 8). Concurrent with this reduction of individuals who die of natural deaths is the absence of substantial size differences between individuals with a MOD of natural death compared to those who died of an accident or homicide (see Figure 6). To be clear, while there are larger numbers of individuals with a natural death in the SA sample, the actual size differences are not present among the MODs. In the US sample, which has overall larger number of individuals less than 2 years of age, it is clear that the loess lines have similar trajectories. However, it is also notable that the few individuals (~4) that exhibit small bone lengths per age and are less than 2 years of age, all died of natural deaths. Therefore, based on both the SA data and the US data, a deceased individual that is less than 2 years of age is likely to be smaller than the living population (see conversation above) and die from natural causes. Early infancy failure-to-thrive results in overall growth disruption of postcranial elements as a way to preserve vital cranial and spinal structures, and ultimately, can lead to death if health issues remain unresolved (Agarwal, 2016; Kuzawa & Bragg, 2012). While this conclusion is more expected from the biological anthropology community, it is important to emphasize that this trend does not continue in individuals greater than 2 years of age. In modern populations, mortality rates associated with congenital anomalies (i.e., natural MODs) decreases with increased age and admissions in hospitals for injuries increases with age (4% in children under 1 year to 30% for 15 to 17 year olds) (Cunningham et al., 2018; Treves et al., 2011).

Previous research has suggested that individuals who died of an accidental death are more reflective of the “normal” population (Spake & Cardoso, 2018). Our data do not display practical or statistically significant differences among these two MODs, in distribution or size, in either the US or SA samples (Figures 6 and 7). These findings clearly demonstrate that age distributions as well as long bone lengths of individuals with a MOD of accident or homicide are minimally different throughout ontogeny. Furthermore, the results of the US deceased data also suggest there is no difference between long bone length of subadults in the FA or ME setting, which are samples presumed to have different MODs (see Figure 9).

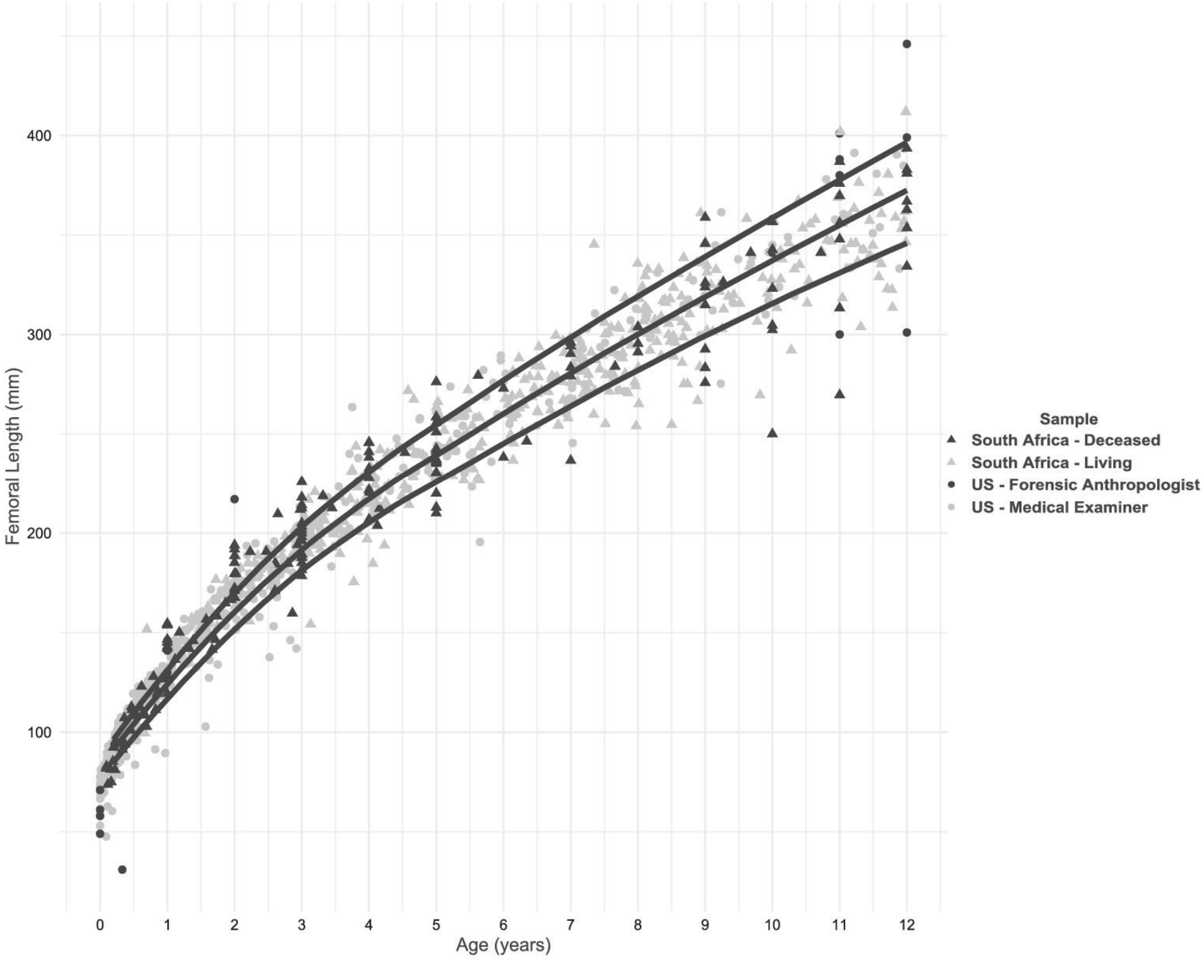

FIGURE 12. Bivariate scatterplot of femur length and age of the United States and South African samples. The lines are the fifth, 50th, and 95th percentiles of (Maresh, 1970)

In addition, if we assumed "homicide" equates to lower SES then we would expect to see larger discrepancies in lower limb lengths between living children and deceased children with a MOD of homicide, especially in the older subadults. These findings were not observed in the SA or the US deceased sample. Furthermore, the mean diaphyseal length for all bones, in both countries, for living and deceased, fall within the range of normal human long bone variation (± 2SD) according to the Denver Child Research Council growth standards commonly used in biological anthropology (Figure 12) (Maresh, 1970). Individuals falling outside of the fifth and 95th percentiles originate from all samples - deceased, living, SA, and the US – and do not show any particular trend. The absence of any clear pattern or trend for these differences means we cannot conclude that living or deceased individuals necessarily come from better or worse environments, respectively, or that subadult victims of homicide are systematically smaller for age than victims of accidents or natural deaths (with the exception of <2-year-olds), or even that disparate SES levels significantly manifest themselves skeletally in these samples. Therefore, associating deceased children in the forensic record with a predominant MOD of homicide, and/or with shorter stature and/or with low SES, does not seem justified.

8.3 Revisiting reference samples and the Osteological paradox

Based on these findings, it seems that reference data do not need to explicitly include living individuals if applied to a living individual nor do they need to explicitly include deceased individuals if applied to a deceased individual. In either situation, the reference sample can include individuals that are either deceased or living as there are no quantifiable biological differences in contemporary children. The one potential caveat to be argued for when considering the composition of the reference and target sample is that methods using diaphyseal dimensions developed from a medico-legal setting that include deceased individuals less than 2 years of age with a natural cause of death are potentially inappropriate when working with living individuals. This brings us back to the complexities of reference samples, specifically when working with the very young. If a forensic anthropological analysis has been requested of an individual that is less than 2 years of age, and it is not specific to trauma analysis, then based on the data available both from the current study and others, there is a high probability the deceased child would be associated with a failure-to-thrive situation. It is very unlikely to ever acquire large samples of very young individuals unless they were from a medicolegal or cemetery setting. In both instances, the reference population would actually reflect the target population for the forensic anthropologist. However, the purview of the forensic anthropologist is expanding and there may be situations, such as with natural or mass disasters (i.e., presuming a MOD of accidental), where a deceased individual is not linked to a failure-to-thrive situation. In this context, the medicolegal and cemetery settings would be an inappropriate reference sample for individuals <2 years of age in an evaluation of long bone lengths.

We would expect to see smaller long bones in a deceased sample, especially in specific age ranges/life history stages highly sensitive to growth delays or stunting, such as infancy and early childhood (birth to ~4–5 years). Yet, our findings echo similar statements made by previous authors: mortality risks and mortality are strongly age-dependent, and instances of mortality and growth stunting are greatest between birth and 2 years, as this is when growth rate is directly informed by nutritional intake and sufficiency (Kuzawa & Bragg, 2012; Prendergast & Humphrey, 2014; S. R. Saunders & Hoppa, 1993). Early environmental conditions interact with fetal programming and the genome to define boundaries of birth weight and future linear growth and development for a given individual (Cooper et al., 2006, p. 201). The influence of the prenatal environment and the maternal-fetal nexus is crucial to an individual's health and participate in defining their ontogenetic outcome, even if their impact can be somewhat mediated postnatally. Therefore, following the osteological paradox theory (Wood et al., 1992), poor prenatal, and early postnatal environments could possibly be linked to growth stunting and higher mortality rates, and could explain why shorter long bones were observed for the infants of the SA sample (birth to 2 years) who died of natural causes. However, this constitutes a notable exception to the trends observed in both samples. The majority of these individuals who died <1 year were classified as natural or undetermined deaths (Figure 8) and the natural deaths were overwhelmingly linked to congenital diseases and viral and bacterial infections (i.e., pneumonia, gastroenteritis, etc.). The number of undetermined deaths could point to possible undiagnosed or untreated prenatal issues experienced by the fetus and/or its mother during gestation, which in turn could have caused stress to the fetus for an undetermined amount of time. Growth failure usually begins in utero, is pronounced during the first year of life, and is at its peak at 2 years of age (Martorell & Zongrone, 2012; Prendergast & Humphrey, 2014; Richard et al., 2012). Alternatively, stress could also have risen within that first post-natal year. For these cases, there is no reliable way of knowing when stress happened and how long it lasted before death occurred. Consequently, interpreting short long bone length as an effect of stress, be it pathological, nutritional, or other, is not straightforward. Short stature is ultimately a response to a stressor; skeletal plasticity allows for different parts of the skeleton to be impacted with varying amplitudes, depending on the life history or developmental stage of the affected individual (Agarwal, 2016).

The osteological paradox would also imply that the magnitude of size differences between the deceased (presumably unhealthy) and living (presumably healthy) would continue as age increases. Following this early age period (< 2 years), there is no evidence of selective mortality, or simply, no differences between the dental and skeletal development of deceased and living individuals above the age of > = 2 years in either of our samples. The osteological paradox is not generalizable to any and every contemporary or past deceased subadult assemblage. Other authors have indeed promoted caution when making health or mortality interpretations based on skeletal and dental indicators of stress in subadults in past populations because stress does not equate to health (DeWitte & Stojanowski, 2015; Saunders & Hoppa, 1993). Similarly, the current study suggests we should take precautions in making inferences from seemingly shorter stature in forensic and clinical samples of current populations. If it is difficult, and sometimes even impossible, to identify a specific cause of death for infants in a forensic sample for which demographic information and pathological records are available, it is also difficult to infer in bioarcheological samples with close to no verified demographic or contextual/individual information available in a majority of cases.

The populations from which the deceased individuals in our samples are taken is made up of an unknown mixture of individuals who varied in their underlying frailty or susceptibility to disease and death. Hidden heterogeneity in risks means mortality is selective for frailty and the distribution shifts with age. Therefore, frailty – and the risk of death – decreases as age increases, making it impossible to interpret aggregate-level age-specific mortality rates in terms of individual risks at death (Wood et al., 1992). In the presence of hidden heterogeneity, population-mortality patterns tell us very little about individual risks of death. We would need to know the distribution of frailty in the living population (Wood et al., 1992) to assess if levels of frailty are similar in deceased and living samples, or if the deceased sample presents with more frailty than the living, assuming that deceased individuals are more frail than living ones.

To expand beyond the modern samples and consider implications for bioarchaeology, it is similarly difficult to understand past levels of frailty within ancient skeletal samples. Further, these may be small populations with unknown levels of migration (i.e., demographic nonstationarity), and therefore, may be relatively homogeneous. Attempts have been made to quantify a skeletal frailty index to explore levels of stress and health and survivorship (e.g., Marklein et al., 2016; Marklein & Crews, 2017); however, these require relatively complete skeletons and can still be an imperfect representation of past frailty. Paleodemographic analyses that incorporate various skeletal indicators of stress or disruptions to growth and development can also be used to assess frailty (e.g., DeWitte, 2014; Yaussy & DeWitte, 2018); however, given the complex etiology of these skeletal indicators, it can be difficult to interpret their meaning in the past (Pilloud & Schwitalla, 2020).

Depending on the magnitude, cause, and length of a certain stress episode, both deceased and living individuals (ancient or recent) may or may not exhibit certain skeletal indicators, such as stunted skeletal growth. In other words, there is no way to tease apart whether stress and mortality are associated, or the product of correlation without causation. Therefore, while skeletal indicators of stress, such as shorter stature, may be used to indicate adverse conditions, caution should be taken to avoid unfounded correlations between skeletal variables and mortality or morbidity. In sum, while adverse environmental conditions may be one explanation of variability between samples (such as MOD or living versus deceased), these differences should not be assumed to account for other factors affecting growth variability, including genetic growth potential, epigenetic processes, and secular change. Moreover, the formation of an appropriate reference sample is dependent on accounting for all forms of variation that either a forensic anthropologist or bioarchaeologist may encounter. For this reason, the current study suggests *not truncating reference samples* based on the presumption that a separation must exist between a “healthy” living population and those individuals that become part of the work of a forensic anthropologist or a bioarchaeologist (Cook, 1981; S. R. Saunders & Hoppa, 1993; B. Wood, 1992).

8.4 The impact of SES in contemporary populations

It has been argued that the differences in growth between “forensic” and “normal” populations would be of a greater magnitude in countries where social inequality is greater (Spake & Cardoso, 2018). It is also plausible that individuals from a country with an overall low SES could present with smaller differences in long bone lengths between living and deceased children because a larger number of the population would be affected by higher than average poverty rates, leading to a downward shift in the overall height distribution. The US and SA do have contrasting Human Development Indices (HDI), which can be used to interpret country-level or city-level SES. The US has a very high HDI of 0.920 and SA has a medium HDI of 0.705 (United Nations Development Program, 2018). Because HDI does not directly capture social inequality, the Gini coefficient may be a more appropriate proxy for SES, as this is a measure of inequality of wealth within a country (Bogin, 1999; World Bank, Development Research Group, 2019). Countries with high inequality typically contain broad discrepancies in access to economic resources, including access to quality food, housing and healthcare, as well as limited upward mobility, all of which are factors that can influence individual and population growth and development. Based on these discrepancies, their populations are argued to be overall shorter compared to those of countries with greater equality (Bogin, 1999; Sahn & Younger, 2005). The US has a medium high (41.1) Gini coefficient while SA has a very high Gini coefficient (63.4). Subsequently, it could be assumed that even the deceased individuals in the US would present with larger long bone lengths than either the living or deceased South Africans. Figure 13 illustrates unremarkable differences between the loess lines of the two populations for all long bones. There are no apparent differences in the long bones of individuals between the two countries, even when using one of the more theoretically stress-sensitive age indicators (i.e., diaphyseal lengths) (Figure 13). Both the upper and lower limbs demonstrate minimal differences between the SA and US samples, but there is a slight difference that becomes apparent as age increases in the upper limb bones (around 10 years of age), but is most apparent in the lower limb bones (around 8 years of age). The pattern of divergence is likely because of positive allometry that mainly affects the lower limbs and has been shown to present with population variation in both its onset and magnitude (Bogin & Varela-Silva, 2010; Meadows Jantz & Jantz, 1999). The leg being more sensitive than the arm, we would expect to see larger differences between the populations for these two appendages. Because long bone lengths are largely comparable through the majority of ontogeny, the size differences that are apparent in adults and the divergence visible here around the prepubertal period may be more associated with divergence in the timing of epiphyseal fusion and the timing and magnitude of the peak height velocity than the growth of the long bones prior to puberty.

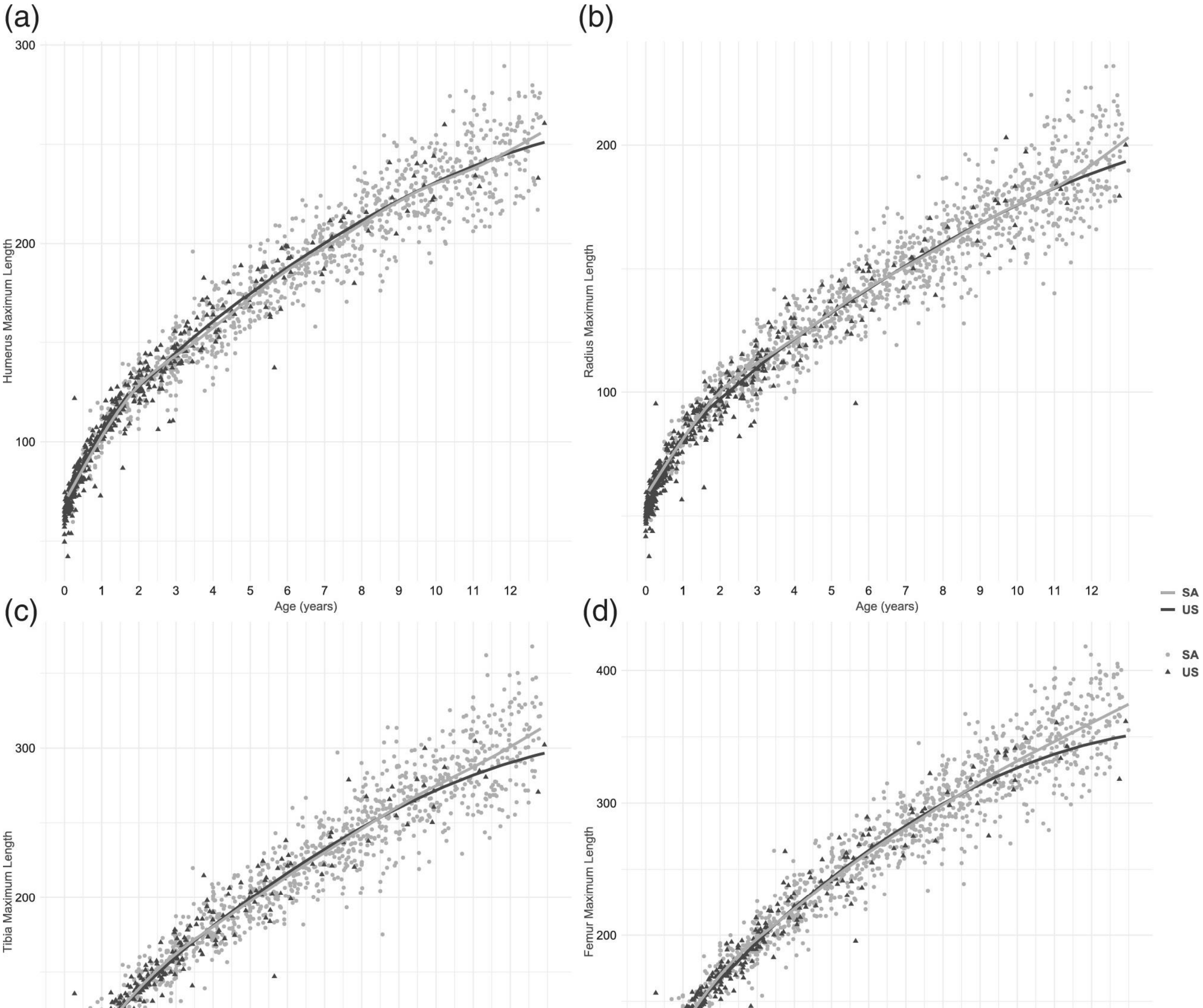

FIGURE 13. Bivariate scatterplots of long bone lengths and age for both the United States and South African samples. Loess lines illustrate the similar trends per country

In sum, our results show that subadults in the forensic record do not show any evidence of shorter length compared to their living peers within the same population, meaning we cannot infer higher levels of stress or lower SES for the former based on long bone length alone. In addition, there appears to be no significant differences in pre-pubertal long bone length between US and South African subadults, living or deceased, meaning the impact of SES could be mitigated, at least up to a certain point, by growth canalization. In terms of anthropological research, this also means we could use larger numbers of subadults from various genetic and environmental backgrounds and contexts to include more variability in samples when building biological or forensic anthropological methods for individuals who have not yet reached puberty.

9 CONCLUSIONS

For decades, paleodemographic researchers have grappled with the inherent biases of the skeletal record as they attempt to make statements about past population structure based on contemporary population structure (e.g., Bocquet-Appel & Masset, 1982; Konigsberg & Frankenberg, 1992; Milner et al., 2008; S. R. Saunders & Hoppa, 1993; J. Wood et al., 1992). Inferences about these past populations inherently require comparative analyses, but that leads to concerns regarding the appropriateness of contemporary samples, and if we should be comparing the historic populations to contemporary deceased or living populations. With these concerns in mind, it seems appropriate to first investigate if growth and development differ between living and deceased children in modern samples from the same population.

This study showed that skeletal measurements and dental development stages, two of the most widely used subadult age indicators, did not show any statistically significant differences between living and deceased individuals or according to MOD in two contemporary samples from SA and the US. However, in the current study, trends of practical importance were identified in individuals younger than 2 years of age compared to older individuals in the sample (>2 years of age). Substantial size differences existed between long bone lengths of living and deceased individuals less than 2 years of age and there were larger numbers of natural deaths in individuals less than 2 years of age. To our knowledge, the current study is one of few to compare dental development and skeletal dimensions in living individuals and non-survivors. These results are expected in the framework of previous studies (e.g., Holman et al., 2004; Saunders & Hoppa, 1993, 1993; Spake & Cardoso, 2019b), and refute fundamental theories associated with expected phenotypic differences between living individuals and non-survivors through all of ontogeny. According to these theories, non-survivors should consistently present with shorter long bone lengths, which was not the case here beyond 1.99 years of age. The notable exception of the diaphyseal length differences observed between deceased and living individuals under the age of 2 years merits advanced research in human biology, pertaining to prenatal ontogeny and early canalization to understand the factors behind these particular findings.

While it may be appropriate to consider population differences and population-specific models in biological anthropology method development and validation when estimating ancestry and sex in adults, the results presented here indicate that it may be unnecessary for subadult age estimation based on long bone growth and dental development indicators, if the parsing of populations is based on expectations related to mortality bias. The assumption that deceased children *should* display different growth patterns to that of living children and should, therefore, be privileged as reference data for past populations is refuted by our current analyses. Variation in contemporary subadult skeletal and dental growth is comparable regardless of whether the individuals are deceased or living, or if deceased, how they died. Our results concur with the conclusions of Saunders and Hoppa (1993) 30 years ago, reprised by DeWitte and Stojanowski (2015): if biological mortality bias exists in subadult skeletal samples, the effects are too small to detect and higher mortality risk should not be inferred based on shorter stature or long bone length, particularly when over the age of 2 years.

We do not argue that differences between non-survivors and living individuals could have existed in past populations, as there is no way of verifying it with the present study. However, the differences are insignificant in the two contemporary samples used for these analyses. Their respective populations of origin are composed of individuals from various genetic and environmental backgrounds, and living in countries of arguably different socioeconomic backgrounds, as reflected by their different levels of HDI and the Gini index. Results suggest some of the principles of the osteological paradox should not be generalized in contemporary populations and subadult reference samples should be all-encompassing and inclusive of populations from various backgrounds. By “various backgrounds” we do not limit this to intra-population variation, but to extend the sample beyond national boundaries. This is of particularly timely relevance as the scope of forensic anthropological casework is shifting on local and global scales toward individuals of various and unknown backgrounds, stemming from socioeconomic globalization and worldwide increase in migrations. In addition, forensic anthropologists are now increasingly called to assist in disaster victim identification associated with mass and natural disasters. The increase of varied work settings results in contexts with considerable skeletal variation. For example, individuals who die during natural or mass disasters are potentially a different subset of the population than the forensic community and are more likely to be representative of “normal” or “healthy” populations of different genetic and environmental backgrounds. The present results demonstrate that a “forensic” sample may actually be just as reflective of normal growth and development as previously believed. Considering the evolution of the forensic anthropologists' role, the diversification of forensic anthropological contexts, and the absence of significant differences in skeletal and dental indicators of growth and development between deceased and living subadults, reference samples, or populations used to develop forensic methods should ideally include subjects from as many backgrounds and contexts as possible. This would allow the practitioner to capture more variation in sampled populations, increase sample size, and ultimately provide the anthropology community with more applicable and robust methods.

ACKNOWLEDGMENTS

Numerous individuals helped to strengthen this manuscript: Hassem Geha provided the orthopantomograms, Richard Jantz provided the FDB data, and Michael Price offered theoretical discussions and comments prior to submission. The data were available because of funding by the National Institute of Justice (NIJ Awards 2015-DN-BX-K409 and 2017-DN-BX-K409) and the National Science Foundation (BCS-13 1551913). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Justice or National Science Foundation.

AUTHOR CONTRIBUTIONS

Kyra Stull: Conceptualization; data curation; formal analysis; funding acquisition; investigation; methodology; project administration; supervision; visualization; writing-original draft; writing-review and editing. **Christopher Wolfe:** Formal analysis; methodology; visualization; writing-original draft; writing-review and editing. **Louise Corron:** Data curation; funding acquisition; methodology; supervision; writing-original draft; writing-review and editing. **Kelly Heim:** Data curation; methodology; writing-review and editing. **Cortney Hulse:** Data curation; methodology; writing-review and editing. **Marin Pilloud:** Conceptualization; data curation; supervision; writing-original draft; writing-review and editing.

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

REFERENCES

Agarwal, S. C. (2016). Bone morphologies and histories: Life course approaches in bioarchaeology. *American Journal of Physical Anthropology*, 159(S61), 130– 149. https://doi.org/10.1002/ajpa.22905

Albanese, J. (2003). Identified skeletal reference collections and the study of human variation. Canada: McMaster University.

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, 481– 490.

Bairagi, R., & Chowdhury, M. K. (1994). Socioeconomic and anthropometric status, and mortality of young children in rural Bangladesh. *International Journal of Epidemiology*, 23(6), 1179– 1184. https://doi.org/10.1093/ije/23.6.1179

Berry, S. D., & Edgar, H. J. H. (2020). Extracting and standardizing medical examiner data to improve health. *AMIA Summits on Translational Science Proceedings*, 2020, 63– 70.

Bocquet-Appel, J.-P., & Masset, C. (1982). Farewell to paleodemography. *Journal of Human Evolution*, 11(4), 321– 333. https://doi.org/10.1016/S0047-2484(82)80023-7

Bogin, B. (1999). Patterns of human growth (Vol. 2nd). Cambridge: Cambridge University Press.

Bogin, B., & Varela-Silva, M. (2010). Leg length, body proportion, and health: A review with a note on beauty. *International Journal of Environmental Research and Public Health*, 7, 1047– 1075.

Boldsen, J., Milner, G., Konigsberg, L., & Wood, J. (2002). Transition analysis: A new method for estimating age from skeletons. In R. Hoppa & J. Vaupel (Eds.), Paleodemography: Age distributions from skeletal samples (pp. 73– 106). Cambridge: Cambrige University Press.

Bosworth, B. (2018). Increasing disparities in mortality by socioeconomic status. *Annual Review of Public Health*, 39(1), 237– 251. https://doi.org/10.1146/annurev-publhealth-040617-014615

Braudt, D. B., Lawrence, E. M., Tilstra, A. M., Rogers, R. G., & Hummer, R. A. (2019). Family socioeconomic status and early life mortality risk in the United States. *Maternal and Child Health Journal*, 23(10), 1382– 1391. https://doi.org/10.1007/s10995-019-02799-0

Buikstra, J. E., & Cook, D. C. (1980). Palaeopathology: An American account. *Annual Review of Anthropology*, 9(1), 433– 470. https://doi.org/10.1146/annurev.an.09.100180.002245

Cameron, N. (2012). The human growth curve, canalization and catch-up growth. In N. Cameron & B. Bogin (Eds.), Human growth and development (2nd ed., pp. 1– 22). New York: Elsevier.

Cardoso, H. (2007a). Differential sensitivity in growth and Developement of dental and skeletal tissue to environmental quality. *Arquivos De Medicina*, 21(1), 19– 23.

Cardoso, H. F. V. (2007b). Environmental effects on skeletal versus dental Developement: Using a documented subadult skeletal sample to test a basic assumption in human Osteological research. *American Journal of Physical Anthropology*, 132, 223– 233.

Cardoso, H. F. V., Vandergugten, J., & Humphrey, L. (2017). Age estimation of immature human skeletal remains from the metaphyseal and epiphyseal widths of the long bones in the post-natal period. *American Journal of Physical Anthropology*, 162(1), 19– 35. https://doi.org/10.1002/ajpa.23081

Chen, L. C., Chowdhury, A., & Huffman, S. L. (1980). Anthropometric assessment of energy-protein malnutrition and subsequent risk of mortality among preschool aged children. *The American Journal of Clinical Nutrition*, 33(8), 1836– 1845. https://doi.org/10.1093/ajcn/33.8.1836

Colman, K. L., Dobbe, J. G. G., Stull, K. E., & Ruijter, J. M. (2017). The geometrical precision of virtual bone models derived from clinical computed tomography data for forensic anthropology. *International Journal of Legal Medicine*, 131(4), 1155– 1163.

Colman, K. L., van der Merwe, A. E., Stull, K. E., & Dobbe, J. G. G. (2019). The accuracy of 3D virtual bone models of the pelvis for morphological sex estimation. *International Journal of Legal Medicine.*, 133, 1853, 1860

Conceição, E., & Cardoso, H. (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

Cook, D. (1981). Mortality, age-structure, and status in the interpretation of stress indicators in prehistoric skeletons: A dental example from the lower Illinois Valley. In I. Chapman, I. Kinnes, & K. Randsborg (Eds.), The archaeology of death (pp. 134– 144). Cambridge: Cambridge University Press.

Cooper, C., Westlake, S., Harvey, N., Javaid, K., Dennison, E., & Hanson, M. (2006). Review: Developmental origins of osteoporotic fracture. *Osteoporosis International*, 17(3), 337– 347. https://doi.org/10.1007/s00198-005-2039-5

Corron, L., Marchal, F., Condemi, S., & Chaumoître, K. (2017). Evaluating the consistency, repeatability, and reproducibility of Osteometric data on dry bone surfaces, scanned dry bone surfaces, and scanned bone surfaces obtained from living individuals. *Bulletins et Mémoires de La Société d'Anthropologie de Paris*, 29(1), 33– 53. https://doi.org/10.1007/s13219-016-0172-7

Cunningham, R. M., Walton, M. A., & Carter, P. M. (2018). The major causes of death in children and adolescents in the United States. *The New England Journal of Medicine*, 379(25), 2468– 2475. https://doi.org/10.1056/NEJMsr1804754

Demirjian, A., Goldstein, H., & Tanner, J. M. (1973). A new system of dental age assessment. *Human Biology*, 45(2), 211– 227.

DeWitte, S. N. (2014). Mortality risk and survival in the aftermath of the medieval Black death. *PLoS One*, 9(5), e96513. https://doi.org/10.1371/journal.pone.0096513

DeWitte, S. N., & Stojanowski, C. M. (2015). The Osteological paradox 20 years later: Past perspectives, future directions. *Journal of Archaeological Research*, 23(4), 397– 450. https://doi.org/10.1007/s10814-015-9084-1

DeWitte, S. N., & Wood, J. W. (2008). Selectivity of black death mortality with respect to preexisting health. *Proceedings of the National Academy of Sciences of the United States of America*, 105(5), 1436– 1441. https://doi.org/10.1073/pnas.0705460105

Dhavale, N. P., Halcrow, S. E., Buckley, H., Tayles, N., Domett, K., & Gray, A. R. (2017). Linear and appositional growth in infants and children from the prehistoric settlement of ban non Wat. Northeast Thailand: Evaluating biological responses to agricultural intensification in Southeast Asia. https://doi.org/10.1016/J.JASREP.2016.12.019

Don, S. (2004). Radiosensitivity of children: Potential for overexposure in CR and DR and magnitude of doses in ordinary radiographic examinations. *Pediatric Radiology*, 34(S3), S167– S172. https://doi.org/10.1007/s00247-004-1266-9

Elder, T. E., Goddeeris, J. H., & Haider, S. J. (2016). Racial and ethnic infant mortality gaps and the role of socio-economic status. *Labour Economics*, 43, 42– 54. https://doi.org/10.1016/j.labeco.2016.04.001

Fazekas, I. G., & Kósa, F. (1978). Forensic Fetal Osteology. Budapest, Hungary: Akademiai Kiado.

Garvin, H. M., & Stock, M. K. (2016). The utility of advanced imaging in forensic anthropology. *Academic Forensic Pathology*, 6(3), 499– 516. https://doi.org/10.23907/2016.050

Getz, S. M. (2020). The use of transition analysis in skeletal age estimation. *WIREs Forensic Science*, 2, e1378. https://doi.org/10.1002/wfs2.1378

Goodman, A. H., Thomas, R. B., Swedlund, A. C., & Armelagos, G. J. (1988). Biocultural perspectives on stress in prehistoric, historical, and contemporary population research. *American Journal of Physical Anthropology*, 31(S9), 169– 202. https://doi.org/10.1002/ajpa.1330310509

Haider, S. J. (2014). Racial and ethnic infant mortality gaps and socioeconomic status. *Focus*, 1, 3.

He, X., Akil, L., Aker, W. G., Hwang, H.-M., & Ahmad, H. A. (2015). Trends in infant mortality in United States: A brief study of the southeastern States from 2005–2009. *International Journal of Environmental Research and Public Health*, 12(5), 4908– 4920. https://doi.org/10.3390/ijerph120504908

Heim, K. (2018). Population Variation in Dental Development and Its Effect on Forensic Age Estimation (dissertation). Reno, Reno, NV: University of Nevada.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6(2), 65– 70.

Holman, D., O'Connor, K. A., & Jones, R. E. (2004). Assessing biological mortality bias in deciduous tooth emergence. CSDE Working Papers: 04–10 (pp. 1–17). Seattle: University of Washington.

Hoppa, R. D. (2000). What to do with long bones: Toward a progressive paleoauxology. *Anthropologie (1962-)*, 38(1), 23– 32.

Jantz, R. L., & Moore-Jansen, P. H. (1988). A Data Base for Forensic Anthropology: Structure content and analysis (no. 47). Knoxville: Department of Anthropology, The University of Tennessee.

Johnston, F. (1962). Growth of the long bones of infants and young children at Indian knoll. *American Journal of Physical Anthropology*, 20, 249– 254.

Kim, D., & Saada, A. (2013). The social determinants of infant mortality and birth outcomes in Western developed nations: A cross-country systematic review. *International Journal of Environmental Research and Public Health*, 10(6), 2296– 2335. https://doi.org/10.3390/ijerph10062296

Kimmerle, E. H. (2014). Forensic anthropology in long-term investigations: 100 cold years. *Annals of Anthropological Practice*, 38(1), 7– 21. https://doi.org/10.1111/napa.12039

Kimmerle, E. H., Jantz, R. L., Konigsberg, L. W., & Baraybar, J. P. (2008). Skeletal estimation and identification in American and east European populations. *Journal of Forensic Sciences*, 53(3), 524– 532. https://doi.org/10.1111/j.1556-4029.2008.00708.x

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

Konigsberg, L. W., & Frankenberg, S. R. (1992). Estimation of age structure in anthropological demography. *American Journal of Physical Anthropology*, 89(2), 235– 256. https://doi.org/10.1002/ajpa.1330890208

Konigsberg, L. W., & Frankenberg, S. R. (1994). Paleodemography: “Not quite dead.”. *Evolutionary Anthropology: Issues, News, and Reviews*, 3(3), 92– 105. https://doi.org/10.1002/evan.1360030306

Kuzawa, C. W., & Bragg, J. M. (2012). Plasticity in human life history strategy: Implications for contemporary human variation and the evolution of genus homo. *Current Anthropology*, 53(S6), S369– S382. https://doi.org/10.1086/667410
Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159– 174. https://doi.org/10.2307/2529310

Lehmann, E. (2006). Nonparametrics: Statistical methods based on ranks. New York: Springer-Verlag. Retrieved from. https://www-springer-com.uplib.idm.oclc.org/gp/book/9780387352121

Maresh, M. (1970). Measurements from roentgenograms. In R. McCammon (Ed.), Human growth and development (pp. 157– 200). Springfield, IL: CC. Thomas.

Marklein, K. E., & Crews, D. E. (2017). Frail or hale: Skeletal frailty indices in medieval London skeletons. *PLoS One*, 12(5), e0176025. https://doi.org/10.1371/journal.pone.0176025

Marklein, K. E., Leahy, R. E., & Crews, D. E. (2016). In sickness and in death: Assessing frailty in human skeletal remains. *American Journal of Physical Anthropology*, 161(2), 208–225. https://doi.org/10.1002/ajpa.23019

Martorell, R., & Zongrone, A. (2012). Intergenerational influences on child growth and undernutrition. *Paediatric and Perinatal Epidemiology*, 26(Suppl 1), 302– 314. https://doi.org/10.1111/j.1365-3016.2012.01298.x

McDonald, C. M., Olofin, I., Flaxman, S., Fawzi, W. W., Spiegelman, D., Caulfield, L. E., … Nutrition Impact Model Study. (2013). The effect of multiple anthropometric deficits on child mortality: Meta-analysis of individual data in 10 prospective studies from developing countries. *The American Journal of Clinical Nutrition*, 97(4), 896– 901. https://doi.org/10.3945/ajcn.112.047639

McKern, T., & Stewart, T. (1957). *Skeletal Age changes in young American males*: *Analysed from the standpoint of age identification.* (Environmental Protection Research Division No. EP-45). Natick, MA: Headquarters Quartermaster Research and Development Command.

Meadows Jantz, L., & Jantz, R. L. (1999). Secular change in long bone length and proportion in the United States. *American Journal of Physical Anthropology*, 110, 57– 67.

Merrill, C., & Owens, P. (2007). Reasons for Being Admitted to the Hospital through the Emergency Department for Children and Adolescents, 2004 (no. statistical brief #33). Rockville, MD: Agency for Healthcare Research and Quality. Retrieved from Agency for Healthcare Research and Quality website: https://www-ncbi-nlm-nih-gov.uplib.idm.oclc.org/books/NBK61976/

Milner, G. R., Wood, J. W., & Boldsen, J. L. (2008). Advances in Paleodemography. In M. A. Katzenberg & S. R. Saunders (Eds.), Biological anthropology of the human skeleton (pp. 561– 600). Hoboken, NJ, USA: John Wiley & Sons, Inc.. https://doi.org/10.1002/9780470245842.ch18

Moorrees, C., Fanning, E., & Hunt, E. (1963). Age variation of formation stages for ten permanent teeth. *Journal of Dental Research*, 42(6), 1490– 1502.

Mummert, A., Esche, E., Robinson, J., & Armelagos, G. J. (2011). Stature and robusticity during the agricultural transition: Evidence from the bioarchaeological record. *Economics & Human Biology*, 9(3), 284– 301. https://doi.org/10.1016/j.ehb.2011.03.004

Olofin, I., McDonald, C. M., Ezzati, M., Flaxman, S., Black, R. E., Fawzi, W. W., … Nutrition Impact Model Study (anthropometry cohort pooling). (2013). Associations of suboptimal growth with all-cause and cause-specific mortality in children under five years: A pooled analysis of ten prospective studies. *PloS One*, 8(5), e64636. https://doi.org/10.1371/journal.pone.0064636

Ousley, S. (2013). A Radiographic Database for Estimating Biological Parameters in Modern Subadults. *Final Technical Report, National Institute of Justice award number 2008-DN-BX-K152. Available at*: https://ncjrs.gov/pdffiles1/nij/grants/242697.pdf. (Grant No. Award No. 2008-DN-BX-K152). Erie, PA: Mercyhurst University.

Pilloud, M. A., & Schwitalla, A. W. (2020). Re-evaluating traditional markers of stress in an archaeological sample from Central California. *Journal of Archaeological Science*, 116, 105102. https://doi.org/10.1016/j.jas.2020.105102

Pinhasi, R., Timpson, A., Thomas, M., & Slaus, M. (2014). Bone growth, limb proportions and non-specific stress in archaeological populations from Croatia. *Annals of Human Biology*, 41(2), 127– 137. https://doi.org/10.3109/03014460.2013.835443

Prendergast, A. J., & Humphrey, J. H. (2014). The stunting syndrome in developing countries. *Paediatrics and International Child Health*, 34(4), 250– 265. https://doi.org/10.1179/2046905514Y.0000000158

Preston-Martin, S., & White, S. C. (1990). Brain and salivary gland tumors related to prior dental radiography: Implications for current practice. *The Journal of the American Dental Association*, 120(2), 151– 158.

R Core Team. (2019). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from. https://www.R-project.org/

Richard, S. A., Black, R. E., & Checkley, W. (2012). Revisiting the relationship of weight and height in early childhood. *Advances in Nutrition*, 3(2), 250– 254. https://doi.org/10.3945/an.111.001099

Sahn, D. E., & Younger, S. D. (2005). Improvements in children's health: Does inequality matter? *The Journal of Economic Inequality*, 3(2), 125– 143. https://doi.org/10.1007/s10888-005-4494-9

Saunders, S. (2008). Juvenile skeletons and growth-related studies. In M. Katzenberg & S. Saunders (Eds.), Biological anthropology of the human skeleton (2nd ed., pp. 117– 147). Hoboken: John WIley & Sons.

Saunders, S. R., & Hoppa, R. D. (1993). Growth deficit in survivors and non-survivors: Biological mortality bias in subadult skeletal samples. *American Journal of Physical Anthropology*, 36(S17), 127– 151. https://doi.org/10.1002/ajpa.1330360608

Smedman, L., Sterky, G., Mellander, L., & Wall, S. (1987). Anthropometry and subsequent mortality in groups of children aged 6-59 months in Guinea-Bissau. *The American Journal of Clinical Nutrition*, 46(2), 369– 373. https://doi.org/10.1093/ajcn/46.2.369

Smith, B. H. (1991). Standards of human tooth formation and dental age assessment. In Advances in dental anthropology (pp. 143– 168). New York: Wiley-Liss.

Sohn, K. (2016). Now, the taller die earlier: The curse of cancer. *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences*, 71(6), 713– 719. https://doi.org/10.1093/gerona/glv065

Spake, L., & Cardoso, H. (2019a). Is there mortality/survival bias in the timing of dental formation? Annual Meeting of the Human Biology Association. Cleveland, OH: Presented at the Human Biology Association.

Spake, L., & Cardoso, H. F. V. (2018). Are we using the appropriate reference samples to develop juvenile age estimation methods based on bone size? An exploration of growth differences between average children and those who become victims of homicide. *Forensic Science International*, 282, 1– 12. https://doi.org/10.1016/j.forsciint.2017.10.041

Spake, L., & Cardoso, H. F. V. (2019b). Indirect evidence for biological mortality bias in growth from two temporo-spatially distant samples of children. *Anthropologischer Anzeiger; Bericht Uber Die Biologisch-Anthropologische Literatur.*, 76, 379– 390. https://doi.org/10.1127/anthranz/2019/0957

Spencer, N. (2004). The effect of income inequality and macro-level social policy on infant mortality and low birthweight in developed countries—A preliminary systematic review. *Child: Care, Health and Development*, 30(6), 699– 709. https://doi.org/10.1111/j.1365-2214.2004.00485.x

Spradley, M. K. (2014). Toward estimating geographic origin of migrant remains along the United States–Mexico border. *Annals of Anthropological Practice*, 38(1), 101– 110. https://doi.org/10.1111/napa.12045

Steckel, R. H. (2005). Young adult mortality following severe physiological stress in childhood: Skeletal evidence. *Economics & Human Biology*, 3(2), 314– 328. https://doi.org/10.1016/j.ehb.2005.05.006

Steenland, K., Hu, S., & Walker, J. (2004). All-cause and cause-specific mortality by socioeconomic status among employed persons in 27 US States, 1984–1997. *American Journal of Public Health*, 94(6), 1037– 1042.

Stock, M. K., Garvin, H. M., Corron, L. K., Hulse, C. N., Cirillo, L. E., Klales, A. R., … Stull, K. E. (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, 110232. https://doi.org/10.1016/j.forsciint.2020.110232

Stock, M. K., Stull, K. E., Garvin, H. M., & Klales, A. R. (2016). Development of modern human subadult age and sex estimation standards using multi-slice computed tomography images from medical examiner's offices. *Proc. SPIE*, 9967, 99670E. https://doi.org/10.1117/12.2237180

Stull, K. E., L'abbé, E. N., & Ousley, S. D. (2014). Using multivariate adaptive regression splines to estimate subadult age from diaphyseal dimensions. *American Journal of Physical Anthropology.*, 154, 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, 780– 786. https://doi.org/10.1002/ca.22225

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

Treves, S. T., Baker, A., Fahey, F. H., Cao, X., Davis, R. T., Drubach, L. A., … Zukotynski, K. (2011). Nuclear medicine in the first year of life. *Journal of Nuclear Medicine*, 52(6), 905– 925. https://doi.org/10.2967/jnumed.110.084202

United Nations Development Programme. (2018). Human Development Indices and Indicators: *2018 Statistica Update* [Human Development Index (HDI)].

Witt, W., Weiss, A., & Elixhauser, A. (2016). Overview of Hospital Stays for Children in the United States, 2012 #187 (no. statistical brief #187). Agency for Healthcare Research and Quality, Rockville, MD: Healthcare Cost and Utilization Project (HCUP). Retrieved from Healthcare Cost and Utilization Project (HCUP) website: www.hcup-us.ahrq.gov/reports/statbriefs/sb187-Hospital-Stays-Children-2012.jsp

Wood, B. (1992). Origin and evolution of the genus homo. *Nature*, 355, 783– 790.

Wood, J., Milner, G., Harpending, H., & Weiss, K. (1992). The Osteological paradox: Problems of inferring prehistoric health from skeletal samples. *Current Anthropology*, 33(4), 343– 370.

World Bank, Development Research Group. (2019). *World Development Indicators*. Retrieved from https://datacatalog.worldbank.org/public-licenses#cc-by

Yaussy, S. L., & DeWitte, S. N. (2018). Patterns of frailty in non-adults from medieval London. *International Journal of Paleopathology*, 22, 1– 7. https://doi.org/10.1016/j.ijpp.2018.03.008