A multifaceted exploration of ontogenetic variation in vertebral neural canal size across contemporary populations

Louise K. Corron[1,*], Christopher A. Wolfe[1], Kyra E. Stull[1,2]

[1]Department of Anthropology, University of Nevada, Reno, Nevada, USA

[2]Department of Anatomy, Faculty of Health Sciences, University of Pretoria, Pretoria, South Africa

*Correspondence: Louise Corron, Department of Anthropology, University of Nevada, Reno, 1664 N Virginia St MS 0096, Reno, NV 89557, USA. Email: lcorron@unr.edu

Funding information: National Institute of Justice, Grant/Award Numbers: NIJ 2015-DN-BX-K409, NIJ 2017-DN-BX-0144; National Science Foundation, Grant/Award Number: BCS-1551913

Abstract

Objectives: Vertebral neural canal (VNC) dimensions are considered a reliable indicator of childhood stress. However, no study has characterized variation in VNC size or shape or the impact of extrinsic or intrinsic factors on their range of variation. The present study explores VNC dimensions of subadult samples varying in chronology, population of origin, geography, and socioeconomic backgrounds.

Materials and Methods: Antero-posterior (AP) and transverse (TR) diameters were measured on the tenth thoracic to the fifth lumbar vertebrae of 1404 contemporary individuals aged between birth and 22 years from Colombia ($N = 28$), France ($N = 484$), the Netherlands ($N = 23$), Taiwan ($N = 31$), and the United States ($N = 838$), and compared to lumbar diameters of subadults from the Spitalfields collection ($N = 84$) and the East Smithfield cemetery ($N = 65$). VNC variation was evaluated with skeletal growth profiles, principal component analyses (PCA), MANOVAs and ANOVAs.

Results: AP diameter growth ends during childhood, while TR diameter growth progressively slows before ending in adolescence. The Colombian sample presented the smallest VNC diameters compared to the other contemporary and historic samples. VNC shape (AP/TR ratio) was similar in contemporary samples. MANOVAs and ANOVAs revealed significant differences in VNC size according to country of origin and socio-economic status, primarily differentiating the Colombian sample.

Discussion: The overall consistency in size and shape among groups is remarkable. While physiological stress may contribute to variability in VNC size, intrinsic ontogenetic processes and other individual and environmental factors also influence variability in VNC size.

Keywords: allometry, childhood stress, developmental plasticity, growth profiles, socio-economic status

1 INTRODUCTION

Non-specific indicators of stress in skeletal samples are often interpreted as evidence of biological growth disruption and/or physiological stress during childhood. These indicators

manifest across human dental or skeletal tissue through various markers, such as linear enamel hypoplasia (LEH), porotic hyperostosis (PH), cribra orbitalia (CO), small-for-age long bone lengths or stature, or small vertebral neural canal (VNC) size (Clark, 1988; Clark et al., 1985; Clark et al., 1986; Klaus, 2014; Klaus & Tam, 2009; O'Donnell et al., 2020; Saunders & Hoppa, 1993; Steckel et al., 2002; Temple et al., 2014; Walker et al., 2009). The magnitude of biological stress manifested by the degree of expression of these skeletal markers is linked to its severity and duration (Bogin, 1997; Hoppa & Fitzgerald, 1999; Konigsberg & Holman, 1999; Watts, 2011). However, the physiological and physical impacts of the stress event are unique to each individual (i.e., hidden heterogeneity), and this subsequently factors into differential expression for non-specific skeletal markers of stress (De Witte & Woods, 2008; Saunders & Hoppa, 1993; Vercellotti et al., 2014; Watts, 2013; Wood et al., 1992). The current study specifically explores the possible expression of physiological stress in growth, size, and shape of the VNC in a diverse sample of subadult individuals.

A wide variety of additional factors can impact an individual's response to stress, and their recovery capacity from said stress. These include nutrition, genetics, and socio-economic factors, among others (Steckel, 2012). The impact of such factors during early development has led to the Developmental Origins of Health and Disease (DOHaD) hypothesis, according to which exposure to environmental insults in early life during critical periods of development can contribute to short-term and long-term negative health outcomes (Barker, 2007). These intrinsic and extrinsic individual-level factors make recognizing the presence and extent of growth disruptions and, by extension, past episodes of stress, difficult. Beyond individual-level health, non-specific indicators are used to interpret population-level demographic factors, such as morbidity and mortality (e.g., Amoroso et al., 2014; Armelagos et al., 2009; De Witte & Woods, 2008; Saunders & Hoppa, 1993; Spake & Cardoso, 2019). However, these relationships have been questioned by several authors (De Witte & Stojanowski, 2015; Saunders & Hoppa, 1993; Wood et al., 1992) because of the lack of information regarding their variability in expression, their frequency in past and present populations, and the effects of secular trends. The same authors advised caution before interpreting them as indicators of biological stress without documented evidence of stress or disease and in the absence of reliable life history information, as is often the case in bioarcheological samples.

A few recent studies have cast some light on the etiology of non-specific skeletal indicators of stress by analyzing the degree of expression of these markers in contemporary reference samples in relation to other documented individual parameters. Beatrice and Soler (2016) and Beatrice et al. (2021) compared the prevalence of PH, CO, and LEH in adult American forensic samples, documented migrants, and undocumented migrants from Central America to examine their association with sex, age, and location. They found higher prevalence of PH compared to CO and LEH, especially in undocumented migrants, and conjectured that these findings could be linked to the individuals having experienced biological stress during childhood, although this could not be corroborated by verified demographic information or the individuals' life histories (Beatrice et al., 2021; Beatrice & Soler, 2016). To understand the impact of physiological stress on individual human skeletons, we must first document the range of variation present in the osteological response (e.g., LEH, PH, or CO) or the range of variation in size (e.g., long-bone-length/stature, or VNC size) across stressed and unstressed individuals or samples. Further, it is particularly important to document such variation during the potential period of onset of the stress marker—that is, when an individual is still developing—to gain insight into the full extent of variation. Unfortunately, to the authors' knowledge, there are few to no studies documenting the extent of variation in VNC size or shape in samples from diverse genetic and/or environmental contexts. The current study addresses this gap and documents

patterns of ontogenetic variation and the influence of intrinsic and extrinsic factors on the size and shape of the VNC in order to ascertain if VNC dimensions may indeed represent an appropriate metric of non-specific stress in the human skeleton.

Contrary to adult individuals who have achieved their final mature form, immature organisms—i.e., non-adults or subadults—have the ability to adapt their energetic needs and ontogenetic processes in response to environmental conditions because of developmental plasticity (Agarwal, 2016; Lasker, 1969; Roberts, 1995). Indeed, the window of highest activity for developmental plasticity ranges from conception to early childhood (Bogin, 1997), although some authors say it can even extend to the period of transition from juvenility to adolescence (Hochberg, 2012). McPherson's (2021) recent review explored the role of timing in the interpretation of non-specific skeletal markers of stress during ontogeny. She highlighted the importance of specific developmental periods called sensitive developmental windows (SDWs), during which an element of the skeleton is more highly susceptible to stress and more likely to exhibit developmental plasticity in response, resulting in a temporary or permanent phenotypic expression of stress measured by non-specific skeletal indicators. McPherson argues that "[…] *integrating the SDW concept into existing developmental origins of health and disease (DOHaD) approaches in skeletal biology will enhance our ability to interpret both the patterning of stress biomarkers—the artifacts of plastic responses to environmental signals—and their relationship to phenotypic development over the life course*." (McPherson, 2021, p. 165). We argue that a comprehensive appreciation of the range of variability in the expression of these markers across ontogeny is another valuable tool to understand and interpret the origin and development of skeletal biomarkers of stress.

In one of the few studies looking at non-specific skeletal markers of stress in contemporary subadult individuals (a subsample of individuals aged between birth and 15 years from the New Mexico Decedent Image Database/NMDID), O'Donnell et al. (2020) found that individuals with a cause of death likely linked to chronic respiratory infection had higher rates of CO and PH lesions, while individuals with chronic or congenital heart conditions showed significantly higher rates of CO but not PH. CO and PH frequencies in that particular medico-legal sample were comparable to those observed in the bioarcheological record, were not correlated with age, and were higher in males than females. The documentation associated with the individuals' health and death records indicated that the individuals likely had to be sick for several weeks at least to develop detectable CO or PH. However, no mortality bias according to sex or age was established for the prevalence of either lesion (O'Donnell et al., 2022). O'Donnell and colleagues bring invaluable information on the prevalence and etiology of CO and PH, which was previously lacking to validate the use of these lesions as non-specific skeletal markers of childhood stress or disease in both contemporary and past populations.

Studies looking at understanding the underlying causes, covariates, and general patterns of expression of non-specific skeletal markers of stress that are reflective of growth disruptions, such as small VNC size, are scarce (Clark, 1985; Clark et al., 1985; Clark et al., 1986; McPherson, 2021; Steckel et al., 2002). VNC size is typically evaluated by its antero-posterior (AP) and transverse (TR) diameters (Figure 1). Vertebral growth begins around 12 weeks of gestation with the formation of the centrum and two halves of the neural arch. The endochondral fusion of the two halves of the neural arch occurs during the first two years of postnatal life, followed by fusion of the neural arch to the centrum/vertebral body at the neurocentral synchondrosis between the ages of three and five years. Fusion of the two halves of the neural arch initiates in the lower thoracic and upper lumbar regions in the latter part of the first postnatal year and progresses cranially and caudally so that the cervical arches and the

lowest lumbar arches fuse between the second and fifth year. The pattern for neurocentral fusion differs; it starts in the lumbar region, followed closely by the cervical region, and finally the thoracic region (Baker et al., 2005; Reichmann & Lewin, 1971; Scheuer & Black, 2004). Differential growth and development patterns have been observed in timing and fusion sequences of the vertebrae (Scheuer & Black, 2004), but only a small amount (5%) of variation has been recorded in the timing of complete fusion of the VNC for a given vertebra or vertebral segment (Baker et al., 2005; Newman & Gowland, 2015; Scheuer & Black, 2004), though some cases of unfused thoracic neurocentral synchondroses in adolescents up to 17 years of age have been recorded (Blakemore et al., 2018).

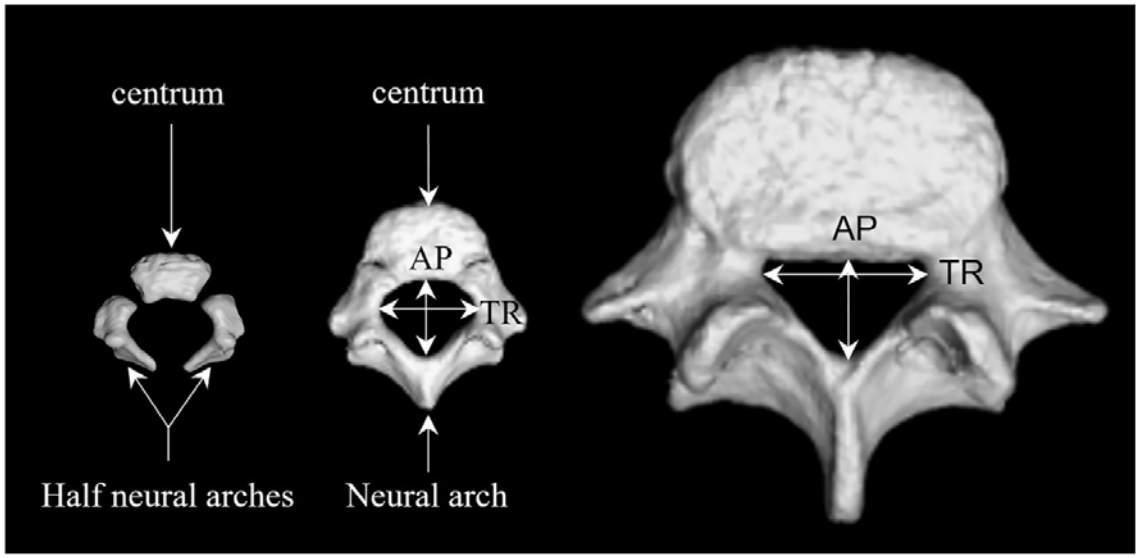

FIGURE 1. Three vertebral ossification centers of the fifth lumbar vertebra (L5): Separate vertebral body or centrum and two half-neural arches present at birth (left), fused neural arch and fusing centrum of a one-year-old individual (center), and fused vertebra of a 14-year-old individual (right). AP, antero-posterior diameter; TR, transverse diameter

Because of these specific ontogenetic patterns and the relative short duration of VNC growth, small VNC size can only appear as a result of an episode of stress spanning infancy and early childhood, when growth is most active (Baker et al., 2005; Scheuer & Black, 2004). Indeed, McPherson and other authors have identified that VNCs grow within the earliest SDW compared to other skeletal markers, ranging from 6 weeks in utero to approximately 4 years of age. Further, McPherson also suggests that this is the period during which VNC exhibits its highest levels of developmental plasticity (McPherson, 2021). This roughly 5-year period of high sensitivity to stress corresponds to a period during which the VNC reaches between 70 and 90% of its adult size (Newman & Gowland, 2015; Papp et al., 1994; Watts, 2013). The clinical literature has evaluated small adult cervical, thoracic, and lumbar VNC diameters and morphology as diagnostic tools for pathologies, such as spinal stenosis or general back pain (Eisenstein, 1977; Legg & Gibbs, 1984; Rapała et al., 2009; Santiago et al., 2001). However, there is no consensus regarding the influence of covariates such as age, sex, or population of origin on VNC size variation, and limited documented evidence on the relationship between stress and VNC size, particularly in subadults. Most studies based on VNC diameters obtained from clinical or historical samples were often based on a limited number of adult vertebrae, individuals, or measurements, resulting in inconsistent results and interpretations of small VNC size in relation to age, sex, disease, morbidity, or mortality (Aly & Amin, 2013; Amonoo-Kuofi, 1982; Amoroso & Garcia, 2018; Griffith et al., 2016; Hermann et al., 1993; Janjua &

Muhammad, 1989; Piera et al., 1988; Pierro et al., 2017; Postacchini et al., 1983; Tacar et al., 2003; Twomey & Taylor, 1988). These studies have also shown that all vertebrae can potentially express small VNC size following episodes of stress, but that the effects are much more pronounced on thoracic (especially lower thoracic) and lumbar segments compared to upper thoracic and cervical elements, likely because of the sequential fusion patterns of the two half-neural arches at the spinous process and of the neural arch with the neurocentrum (Newman & Gowland, 2015; Watts, 2013, 2015). Consequently, AP and TR diameters present with slightly different growth patterns (Clark, 1988; Reichmann & Lewin, 1971; Watts, 2013). Based on these observations, authors theorized that the timing of stress events could be inferred from the size of each VNC diameter and interpreted that small AP diameters were an indication that stress occurred during infancy or early childhood, whereas a small TR diameter indicated that stress occurred during late childhood or even early adolescence (Hinck et al., 1966; Watts, 2013, 2015).

Most of the studies above highlight the existence of variation in VNC size between vertebral regions (cervical, thoracic, or lumbar) and VNC dimensions (AP or TR diameters). However, the lack of data on the range of natural variation for VNC size limits our ability to discern the true etiology of small VNC size. Moreover, studies rarely mention processes such as developmental plasticity, growth canalization, or allometry, as modulating intrinsic responses to biological stress. As a result, we need more insight into the manifestation of phenotypic plasticity to confirm whether "small" VNC size is indeed a reliable skeletal marker of early growth disruptions possibly caused by episodes of biological stress (Galbusera, 2018; Steckel, 2012; Temple, 2018; Wells, 2016). The goal of the present study is to provide such insight by conducting a large-scale quantification of VNC variability across ontogeny in a large, varied, and contemporary reference sample of subadults. The specific objectives of the present study are three-fold: (1) evaluate the range of variation of VNC size in contemporary individuals aged between birth and 22 years representing varied population, geographical, and socio-economic backgrounds; (2) expose any patterns to the variation in terms of the covariates (i.e., socioeconomic status and population history/population of origin); and (3) interpret these patterns of variation in light of our knowledge on developmental plasticity, canalization, and allometry.

2 MATERIAL AND METHODS

2.1 Contemporary samples

A large sample of individuals ($n = 1404$) between the ages of birth and 22 years from Colombia, France, the Netherlands, Taiwan, and the United States (Tables 1 and 2, Figure 2). The individuals comprise the Subadult Virtual Anthropology Database (SVAD), a freely available repository of contemporary reference data obtained from anonymized computed tomography (CT) scans, radiographic images, Lodox Statscan, or dry bones of either living or deceased male and female individuals aged between birth and 22 years from eight different countries (Stull & Corron, 2022). Previous research exposed similar skeletal and dental growth and development patterns in living and deceased individuals; though the explorations were not specific to VNC growth, the authors feel confident the samples can be compared (Stull et al., 2021).

TABLE 1. Demographics and descriptive socio-economic indicators for the five SVAD samples, the Spitalfields sample, and the East Smithfield cemetery sample

Population (city)	Modality	Sample size	Age range	HDI[a]		Gini[b]	
				Raw	Category	Raw	Category
Colombia (Medellin)	Dry bone (deceased sample)	28 (5 F, 23 M)	0–22 years	0.766	High	49.7	High
France (Marseille)	CT scan (living sample)	484 (215 F, 269 M)	0–15 years	0.890	Very High	32.7	Low/Very Low
Netherlands (Amsterdam)	CT scan (living sample)	23 (11 F, 12 M)	0–15 years	0.949		28.2	
Taiwan (Taipei)	CT scan (living sample)	31 (13 F, 18 M)	0–16 years	0.882		33.8	
United States (Albuquerque, NM)	CT scan (deceased sample)	838 (354 F, 484 M)	0–20 years	0.920		41.5	Medium
Spitalfields collection (London)	Dry bone (deceased sample)	84 (sex ratio unreported)	1–16 years	–	–	–	–
East Smithfield Black Death cemetery (London)	Dry bone (deceased sample)	65 (sex unknown)	3–17 years	–	–	–	–

[a] Data obtained from the United Nations Development Programme, 2019 report, 2018 numbers.

[b] World Bank estimates, 2015–2018.

TABLE 2. Number of individuals with VNC measurements in the contemporary samples by age group and country of origin

Age group	Country of origin/sample	Number of individuals/measurements
0–2.9 years	Colombia	3
	France	26
	The Netherlands	0
	Taiwan	1
	U.S.	120
3.0–5.9 years	Colombia	3
	France	51
	The Netherlands	1
	Taiwan	0
	U.S.	57
6.0–10.9 years	Colombia	3
	France	125
	The Netherlands	3
	Taiwan	2
	U.S.	46
11.0–14.9 years	Colombia	2
	France	123
	The Netherlands	6
	Taiwan	7
	U.S.	91
15.0–17.9 years	Colombia	4
	France	49
	The Netherlands	0

	Taiwan	3
	U.S.	141
18.0–22.0 years	Colombia	2
	France	6
	The Netherlands	0
	Taiwan	0
	U.S.	206

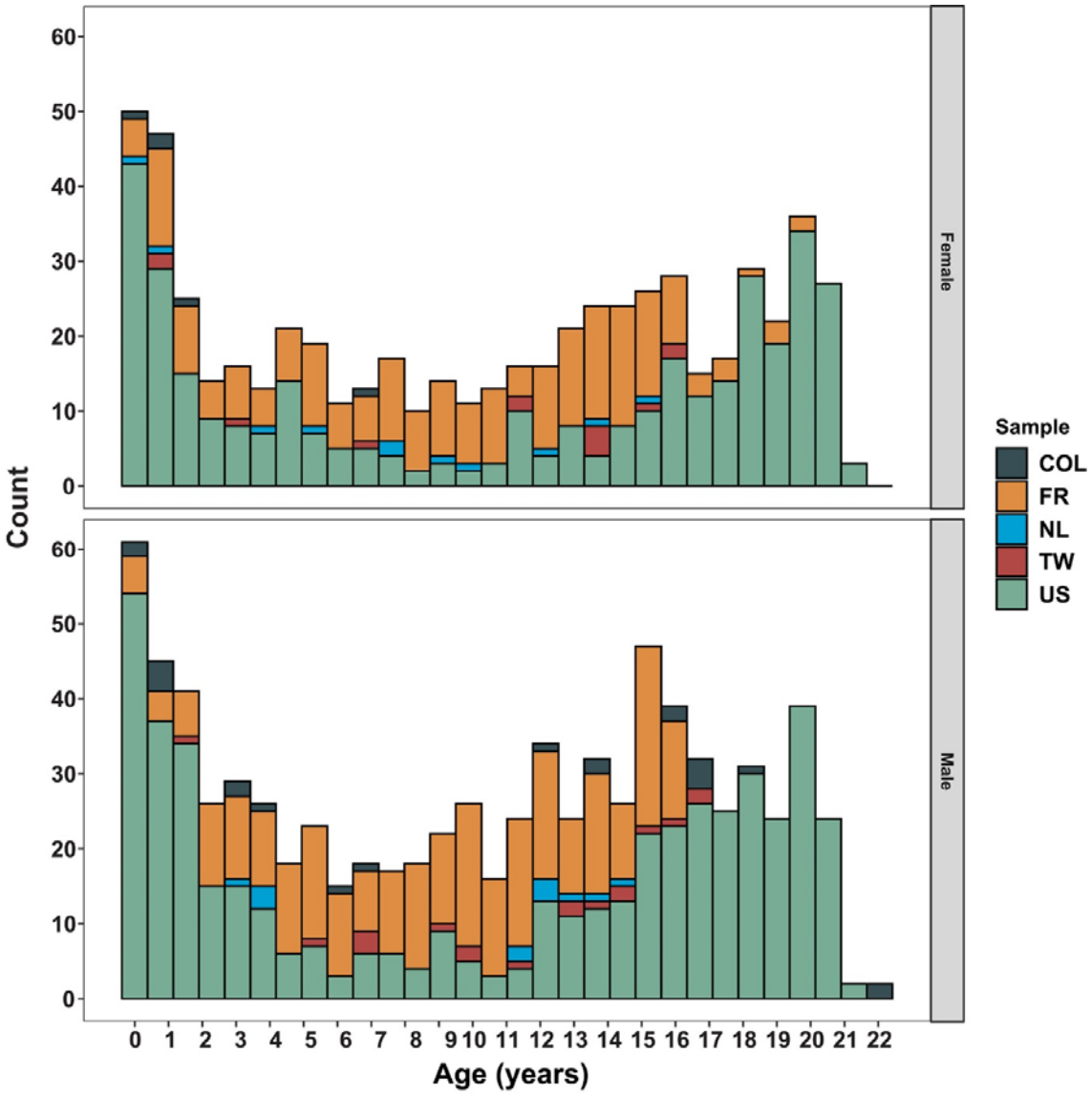

FIGURE 2. Age and sex distributions of the five contemporary SVAD samples. COL, Colombia; FR, France; NL, Netherlands; TW, Taiwan; US, United States

CT scans were obtained from public hospitals (France: Assistance Publique des Hopitaux de Marseille, the Netherlands: Amsterdam Medical Center, and Taiwan: National Taiwan University Hospital) and from a medical examiner's office (United States—Office of the Chief Medical examiner, University of New Mexico Health Sciences Center, now a part of the NMDID repository), while the Colombian sample consisted in a forensic skeletal sample housed at the Universidad de Antioquia, in Medellin. All samples were retrospectively assembled for the creation of SVAD in respect of each partner institution's ethical standards

and research practices (Stull & Corron, 2022). Dates of birth for all individuals ranged between the mid-to-late 1980's and early 2010's, while dates of death for deceased individuals (samples from the U.S. and Colombia) and dates of acquisition of the medical images for living individuals (samples from France, the Netherlands, and Taiwan) ranged between the late 1990's and 2018. The SVAD samples are meant to be reflective of the range of variation in growth and developmental patterns for contemporary subadults in various countries though there are sampling biases associated with each type of collaborating institution (Stull et al., 2021; Stull & Corron, 2022). Pathologies that could affect growth and development were not always controlled for in the deceased sample from the U.S. as they were often related to a "natural" manner of death, especially for infants (Stull & Corron, 2022). However, previous research on that same sample showed that manner of death ("natural," "homicide," "accident," "suicide," or "undetermined") did not significantly influence long bone dimensions or dental development for individuals older than two (Stull et al., 2021). For the current study, individuals with vertebral trauma or anomalies such as supernumerary vertebrae, spina bifida, or spondylolysis were excluded from the analyses, as these might affect the shape of the vertebrae examined in the present study (Barnes, 1994; Waldron, 2021). While no prenatal individuals are included in this study, we recognize their inclusion would more appropriately cover the entire period of vertebral development (Bagnall et al., 1977a, 1977b; Fazekas & Kosa, 1978). However, it is extremely difficult—both ethically and practically—to obtain contemporary skeletal data from prenatal individuals (Scheuer & Black, 2000). With this limitation aside, the SVAD samples include very young infants (less than 6 months of age), children, and adolescents and are therfore considered adequate in capturing crucial periods of sensitivity for VNC growth.

2.2 Comparative historic samples

Data from two previous publications examining VNC dimensions of historic samples were used as comparative material (Papp et al., 1994; Watts, 2013). The first sample consists of 65 children excavated from the medieval East Smithfield Black Death cemetery in London (Watts, 2013). These individuals died very suddenly and quickly (within a 2-year period) from the Plague (1348–1350), thus potentially capturing the range of normal and pathological variation in VNC size for that population. Some selection linked to pre-existing frailty was probable, although not as strong as natural mortality (De Witte & Woods, 2008).

The second historic sample consists of 84 children aged 1 year to 16 years from the Spitalfields collection (Papp et al., 1994). This sample includes individuals of known age and sex who lived in London between the mid-17th and mid-19th centuries. Molleson and Cox (1993) suggested some individuals suffered from acute or chronic episodes of stress during their lifetime based on short-for-age long bones and the prevalence of dental and skeletal lesions for both children and adults in the sample. The individuals who presented evidence of pathologies on the vertebrae were excluded from the original study (see Papp et al., 1994) and are therefore not included herein.

2.3 VNC measurements

AP and TR VNC diameters of the 10th thoracic (Th10) vertebra to the fifth lumbar (L5) vertebra were measured for individuals in all five SVAD samples following a standardized protocol (Figure 1) (Stull & Corron, 2021). Diameters were collected from virtual surfaces reconstructed from computed tomography (CT) scans and dry skeletal remains. Vertebrae were virtually reconstructed using the *Volume Rendering* tool of the Amira® software with a threshold of around 200 Hounsfield units to exclusively visualize bony elements. All

measurements on the CT samples were done with the Amira® software measuring tool and recorded with 0.01 mm precision following a standardized protocol with the superior face of the vertebral body oriented toward the observer, parallel to the computer screen (Stull & Corron, 2021, 2022). One of the authors (Louise K. Corron) measured the AP and TR diameters on the virtual surfaces of the individuals from the French, Taiwanese, and part of the U.S. samples, while another author (Christopher A. Wolfe) took VNC measurements of the individuals from the Netherlands and most of the U.S. sample.

VNC measurements for the dry bone Colombian sample were collected by one author (Louise K. Corron) using a digital sliding caliper and were recorded at the nearest 0.1 mm. AP diameters were only collected if the neural arch was actively fusing or completely fused to the vertebral body, which could be as early as a few months post-birth for both the CT samples and the dry bone sample. TR diameters were collected on completely fused neural arches, whether they were fused or unfused to the vertebral body, which could be done since birth in all samples, for the lumbar vertebrae in particular. AP/TR ratios were also calculated for each vertebra (Th10 to L5) in the contemporary samples as a proxy for vertebral shape (AP-TR ratio = AP diameter/TR diameter).

VNC data for the East Smithfield Black Death cemetery sample consisted in the mean values of the AP and TR diameters of lumbar vertebrae (L1-L5) for each of the four age groups referenced in the original publication by Watts (2013). Watts (2013) estimated age based on dental development and ossification/epiphyseal fusion and individuals were categorized into one of the following age categories: 3.0 to 5.9 years, 6.0 to 10.9 years, 11.0 to 14.9 years, and 15.0 to 17.0 years. Diameters were collected directly on the dry bones by Watts using a sliding caliper with 0.01 mm precision.

AP and TR diameters of all completely fused lumbar vertebrae (L1–L5) were collected by previous authors for the subadults in the Spitalfields sample (Papp et al., 1994). Data were collected from unmagnified, undistorted photographs of the "upper most aspect of the lumbar vertebral canals" (Papp et al., 1994, p 2770) using a digital measuring software with 0.01 mm precision. Only the mean values for AP and TR diameters of vertebrae L1 to L5 by biannual age category (1–2 years, 3–4 years, 5–6 years, 7–8 years, 9–10 years, 11–12 years, 13–14 years, 15–16 years, adult) were made available in the original publication and were used in the present study.

2.4 Modality type and observer errors

Several publications have assessed that size and shape of virtually reconstructed bones from CT scans using semi-automatic segmentation and volume reconstruction tools are comparable to their dry bone counterparts (Colman et al., 2017; Colman et al., 2019) and that measurements taken on CT scans are comparable to all measurements taken on the same dry bones (Brough et al., 2012; Corron et al., 2017; Hildebolt et al., 1990; Spake et al., 2020; Stull et al., 2014). Furthermore, a previous study of eight measurements obtained on virtual fifth lumbar vertebral body using the Amira® measuring tool found that intra-and inter-observer errors were on average less than 1 mm and were comparable to error rates obtained on their dry bone counterparts (Corron et al., 2017).

Two co-authors measured AP and TR diameters once (Christopher A. Wolfe) and twice (Louise K. Corron) on the Th10 to L5 vertebrae of a random subset of 10 individuals between the ages of birth and 18 years from the U.S. SVAD sample (80 AP diameters and 80 TR

diameters) to assess intra- and inter-observer errors. Absolute Technical Error of Measurement ($TEM = \sqrt{\sum D^2/2\,N}$, where D is the difference between repeated measurements and N is the number of individuals measured) and relative TEM (%TEM = TEM/mean × 100) (Perini et al., 2005; Ulijaszek & Kerr, 1999) were calculated.

2.5 Covariates

Socio-economic and biological descriptors, collectively recognized as covariates, were used to explore variation in the sample. These included human development index (HDI), Gini coefficient, population/country of origin, age group, and stature. The socio-economic descriptors of HDI and the Gini coefficient were selected to attempt to capture the multifaceted concept of 'environment' (Bogin et al., 2017; Grasgruber et al., 2016) and its influence on skeletal indicators. The HDI is a composite index based on life expectancy at birth, the ability to acquire knowledge measured by mean age of schooling and the number of years expected in school, and the gross national income per capita (Bogin et al., 2017; Grasgruber et al., 2016; Jahan, 2015). HDI categories are based on fixed cut-off points derived from the quartiles of distributions of component indicators: low (HDI less than 0.550), medium (HDI between 0.550 and 0.699), high (HDI between 0.700 and 0.799), and very high (HDI of 0.800 or greater) (Jahan, 2015); the higher the HDI, the higher the quality of life. Because there could be large disparities between country level and city level values, the HDI of the cities the samples originated from were recorded to better reflect local conditions (United Nations Development Programme and report, 2019). For the current samples, each city HDI value was within the same category as their respective country's HDI value. The Gini coefficient represents the level of social inequality within a population; its value quantifies the distribution of income in a population and specifically how it deviates from being perfectly equal (DataBank, 2020). Lower Gini values (closest to 0 and furthest from 100), indicate greater equality.

Population of origin or country of origin was based on where the collaborating institution was located (e.g., Colombia, France, the Netherlands, Taiwan, and the U.S.), with no knowledge of personal identity or nationality. Individuals were assigned to a particular age group according to their known chronological age following the thresholds associated with Bogin's biological life history stages (Bogin et al., 2007; Cameron & Bogin, 2012). This covariate was primarily used to mitigate the wide age range covered by the sample as well as the uneven age distribution. The age groups were defined as infancy (birth to 2.99 years), childhood (3.00–7.99 years), juvenile (8.00–11.99 years), and adolescence/early adulthood (12.00–22 years). While the authors would have preferred to use developmental milestones to partition individuals into particular life history stages, not all individuals had the required developmental data to do so. Therefore, maximizing variability and sample size outweighed developmentally determining life history stages.

Allometric relationships among skeletal elements could also mean that smaller VNC size could be linked to smaller overall stature (Bogin et al., 2017; Pfeiffer & Harrington, 2011). Health studies and official reports show that populations in Central and South America have smaller average adult height than European or North American populations and therefore this also was a covariate that needed to be considered (NCD Risk Factor Collaboration, 2016). Amoroso and Garcia (2018) published an approach that enables a reliable way of normalizing/scaling VNC data against a reference standard, in their case femoral length, to assess sample variation and estimate whether individuals deviate from a normal growth trajectory. Since individual femoral length was not available for all individuals, allometric relationships were evaluated using overall stature. In the U.S. sample, the AP and TR diameters were divided by the individual's

stature (in mm) as recorded in the autopsy report. However, since individual stature was unknown for the subjects from France, the Netherlands, Taiwan, and Colombia, allometric relationships were also evaluated by dividing the AP and TR diameters of individuals in all contemporary samples (Colombia, France, the Netherlands, Taiwan, and the U.S.) by the average adult male or female statures (in mm) reported in the NCD Risk Factor Collaboration (NCD Risk Factor Collaboration, 2016) according to sex and country of origin (Table 3).

TABLE 3. Average adult stature in the SVAD samples per the NCD Risk Factor Collaboration, 2016

Population of origin	Average adult male stature (cm)	Average adult female stature (cm)
Colombia	169.50	156.85
France	179.74	164.88
The Netherlands	182.53	168.72
Taiwan	174.52	161.45
U.S.	177.13	163.54

2.6 Statistical analyses

Several statistical approaches were implemented to explore the relationship between VNC diameters and covariates. VNC growth profiles were developed by plotting AP and TR diameters of all eight vertebrae (Th10 to L5) against age. Loess (locally estimated scatterplot smoothing) lines were used to visualize the relationship between VNC diameters and age for each sample. Loess is commonly used in non-parametric regression analyses as a strategy for fitting a smooth curve to noisy data values, sparse data points, weak interrelationships, or scatterplots showing local non-linear relationships (Cleveland, 1979; Fox & Weisberg, 2019). Relationships among VNC diameters for all seven vertebrae (Th10 to L5) in the five contemporary samples were also assessed using Pearson correlations and visualized with the "corrplot" package (Wei & Simko, 2021).

Because data from the historic samples were limited and based on different and sometimes inconsistent age categories (means and percentiles in four age categories for the East Smithfield Black Death cemetery and means in eight age categories for Spitalfields), both scatterplots and bar plots were used to examine and compare vertebral growth profiles to the contemporary data. Scatterplots of VNC diameters against age were visualized for the five contemporary samples (Figure 4). Barplots of the mean, minimum, and maximum of VNC diameters were visualized by age group (3.0–5.9 years, 6.0–10.9 years, 11.0–14.9 years, and 15.0–17.0 years) for the five contemporary samples and the East Smithfield Black Death sample, while the means for the Spitalfields sample age groups were only plotted when they corresponded to the East Smithfield Black Death sample (Figure 5) [Correction added on 27th Jan 2023, after first online publication: Textual corrections have been made.]. The discrepancies between age groups and the small sizes of age subsets across samples prevented the application of statistical tests to compare the means between age groups and samples.

Multivariate Analysis of Variance (MANOVA) was used to test the relationship between each sample (country of origin) and the vector of all AP and TR diameters. Afterwards, a univariate Analysis of Variance (ANOVA) was used to test if each single dimension (AP or TR diameter) differed between each sample. Tukey's Honestly Significant Difference (Tukey's HSD) was run after the ANOVA to evaluate whether statistically significant differences between each univariate VNC dimension existed according to country of origin.

Principal component analysis (PCA) was employed to better understand the combined VNC variation. The raw VNC data were used in the PCA and country of origin, HDI, Gini, sex, and age group were used as illustrative categorical variables. The percentage of missing data ranged from 2% (L5 AP diameter) to 18% (L2 TR diameter), with most of the missing data being around 10%. To avoid bias in the analyses, missing data were imputed using a robust sequential algorithm that did not require subsetting the sample (Todorov, 2020; Verboven et al., 2007) in the "rrcovNA" package (Todorov, 2020). Raw AP and TR variables, and AP/TR ratios were centered and scaled in the PCA to mitigate size differences associated with age and growth. PCA was completed using the "FactoMineR" package (Lê et al., 2008) and "factoextra" package (Kassambara & Mundt, 2020).

The PC scores from the PCA using the AP and TR diameters (i.e., "raw VNC data") were then compared to the PC scores obtained from the PCA done on the AP and TR diameters scaled by stature (i.e., the "scaled VNC data") using multi-group principal components analysis (MGPCA) (Eslami et al., 2013; Krzanowski, 1984). The scaled VNC data were the diameters following the division by either individual stature (U.S. sample) or by the average stature in a given population and sex per country (Taiwan, Netherlands, Colombia, and France samples). A matrix of common loadings between each group (i.e., raw vs. scaled) is derived from the eigen analysis of the within-groups variance–covariance matrix. The results include a similarity index ranging between 0 and 1, with 1 being perfect agreement between the common vectors of loading. In the current study, a similarity index approaching 1 would suggest that the pattern of VNC variation present is not the result of differences in overall body size. Groups that share a similar variance–covariance matrix structure appear almost transposed when compared visually.

All analyses were completed in the R programming environment (R Core Team, 2021).

3 RESULTS

The mean, standard deviation, minimum, and maximum of VNC AP and TR diameters are provided for each vertebra and sample in the Supporting Information S1. Absolute TEM values ranged from 0.02 mm (Th10 TR diameter) to 0.18 mm (L3 AP diameter) for intra-observer error and 0.04 mm (Th10 TR diameters) to 0.29 mm (L2 AP diameter) for inter-observer error. Relative TEM (%TEM) ranged from 0.22% (L2 TR diameter) to 1.10% (L3 AP diameter) for intra-observer error and from 0.17% (L5 TR diameter) to 1.77% (L2 AP and L3 TR diameters) for inter-observer error. %TEMs are well within the upper limit for acceptable measurement error set a 2% (Ulijaszek & Kerr, 1999).

3.1 Correlation coefficients of VNC diameters

Raw TR and AP diameters are all strongly correlated with one another, especially within the same vertebrae (Figure 3). Correlations are strongest between homologous variables (TR or AP diameters) and between adjacent vertebrae (e.g., Th10, Th11, and Th12). All correlation coefficients are consistently greater than or equal to 0.39 with TH10 AP diameter and L5 AP diameter exhibiting the weakest relationship (0.39). Correlations of the ratios follow a similar pattern, although the correlation coefficients are lower than those for the raw diameters ($0.27 < r < 0.78$, see Supporting Information S1).

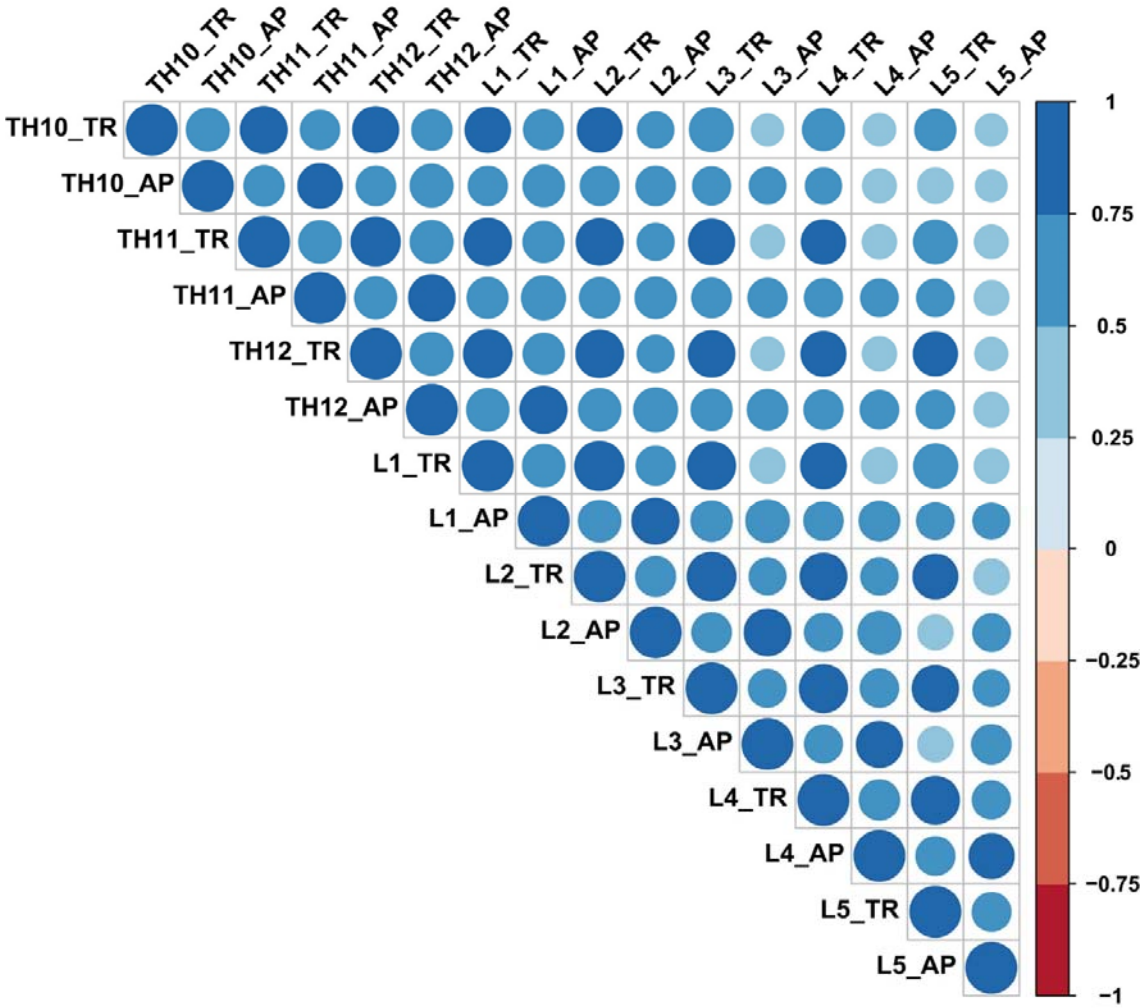

FIGURE 3. Correlation matrix of AP and TR diameters for all eight vertebrae (tenth thoracic/Th10 to fifth lumbar/L5). AP, antero-posterior diameter; TR, transverse diameter

3.2 VNC growth profiles and ranges of variation

Visualizations of the AP and TR diameters of all vertebrae plotted against age demonstrate the same trend: although slight population differences can be observed in growth profiles, individuals from France, the Netherlands, Taiwan, and the U.S. have comparable ranges of variation, subadults from the Spitalfields collection are on the lower end of contemporary ranges of variation, and the Colombian individuals appear to have substantially smaller dimensions than all the other contemporary and historic samples (Figures 4 and 5). Supporting Information S1 provides additional illustrations of the growth profiles for each VNC dimension by country of origin and summary statistics (mean, standard deviation, minimum and maximum values) for each measurement by annual age group and country of origin.

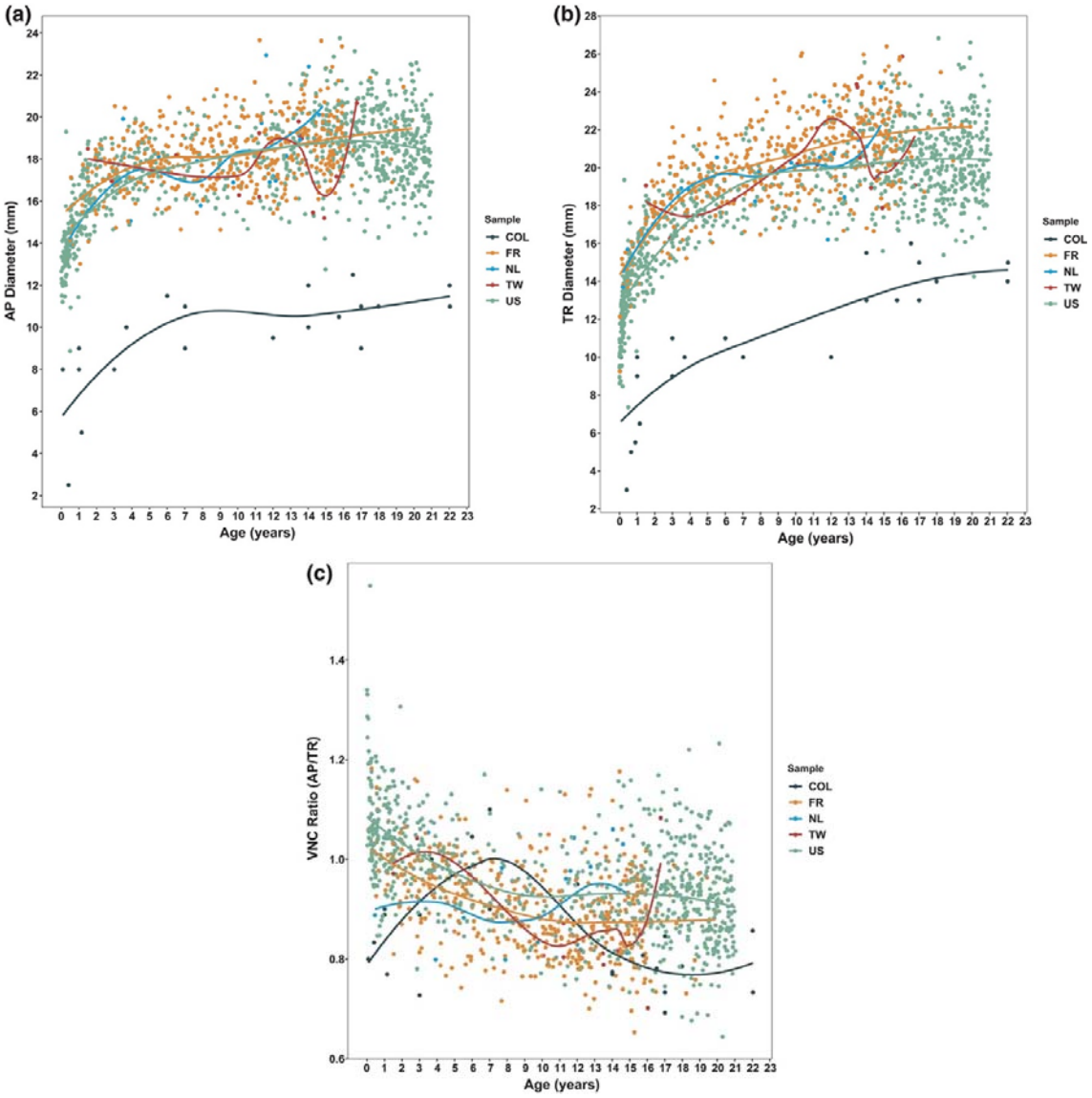

FIGURE 4. VNC dimensions plotted against age in the five contemporary samples—Example of the twelfth thoracic vertebra (Th12). (a) AP diameter against age; (b) TR diameter against age; (c) AP/TR ratio against age. AP, antero-posterior diameter; TR, transverse diameter; VNC, vertebral neural canal

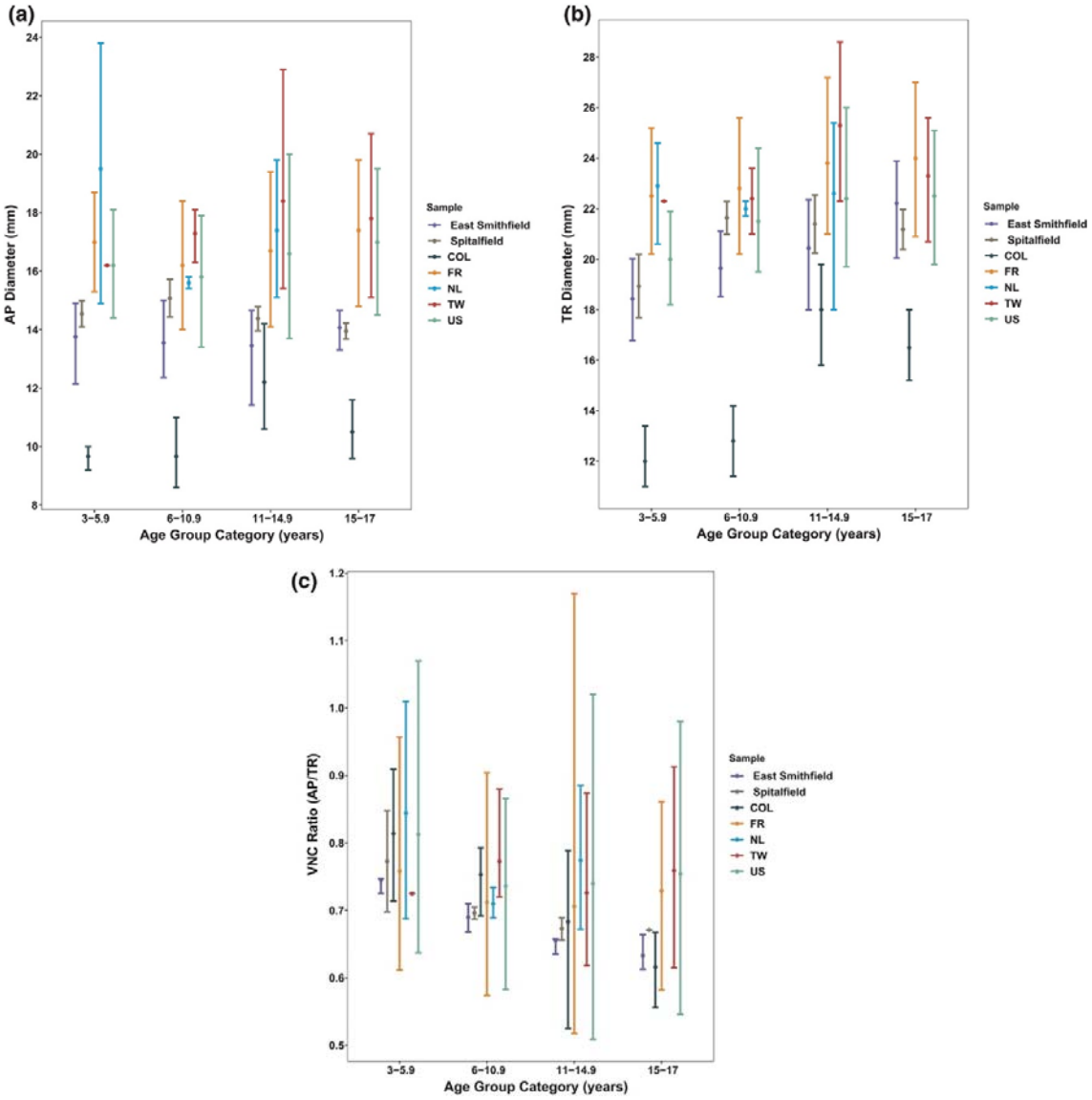

FIGURE 5. Ranges of variation (mean, minimum, and maximum) of VNC dimensions for the fourth lumbar vertebra (L4) plotted against age in the five contemporary SVAD samples and two historic samples from the Spitalfields collection and the East Smithfield Black death cemetery. (a) Plot of L4 AP diameter by age group; (b) Plot of L4 TR diameter by age group; (c) Plot of L4 AP/TR ratio by age group. AP, antero-posterior diameter; TR, transverse diameter; VNC, vertebral neural canal

The loess lines visualized on the growth profiles for the AP diameter display rapid growth in the first 2 to 3 years, captured by the steep change in VNC size, and active growth until the ages of 4 to 5 years, followed by a progressive decrease in the slope until early adolescence (12–13 years) along with a progressive increase of inter-individual variation in VNC size from early adolescence onwards (Figure 4a). TR diameter stabilizes later than AP diameter, with the steepest part of the slope for the TR growth profile covering the period between birth and ages 5–6 (Figure 4b), followed by a slower but ongoing increase in size until late adolescence (15–16 years) along with an increase in the range of variation. Population differences are more marked for TR diameters than for AP diameters; the loess line for the French sample is consistently above the line for the U.S. sample. Because the Taiwan and Netherlands samples are small ($n < 35$), it is unwise to infer specific trends in the growth profiles for these two

samples. Even with the small sample for the Netherlands, the consistency in VNC size in that sample is remarkable and is less variable than in the Taiwanese sample. Although the Colombian sample falls below the others in terms of its range of variation, growth profiles of both AP and TR diameters show similar trajectories to the other four contemporary samples (Figure 4).

When comparing ranges of variation across age groups—namely early childhood (3–5.9 years), late childhood (6–10.9 years), early adolescence (11–14.9 years), and late adolescence (15–17 years)—population differences in AP and TR diameters emerge more clearly (Figure 5): contemporary samples from France, Netherlands, Taiwan, and the U.S. share considerable overlap in their ranges of variation within and across all age groups for both AP and TR diameters. The AP and TR diameters present with more variability for the historic samples: there is considerable overlap between the contemporary populations and the Spitalfields and East Smithfield samples for TR diameters (Figure 5b), however the AP diameters for the Spitalfields and East Smithfield samples either fall below or within the lower range of contemporary variation for all age groups (Figure 5a). Remarkably, the range of variation for both AP and TR diameters in the Colombian sample falls well below the other samples, except for the individuals in the 11–14.9-year age group who present overlap with the lower end of variation of the other samples for the TR diameter and aligns with the range of variation of the East Smithfield cemetery sample for the AP diameter in that same age group.

Unlike VNC diameters, VNC ratios (AP/TR diameters) follow comparable patterns and ranges of variation (minimum = 0.6 and maximum = 1.4) across all samples throughout ontogeny (Figures 4c and 5c).

3.3 Influence of covariates on VNC size

MANOVA and ANOVA results corroborate the differences highlighted in the visualizations. That is, both combined (AP and TR) and single (AP or TR) vertebral measurements all show statistically significant differences between the samples. Specifically, the post-hoc Tukey's highlights Colombia as being different from the remainder of the contemporary samples for both AP and TR diameters. Mean differences ranged between 6.78 mm (L1 TR diameter, between Colombia and the U.S.) and 12.21 mm (L5 TR diameter, between Colombia and Taiwan) (see Supporting Information S1 for all mean differences and 95% confidence intervals). Significant differences were also noted between France and the U.S. for the AP and TR diameters of Th10, Th11, Th12 and L1 (after ANOVA and post-hoc Tukey) although differences were less than 1 mm for AP diameters and less than 3 mm for TR diameters (Table 4). The U.S. and Taiwan also exhibited significant differences for TR diameters of L3 to L5, with differences ranging between 1 mm and 6 mm (Tukey's HSD $p < 0.001$). In these cases, the U.S. VNC diameters were always larger than the French and Taiwanese VNC diameters.

TABLE 4. Results of the MANOVAs run for L2 and significant Tukey's HSD for AP and TR diameters for all vertebrae across the contemporary samples. Significant differences were also noted in the Tukey's HSD between Colombia and all other samples ($p < 0.001$) but they were not included in the table

MANOVA—example of L2					
Comparison	**Df**	**Sum of squares**	**Mean square**	**F**	***p*-value**
AP diameter					
Sample	4	1659.6	414.90	120.24	<2.2 e-16
Residuals	1302	4492.6	3.45		
TR diameter					
Sample	4	2362.2	590.55	80.733	<2.2 e-16
Residuals	1302	9523.9	7.31		
Significant differences ($p < 0.001$) between France, the Netherlands, Taiwan, and the United States samples					
Vertebrae		**VNC dimension**		**Populations, mean differences, (lower: upper) 95% confidence interval**	
Th10		AP diameter		US and France: −0.56 (−0.84: −0.27) mm	
		TR diameter		US and France: −1.63 (−2.01: −1.25) mm	
Th11		AP diameter		US and France: −0.59 (−0.90: −0.29) mm	
		TR diameter		US and France: −1.70 (−2.13: −1.27) mm	
Th12		AP diameter		US and France: −0.62 (−0.95: −0.29) mm	
		TR diameter		US and France: −1.97 (−2.46: −1.47) mm	
L1		AP diameter		US and France: −0.70 (−1.02: −0.38) mm	
		TR diameter		US and France: −2.08 (−2.56: −1.59) mm US and Taiwan: −2.55 (−4.70: −0.40) mm	
L2		AP diameter		N/A	
		TR diameter		US and France: −2.00 (−2.47: −1.53) mm US and Taiwan: −2.54 (−4.49: −0.57) mm	
L3		AP diameter		US and France: −0.43 (−0.73: −0.14) mm US and Taiwan: −1.63 mm (−2.75 mm: −0.51 mm)	
		TR diameter		US and France: −2.06 mm (−2.52 mm: −1.59 mm) US and Taiwan: −3.07 mm (−4.91 mm: −1.23 mm)	
L4		AP diameter		N/A	
		TR diameter		US and France: −2.18 mm (−2.70 mm: −1.66 mm) US and Taiwan-2.94 mm (−4.62 mm: −1.25 mm)	
L5		AP diameter		N/A	
		TR diameter		US and France: −2.76 mm (−3.42 mm: −2.11 mm) US and Taiwan: −3.84 mm (−5.88 mm: −1.81 mm)	

The PCA results of the centered and scaled raw VNC diameters show that PC1 and PC2 contribute to 74.5% and 8.9% of variation, respectively (Table 5). All loadings are positive along PC1, with TR diameters of Th11 to L4 contributing the most to the variation captured in the dimension. The contribution of the vertebrae is similar, ranging from 7% to 7.6%. AP

diameters of TH11, TH12, and L1 also contribute to PC1, but less than most of the TR diameters. The second PC captures differences between the AP diameters and TR diameters; all AP diameters have positive loadings and all TR diameters have negative loadings. AP diameters for L2 to L5 contribute the most to PC2, each ranging from 10.5% to 20.5%. Lumbar variables, especially L2–L4, tend to contribute more than thoracic variables for both PC1 and PC2 (Figure 6).

TABLE 5. Summary of PCA results based on VNC diameters and VNC ratios and the five illustrative covariates in the five contemporary samples

PCA	**Dim 1**	**Dim 2**	**Country of origin**	**Sex**	**Gini**	**HDI**	**Age groups**
AP and TR diameters	74.5%	8.9%	Colombia clearly separates from France, the Netherlands, Taiwan, and the U.S.	Overlap	High category separates from the Low and Medium categories	High category separates from the Very high category	General overlap with slight divergence of infancy
Variable contribution	Mostly TR	Mostly AP					
AP/TR ratios	68.09%	10.91%	Overlap	Overlap	Overlap	Overlap	Overlap
Vertebral contribution	Th11 to L4	Th10, Th11, L4 and L5					

Abbreviations: AP, antero-posterior diameter; HDI, human development index; L, lumbar; Th, thoracic; TR, transverse diameter.

Once the PCA was conducted the covariates were used in the visualizations to expose any obvious patterning in the data. Figure 6 reveals the Colombian sample clustering separately from other countries, similar to patterns exposed in Figures 4 and 5. There are no noticeable differences in variation according to biological sex as males and females overlap. Infants form a distinct cluster compared to children, juveniles, and adolescents. There are distinct clusters between the high and very high HDI categories, and between the high and medium/low Gini categories (Figure 6).

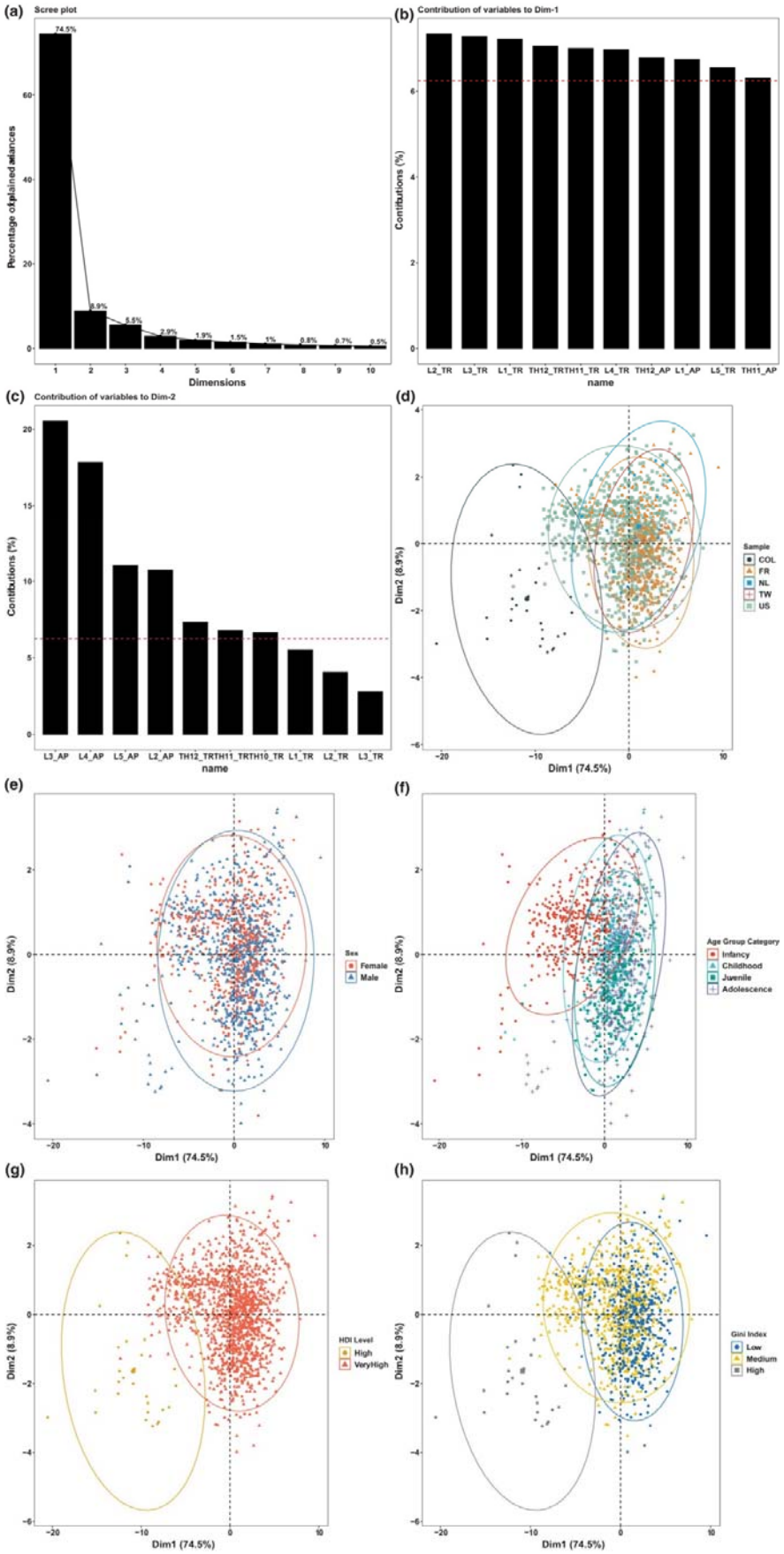
(a)
Scree plot
Percentage of explained variances
Dimensions
74.5%
8.9%
5.5%
2.9%
1.9%
1.5%
1%
0.8%
0.7%
0.5%
(b)
Contribution of variables to Dim-1
Contributions (%)
name
L2_TR
L3_TR
L1_TR
TH12_TR
TH11_TR
L4_TR
TH12_AP
L1_AP
L5_TR
TH11_AP
(c)
Contribution of variables to Dim-2
Contributions (%)
name
L3_AP
L4_AP
L5_AP
L2_AP
TH12_TR
TH11_TR
TH10_TR
L1_TR
L2_TR
L3_TR
(d)
Dim2 (8.9%)
Dim1 (74.5%)
Sample
COL
FR
NL
TW
US
(e)
Dim2 (8.9%)
Dim1 (74.5%)
Sex
Female
Male
(f)
Dim2 (8.9%)
Dim1 (74.5%)
Age Group Category
Infancy
Childhood
Juvenile
Adolescence
(g)
Dim2 (8.9%)
Dim1 (74.5%)
HDI Level
High
VeryHigh
(h)
Dim2 (8.9%)
Dim1 (74.5%)
Gini Index
Low
Medium
High

FIGURE 6. PC plots based on AP and TR diameters of VNC for vertebrae Th10 to L5. (a) scree plot of eigenvalues, (b) variables contributing to the 1st PC, (c) variables contributing to the 2nd PC, and plots of PC1 and PC2 grouping by (d) country of origin, (e) sex, (f) age group, (g) HDI, category, (h) Gini category. The dotted red line corresponds to the expected value under uniform contribtuions of the variables to PC1. AP, antero-posterior diameter; TR, transverse diameter; VNC, vertebral neural canal

The PCA using VNC ratios revealed that PC1 and PC2 account for 68.09% and 10.91% of variance, respectively (Figure 7, Table 5). When visualizing with the covariates, there is no clear separation of the samples according to sex, country of origin, HDI, or Gini categories. PC plots show overlap of all samples and categories. The only covariate with any type of distinctness was the age group variable. Similar to the raw diameters, infants are slightly separated compared to the other three age groups (Figure 7).

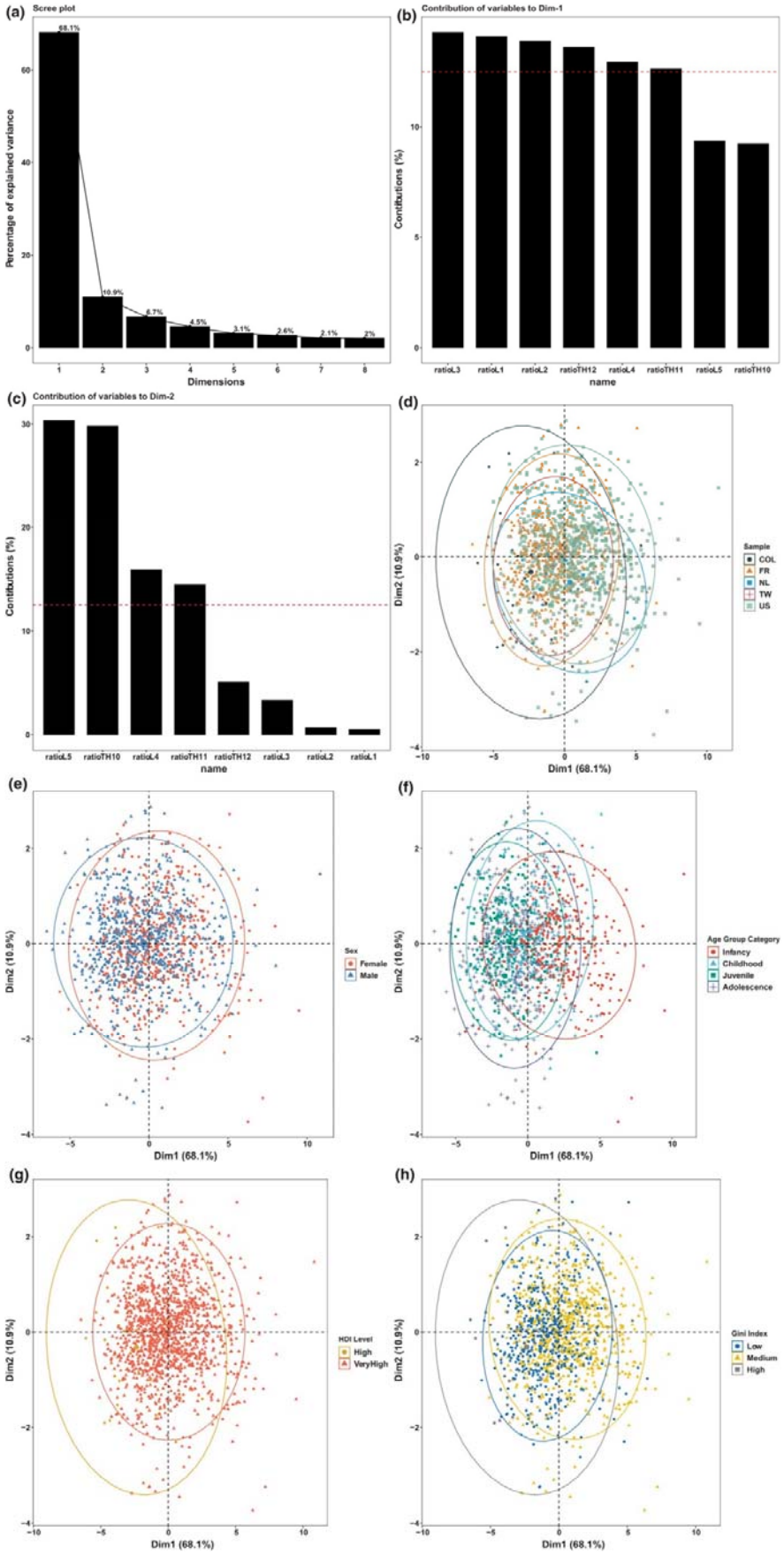
(a)
Scree plot
Percentage of explained variance
Dimensions
68.1%
10.9%
6.7%
4.5%
3.1%
2.6%
2.1%
2%
(b)
Contribution of variables to Dim-1
Contributions (%)
name
ratioL3
ratioL1
ratioL2
ratioTH12
ratioL4
ratioTH11
ratioL5
ratioTH10
(c)
Contribution of variables to Dim-2
Contributions (%)
name
ratioL5
ratioTH10
ratioL4
ratioTH11
ratioTH12
ratioL3
ratioL2
ratioL1
(d)
Dim2 (10.9%)
Dim1 (68.1%)
Sample
COL
FR
NL
TW
US
(e)
Dim2 (10.9%)
Dim1 (68.1%)
Sex
Female
Male
(f)
Dim2 (10.9%)
Dim1 (68.1%)
Age Group Category
Infancy
Childhood
Juvenile
Adolescence
(g)
Dim2 (10.9%)
Dim1 (68.1%)
HDI Level
High
VeryHigh
(h)
Dim2 (10.9%)
Dim1 (68.1%)
Gini Index
Low
Medium
High

FIGURE 7. PC plots based on VNC ratios of vertebrae Th10 to L5. (a) Scree plot of eigenvalues, (b) variables contributing to the 1st PC, (c) variables contributing to the 2nd PC, and plots of PC1 and PC2 grouping by (d) country of origin, (e) sex, (f) age group, (g) HDI category, (h) Gini category. HDI, human development index; VNC, vertebral neural canal

Results show that PC scores do not differ significantly between the raw VNC data and the VNC data after it was scaled using population average or individual stature (Figure 8). The index of similarity is equal to 1, suggesting perfect agreement between the raw and scaled metrics. These results highlight that the smaller VNC size in the Colombian sample does not appear to be informed by the inherently smaller stature presented by this sample.

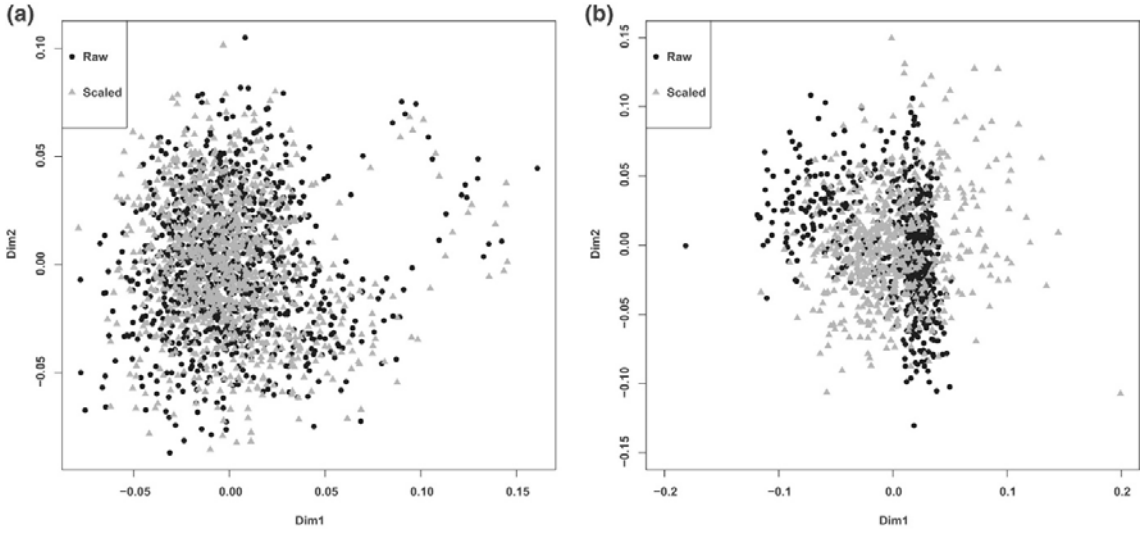

FIGURE 8. Comparison of principle component distributions between scaled (gray) and raw (black) VNC data. (a) Raw VNC diameters compared to scaled VNC diameters using average adult stature according to sex and country of origin. (b) Raw VNC diameters compared to scaled VNC diameters using individual stature according to sex in the U.S. sample. VNC, vertebral neural canal

4 DISCUSSION

Some colleagues (e.g., Agarwal, 2016; Temple, 2018; Vercellotti et al., 2014) have argued that assessing the "normal" range of variation or expression of skeletal markers in relation to contextualized biosocial parameters and conducting analyses based on life history approaches, appropriate statistics, and large, varied, and balanced documented skeletal samples are crucial to verify whether "traditional" non-specific skeletal indicators can indeed be considered informative of individual and population life histories. Following these guidelines and using a large, diverse (in time and geography) subadult reference sample revealed the overwhelming similarity among all groups in VNC shape (VNC ratios) regardless of time, living conditions, and geography and allowed us to quantify the rage of natural variation of VNC dimensions across ontogeny. VNC size (AP and TR diameters) displays more variability, but also remarkable consistency in growth trajectories across all groups. The TR diameter displays less differences in size among groups compared to the AP diameter. Differences between AP and TR diameters were also observed in the PC results, with each type of measurement equally contributing to the variation captured by the PCA, but TR diameters contributing more to PC1, while AP diameters contributed more to PC2. Both types of variables also presented with positive or negative loadings along PC2, respectively, with lumbar diameters contributing more

than thoracic diameters. Specifically, because the two dimensions exhibit differences in their growth trajectories, they likely present with different sensitive developmental windows during which they can potentially capture episodes of stress. The different patterns of growth and variation for these two types of measurements should therefore be considered in future research and conclusions. Indeed, prolonged TR diameter growth may capture stress events for a longer period of time compared to AP diameters, but also allow for possible catch-up growth until growth ceases much later in adolescence. This could also explain why the rnage of variation for TR diameters is greater than for AP diameters, and support the hypothesis that AP diameters capture early growth disruptions—during childhood—while TR diameters can capture later growth disruptions (Newman & Gowland, 2015; Wren, 2017).

4.1 Trends in ontogenetic variation of VNC size

Numerous studies have interpreted smaller VNC size in adults in past and present populations as a non-specific skeletal indicator of stress during early childhood (Clark et al., 1986; Rewekant, 2001; Watts, 2011; Watts, 2013) but have not always examined the extent of normal variation or secular trends in VNC growth profiles. To the authors' knowledge, no studies have analyzed VNC size and shape in subadults from historic and contemporary populations to effectively quantify the range of normal human variation and compare the range of dimensions. Figures 4 and 5 display overlap in the range of variation for AP and TR diameters for the five contemporary samples and the two historic samples—and the Spitalfields skeletal sample in particular—while the Colombian sample presented consistently smaller dimensions. The ranges for VNC size (both AP and TR diameters) and intra-vertebral associations of correlation between VNC dimensions across vertebrae observed in the contemporary samples from France, the Netherlands, Taiwan, and the U.S. after growth completion or once growth slows down (Figures 3 and 4) were comparable to those previously found by Ruhli on subadult and adult prehistoric and modern samples from Switzerland and Germany (Ruhli, 2003) and for the adult individuals from the Luis Lopes osteological collection (Amoroso & Garcia, 2018). Consistent with these previous findings, VNC ratios were comparable for the five contemporary samples (Figure 4c) regardless of country of origin. An additional result stemming from these comparisons and growth profiles is the confirmation that the modality used to obtain the measurements (skeletal material or CT scan) does not significantly influence the range of VNC values, as illustrated by the overlap in VNC size for contemporary CT samples and historic dry bone samples (Figure 5). Therefore, the differences in VNC size between samples are hihgly unlikely to be linked solely to selective mortality (deceased versus living) or to differences in data collection modality (dry bone versus virtual renditions of the vertebrae). While the Colombian population may seem to stand out compared to the other four contemporary samples, the incorporation of additiona samples from diverse locations, or sampling more individuals from Colombia could reveal the true continuum of growth and variation of VNC size.

Since 150 to 650 years separate the contemporary samples from the Spitalfields and the East Smithfield Black Death cemetery samples, respectively, secular trends in growth rates and overall size could explain the somewhat larger VNC size observed in the contemporary samples (Schizas et al., 2014). However, since both samples exhibit ranges of variation in TR diameters somewhat comparable to the contemporary samples from France, the Netherlands, Taiwan, and the U.S., but substantially smaller AP diameters (Figure 5), differences could also be because of biological stress encountered by some individuals from the East Smithfield Black Death cemetery during early childhood or even prior to birth. Indeed, differences only appeared significant for the AP diameter, whose growth ceases or slows significantly after 2 or 3 years

of age (Figure 5) (Griffith et al., 2016; LoPresto, 2020). Skeletal evidence of early childhood stress, including CO and LEH was recorded for subadult individuals in the East Smithfield sample, which are both more common non-specific indicators of stress (Grainger, Museum of London, & English Heritage, 2008). We cannot exclude that both factors (secular change and biological stress) contributed to the observed differences in AP diameters. However, it seems highly unlikely that all 65 individuals from the East Smithfield Black Death cemetery experienced early childhood stress at the same magnitude that led to reduced AP diameters. This seems even less likely if we consider the Colombian sample compared to all the other samples (historic and contemporary) for both AP and TR diameters (Figures 4 and 5). The variety of analytical techniques in the current study showed that the Colombian sample differed from the other four contemporary samples across all AP and TR diameters for vertebrae Th10 to L5 and were often distributed along the lower range of variation or even lower than that of the two historic samples (Figure 5). This trend is present in the youngest individuals and persists throughout ontogeny, though the magnitude of the differences in TR diameters decreases as age increases (Figure 4a,b). Results indicate that VNC size was likely already smaller in utero. In contrast, VNC shape was similar across ontogeny for all contemporary samples, meaning that the overall morphological integrity of the VNC is preserved for these samples.

4.2 Disentangling the factors behind small VNC size

While the Colombian sample is quite small ($n = 28$), the magnitude of VNC size difference as compared to other components of the sample is notable. According to contemporary research related to VNC size and by comparing the Colombian sample to the other four contemporary samples of the present study, a natural conclusion of such results would be that Colombian subadults may have been subject to external stressors during early life that impacted VNC growth and final adult size (Clark, 1988; Clark et al., 1986; Klaus, 2014; Klaus & Tam, 2009; O'Donnell et al., 2020; Saunders & Hoppa, 1993; Steckel et al., 2002; Temple et al., 2014; Walker et al., 2009). However, this conclusion ignores several factors related to sample size and composition, known covariates, environmental backgrounds, and individual versus population heterogeneity. Specifically, the Colombian sample is composed of deceased individuals, it is the only sample representing one of two HDI categories, and one of the three Gini categories (Table 1). Further, the sample presents a unique population history intertwined with the legacy of colonialism across South America. Each of these factors (selective mortality/hidden heterogeneity, body size, HDI/Gini and environmental background, and population of origin/genetic background) are explored to infer any reasoning as to *why* VNC diameters are so remarkably small in the Colombian sample.

The contemporary samples from France, the Netherlands, Taiwan, and the U.S. overlap according to the modality used for data collection (CT scan) and one or several covariates, be it their population of origin, HDI category, or Gini category. These samples show significant overlap in the PCA plots and have ranges for VNC AP and/or TR diameters that are comparable to clinical data (Griffith et al., 2016; Hinck et al., 1966; Papp et al., 1994). Although the Colombian sample is the only contemporary dry bone sample, several publications have shown that virtual bone measurements and other non-metric skeletal indicators showed no significant or systematic differences compared to their dry bone counterparts (Abegg et al., 2021; Anderson et al., 2021; Colman, de Boer, et al., 2019; Colman, van der Merwe, et al., 2019; Corron et al., 2017; Spake et al., 2020; Stull et al., 2014). The same individual (Louise Corron) collected the VNC data on dry bone and CT images, with minimal intra-observer error rates for the latter (see Results Section 3) and had verified the consistency between vertebral body

measurements obtained on both dry bone and CT scan of immature fifth lumbar vertebrae in a previous study (Corron et al., 2017). Additionally, ranges for mean VNC diameters in the East Smithfield cemetery sample and the Spitalfields collection were often overlapping with those of the contemporary CT samples, albeit sometimes in the lower ranges (Figure 5). Therefore, a difference in data collection modality is a highly unlikely explanatory factor of smaller VNC size in the Colombian sample (Figure 5).

The Colombian sample is composed of deceased individuals and there are potential arguments to be made regarding their small size as a result of mortality bias (Saunders & Hoppa, 1993; Spake & Cardoso, 2019). However, this is also a highly unlikely reason for their systematic size differences. A recent study comparing skeletal and dental markers collected from deceased individuals from the United States SVAD sample and living individuals from another SVAD sample from South Africa showed no significant differences in skeletal growth or dental development between individuals over the age of 2 years (Stull et al., 2021). Similarly in the present study, VNC size did not appear to exhibit any significant differences between the deceased (U.S.) sample and the other three living (France, the Netherlands, and Taiwan) samples who also presented with significant overlap with the Spitalfields and East Smithfield cemetery ranges (Figure 5). This eliminates – or at least reduces – mortality bias as a causal factor alone for the differences observed with the Colombian and historic samples.

Smaller VNC size of both AP and TR diameters do not appear to be related to the overall smaller stature of the Colombian population (Figure 8). Average adult stature is smallest for the Colombian population compared to the other four samples for both males and females (Table 3). However, the distribution in stature among the samples is large, and other countries—such as between Netherlands and Taiwan—also present with significant differences in size. All VNC diameters for the other four contemporary samples were comparable, be they scaled or unscaled, or whether individual stature or population averages were used (Figure 8). Therefore, it seems reasonable to exclude body size as a possible factor behind significantly smaller VNC diameters in the Colombian sample, which other studies have also ascertained (Amoroso & Garcia, 2018; Clark, 1985; Clark et al., 1985).

The Colombian sample does not share the same HDI (High) or Gini (High) categories with any of the other four contemporary samples and could effectively be qualified as a lower-SES population in comparison (Hicks & Leonard, 2014; United Nations Development Programme and Report, 2019). Furthermore, this sample and the Taiwanese sample present a different population history than the other three contemporary samples from Frnce, the Netherlands, and the U.S. (Table 1), although the latter did not show significant differences in VNC size with the other three. The French, Dutch, and the U.S. samples present with largely European-based backgrounds while the Taiwanese sample likely presents with primarily Han Chinese ethnic backgrounds (https://www.cia.gov/the-world-factbook/countries/taiwan/#people-and-society; https://www.census.gov/data/tables/2000/dec/phc-t-43.html). While there is indeed dominant European influence resulting from colonization (75%), the Medellin population history is complex with influence from the local indigenous groups (18%) and Africa (7%) (Conley et al., 2017). The distinctive pattern of VNC variation could be because of socio-economic and/or population of origin, as well as other factors that we did not consider as covariates. Therefore, a more in-depth analysis is required prior to inferring a causal relationship between smaller VNC diameters and stress.

The fact that no significant differences were found for VNC shape (AP/TR ratio) compared to VNC size in the PCA plots according to population of origin, HDI, or Gini category,

corroborates arguments made for VNC shape having a stronger degree of morphological integration (O'Higgins et al., 1997) (Figures 6 and 7). It is hypothesized that strong genetic and developmental constraints or canalization and natural selection changed patterns of integration in humans in the traits associated with bipedalism, such as the shape of the vertebral column and of the neural canal (Arlegi et al., 2018; Arlegi et al., 2020). Essentially, VNC shape has two main functions, which are protection of the spinal cord and bipedalism (Galbusera, 2018). Therefore, the size of the vertebral body and spinous processes may vary, but the shape is retained so that functionality is not compromised (Arlegi et al., 2018; Shapiro & Kemp, 2019). In effect, growth could be negatively affected by a stressor in either the AP or TR direction, depending on when it occurred during growth and development, with earlier stress being more likely to permanently affect AP diameters, while later or more prolonged stress could largely impact TR diameters (Hinck et al., 1966; Watts, 2013, 2015). Moreover, internal biological processes could regulate the counter dimension's growth to preserve the affected dimension and, by extension, the integrity of VNC shape, suggesting strong canalization of VNC size, especially for TR diameters and for the lumbar vertebrae (McPherson, 2021).

4.3 Small VNC size and stress in contemporary samples: Individual versus population levels of variation

According to the World Health Organization (WHO), individuals can be classified as short-for-age or stunted if they fall below the 5th percentile (two standard deviations) of a given growth standard (WHO Multicentre Growth Reference Study Group, 2006). By plotting the 95th, 50th, and 5th percentiles of the growth profile of the AP diameter of L4 for the U.S. sample (Figure 9), we can see that 98 individuals fall below the 5th percentile across the entire age range. Of these 98 individuals, 78 are aged between birth and 7 years, while the remaining 20 individuals are aged between 10 and 19 years. VNC growth profiles in our contemporary samples showed that growth was still ongoing for AP diameters up until around 5–6 years of age. Therefore, individuals younger than six could still have potential for catch-up growth for both AP and TR dimensions (Cameron, 2012). An examination of the demographic information, autopsy reports, and health information for the 20 older individuals who presented with small AP diameters for L4 showed no consistent pattern in terms of sex or manner of death, nor did they systematically present with other skeletal indicators of growth disruption or stress such as stunting, CO, PH, or Harris lines according to their autopsy reports. However, small VNC size in the U.S. sample has shown to be associated with perinatal disorders and low birthweight (O'Donnell et al., n.d.), indicating that these older individuals could have experienced growth disruption during early childhood, which could have led to small AP diameters. Our findings seem to point toward low quality of life/HDI as a more likely cause for smaller VNC size in the Colombian sample, because of its different SES level compared to all the others. By comparison, the Colombian individuals seem to be more within the range of the U.S. outliers mentioned previously. If we were to use the VNC growth profiles and/or the AP or TR diameters of the vertebrae measured for the U.S. sample as reference data to identify possible biological outliers in the Colombian sample, it would be tempting to conclude that since the Colombian measurements fall well below the 5th percentile of the U.S. sample's growth profile and are close to or lower than the minimum value for U.S. individuals of the same age group, they could have experienced stress during childhood and even in utero (see Supporting Information S1 for tables recording the mean, standard deviation, minimum, and maximum AP and TR diameters for each vertebrae, by annual age group and country of origin). However, the Colombian sample is relatively small ($N = 28$), which means only a few individuals provided VNC measurements for each age group, thus masking the true variation in VNC size for that population. Moreover, it also begs the question of the validity behind using

the U.S. sample as comparative data for assessing growth disruptions in samples or individuals from different populations, knowing they present with different socio-economic, geographic, and genetic backgrounds. Comparative studies or the detection of outliers or growth disruptions might only be applicable to individuals from the same population, or other populations who share a certain level of genetic, geographic, and/or socio-economic proximity. For instance, although there were some differences in the growth profiles (Figure 4), the ranges of variation for the U.S., France, and the Netherlands showed significant overlap, which could be due to their somewhat similar backgrounds.

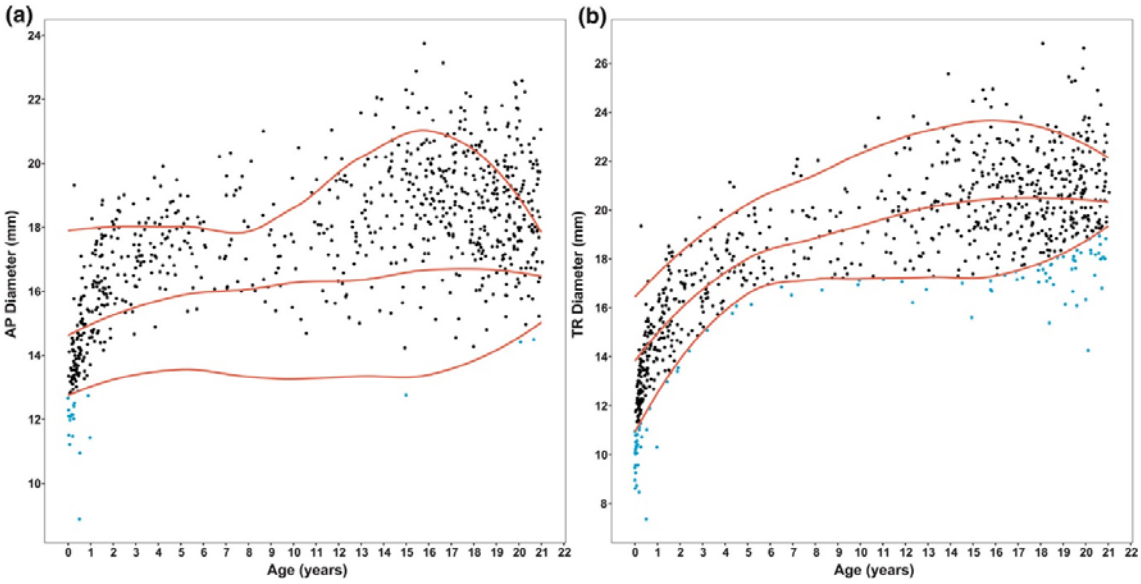

FIGURE 9. Percentiles (5th, 50th, and 95th-lower, middle, and upper red lines, respectively) of VNC dimensions of the fourth lumbar vertebra (L4) plotted against age in the United States sample—Biological outliers (individuals below the 5th percentile) are highlighted in blue. AP, anteroposterior; TR, transverse; VNC, vertebral neural canal

Even in contemporary, large, and varied reference samples, the effects of biocultural factors are difficult to attribute to one factor and/or another without verified and differentiated individual life histories, and genetic and environmental backgrounds that are not based on speculative health and/or demographic information. Interpreting metric skeletal indicators such as VNC diameters as evidence of biological stress is particularly challenging, as the skeletal expression of stress is in fact a reflection of that event across multiple dimensions. In this regard, Vercelotti and colleagues proposed an interpretation of small for age stature that can be generalized to any bodily dimension used as a marker of stress: "*The complexity of growth and its disruptions implies that […] the terminal, measurable outcome of a prolonged and mutable process, may not always allow us to infer growth conditions experienced by individuals and populations*" (Vercellotti et al., 2014, p. 230). Indeed, the "complexity of growth" includes the timing of the occurrence of the disruption event and how it relates to an element's SDW, the anatomical directionality of the element's remodeling, its level of plasticity and potential for catch-up growth, and the varying levels of response and analyses of stress at both the population and individual levels. The first two dimensions (timing and directionality) will conjointly condition the extent to which the element will be affected and how it will express a skeletal marker. This complexity is clearly problematic for evaluating growth and development disruptions and past episodes of stress in archeological or historic human groups, typically characterized by limited relevant contextual information on nutrition and disease throughout the life course, with a few exceptions (Pilloud & Schwitalla, 2020). The third dimension

(individual versus population levels) takes into account levels of variation that can only be appreciated and interpreted with the use of extensively documented skeletal samples of subadults and adults of the same population, which are extremely rare resources in the bioanthropological field (Albanese, 2003; Berry & Edgar, 2021; Stull & Corron, 2022). Clinical samples and longitudinal data would be ideal for such a purpose, as they can provide concrete and reliable information on an individual's life history and allow the visualization of episode of stress and catch-up growth, following a life course approach specifically catered to subadult specimens (Agarwal, 2016; Vercellotti et al., 2014).

This study is the first to evaluate the growth profiles of VNC dimensions on a large sample of contemporary subadult individuals and explore the range of variation and growth patterns of this non-specific skeletal marker of stress across postnatal ontogeny. Further, this study contextualizes the results based on a diverse suite of biosocial covariates including age, biological sex, body size, geographic origin (Colombia, France, the Netherlands, Taiwan, and the United States), and economic/quality of life parameters (medium to high Human Development Indices, low to high inequality levels). VNC shape is preserved across age, sex, geography, and socio-economic level, a trend that will likely be corroborated in future studies with different samples. This commonality in VNC ratios across samples points to a stronger genetic/functional control or growth canalization that likely stems from modifications associated with bipedalism. In contrast, it is possible for populations/samples to present with varying VNC diameters, and likely a greater variability in AP diameters compared to TR diameters. This understanding of variation is necessary for any downstream interpretations related to phenotypic variation or physiological stress. While the findings indicate that quality of life (i.e., HDI) may be a driving factor influencing VNC dimensions, the causal relationship between the two, or even correlation between the two, cannot yet be ascertained. The Taiwanese sample yields the greatest support that HDI *is contributing* to the differentiation in VNC dimensions. Like the Colombian sample, Taiwan has a different population history and smaller body size, but unlike the Colombian sample, the Taiwanese sample has a comparable HDI to the other contemporary groups. The visualizations and analyses refute that population or body size contribute to the differences in VNC size. The hesitancy by the current authors is primarily because only one sample categorized in the lower HDI group. Additional samples with varying HDI categories are needed to accept or refute the claim that HDI/quality of life is strongly correlated with VNC size, and that low HDI categories would lead to disruptions of VNC growth. This study was also the first to directly compare VNC dimensions of past populations with contemporary populations to expose the strength of the claims regarding the association of smaller VNC sizes with stressed populations. Historic samples did not necessarily fall far below the range of variation for VNC dimensions in contemporary samples, even though they are thought to have experienced greater biological stress because of extensive periods of poor health, disease, and possible malnutrition (Amoroso & Garcia, 2018; Molleson & Cox, 1993; Newman & Gowland, 2015; Watts, 2013; Wren, 2017). The visualizations suggest that secular trends may in fact be another driving force, further affecting our ability to disentangle the exact cause or contributions of the factors behind VNC size variation.

We caution that skeletal markers, such as VNC size, should not be unquestionably interpreted as indicators of biological stress without a thorough assessment of their “normal” variation and the examination of other possible sources of variation. In an effort to achieve this, we took inspiration from the human life history approach in an attempt to better understand how, when, and why skeletal markers of stress occur in contemporary populations, which is necessary before interpreting their presence in past populations (Agarwal, 2016; Temple, 2018). There is no doubt that there is a uniqueness to the VNC dimensions of the Colombian sample, and that

from our comparisons/analyses, HDI could be a larger driving force as compared to population variation, sex, body size, or the distribution of wealth in a country. Although VNC size is commonly used as a skeletal indicator of biological stress in past populations, this relationship is far from straightforward.

AUTHOR CONTRIBUTIONS

Louise K. Corron: Conceptualization (equal); data curation (supporting); formal analysis (equal); funding acquisition (supporting); investigation (equal); methodology (equal); project administration (equal); software (equal); supervision (supporting); validation (equal); visualization (equal); writing – original draft (equal); writing – review and editing (lead). **Christopher A. Wolfe:** Data curation (supporting); formal analysis (lead); investigation (equal); methodology (supporting); visualization (lead); writing – original draft (equal); writing – review and editing (equal). **Kyra E. Stull:** Conceptualization (equal); data curation (equal); formal analysis (equal); funding acquisition (lead); investigation (supporting); methodology (supporting); project administration (lead); resources (lead); supervision (equal); validation (equal); writing – review and editing (equal).

ACKNOWLEDGMENTS

Thank you to all our collaborators who accepted to share their resources and build the SVAD: Dr Kathia Chaumoitre, AP-HM, Aix-Marseille University, Dr Kerri Colman, University of Amsterdam, Dr Rick van Rijn, University of Amsterdam, Dr Timisay Monsalve-Vargas, Universidade de Antioquia, Dr An-Di Yim, Truman State University. This research was funded by the National Institute of Justice (NIJ Awards 2015-DN-BX-K409 and 2017-DN-BX-0144) and the National Science Foundation (BCS-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 the National Science Foundation.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in the SVAD Zenodo Community repository at https://zenodo.org/communities/svad/ (DOI: 10.5281/zenodo.6342097).

REFERENCES

Abegg, C., Balbo, I., Dominguez, A., Grabherr, S., Campana, L., & Moghaddam, N. (2021). Virtual anthropology: A preliminary test of macroscopic observation *versus* 3D surface scans and computed tomography (CT) scans. *Forensic Sciences Research*, 6(1), 34–41. https://doi.org/10.1080/20961790.2020.1817270

Agarwal, S. C. (2016). Bone morphologies and histories: Life course approaches in bioarchaeology. *The American Journal of Physical Anthropology*, 159, S130–S149.

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

Aly, T., & Amin, O. (2013). Geometrical dimensions and morphological study of the lumbar spinal canal in the normal Egyptian population. *Orthopedics*, 36(2), e229–e234. https://doi.org/10.3928/01477447-20130122-27

Amonoo-Kuofi, H. S. (1982). Maximum and minimum lumbar interpedicular distances in normal adult Nigerians. *Journal of Anatomy*, 135(2), 225–233.

Amoroso, A., & Garcia, S. (2018). Can early-life growth disruptions predict longevity? Testing the association between vertebral neural canal (VNC) size and age-at-death. *International Journal of Paleopathology*, 22, 8–17.

Amoroso, A., Garcia, S., & Cardoso, H. (2014). Age at death and linear enamel hypoplasias: Testing the effects of childhood stress and adult socioeconomic circumstances in premature mortality. *American Journal of Human Biology*, 26(4), 461–468.

Anderson, A. S., Sutherland, M. L., O'Donnell, L., Hill, E. C., Hunt, D. R., Blackwell, A. D., & Gurven, M. D. (2021). Do computed tomography findings agree with traditional osteological examination? The case of porous cranial lesions. *International Journal of Paleopathology*, 33, 209–219. https://doi.org/10.1016/j.ijpp.2021.04.008

Arlegi, M., Gómez-Robles, A., & Gómez-Olivencia, A. (2018). Morphological integration in the gorilla, chimpanzee, and human neck. *American Journal of Physical Anthropology*, 166(2), 408–416. https://doi.org/10.1002/ajpa.23441

Arlegi, M., Veschambre-Couture, C., & Gómez-Olivencia, A. (2020). Evolutionary selection and morphological integration in the vertebral column of modern humans. *American Journal of Physical Anthropology*, 171(1), 17–36. https://doi.org/10.1002/ajpa.23950

Armelagos, G. J., Goodman, A. H., Harper, K. N., & Blakey, M. L. (2009). Enamel hypoplasia and early mortality: Bioarcheological support for the Barker hypothesis. *Evolutionary Anthropology*, 18, 261–271.

Bagnall, K. M., Harris, P. F., & Jones, P. R. M. (1977a). A radiographic study of the human fetal spine. 1. The development of the secondary cervical curvature. *Journal of Anatomy*, 123(3), 777–782.

Bagnall, K. M., Harris, P. F., & Jones, P. R. M. (1977b). A radiographic study of the human fetal spine. 2. The sequence of development of ossification centres in the vertebral column. *Journal of Anatomy*, 124(3), 791–802.

Baker, B. J., Dupras, T. L., & Tocheri, M. W. (2005). The osteology of infants and children. Texas A and M University Press, 178p.

Barker, D. J. P. (2007). The origins of the developmental origins theory. *Journal of Internal Medicine*, 261(5), 412–417. https://doi.org/10.1111/j.1365-2796.2007.01809.x

Barnes, E. (1994). Developmental defects of the axial skeleton in paleopathology. University Press of Colorado.

Beatrice, J. S., & Soler, A. (2016). Skeletal indicators of stress: A component of the biocultural profile of undocumented migrants in southern Arizona. *Journal of Forensic Sciences*, 61(5), 1164–1172. https://doi.org/10.1111/1556-4029.13131

Beatrice, J. S., Soler, A., Reineke, R. C., & Martínez, D. E. (2021). Skeletal evidence of structural violence among undocumented migrants from Mexico and Central America. *American Journal of Physical Anthropology*, 176(4), 584–605. https://doi.org/10.1002/ajpa.24391

Berry, S. D., & Edgar, H. J. H. (2021). Announcement: The New Mexico decedent image database. *Forensic Imaging*, 24, 200436.

Blakemore, L., Schwend, R., Akbarnia, B. A., Dumas, M., & Schmidt, J. (2018). Growth patterns of the neurocentral synchondrosis (NCS) in immature cadaveric vertebra. *Journal of Pediatric Orthopaedics*, 38(3), 181–184. https://doi.org/10.1097/BPO.0000000000000781

Bogin, B. (1997). Evolutionary hypotheses for human childhood. *Yearbook of Physical Anthropology*, 40, 63–89.

Bogin, B., Scheffler, C., & Hermanussen, M. (2017). Global effects of income and income inequality on adult height and sexual dimorphism in height. *American Journal of Human Biology*, 29(2), 1–11.

Bogin, B., Varela Silva, M. I., & Rios, L. (2007). Life history trade-offs in human growth: Adaptation or pathology? *American Journal of Human Biology*, 19, 631–642.

Brough, A., Rutty, G., Black, S., & Morgan, B. (2012). Post-mortem computed tomography and 3D imaging: Anthropological applications for juvenile remains. *Forensic Science Medicine and Pathology*, 8(3), 270–279. https://doi.org/10.1007/s12024-012-9344-z

Cameron, N. (2012). Human growth curve, canalization, and catch-up growth. In N. Cameron (Ed.), Human growth and development (pp. 1–22). Elsevier Academic Press.

Cameron, N., & Bogin, B. (2012). Human growth and development (2nd ed.). Academic Press.

Clark, G. A. (1985). Heterochrony, allometry, and canalization in the human vertebral column: Examles from the prehistoric Amerindian populations. Amherst: University of Massachussets, Ann Harbor.

Clark, G. A. (1988). New method for assessing changes in growth and sexual dimorphism in paleoepidemiology. *American Journal of Physical Anthropology*, 77(1), 105–116.

Clark, G. A., Hall, N. R., Armelagos, G. J., Borkman, G. A., Panjabi, M. M., & Wetzel, F. T. (1986). Poor growth prior to early childhood: Decreased health and life-span in the adult. *American Journal of Physical Anthropology*, 70, 145–160.

Clark, G. A., Panjabi, M. M., & Wetzel, F. T. (1985). Can infant malnutrition cause adult vertebral stenosis? *Spine*, 10(2), 165–170. https://doi.org/10.1097/00007632-198503000-00012

Cleveland, W. S. (1979). Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association*, 74(368), 829–836. https://doi.org/10.1080/01621459.1979.10481038

Colman K. L., Dobbe J. G. G., Stull K. E., Ruijter J. M., Oostra R. J., van Rijn R. R., van der Merwe A. E, de Boer H. H., Streekstra G. J. (2017). The geometrical precision of virtual bone models derived from clinical computed tomography data for forensic anthropology. *Int J Legal Med.* 131(4), 1155–1163. https://doi.org/10.1007/s00414-017-1548-z.

Colman K. L., de Boer H. H., Dobbe J. G. G., Liberton N. P. T. J., Stull K. E., van Eijnatten M., Streekstra G. J., Oostra R. J., van Rijn R. R., van der & Merwe A. E. (2019). Virtual forensic anthropology: The accuracy of osteometric analysis of 3D bone models derived from clinical computed tomography (CT) scans. *Forensic Sci Int.* 304, 109963. https://doi.org/10.1016/j.forsciint.2019.109963.

Colman, K. L., van der Merwe, A. E., Stull, K. E., Dobbe, J. G., Streekstra, G. J., van Rijn, R. R., … de Boer, H. H. (2019). The accuracy of 3D virtual bone models of the pelvis for morphological sex estimation. *International Journal of Legal Medicine*, 133(6), 1853–1860.

Conley, A. B., Rishishwar, L., Norris, E. T., Valderrama-Aguirre, A., Mariño-Ramírez, L., Medina-Rivas, M. A., & Jordan, I. K. (2017). A comparative analysis of genetic ancestry and admixture in the Colombian populations of Chocó and Medellín. *G3*, 7(10), 3435–3447. https://doi.org/10.1534/g3.117.1118

Corron, L., Marchal, F., Condemi, S., Chaumoître, K., & Adalian, P. (2017). Evaluating the consistency, repeatability and reproducibility of osteometric data on dry bone surfaces, scanned dry bone surfaces and scanned bone surfaces from living individuals. *Bulletins et Mémoires de la Société d'Anthropologie de Paris*, 29(1–2), 33–53.

DataBank, D. R. G. (2020). World development indicators. DataBank. https://databank.worldbank.org/source/world-development-indicators

De Witte, S., & Woods, J. W. (2008). Selectivity of Black death mortality with respect to preexisting health. *PNAS*, 105(5), 1436–1441.

De Witte, S. N., & Stojanowski, C. M. (2015). The osteological paradox 20 years later: Past perspectives, future directions. *Journal of Archaeological Research*, 23, 397–450.

Eisenstein, S. (1977). The morphometry and pathological anatomy of the lumbar spine in south African negroes and Caucasoids with specific reference to spinal stenosis. *The Journal of Bone and Joint Surgery*, 59(2), 173–180.

Eslami, A., Qannari, E., Kohler, A., & Bougeard, S. (2013). General overview of methods of analysis of multi-group datasets. Advances in theory and applications of high dimensional and symbolic data. *Analysis*, RNTI-E-25, 108–123.

Fazekas, I., & Kosa, F. (1978). Forensic fetal osteology. Akadémiai Kiadó, Budapest. 413p.

Fox, J., & Weisberg, S. (2019). An R companion to applied regression (3rd ed.). SAGE.

Galbusera, F. (2018). The spine: Its evolution, function, and shape. In F. Galbusera & H.-J. Wilke (Eds.), Biomechanics of the spine: Basic concepts, spinal disorders and treatments. Academic Press.

I. Grainger, & Museum of London, & English Heritage (Eds.). (2008). The Black death cemetery, East Smithfield, London. Museum of London Archaeology Service.

Grasgruber, P., Sebera, M., Hrazdira, E., Cacek, J., & Kalina, T. (2016). Major correlates of male height: A study of 105 countries. *Economics and Human Biology*, 21, 172–195.

Griffith, J. F., Huang, J., Law, S.-W., Xiao, F., Leung, J. C. S., Wang, D., & Shi, L. (2016). Population reference range for developmental lumbar spinal canal size. *Quantitative Imaging in Medicine and Surgery*, 6(6), 671–679. https://doi.org/10.21037/qims.2016.12.17

Hermann, A. P., Brixen, K., Andresen, J., & Mosekilde, L. (1993). Reference values for vertebral heights in Scandinavian females and males. *Acta Radiologica*, 34(1), 48–52. https://doi.org/10.1177/028418519303400111

Hicks, K., & Leonard, W. R. (2014). Developmental systems and inequality: Linking evolutionary and political-economic theory in biological anthropology. *Current Anthropology*, 55(5), 523–550. https://doi.org/10.1086/678055

Hildebolt, C. F., Vannier, M. W., & Knapp, R. H. (1990). Validation study of skull three-dimensional computerized tomography measurements. *American Journal of Physical Anthropology*, 40, 283–294.

Hinck, V. C., Clark, W. M., & Hopkins, C. E. (1966). Normal interpedicular distances (minimum and maximum) in children and adults. *American Journal of Roentgenology*, 97, 141–153.

Hochberg, Z. (2012). Evo-devo on child growth and human evolution. Wiley-Blackwell.

Hoppa & Charles M. Fitzgerald. (1999) Human Growth in the Past: Studies from Bones and Teeth. Cambridge; New York: Cambridge University Press. 315p.

Jahan, S. (2015). Human development report 2015—Work for human development (p. 288). United Nations Development Programme.

Janjua, M. Z., & Muhammad, F. (1989). Measurements of the normal adult lumbar spinal canal. *The Journal of the Pakistan Medical Association*, 39(10), 264–268.

Kassambara, A., & Mundt, F. (2020). Factoextra: Extract and visualize the results of multivariate data analyses. https://CRAN.R-project.org/package=factoextra.

Klaus, H. D. (2014). Frontiers in the bioarchaeology of stress and disease: Cross-disciplinary perspectives from pathophysiology, human biology, and epidemiology. *American Journal of Physical Anthropology*, 155(2), 294–308.

Klaus, H. D., & Tam, M. E. (2009). Contact in the Andes: Bioarchaeology of systemic stress in colonial Morrope, Peru. *American Journal of Physical Anthropology*, 138, 356–368.

Konigsberg LW. & Holman D. (1999) Estimation of age-at-death from dental emergence and implications for studies of prehistoric somatic growth, In: Robert D. Hoppa and Charles M. Fitzgerald. Human Growth in the Past: Studies from Bones and Teeth. Cambridge; New York: Cambridge University Press. pp. 264–289.

Krzanowski, W. J. (1984). Principal component analysis in the presence of group structure. *Applied Statistics*, 33(2), 164. https://doi.org/10.2307/2347442

Lasker, G. W. (1969). Human biological adaptability. The ecological approach in physical anthropology. *Science*, 166, 1480–1486.

Lê, S., Josse, J., & Husson, F. (2008). FactoMineR: An R package for multivariate analysis. *Journal of Statistical Software*, 25(1), 1–18. https://doi.org/10.18637/jss.v025.i01

Legg, S. J., & Gibbs, V. (1984). Measurement of the lumbar spinal canal by echo ultrasound. *Measurement of the Lumbar Spinal Canal by Echo Ultrasound*, 9, 79–82.

LoPresto, S. (2020). Vertebral neural canal growth and developmental stress: A case study from the American southwest. George Mason University.

McPherson, C. B. (2021). Examining developmental plasticity in the skeletal system through a sensitive developmental windows framework. *American Journal of Physical Anthropology*, 176(2), 163–178. https://doi.org/10.1002/ajpa.24338

Molleson, T., & Cox, M. (1993). The middling sort (p. 231). Council for British Archaeology.

NCD Risk Factor Collaboration. (2016). A century of trends in adult human height. *ELife Epidemiology and Global Health*, 5, e13410.

Newman, S. L., & Gowland, R. L. (2015). Brief communication: The use of non-adult vertebral dimensions as indicators of growth disruption and non-specific health stress in skeletal populations. *American Journal of Physical Anthropology*, 158, 155–164.

O'Donnell, L., Corron, L., & Hill, E. C. (n.d.). Perinatal disorders and small birthweight are significantly associated with small vertebral neural canal size in a contemporary pediatric autopsy sample. *Spine*.

O'Donnell, L., Hill, E. C., Anderson, A. S., & Edgar, H. J. H. (2022). A biological approach to adult sex differences in skeletal indicators of childhood stress. *American Journal of Biological Anthropology*, 177(3), 381–401. https://doi.org/10.1002/ajpa.24424

O'Donnell, L., Hill, E. C., Anderson, A. S. A., & Edgar, H. J. H. (2020). Cribra orbitalia and porotic hyperostosis are associated with respiratory infections in a contemporary mortality sample from New Mexico. *American Journal of Physical Anthropology*, 173(4), 721–733. https://doi.org/10.1002/ajpa.24131

O'Higgins, P., Milne, N., Johnson, D. R., Runnion, C. K., & Oxnard, C. E. (1997). Adaptation in the vertebral column: A comparative study of patterns of metameric variation in mice and men. *Journal of Anatomy*, 190(1), 105–113. https://doi.org/10.1046/j.1469-7580.1997.19010105.x

Papp, T., Porter, R. W., & Aspden, R. M. (1994). The growth of the lumbar vertebral canal. *Spine*, 19(24), 2770–2773.

Perini, T. A., de Oliveira, G. L., Ornellas, J. S., & Oliveira, F. P. (2005). Technical error of measurement in anthropometry. *Revista Brasileira de Medicina do Esporte*, 11(1), 81–85.

Pfeiffer, S., & Harrington, L. (2011). Bioarchaeological evidence for the basis of small adult stature in southern Africa: Growth, mortality, and small stature. *Current Anthropology*, 52(3), 449–461.

Piera, V., Rodriguez, A., Cobos, A., Hern, R., & Cobos, P. (1988). Morphology of the lumbar vertebral canal. *Cells, Tissues, Organs*, 131(1), 35–40. https://doi.org/10.1159/000146482

Pierro, A., Cilla, S., Maselli, G., Cucci, E., Ciuffreda, M., & Sallustio, G. (2017). Sagittal normal limits of lumbosacral spine in a large adult population: A quantitative magnetic resonance imaging analysis. *Journal of Clinical Imaging Science*, 7, 35. https://doi.org/10.4103/jcis.JCIS_24_17

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.

Postacchini, F., Ripani, M., & Carpano, S. (1983). Morphometry of the lumbar vertebrae: An anatomic study in two Caucasoid ethnic groups. *Clinical Orthopaedics and Related Research*, 172, 296–303.

R Core Team. (2021). R: A language and environment for statistical computing. Vienna, Austria.

Rapała, K., Chaberek, S., Truszczyńska, A., Łukawski, S., & Walczak, P. (2009). Assessment of lumbar spinal canal morphology with digital computed tomography. *Ortopedia, Traumatologia, Rehabilitacja*, 11(2), 156–163.

Reichmann, S., & Lewin, T. (1971). Growth processes in the lumbar neural arch. *Zeitschrift Fur Anatomie Und Entwicklungsgeschichte*, 133, 89–101.

Rewekant, A. (2001). Do environmental disturbances of an individual's growth and development influence the later bone involution processes? A study of two mediaeval populations. *Int. J. Osteoarchaeol.*, 11, 433–443. https://doi.org/10.1002/oa.584

Roberts, D. (1995). The pervasiveness of plasticity. In C. G. N. Mascie-Taylor & G. E. H. Mohamed (Eds.), Human variability and plasticity (pp. 1–17). Cambridge University Press.

Ruhli, F. (2003). Osteometric variation of the human spine in Central Europe by historic time period and its microevolutionary implications. The University of Adelaide.

Santiago, F., Milena, G., Herrera, R., Romero, P., & Plazas, P. (2001). Morphometry of the lower lumbar vertebrae in patients with and without low back pain. *European Spine Journal*, 10(3), 228–233. https://doi.org/10.1007/s005860100267

Saunders, S. R., & Hoppa, R. D. (1993). Growth deficit in survivors and non-survivors: Biological mortality bias in subadult skeletal samples. *Yearbook of Physical Anthropology*, 36, 127–151.

Scheuer, L., & Black, S. (2000). Developmental juvenile osteology. Gray Publishing. San Diego, 587p.

Scheuer, L., & Black, S. (2004). The juvenile skeleton. Elsevier Academic Press, London, 485p.

Schizas, C., Schmit, A., Schizas, A., Becce, F., Kulik, G., & Pierzchala, K. (2014). Secular changes of spinal canal dimensions in western Switzerland: A narrowing epidemic? *Spine*, 39(17), 1339–1344.

Shapiro, L. J., & Kemp, A. D. (2019). Functional and developmental influences on intraspecific variation in catarrhine vertebrae. *American Journal of Physical Anthropology*, 168(1), 131–144. https://doi.org/10.1002/ajpa.23730

Spake, L., & Cardoso, H. (2019). Indirect evidence for biological mortality bias in growth from two temporo-spatially distant samples of children. *Anthropologischer Anzeiger*, 76(5), 379–390.

Spake, L., Meyers, J., Blau, S., Cardoso, H., & Lottering, N. (2020). A simple and software-independent protocol for the measurement of post-cranial bones in anthropological contexts using thin slab maximum intensity projection. *Forensic Imaging*, 20, 200354.

Steckel, R. H. (2012). Social and economic effects on growth. In Human growth and development (pp. 225–244). Elsevier. https://doi.org/10.1016/B978-0-12-383882-7.00009-X

Steckel, R. H., Rose, J. C., Spencer Larsen, C., & Walker, P. L. (2002). Skeletal health in the Western hemisphere from 4000 B.C. to the present. *Evolutionary Anthropology: Issues, News, and Reviews*, 11(4), 142–155. https://doi.org/10.1002/evan.10030

Stull, K., & Corron, L. (2022). The subadult virtual anthropology database (SVAD): An accessible repository of contemporary subadult reference data. *Forensic Sciences*. 2(1), 20–36 https://doi.org/10.3390/forensicsci2010003

Stull, K. E., & Corron, L. K. (2021). Subadult virtual anthropology database (SVAD) data collection protocol: Epiphyseal fusion, diaphyseal dimensions, dental development stages, vertebral neural canal dimensions. *Zenodo*. https://doi.org/10.5281/ZENODO.5348392

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.

Stull, K. E., Wolfe, C. A., Corron, L. K., Heim, K., Hulse, C. N., & Pilloud, M. (2021). A comparison of subadult skeletal and dental development based on living and deceased samples. *American Journal of Physical Anthropology*, 175(1), 36–58.

Tacar, O., Demirant, A., Nas, K., & Altindag, O. (2003). Morphology of the lumbar spinal canal in normal adult Turks. *Yonsei Medical Journal*, 44, 679–685.

Temple, D. H. (2018). Bioarchaeological evidence for adaptive plasticity and constraint: Exploring life-history trade-offs in the human past. *Evolutionary Anthropology*, 28(1), 1–13.

Temple, D. H., Bazaliiskii, V. I., Goriunova, O. I., & Weber, A. W. (2014). Skeletal growth in early and late Neolithic foragers from the cis-Baikal region of eastern Siberia. *American Journal of Physical Anthropology*, 153, 377–386.

Todorov, V. (2020). RrcovNA: Scalable robust estimators with high breakdown point for incomplete data. https://CRAN.R-project.org/package=rrcovNA

Twomey, L., & Taylor, M. B. (1988). Age changes in the lumbar spinal and intervertebral canals. *Paraplegia*, 26, 238–249.

Ulijaszek, S. J., & Kerr, D. A. (1999). Anthropometric measurement error and the assessment of nutritional status. *British Journal of Nutrition*, 82, 165–177.

United Nations Development Programme and Report. (2019). UNDP Human Development Report 2019. Beyond income, beyond averages, beyond today: Inequalities in human development in the 21st century. New York.

Verboven, S., Vanden Branden, K., & Goos, P. (2007). Sequential imputation for missing values. *Computational Biology and Chemistry*, 31(5–6), 320–327.

Vercellotti, G., Piperata, B. A., Agnew, A. M., Wilson, W. M., Dufour, D. L., Reina, J. C., … Sciulli, P. W. (2014). Exploring multidimensionality of stature variation in the past through comparisons of archaeological and living populations. *American Journal of Physical Anthropology*, 155, 229–242.

Waldron, T. (2021). Palaeopathology (2nd ed.). Cambridge University Press.

Walker, P., Bathurst, R., Richman, R., Gierdrum, T., & Andrushko, V. A. (2009). The causes of porotic hyperostosis and cribra orbitalia: A reappraisal of the iron-deficiency-anemia hypothesis. *American Journal of Physical Anthropology*, 139(2), 109–125.

Watts, R. (2011). Non-specific indicators of stress and their association with age at death in Medieval York: Using stature and vertebral neural canal size to examine the effects of stress occurring during different periods of development. *Int. J. Osteoarchaeol.*, 21, 568–576. https://doi.org/10.1002/oa.1158

Watts, R. (2013). Lumbar vertebral canal size in adults and children: Observations from a skeletal sample from London, England. *Homo - Journal of Comparative Human Biology*, 64, 120–128.

Watts, R. (2015). The long-term impact of developmental stress. Evidence from later medieval and post-medieval London (AD 1117-1853). *American Journal of Physical Anthropology*, 158, 569–580.

Wei, T., & Simko, V. (2021). R package "corrplot": Visualization of a correlation matrix. https://github.com/taiyun/corrplot.

Wells, J. (2016). Worldwide variability in growth and its association with health: Incorporating body composition, developmental plasticity, and intergenerational effects. *American Journal of Human Biology*, 29(2), 1–16.

WHO Multicentre Growth Reference Study Group. (2006). WHO child growth standards. Length, height for-age, weight-for-age, weight-for-length and body mass index-for age. Methods and development. Geneva: World Health Organization.

Wood, J. W., Milner, G. R., Harpending, H. C., Weiss, K. M., Cohen, M. N., Eisenberg, L. E., … Wilkinson, R. G. (1992). The osteological paradox: Problems of inferring prehistoric health from skeletal samples. *Current Anthropology*, 33(4), 343–370.

Wren, K. (2017). The effects of racialization on European American stress in the nineteenth and twentieth centuries (PhD thesis). University of Tennessee.