Selecting an appropriate reference sample for juvenile age estimation methods in a forensic context

by

Laure Spake

B.A., McGill University, 2013

Thesis Submitted in Partial Fulfillment of the

Requirements for the Degree of

Master of Arts

in the

Department of Archaeology

Faculty of Environment

SIMON FRASER UNIVERSITY

Spring 2016

Approval

Name: **Laure Spake**

Degree: **Master of Arts**

Title: ***Selecting an appropriate reference sample for juvenile age estimation methods in a forensic context***

Examining Committee: **Chair:** George Nicholas
Professor

Hugo Cardoso
Senior Supervisor
Assistant Professor

Tosha Dupras
Supervisor
Professor
Department of Anthropology
University of Central Florida

Daniel Franklin
External Examiner
Associate Professor
School of Anatomy, Physiology and Human Biology
University of Western Australia

Date Defended/Approved: April 18, 2016

Ethics Statement

The author, whose name appears on the title page of this work, has obtained, for the research described in this work, either:

a. human research ethics approval from the Simon Fraser University Office of Research Ethics

or

b. advance approval of the animal care protocol from the University Animal Care Committee of Simon Fraser University

or has conducted the research

c. as a co-investigator, collaborator, or research assistant in a research project approved in advance.

A copy of the approval letter has been filed with the Theses Office of the University Library at the time of submission of this thesis or project.

The original application for approval and letter of approval are filed with the relevant offices. Inquiries may be directed to those authorities.

Simon Fraser University Library
Burnaby, British Columbia, Canada

Update Spring 2016

Abstract

The population on which forensic juvenile skeletal age estimation methods are applied has not been critically considered. Previous research suggests that child victims of homicide tend to be from socioeconomically disadvantaged contexts, and that these contexts impair growth. Thus, juvenile skeletal remains examined by forensic anthropologists may be short for age.

Cadaver lengths were obtained from records of autopsies of 1256 individuals, aged birth to eighteen years at death, conducted between 2000 and 2015 in Australia, New Zealand, New Mexico, New York City, and Cuyahoga County. Growth status of the forensic population, represented by homicide victims, and general population, represented by accident victims, were compared using height for age Z-scores and independent sample t-tests. Cadaver lengths of the accident victims were evaluated against growth references using one sample t-tests to evaluate whether accident victims reflect the general population.

Homicide victims are shorter for age than accident victims in samples from the United States, but not in Australia and New Zealand. Accident victims are more representative of the general population in Australia and New Zealand. Different results in Australia and New Zealand as opposed to the United States may be linked to higher socioeconomic inequality in the United States.

These results suggest that physical anthropologists should critically select reference samples when devising forensic juvenile skeletal age estimation methods. Children examined in forensic investigations may be short for age, and thus methods developed on normal healthy children may yield inaccurate results.

Keywords: Forensic anthropology; linear growth; juvenile; skeletal age; age estimation methods

Acknowledgements

This project could not have been realized without the support of several people working within several institutions. Access to data was kindly provided by the following institutions: the **Cuyahoga County Medical Examiner**, the **New York City Office of the Chief Medical Examiner**, the **New Mexico Office of the Medical Investigator**, and the **National Coronial Information System** (New Zealand and Australian data). At the Cuyahoga County Medical Examiner, I would like to thank **Dr. Thomas Gilson** and **Jodie Schneider**, both of whom dedicated considerable time to communicating with me and compiling data. **Maureen Paulino**, at the New York City Office of the Chief Medical Examiner, facilitated the ethics approval process and succeeded in finding me a workstation during my brief stay in New York. **Dr. Berton Freedman**, at the Office for Vital Statistics, was incredibly responsive to e-mails and facilitated the ethics approval and data acquisition processes when I had particularly tight deadlines. At the New Mexico Office of the Medical Examiner, **Dr. Sarah Lathrop** was kind and responsive. **Jessica Bryan**, at the National Coronial Information System, worked closely with me to meet various deadlines in the ethics clearance processes for data from New Zealand and Australia. To each a sincere thank you.

Dr. Dana Lepofsky pushed us to attack our work from multiple intellectual angles, and nudged us to take on new challenges. **Dr. George Nicholas**' wealth of knowledge and kindness were much appreciated throughout these years. **Dr. Tosha Dupras**' input and dedication were essential to this project and I am thankful to have had her support throughout the last two years. My supervisor**, Dr. Hugo Cardoso,** offered me so many opportunities, pushed me to participate in academic discourse, and was always willing to include me in his projects. I am more grateful than I can express.

These two years have been formative, and I am grateful to a group wonderful people for their friendship. I couldn't ask for a better colleague than **Luísa Marinho**. The residents of **Princess House** took me under their wings and included me in their adventures.

Several others, notably **Stephen Crocker, Daniel Levy**, **Miriam Dyberg**, **Connor Harbin**, and **Alex Reedy** were incredible sources of friendship and support. I owe **András Szabo** my newfound love of early mornings and the outdoors.

Finally, my family has been incredibly understanding of the demands of these two years, and of my occasionally prolonged absences. Un gros merci à vous trois.

Thank you once more to all.

Table of Contents

List of Tables

List of Figures

Chapter 1. Introduction

In the last thirty years, physical anthropologists have acknowledged that the ways in which skeletal estimation methods are developed influence their performance. Large-scale forensic anthropology operations following civil conflict in the former Yugoslavia, Rwanda, and Latin America have forced researchers to confront the large spectrum of modern human variation and its implication for forensic identification (Komar 2003). Research regarding the validity of inter- population application of age, sex, and stature estimation methods has increased (e.g., Đurić et al. 2005, Kimmerle et al. 2008, Galić et al. 2011), as has the development of new, population-specific methods (e.g., Prince and Koningsberg 2008). Throughout this discussion of methods and their performance, three main issues are at play: the reference sample to be used, the statistics employed in developing the method, and both intra- and inter-population variation.

1.1. Problems with skeletal estimation methods

There is a general consensus that the reference sample on which methods are developed should be critically selected to reflect the full range of normal human skeletal variation. Many of the classic methods employed in forensic anthropology were developed using skeletal collections amassed in the 20th century, such as the Terry or Hamann-Todd collections, which include a large proportion of indigent individuals. Critical examinations of these samples suggest that they are not demographically representative of the American population (Ericksen 1982, Komar and Grivas 2008), and thus may not fully represent the range of biological variation expressed in the United States (Hunt and Albanese 2005). This potentially affects the performance of age, sex, stature and ancestry estimation methods developed from these collections. Additionally, the collection of indigent individuals means that demographic information, such as age, could be estimated rather than known (Meindl et al. 1990), affecting the validity of age estimation methods. Although these issues do not preclude the use of skeletal collections in method development, they do prescribe informed and mindful selection of individuals for inclusion in reference samples.

Statistical modeling has become a hot topic as increasingly sophisticated methods become available, and courts of law require increasingly data-driven over professional experience-driven testimony (Dirkmaat et al. 2008, Holobinko 2010). Bayesian analysis is employed by some (e.g.: Koningsberg et al. 1998, Kimmerle et al. 2008, Prince and Koningsberg 2008, Langley-Shirley and Jantz 2010, Ferrante et al. 2015) because it acknowledges the prior assumption inherent in all estimation models, namely that the target individual fits within the distribution of the reference sample (Koningsberg et al. 1998). From simple linear regression to complicated multivariate adaptive regression splines (e.g., Liversidge and Molleson 1999, Robbins Shug et al. 2013, Stull et al. 2014), regression methods are more commonly employed by physical anthropologists. Although linear regression is the simplest and most common of these methods, debates still occur about proper model calibration (Koningsberg et al. 1994, Aykroyd et al. 1997, Cardoso 2014) and appropriate reporting of descriptive statistics, including information to calculate 95% prediction ranges (Soyer and Hogarth 2012, Stull et al. 2014). Linear regression analysis is a tool with its own oft-ignored limitations, and not a foolproof prediction method (Palmer and O'Connell 2009, Taleb and Goldstein 2012). Increasingly complicated statistical techniques are dubiously advantageous: although they produce high explanations of variation and perform well within the test samples, they may be over-fitting the data. This can result in poor performance outside the original reference population (Armstrong 2012), which is especially problematic in physical anthropology due to the introduction of inter-population variation.

While fully representing normal human variation in reference samples can be a challenge for developing estimation methods, the application of these methods outside of the reference population is further complicated by inter-population variation (Jackes 2000, Kemkes-Grottenthaler 2002). Without understanding the specific factors affecting both the reference sample and target individual or population, the cross-population application of estimation methods can yield inaccurate results. Physical anthropologists now recommend the use of “appropriate” skeletal standards, meaning methods developed on reference samples similar to the target individual or sample (Saunders 2000, Cunha et al. 2009, Franklin 2010).

Skeletal estimation methods are inherently problematic, as they aim to use a known sample to estimate the biological parameters of an unknown individual or population. For some parameters, such as non-metric sex indicators, inter-population variation can have less impact, as trait expression is more obvious (Koningsberg and Frankenberg 1994). Differences between populations are more important for age estimation, and especially so in the case of juveniles, as age is the only biological parameter which can reliably be estimated.

1.2. Inter-population variation and juvenile skeletal estimation methods

Juveniles are still developing many or all of the skeletal and dental characteristics readily observed in adult individuals. For this reason, the biological parameters which can be estimated from the skeleton are currently more limited. While some attempts have been made to estimate sex from the juvenile skeleton (e.g.: Weaver 1980, Hunt 1990, Mittler and Sheridan 1992, Schutkowski 1993), these methods have not yielded consistently high accuracy rates. Ancestry is not estimated for juveniles, although some dental morphological variants may suggest a particular ethnic origin (Lewis and Rutty 2003). There is limited research concerning stature estimation (e.g.: Feldesman 1992, Smith 2007, Robbins Schug et al. 2013), however doubts are expressed about the usefulness of this parameter in comparison against missing persons lists, considering that stature changes so quickly during childhood. The most useful and securely estimated juvenile biological parameter is age (Lewis and Rutty 2003).

As opposed to adults, where age is estimated based on the degeneration of skeletal or dental elements, juvenile age is estimated based on their growth and development. The more predictable nature of development over degeneration leads to higher precision and accuracy in estimating age from juvenile skeletal remains. Because this estimation is based on development rather than degeneration, however, physical anthropologists are essentially assessing growth and development (skeletal age, SA, or dental age, DA), to estimate chronological age (CA).

Although growth is more stable and predictable than degeneration, it remains variable. At the individual level, growth is both biologically determined (Klebanoff et al. 1997, Behrman et al. 2009, Towne et al. 2012, Martorell and Zongrone 2012) and environmentally influenced (Schell et al. 2009). At the population level, individual biological effects are obscured by environmental factors (Steckel 2012). Environmental factors that affect growth are linked to socioeconomic status (SES) (Schmeling et al. 2000), and include: nutrition, disease, access to health care, poverty, housing, neglect and abuse, and chronic psychosocial stress (Skuse 1989, Bradley and Corwyn 2002, Fernald and Grantham-McGregor 2002, Block and Krebs 2005, Chilton et al. 2007, Gilbert et al. 2009, Norgan et al. 2012). On both the intra- and inter-population levels, growth can be quite variable, which ultimately complicates age estimation from juvenile skeletal remains.

Variability in growth leads to an incongruence between skeletal age and chronological age for every individual in a population, which produces a range of skeletal ages for the same chronological age within a population (Lampl and Johnston 1996). When aggregating individuals in one sample, some individual variation is obscured, and the resulting group reflects the growth status of the group as a whole (Steckel 2012). This status may not be the same for two populations with different growth environments, however, even when biology is held relatively constant (Cole 2003, Hermanussen et al. 2010). Thus, when physical anthropologists use a method based on the growth of one population to estimate age in another population, consideration must be given to the growth environments of both populations.

Skeletal and dental growth are differentially impacted by environmental effects. Although research shows that dental development is more strongly biologically determined than skeletal development, it is not wholly insulated from the impact of environmental effects (Saunders 2000, Cardoso 2007, Conceição and Cardoso 2011). As much as possible, it is therefore preferable to estimate age from the development of the dentition. Forensic anthropologists, however, cannot discount the study of skeletal growth and maturation as age estimators, as dental elements may not always be recovered in practice.

1.3. Research goals

Inter-population variation in skeletal growth has been recognized as a problem when estimating age of archaeological juvenile skeletal remains. Researchers have developed regionally or temporally specific methods to compensate for this effect (Pfau and Sciulli 1994, Facchini and Veschi 2004, Rissech et al. 2008, 2013, Danforth et al. 2009, Boccone et al. 2010, López-Costas et al. 2012, Primeau et al. 2012, 2016, Stull et al. 2014). However, little consideration has been given to the target population on which forensic anthropologists must employ juvenile skeletal age estimations methods. These unidentified skeletal remains represent a small group of children who died clandestinely and whose remains were undiscovered for some time, and are generally victims of homicide (Snow and Luke 1970, Cattaneo 2007, Simmons 2007, Ross 2011). They will be referred to as the forensic population throughout the rest of this study.

Little is known about the forensic population, the group of children whose deaths are under investigation by the forensic anthropologist, but the victims of juvenile homicide in general are relatively well understood. Juvenile homicide victims tend to come from very specific, disadvantaged socioeconomic contexts, and thus the quality of their growth environment is likely to differ from that of the average population. Their growth and development may be compromised in relation to their more socioeconomically privileged peers. The traditional methods (Maresh 1943, 1955, 1970) employed to estimate skeletal age of the forensic population are generally developed on healthy children. If they are applied to a population which is comparatively growth compromised, they may be yielding inaccurate age estimates, which could in turn slow or prevent victim identification.

Given the difference in socioeconomic makeup between the reference and target populations, and considering the impact that socioeconomic status can have on linear growth, a difference in growth status can be expected between the reference and target populations. The aim of this study is to determine whether children who may become part of the forensic population (as represented by homicide victims) differ in growth status from average children (as represented by accident victims), and in turn to better understand the nature of the juvenile forensic population. Thus, the first research question is: **do differences in growth occur between accident victims, which are meant to represent the normal population, and homicide victims, which are meant to represent the target population?**

Extrapolation from the comparison between accidental death victims and homicide victims to the differences between the forensic and general populations is premised on the appropriateness of the accident victims as proxies for the general population. However, fatal accidents are not evenly distributed across the population; they tend to disproportionately impact the lower end of the socioeconomic spectrum (UCL Institute for Health Equity 2015). Thus, the second research question is: **does the accidental death sample accurately represent the normal population?**

What follows is an investigation of these questions using retrospective collection of cadaver length from autopsy records obtained from coronial bodies in the United States, Australia, and New Zealand

Chapter 2. Literature Review

A review of existing knowledge provides the basis for this study. In order to be accurate, age estimation methods should be derived from a sample population that is appropriate for the individuals it is applied to. Thus, methods to estimate age from long bones available to forensic anthropologists are examined to better understand the samples on which they are developed. Then, the group of children on which these methods may be used is tentatively described by examining the socioeconomic and demographic characteristics of homicide victims. While the composition of the forensic sample is poorly understood, it generally represents a subset of the larger, better-understood group of child homicide victims. The forensic population is likely composed of a greater proportion of children from lower socioeconomic backgrounds, who have grown up in potentially adverse growth environments. Thus, studies of differences in growth across socioeconomic status groups are considered in order to reflect on whether growth differences can be expected between the target and reference populations.

2.1. Reference sample: available juvenile skeletal age estimation methods

Skeletal age estimation methods are constrained by the availability of reference material, which is scarce. They are generally based on the following sources: longitudinal growth studies of the living, known age skeletal collections, or known age forensic or clinical samples.

2.1.1. Longitudinal Studies

The main obstacle to creating juvenile skeletal age estimation methods is the rarity of reference material, especially dry bone collections. Longitudinal studies of growth were conducted throughout the United States in the early and middle 1900s and many of these studies included radiographs of various long bones. These radiographs were then used to measure the long bones, and typical long bone lengths for age charts were published throughout the second half of the 20th century. These were originally designed to assess growth in clinical settings, not for age estimation, but they have been used as such by physical anthropologists. Because the data from these studies stems from studies concluded decades ago, an immediate concern with secular change can be raised. For example, studies of children from the Fels longitudinal study have found that compared to the earliest cohorts, children now reach skeletal maturity four to five months earlier (Duren et al. 2015). Additional concerns about the applicability of these methods to a forensic population are raised by the socioeconomic and demographic compositions of the study populations.

The most widely used source for long bone age estimation, the Maresh data (1943, 1955, 1970), is based on the growth study conducted by the Child Research Council in the University of Colorado Medical School. Children recruited to the study were reported to be of higher than average socioeconomic status (upper middle-class) and education for the area. They were also noted to be larger and heavier than children from other published growth studies (Maresh and Beal 1970).

Similarly to the Maresh data, the ongoing Fels longitudinal study was used by Gindhart (1973) to produce growth standards for the tibia and radius. The Fels population is well understood, and has enrolled children over several generations of the same family. It was based in southern Ohio, and sampled individuals across all socioeconomic status groups until 1939, at which point the lowest SES group became underrepresented. Additionally, the sample is almost completely white, although the local population at the time of data collection included 25% Black citizens (Roche 1992).

Another data source published by Anderson et al. (1964), relied on a growth study conducted by the Harvard Longitudinal Studies on Child Health and Development. This group was exclusively middle class families, whose parents could afford to pay hospital fees associated with the clinic through which they were recruited. Inclusion in the group required that the family be of predominantly Northern European descent, although one grandparent of each of the parents were allowed to be from other parts of Europe. About half of the parents had completed high school, and a quarter had pursued some type of post-secondary education (Stuart et al. 1959).

The last reference data used for age estimation was published by Ghantus (1951), who used the Brush Inquiry longitudinal growth study data and expressly excluded the small percentage of black participants. The Brush sample includes children recruited from various local public and private schools, and included one or two visits from children across the socioeconomic status groups. However, the yearly follow up compliance rate was much higher in the middle and high SES groups (Nelson et al. 2000).

The samples on which the previous age estimation “methods” are based do not accurately represent a modern forensic sample. Reference data published using longitudinal studies are generally based on middle or high socioeconomic status groups, likely due in part to the fact that inclusion in a longitudinal study requires frequent visits to a clinician at regular intervals. In the case of the Colorado study, children were even noted to be larger than published values for other groups. Because these data were published mainly for clinical use, in order to diagnose abnormal growth, researchers sought to include only “normal” children. Longitudinal studies were also conducted in the middle of the last century, and child development has accelerated since that time. However, these methods may be appropriate for use in the forensic population today, as they are based on children who matured slowly compared to today’s average children (Duren et al. 2015).

2.1.2. Known age skeletal collections

One way to combat small sample sizes and inter-population variation at the same time is to pool data from multiple collections into one study sample. Recent studies have done this, aggregating data from several collections to create age estimation methods which are sometimes purported to be applicable in forensic contexts.

Cardoso et al. (2014) pooled cemetery samples from 18th and 19th century Britain (Spitalfields and St. Bride's) as well as 19th and 20th century Portugal (Lisbon). Rissech and colleagues (Rissech et al. 2008, López-Costas 2012, Rissech et al. 2013) used a similar cemetery sample from 18th and 19th century Britain (St. Bride's) as well as 19th and 20th century Portugal (Coimbra CEI and Lisbon) (López-Costas 2012, Rissech et al. 2013), and 20th century Portugal and England (Scheuer collection) (Rissech et al. 2008). Facchini and Veschi (2004) used a cemetery sample of Italian children deceased in Bologna in the early 20th century. In each case, these samples were used to conduct regression analyses of the relationship between age and bone length.

Although these methods constitute an improvement over methods based on longitudinal studies of the living, there are problems associated with applying these methods to modern skeletal remains. Both the British and Portuguese samples are composed of children who lived and died before the onset of the major secular changes in growth and maturation which occurred in their respective countries over the last century (Cardoso et al. 2014). Thus, children from these samples are likely smaller than modern children in their respective countries, as are the Italian samples used by Facchini and Veschi (Cardoso et al. 2014). These methods are also based on small samples with uneven age distribution, as mortality is generally high in the first two years of life and much lower throughout the age range. Although they may be inappropriate for contemporary forensic cases in Portugal, Spain, Great Britain or other developed countries, they may be well suited to children in developing nations because of the differences in growth environment.

2.1.3. Recent clinical or forensic samples

Renewed interest in skeletal age estimation and wider availability of imaging techniques have spurred new research using non-traditional reference samples. Attempting to develop new age estimation methods which address the secular increases in height and quicker maturation, Stull et al. (2014) notably accessed over 1300 full body Lodox radiographs of children born mostly after 2000 from a forensic institute and a hospital in South Africa. However, the forensic institute sample represented less than

30% of cases included in the study, and only a small percentage of these were homicide cases. No information was available as to the socioeconomic composition of this sample.

Pfau and Sciulli (1994) collected radiographs from individuals autopsied in Ohio in 1990 and 1991 in order to develop age estimation methods that could be used in the increasing number of cases of skeletonized juvenile remains recovered in the area at the time. The sample included children of all manners of death autopsied at the Franklin County Coroner. The youngest (0-2 years) and oldest (12-20 years) age ranges are overrepresented compared to the other ages, but this method constitutes the only age estimation method based on sample collected purely from forensic institutions.

Tsai et al. (2016) recently published new skeletal age estimation based on diaphyseal measurements of the fibula. The sample was composed of radiographs of infants who had been evaluated for suspected child abuse, but excluded all confirmed abuse cases, as skeletal development in these children could be impaired. This unfortunately yields a method that expressly excludes the cases included in the forensic population.

Of the above diaphyseal age estimation methods, which represent the large majority of methods available to forensic anthropologists, only two research teams incorporate the forensic population in their sample: Stull et al. (2014) and Pfau and Sciulli (1994), and both did so inadvertently. Neither team reflected on the composition or nature of the forensic population on which these methods were prescribed to be used. Both methods are developed on contemporary children, although Pfau and Sciulli's sample is now potentially outdated. Both samples likely include a large proportion of accident victims, who may not present growth patterns adequate for extrapolation to a homicide population. The bulk of recently published skeletal age estimation methods available to researchers are thus potentially inappropriate for use, while older methods or reference data may actually now be more applicable to forensic use, as they are based on a slower maturing, shorter set of children.

2.2. Target sample: composition of the forensic population

The forensic population, the group of skeletonized remains which are examined by forensic anthropologists, is poorly understood. Little information about this group of cases is available, likely because these are highly sensitive data of interest to relatively few researchers in comparison to epidemiological metadata. The forensic population, which consists of remains recovered in clandestine or suspicious circumstances, is likely a subset of the homicide population, especially for juveniles (Snow and Luke 1970, Cattaneo 2007, Simmons 2007, Ross 2011). While the forensic population is not well defined, the homicide population is frequently studied and its characteristics are relatively well understood. The congruence between the two is not perfect, but understanding the homicide population sheds light on the composition of the forensic population.

Not all segments of the population are at equal risk for homicide, and this trend exists even in the youngest age groups. In high-income, developed nations, there is a consistent inverse relationship between socioeconomic status (SES) and both experience of violence and child mortality (Currie 2007, Sidebotham et al. 2014). In the United States, which has the highest rate of child homicide of the 20 top countries for human development (United Nations Development Program 2015, United Nations Children's Fund 2014), children in the most disadvantaged fifth of the population are at 159% higher risk of homicide than their counterparts in the most affluent fifth (Singh and Kogan 2007). Child homicides are also unevenly distributed geographically: the majority of these events occur in or at the edges of urban centers, which constitute socioeconomically disadvantaged areas (Finkelhor 1997, Finkelhor and Ormrod 2001).

Minority groups represent potentially marginalized populations, who may be subject to varying levels of social disadvantage and therefore potentially more adverse growth environments. They are also at disproportionate risk for child homicide. In Australia for example, 11% of child homicide victims in 2006 were Indigenous Australian, although the Indigenous population represents less than 2% of the overall Australian population (Australian Institute of Criminology 2008). Between 2001 and 2005 in New Zealand, the rate of child homicide for Māori children was 1.34 per

100,000, double the non-Māori rate of 0.60 per 100,000 for the same period (Doolan 2006). Various studies from the United States, where child homicide rates and numbers are extremely high, agree that marginalized and underprivileged minority groups are disproportionally represented in the homicide population (Christoffel 1990, Finklehor 1997, Douglas and Finkelhor 2005, Douglas and Vanderminden 2014).

The homicide population thus is composed of disproportionate numbers of children from disadvantaged contexts. They are likely to be infants or teenagers from urbanized contexts, killed by caretakers in the case of infants, family or close acquaintances in the case of adolescent females, or as part of another crime in the case of adolescent males. While this does not provide direct information as to the composition of the forensic population, that group can be assumed to be a subset of the homicide population. Therefore, the forensic population is likely to be drawn from disadvantaged, low socioeconomic urban backgrounds.

2.3. Effects of socioeconomic status on linear growth

The forensic population is drawn from a lower socioeconomic status group than the general population, and this lower status is associated with a host of adverse health and developmental conditions (Skuse 1989, Bogin 1999, Bradley and Corwyn 2002, Fernald and Grantham-McGregor 2002, Block and Krebs 2005, Chilton et al. 2007, Gilbert et al. 2009, Norgan et al. 2012). Studies of growth in developed and developing nations show that differences in linear growth can occur between high and low socioeconomic groups. Therefore, it is probable that the growth of the forensic sample is compromised in comparison to the average, “normal” population of children.

Studies of linear growth in children of different socioeconomic status (SES) groups commonly find growth differences between SES groups in the developing world (e.g. Christiansen et al. 1975, Julia et al. 2004, Sereebutra et al. 2006, Abubakar et al. 2008), and these findings are echoed in developed nations. Early nationwide studies conducted in the 1980s consistently found that young adults and children from lower SES groups were shorter than their higher SES counterparts (Garn and Clark 1975, Bielicki et al. 1981, Garn et al. 1984, Lasker and Mascie-Taylor 1989, 2005). This difference is present starting at birth, when children are matched for birthweight (Garn et al. 1984), and continues throughout the growth period (Lasker and Mascie-Taylor 1989). In a sample of British children born in 1970, who were stratified into five levels of SES based on occupation of the father, boys from the highest SES group were 11cm taller than boys from the lowest SES group (Lasker and Mascie-Taylor 2005). Not only are children from lower SES groups shorter for age, their maturation is slowed: boys from lower SES groups were reported to continue growing linearly until their mid-twenties (Lasker and Mascie-Taylor 1989).

Recent studies tend to find much subtler differences between socioeconomic status groups than previous work, although they are also conducted on smaller geographic scales. In Hamilton, Ontario (Canada), Moffat and colleagues (2005 and 2007) found a statistically significant difference in height for age between high and low SES schools. Although the difference in mean height for age Z-score between the schools was small (0.29), it was significant ($p<0.05$). In a homogenous poor rural Appalachian community, Crooks (1999) found high rates of low height for age and stunting, as well as an association between some indicators of SES and height for age. This finding was echoed in an urban American setting by Grimberg and colleagues (2009), who found that children insured by Medicaid, an indicator of lower SES, were more likely to exhibit growth faltering. Growth outcomes have also been assessed for the effects of poverty. In the Quebec Longitudinal Study, children who had lived through two episodes of poverty were more likely to exhibit growth delay then their peers who had gone through one, or no episodes of poverty (Ehouxou et al. 2009). The timing of these events was important: a very recent poverty event showed no effect on growth, and a single distant event showed some impact on growth, which was less severe and consistent than the impact from multiple events (Ehouxou et al. 2009). However, these studies are conducted on a small geographic scale as opposed to previous nationwide studies, and it is possible that children in these contexts are all affected by socioeconomic status. For example, 78% of the children in Crooks' (1999) study qualified for free or reduced priced lunch, meaning that SES differences in anthropometrics were found even within a largely homogenously low SES community.

Socioeconomic differences in growth provide indirect evidence for compromised growth in the forensic population. More direct evidence comes from cases where growth outcomes have been assessed for the effects of child abuse and neglect. Karp and colleagues (1989) found twice the prevalence in stunting in abused children as opposed to non-abused children (11.6% versus 5.6%) in an already underprivileged city in New Jersey (U.S.A), although it was not as strong as the difference in wasting. In a sample of 260 victims of child abuse, Taitz and King (1988) found 39 children who were below -2SD expected height for age, which means that 15% of children fell below the 3rd percentile of normal height variation. Of 91 abuse victims evaluated, Wales et al. (1992) found 31 cases (about 34% of cases) where the victim was considered "short." The effects of abuse and neglect compound over time: Elmer and Gregg (1967) found that abuse victims with longer records of medical visits were more likely to show compromised growth. Although not all juvenile skeletal remains examined by forensic anthropologists will necessarily be abuse or neglect cases, these studies illustrate potential differences in growth status between the average and forensic populations.

Chapter 3. Materials and Methods

3.1. Materials

The present study was conducted retrospectively. Data were obtained from coronial bodies in the Australia, New Zealand, and the United States. Australian states which contributed to the sample were the Northern Territory, Australian Capital Territory, Victoria, Tasmania, Queensland, and New South Wales. New Zealand is represented by a single national coronial institution. The coronial institutions from the United States included in the sample are Cuyahoga County (Ohio), New Mexico, and New York City. In the case of New Mexico and Cuyahoga County, data were extracted from the case management system of the coronial body. For all other bodies, data were obtained directly from autopsy reports.

For each individual, the data obtained included date of birth and death, manner and cause of death, and cadaver length and weight measurements. For each coronial body, deaths which occurred during gestation or which were due to complications from prematurity were excluded, as the factors which affect gestational growth differ from those which affect post-natal growth. Cases where chronic disease impacted growth were excluded from the sample. Also excluded were cases where significant thermal damage or decomposition was noted in the autopsy notes, because both types of post-mortem alteration to the body could influence the accuracy of the measurement of cadaver length.

The sample ranged from birth to 18 years at death. The final data set was comprised of 1256 individuals, having died from homicides (n=499) or accidents (n=757) between 2000 and 2015, which were collected from five coronial institutions. The sample was then divided into four age groups based on changes in growth velocity according to Bogin (1999). Infancy (birth to 3 years) is categorized by a period of high, although declining, growth velocity. Childhood (3 to 7 years) is characterized by a low growth velocity, marked by a growth spurt at the exit of the period. It is followed by the juvenile period (7 to 10 years in females and 7 to 12 years in males), which consists of the period between the late childhood growth spurt and puberty. Adolescence, (10 to 18 years in females and 12 to 21 years in males) begins at puberty and ends with the completion of development. For a sex and age profile of the total sample, see Figure 3.1.

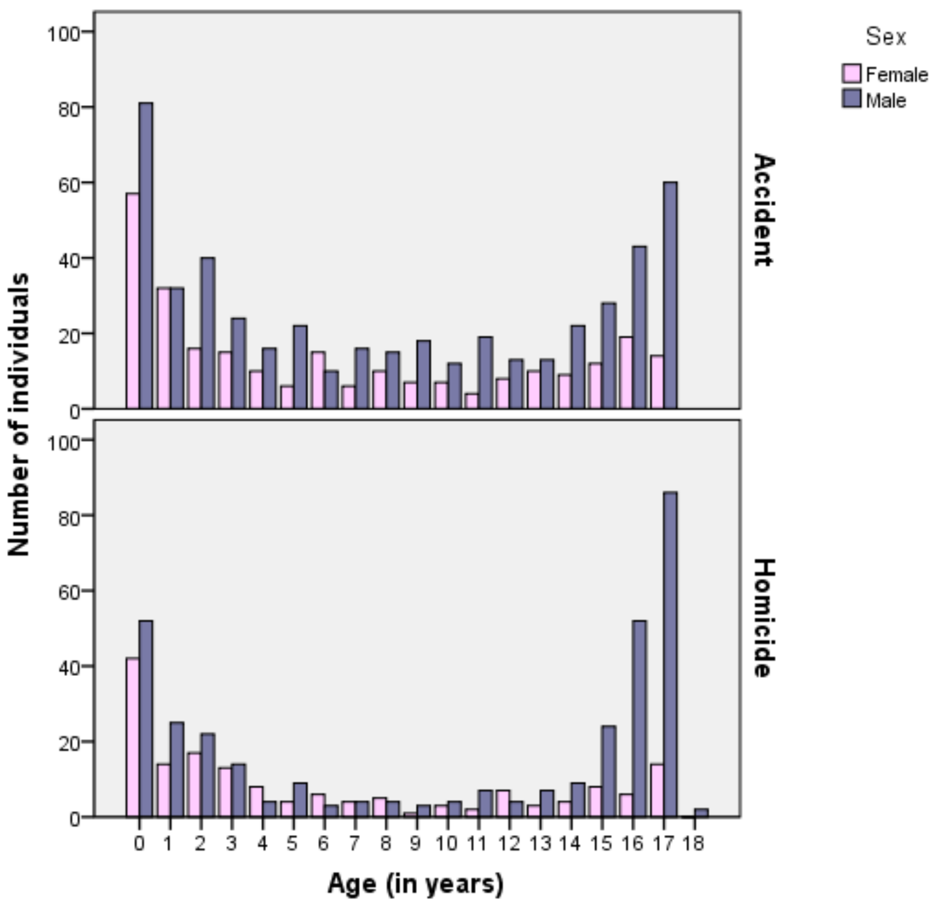

Figure 3.1: Age and sex distribution of the complete sample

3.1.1. State Coroner's Courts, Australia

Australian data was accessed through the National Coronial Information System (NCIS), which is a database designed to manage and store information and documents related to deaths reported to coronial bodies in Australia and New Zealand. Data was collected from the following states: Northern Territory, Australian Capital Territory, Victoria, Tasmania, Queensland, and New South Wales, and were pooled together to represent Australia as a whole. Data were recorded directly from autopsy record for child deaths occurring between 2000 and 2015. Cases were randomly selected.

The final dataset included 185 accidental deaths (57 females and 128 males) and 87 homicide deaths (32 females and 55 males), for a total of 272 cases (89 females and 183 males), primarily from the more populated states of Queensland, Victoria, and New South Wales (age distribution available in Figure 3.2).

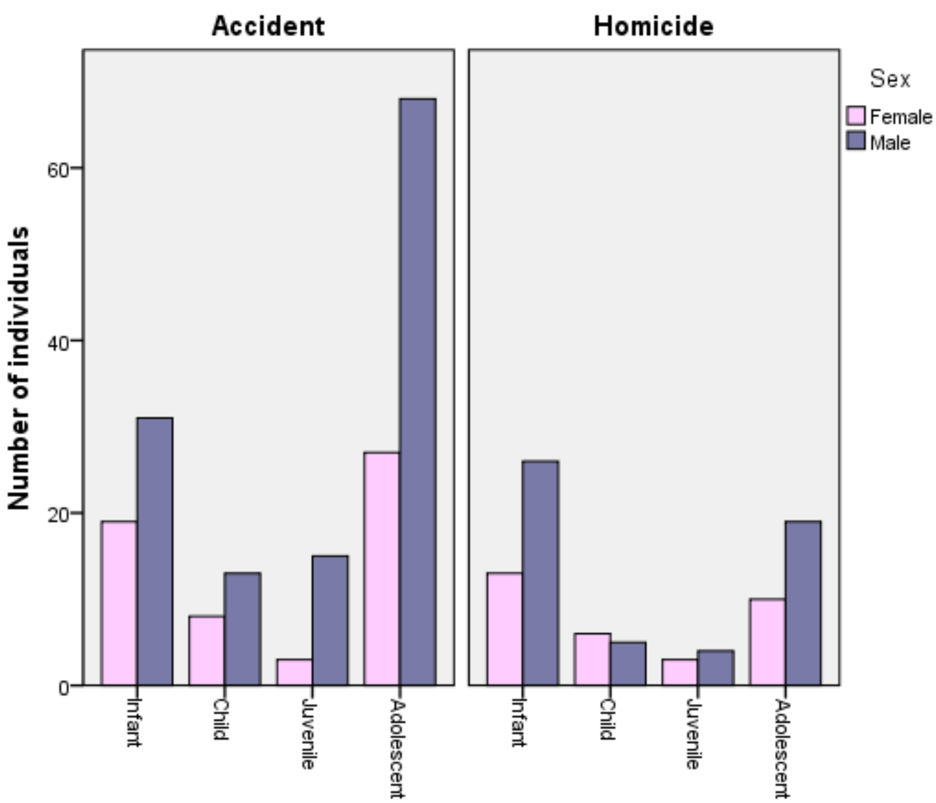

Figure 3.2: Distribution of cases from Australia, by age category and manner of death

Compared to the socioeconomic characteristics of New Zealand and the United States, Australia sits in the middle. It has the lowest unemployment rate of the three, but both its Gini coefficient of inequality and poverty ratio are between the lows of New Zealand and the highs of the United States (Table 3.1).

Table 3.1 Economic and social characteristics of Australia, New Zealand, United States, state of New Mexico, New York City, and Cuyahoga County compared

	Australia.	New Zealand	U.S.A	New Mexico	New York City	Cuyahoga County
OECD unemployment rate (%)*	5.2	6.5	9.6	-	-	-
OECD Gini coefficient of inequality**	0.335	0.327	0.379	-	-	-
OECD poverty ratio***	0.149	0.099	0.179	-	-	-
U.S. unemployment rate (%)	-	-	7.9	7.2	8.8	10.5
U.S. Gini coefficient of inequality	-	-	0.469	0.464	0.535	0.485
% of families with children under 18 who fall under poverty line	-	-	15.7	15.7	23.2	19.9
% of families which received SNAP support in past year	-	-	9.3	10.0	15.8	13.1
% of families with children led by a single female	-	-	24.08	14.00	18.70	36.52

Sources: * (2010) OECD (2016), Unemployment rate (indicator). doi: 10.1787/997c8750-en (Accessed on 03 March 2016)
** (2008-2010) average of all values available over the period of 2008-2012
*** (2012, first year when all three available)
All other indicators: U.S. Census Bureau, 2006-2010 American Community Survey and U.S. Census Bureau, 2010 Census

3.1.2. Coroner's Court of New Zealand

The New Zealand Coroner's Court handles the entirety of sudden and unexpected deaths occurring in New Zealand. Cases were accessed through the NCIS, and data were recorded from records of autopsies of children occurring between 2007 and 2015. All available homicides were recorded, and a random sample of accidents was selected.

The final sample included 97 accidental deaths (26 females and 71 males) and 45 homicides (20 females and 25 males), for a total of 141 cases (45 females and 96 males). The age distribution of the sample is available in Figure 3.3.

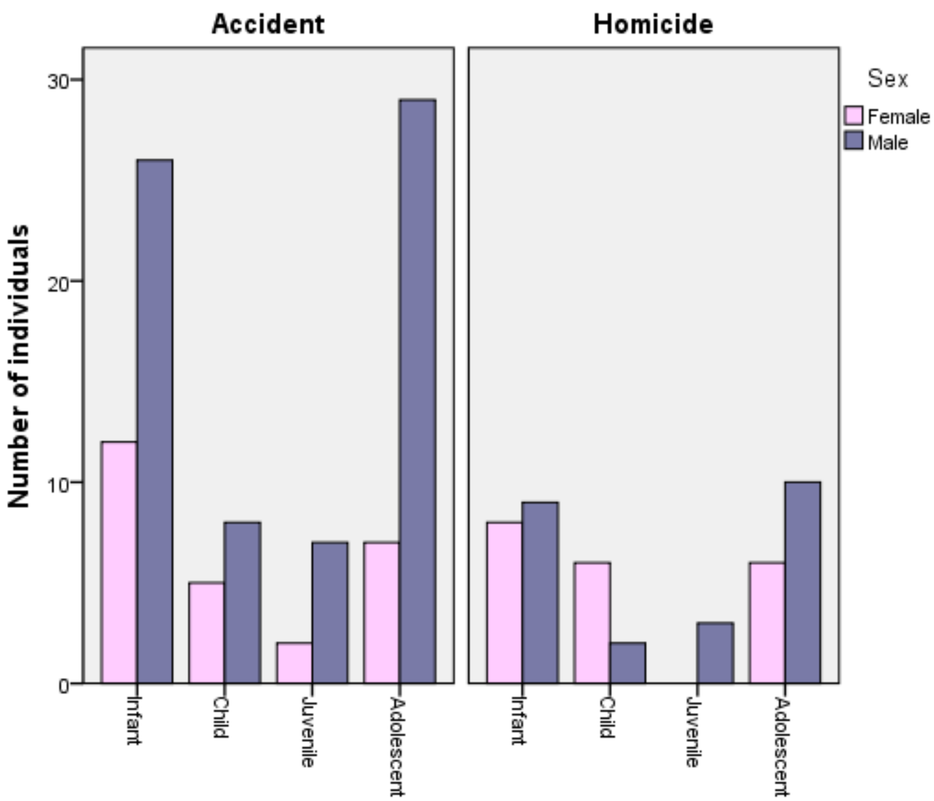

Figure 3.3: Distribution of cases from New Zealand, by age category and manner of death

Compared to the socioeconomic characteristics of the United States and Australia, New Zealand has both the lowest Gini coefficient of inequality and the lowest poverty ratio. New Zealand has higher unemployment than Australia, but these figures are closer to each other than to unemployment in the United States (Table 3.1).

3.1.3. Office of the Medical Investigator, New Mexico, U.S.A.

The Office of the Medical Investigator (OMI) operates within the University of New Mexico, Albuquerque, and is responsible for managing all cases of sudden or unexpected deaths in New Mexico, exclusive of federal or tribal land. Cases were extracted from the case management system for juvenile homicides occurring between 2008 and 2013.

The resulting sample includes 61 accidental deaths (22 females and 39 males) and 51 homicides (19 females and 32 males) for a total of 112 individuals (41 females and 71 males). The sample includes deaths of children aged 0-12 years inclusive at death (see Figure 3.4 for the sample's age distribution).

Compared to the United States as a whole, New Mexico does slightly better on some indicators (lower unemployment rate, lower Gini coefficient, lower percentage of families led by a single mother), but slightly worse on others (lower median household income, higher percentage of families receiving SNAP support), and both geographies have the same poverty rate. Although it is not very different from the United States as a whole, New Mexico does better than Cuyahoga County on all indicators (Table 3.1).

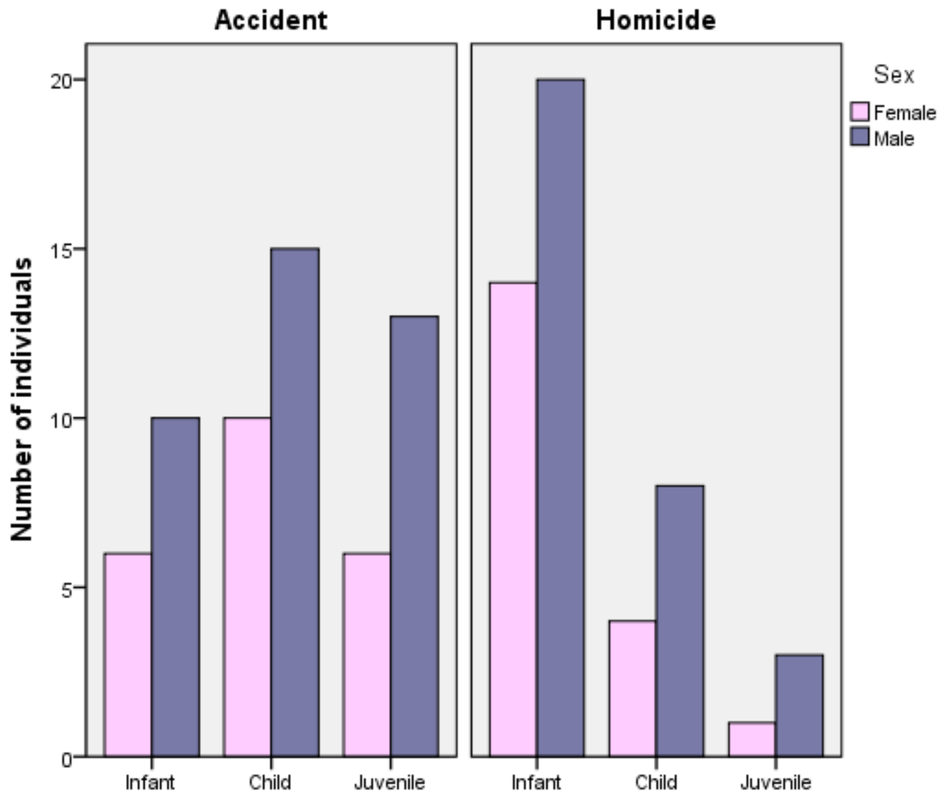

Figure 3.4 Distribution of cases from New Mexico, by age category and manner of death

3.1.4. New York City Office of the Chief Medical Examiner, New York, U.S.A.

The New York City Office of the Chief Medical Examiner (OCME) is responsible for the forensic investigation of all suspicious or unexpected deaths occurring throughout the five boroughs of New York City. Data was recorded directly from records of autopsies of children occurring between 2007 and 2014. Cases were randomly selected, and a further ten homicide in the ranges of 7-12 years were selected as the random sample did not include many children that age group.

Excluding cases where cadaver length could not be obtained, data were collected from 183 accidental deaths (71 females and 113 males) and 154 homicides (35 females and 119 males), for a total of 337 cases (115 females and 222 males). The age distribution of the sample is described in Figure 3.5.

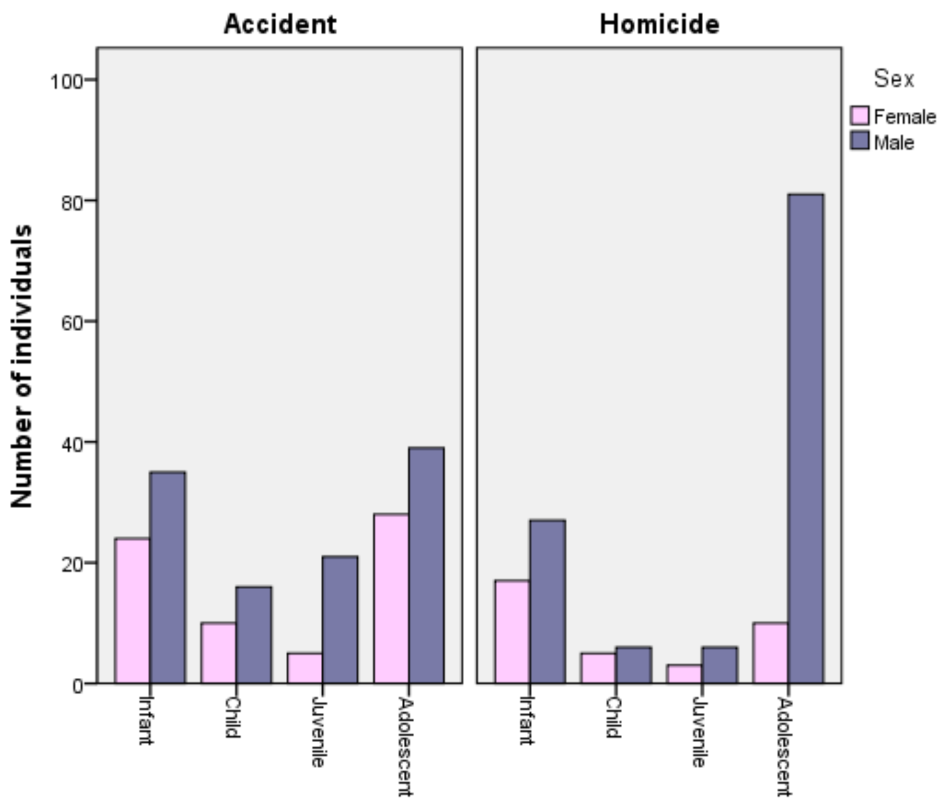

Figure 3.5 Distribution of cases from New York, by age category and manner of death

New York City fares worse than New Mexico on all socioeconomic indicators. It also fares worse than Cuyahoga County in most socioeconomic indicators: it has higher inequality, a higher percentage of families living below the poverty line, a higher percentage of families receiving SNAP assistance, but Cuyahoga County has higher unemployment and a higher percentage of families led by a single female (Table 3.1).

3.1.5. Cuyahoga County Medical Examiner, Ohio, U.S.A.

Case data from Cuyahoga County Medical Examiner's office, whose jurisdiction encompasses the city of Cleveland and its surrounding areas, contributed a major dataset.

Information was extracted from the case management system for autopsies of juveniles occurring between 2002 and 2014.

The final sample included 232 accidental deaths (89 females and 143 males) and 161 homicides (55 females and 106 males), for a total of 393 individuals (144 females and 249 males). The age distribution of the sample is described in Figure 3.6.

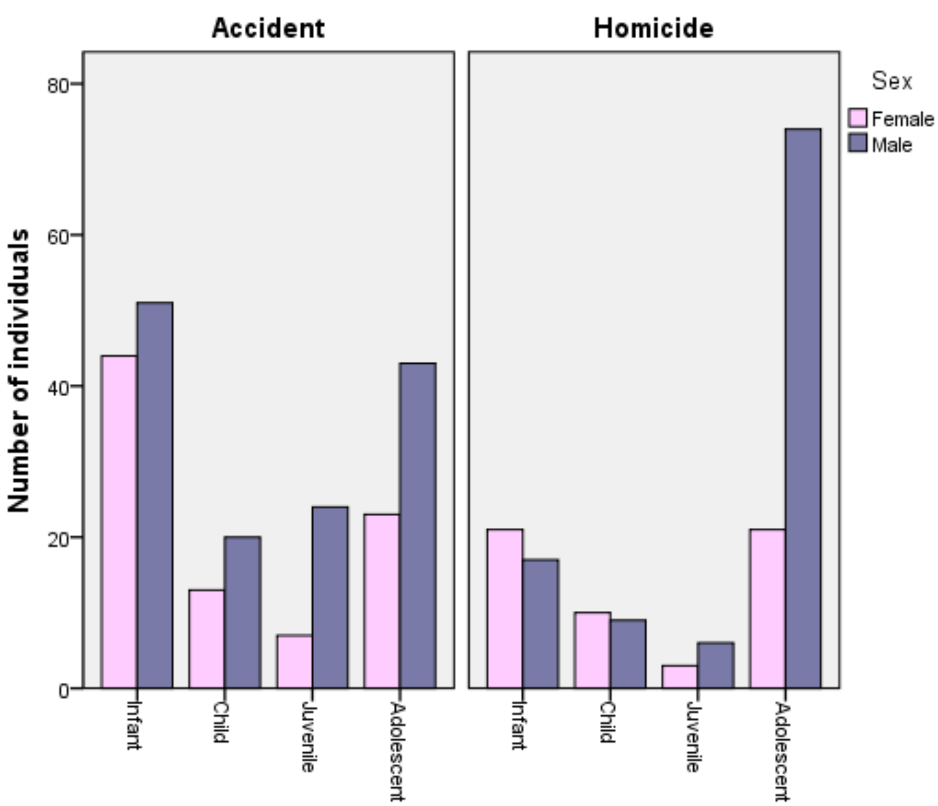

Figure 3.6: Distribution of cases from Cuyahoga County, by age category and manner of death

Compared to the other samples from the United States, Cuyahoga County does better in some respects, and trails in others. It has the highest unemployment rates and highest percentage of families led by single females. However, the rest of its indicators are better than New York City.

3.1.6. Ethical Review

Previous to data collection, ethics approval was obtained from the Simon Fraser University Office of Research Ethics (ORE). Permission to access data was then sought from all coronial institutions, which involved additional ethics review by an internal or external body in addition to ORE approval.

3.2. Methods:

Cadaver length is a measurement of linear growth, and was used as a proxy for measurements of the developing long bones. Once the sample was collected, each case was compared to the Center for Disease Control (CDC-2000) growth reference using Z-scores (Kuczmarski et al. 2000). Cases with potentially unreliable information were excluded. This included some instances of gross incongruence between recorded length and weight, as well as cases with extreme length/height for age Z-scores, which may indicate data entry errors.

3.2.1. Preliminary analysis of differences using ANCOVA

An ANCOVA was used to test whether there were any differences in mean cadaver length between the manner of death groups, with age in months as a covariate. Because the reference parameters employed for the infants are different than those used for older children, one ANCOVA was run for the infant category and a separate one for all other age groups. A separate set of ANCOVAs were used to test whether there were any differences in mean cadaver length between the sexes and the coronial samples, with age as the covariate.

3.2.2. Using Z-scores to assess growth status in samples

Deviation from mean length/height for age and sex was quantified for each individual in the sample using Z-scores, and the LMS parameters for age (in months) and sex provided by the CDC-2000 growth reference (Kuczmarski et al. 2000). By calculating the deviation from the mean for each age and sex group, the Z-score effectively controls for these factors and allows groups of individuals to be pooled and compared irrespective of age or sex. Summary statistics can then be calculated using groups of Z-scores, making these ideal for use in population level growth studies (Jorge and Doris 2014).

In clinical assessments of individuals and populations, children are directly compared to growth references to evaluate individual nutrition status and population health. Selecting the right growth reference is crucial in these situations, as different comparative samples will yield different evaluations of deviation from normal (Eckhardt and Adair 2002, Jorge and Doris 2014). When two populations are compared against each other, however, the growth reference matters less: the obtained Z-scores are compared against the other population rather than against the reference. The growth reference selection criteria become methodological, for example ease of Z-score calculation, sampling strategy used, and smoothing procedures employed. Likewise, because individual measurements are not being compared directly to the growth reference, cadaver length can be converted to height for age z-scores and compared between the accident and homicide groups without correction to stature or even without special considerations about selection of growth reference.

3.2.3. Choosing a growth reference

Regionally specific reference samples were not necessary in this study, as the goal was not to understand whether groups follow typical growth patterns but rather to investigate differences in growth between the homicide and accidents groups. Additionally, using one reference allows for direct comparison between all of the samples. Three major growth standards are currently available for use. They have been published by: the National Center for Health Statistics (NCHS-1977), the Center for Disease Control (CDC-2000), and the World Health Organization (WHO-2006).

The NCHS-1977 reference is composed of data from the National Health Examination Study (NHES), supplemented by the Fels longitudinal study (Hamill et al.

1977). It was prepared as an update on growth standards from the 1940's. The Fels data was collected between 1929 and 1975, which comprised the reference sample for children from 0-36 months. The NHES data was compiled between 1963 and 1974, and comprised the reference sample for children from 2-18 years (Hamill et al. 1977). At the time, these were recommended by the World Health Organization (WHO) for use in monitoring the nutritional status of children internationally (World Health Organization 1983). The advantage of this using this reference would mainly be that it incorporates data from similar cohorts as those included in the Fels longitudinal study data, which was used by Maresh (1943, 1955, 1970) to develop a record of average long bone length for age. After dental development, the Maresh data are one of the main methods recommended for age estimation by forensic anthropologists (Schaefer et al. 2008, Cunha et al. 2009, Franklin 2010).

The CDC-2000 reference was compiled because NCHS-1977 had become inadequate for use. In addition to concerns with secular change, which had accelerated development and increased final stature, methodological concerns were raised. Problems with NCHS-1977 were largely due to the exclusive use of the Fels data for the 0-36 month sample: because the data was collected beginning in 1929, it did not reflect the more modern ranges of birth weights (Kuczmarski et al. 2000, Johnson et al. 2012). Concerns were also raised because the Fels data collection was restricted to a small area of Ohio, where the population was homogenously white, middle-class, formula-fed infants, and because data was collected every three months rather than monthly. Additionally, because the 0-36 month and 2-18 year data were collected from entirely different samples, the 2-3 year group was inconsistently assessed by the two references (Kuczmarski et al. 2000).

Data for the CDC-2000 reference was obtained primarily from two cycles of the National Health Examination Survey (NHES) as well as three cycles of the National Health and Nutrition Examination Survey (NHANES), which occurred between 1963 and 1994 (Kuczmarski et al. 2000). In addition to a more nationally representative sample, the reference also uses more advanced statistical smoothing methods, namely an adapted version of the LMS method (Cole 1990, Kuczmarski et al. 2000). This smoothing method reduces the distortion of Z-scores outside the range of -3 to 3 (Kuczmarski et al. 2000).

In the 1990s, the World Health Organization commissioned the creation of growth standards, forming the Multicenter Growth Reference Study Group (de Onis and Garza 2006). It incorporated measurements of well-fed, healthy children from countries around the world, taken over a 6 year period. These measurements culminated in the WHO-2006 reference sample (WHO Multicenter Growth Reference Study Group 2006a). The sample itself consisted of an international longitudinal study of a group of breastfed children from birth to 24 months, combined with cross-sectional data of children from 18-71 months, all of which were described to be of relatively high socioeconomic status, with growth therefore "unconstrained" by environmental factors (WHO Multicentre Growth Reference Study Group 2006b).

The World Health Organization then explored the possibility of creating a growth reference for children aged 5-19 (de Onis et al. 2007a). Acknowledging the difficulty of controlling the growth environment of older children, they sought to compile existing datasets from multiple countries. Differences in methodology between these studies were too great, and the committee opted to use data collected for the NCHS-1977 reference, supplemented by the WHO-2006 data, to ease the transition from the under 5 years reference to the new curves (de Onis et al. 2007a). Thus, the WHO-2007 reference does not represent a new dataset, but rather one smoothed to fit the WHO-2006 reference at the point of 5 years of age.

Naturally, the multiple references available for use have prompted investigation into the differences between them. Comparing the CDC-2000 with the WHO-2006 references yields the information that the studies vary mainly in weight for age, where the CDC children are heavier, and length/stature for age, where the WHO children are slightly taller and yielded tighter distributions (de Onis et al. 2007b). Differences in the weight for age references are not important to this study. However, greater variability in height for age may be advantageous, as it should reduce the number of children who fall considerably outside the Z-score range of -3 to 3, after which there can be some distortion in calculations.

Comparison of the NCHS-1977 and CDC-2000 growth references found inconsistencies in calculating length/height for age for very small individuals (Eckhardt and Adair 2002). Stunting prevalence was found to be higher in infants when using the NCHS-1977, and inconsistent after 8.5 years of age. Eckhardt and Adair (2002) hypothesize that this is partially due to the updated reference sample, as well as the improved statistical methods employed. Likely, the larger variation in length for age in the CDC-2000 sample over the WHO-2006, due to data collection techniques (de Onis et al. 2007), contributes to the lower rates of stunting observed for infants.

After considering all of the methods available, the CDC-2000 reference was selected for its simplicity of use and the fact that data was readily available for calculating Z-scores. It presented the additional advantages of having separate growth charts for recumbent length (birth to 36 months) and stature (2-18 years), information to calculate Z-scores for each month over the age range, as well as an improved smoothing method and Z-score calculation procedure that minimized the distortion of Z-scores at extremes of the distribution.

The CDC-2000 growth charts are split into two: 0-36 months, during which recumbent length is measured, and 2-18 years where standing height, or stature is measured (Kuczmarski et al 2000). Growth during the second year of life can therefore be assessed using either reference. As the age categories employed in this study classified all children aged 0-3 years in one group, the recumbent length reference was employed for all children aged 0-3 years old, so as to keep continuity through the infant group.

3.2.4. Extreme Z-scores

In each coronial sample, several individuals yielded extreme height for age Z-scores which were biologically implausible. This study was retrospective in design, rendering transcription errors highly possible, and the presence of mistakes was likely.

The World Health Organization recommends a cutoff range of +/- 6.0 Z-score units from 0 for identifying potential data entry errors (World Health Organization 2006). However, for this study, the appropriate Z-score inclusion range should not exclude children presenting atypical growth, such as malnourished children. To better understand the range of Z-scores which would reasonably include these small for age children and exclude potential errors, a review of fatal child neglect cases was conducted. Cases where cadaver length and weight were reported were selected, and Z-scores for these individuals were calculated using the CDC-2000 reference (Wehner et al. 1999, Mimasaka et al. 2000, Fieguth et al. 2002, Kellog and Lukefahr 2005, Madea 2005, Solarino et al. 2012, Dettemeyer et al. 2014, Yamaoka et al. 2015). A total of 20 records of individuals aged 1 month to 8.5 years at death were used for this. Resulting Z-scores ranged from -5.67 to 0.08 (Table 3.2).

Table 3.2: Three highest and three lowest Z-scores for clinical reported cases of fatal child neglect reported in the surveyed literature

Case number	Source	Age	Sex	Z-score
1	Fieguth et al. 2002	3.5 years	Female	0.08
2	Solarino et al. 2012	1.3 years	Female	-0.89
3	Yamaoka et al. 2015	1.5 years	Female	-0.25
4	Mimasaka et al. 2000	6 years	Female	-5.60
5	Yamaoka et al. 2015	1.3 years	Female	-5.42
6	Wehner et al. 1999	5.9 years	Male	-5.67

This review of cases indicates that in cases of extreme child malnutrition and starvation, Z-scores as low as -5.67 can be expected. Thus, the 2006 criterion of +/- 6.0 Z-score units from 0 was in agreement with case findings, and was selected for use in this project. This selection criterion excluded 17 cases out of the total sample of 1256 cases. Only two cases were homicides, and the rest were accidents. Of the 17 cases, 12 were infants. Of the remaining 5 cases (3 children, 1 juveniles and 1 adolescent), three were deaths due to traffic accidents and two were due to unspecified blunt trauma. It is possible that these mechanisms altered or disfigured the body, leading to distorted or estimated rather than measured cadaver lengths, or that they are measurement errors.

3.2.5. Validating the accidental group as representative of normal growth

A key assumption of this study is that the accidental deaths accurately represent the growth of the "normal" population, on which forensic methods are typically developed. However, researchers are increasingly aware that fatal accidents are not evenly distributed across the socioeconomic gradient (Singh and Kogan 2007). Although no socioeconomic data was collected for individual cases, one potential way to assess whether the accidental sample reflects the average population is to compare the sample directly to the growth reference to detect deviation from 0 in Z-score distribution. However, this comparison is complicated by the nature of the collected data, which consists of cadaver length, while the reference consists of recumbent lengths from birth until 36 months, and statures after 36 months. The infant cadaver lengths are directly comparable to the reference, but the cadaver lengths of the other age groups must be corrected into stature before direct comparison to the growth reference.

In adults, recumbent length is known to be greater than standing height (stature) by somewhere between 1 and 4 centimeters on average depending on stature, sex, age, and time of the day (Tyrell et al. 1985, Bidmos 2005, Cardoso et al. 2016). In children, especially in the youngest, an error of a few centimeters can lead to large difference in Z-score. Therefore, using the appropriate correction is essential to understanding whether the accidental sample reflects the growth of the average population. Studies of the relationship between recumbent length and stature are few in children, but all note that the difference between the two increases throughout the growth period (Krishan and Sharma 2002, Buyken et al. 2005). Cadaver lengths of the accident sample were converted to a corresponding stature using a regression method (Krishan and Sharma 2002), except in the case of infants, where the growth reference is composed of recumbent length, and therefore the two measures already match. The Z-score for each accident case was then re-calculated, and the distribution of Z-scores for each collection was compared to 0 using a one sample t-test.

To assess the effect of different correction methods on the assessment of deviation from the reference, cases of accidental deaths of males 17-18 years old, in which linear growth has likely already completed, were selected from each coronial sample. Krishan and Sharma's (2002) regression equation yielded a difference between recumbent length and stature of about 2 cm at the completion of growth. Although adult cadaver length has traditionally been considered to be 2.5 cm longer than standing height (Trotter and Gleser 1952), recent research has suggested that this might be a minimum rather than a mean, and that the difference between cadaver length in an adult male sample (average cadaver length of 169.5 cm) and stature is closer to 4.3 cm (Cardoso et al. 2016). Stature was estimated using two regression methods (Krishan and Sharma 2002, Cardoso et al. 2016), as well as the most commonly used correction factor (Trotter and Gleser 1952). Height for age Z-scores were calculated, for each individual for each of the three obtained statures, and each group of Z-scores was compared against zero using one sample t-tests.

3.2.6. Exploring differences in Z-scores between homicides and accidental deaths

The distributions of Z-scores for each manner of death group within each age group and coronial sample were assessed visually for deviation from normality using histograms and Levene's tests of homoscedasticity were used to test whether variances were equal between the Z-score distributions of the manner of death groups in order to select either the parametric Student's t-test or the non-parametric Mann-Whitney u-test. Differences in mean Z-scores by manner of death groups were explored throughout the age range by using age categories previously described using independent sample t-tests. This was done to reflect both the biologically and socially related shifts in child mortality across the life span, as well as the biological mechanism of growth impairment over the age range, for example how environmental insults on growth compound over time, resulting in increasingly impaired growth in children exposed to chronic hardship as they age.

Chapter 4. Results

4.1. Exploring differences in height for age between homicides and accidental deaths

4.1.1. ANCOVA

ANCOVA results show that length of the infants differed significantly between the coronial samples (SAMP, p=0.00), but not the manners of death (MOD). In the other age groups, height differed significantly between the manner of death groups (p=0.03) (Table 4.1). This suggests that in the older individuals, there is a difference in height for age between accident and homicide victims. This is not the case in the infant group, where the main differences in length for age exist between samples.

Table 4.1: F and p-values for ANCOVAS testing the effects of sample (SAMP) and manner of death (MOD) on length/height for age Z-scores

	Infants		Other groups	
	F	p	F	p
SAMP	3.94	0.00	1.33	0.26
MOD	0.01	0.94	4.57	0.03
SAMP*MOD	1.06	0.38	1.51	0.20

A second set of ANCOVA tests were used to examine differences in height for age between the coronial samples and sexes. A significant effect of sample of length for age of infants (p=0.00) was found, and no significant effect of sex was found although the p-value was borderline significant at the 0.10 level (p=0.11) (Table 4.2). There was no significant effect of sex or sample in the older age categories.

Table 4.2: F and p-values for ANCOVAs testing the effects of sample (SAMP) and sex (SEX) on height for age Z-scores

	Infants		Other groups	
	F	p	F	p
SAMP	5.76	0.00	0.727	0.57
SEX	2.64	0.11	0.01	0.91
SAMP*SEX	1.00	0.41	0.17	0.20

4.1.2. Validating the accidental group as representative of normal growth

Values for the one sample t-tests between the height Z-scores obtained using Krishan and Sharma's (2002) method and zero, for each age group and coronial sample in the accident victims are found in Table 4.3. The groups which have mean height for age Z-scores below zero are: Australian children and adolescent, infants from New Mexico, adolescents from New York City, and infants from Cuyahoga County. However, only the Cuyahoga County infants deviate significantly from the reference mean (p=0.00). Three more groups approach significant deviation from zero at a p level of 0.10: New York City adolescents (p=0.08), New Mexico children (p=0.10), and Cuyahoga County children (p=0.11).

Table 4.3: Test and p-values for one sample t-tests of height for age Z-score against zero for the accidents in each age group from each coronial sample. Standing height was estimated for the children, juveniles, and adolescents using Krishan and Sharma (2002).

	Infant		Child		Juvenile		Adolescent	
	t	p	t	p	t	p	t	p
Australia	0.91	0.37	-0.16	0.88	0.24	0.81	-0.80	0.43
New Zealand	0.46	0.65	1.49	0.16	0.30	0.77	0.19	0.85
New Mexico	-0.70	0.50	1.74	0.10	0.79	0.44	-	-
New York	0.53	0.60	0.68	0.50	0.67	0.51	-1.31	0.19
Cuyahoga County	-4.94	0.00	1.66	0.11	1.36	0.18	0.39	0.70

A comparison of one sample t-test values obtained when correcting cadaver length into standing height using three different methods (Krishan and Sharma et al. 2002, Trotter and Gleser 1952, Cardoso et al. 2016) is shown in Table 4.4. Krishan and Sharma (2002) yielded the smallest overall correction, and Cardoso et al. (2016) the largest. While only one coronial sample yielded a Z-score distribution which deviated significantly from zero (Cuyahoga County), two groups of mean Z-scores emerged: Australia and New Zealand and New York and Cuyahoga County. The second group showed means between -0.46 and -0.72, depending on the correction method, and the first had means between 0.07 and -0.25 (Table 4.6).

Table 4.4: Comparison of the mean Z-scores and t and p-values for one sample t-tests against zero when cadaver length is converted to standing height using three different methods.

	Krishan and Sharma (2002)			Trotter and Gleser (1952)			Cardoso et al. (2016)		
	Mean	t	p	Mean	t	p	Mean	t	p
Australia N=28	-0.08	-0.34	0.74	-0.17	-0.68	0.51	-0.25	-0.85	0.40
New Zealand N=9	0.07	0.20	0.85	-0.01	-0.04	0.97	-0.07	-0.17	0.87
New York N=8	-0.49	-0.70	0.50	-0.58	-0.82	0.44	-0.72	-0.87	0.41
Cuyahoga C. N=15	-0.46	-1.81	0.09	-0.55	-2.31	0.05	-0.68	-2.31	0.04

4.1.3. T-Tests by age group

Group Z-score means for all sample and manner of death groups within the infant age category ranged from -0.83 to 0.42, and from -0.42 to 0.71 for all other age categories (Table 4.5). The mean Z-score for the accident group is generally larger than that of the homicide group, except in six cases: Australian adolescents, New Zealand infants, children and juveniles, New Mexico infants, and Cuyahoga County infants. Comparisons where the homicide group shows smaller height for age Z-scores compared to the accident group sometimes reach statistical significance, while comparisons where

Table 4.5: Number of individuals (N), mean and standard deviation (SD) of length/height for age Z-scores by manner of death within each age category and sample

	Infant			Child			Juvenile			Adolescent		
	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD
Australia												
Accident	50	0.18	1.40	21	0.09	1.55	18	0.28	1.60	95	0.14	1.20
Homicide	39	-0.11	1.56	11	0.31	0.87	7	-0.10	0.70	29	0.33	0.87
New Zealand												
Accident	38	0.12	1.44	13	0.70	1.35	9	0.71	1.27	36	0.27	0.90
Homicide	17	0.42	1.57	8	0.71	1.06	3	1.01	1.31	16	-0.07	1.02
New Mexico												
Accident	16	-0.33	1.92	25	0.66	1.48	19	0.48	1.63	-	-	-
Homicide	34	-0.13	1.73	12	-0.31	1.53	4	-0.01	0.75	-	-	-
New York												
Accident	59	0.11	1.56	26	0.38	1.80	26	0.36	1.32	65	-0.07	1.89
Homicide	44	-0.38	1.58	11	0.22	1.75	9	0.18	1.94	91	-0.15	1.09
Cuyahoga County												
Accident	95	-0.83	1.63	33	0.51	1.29	31	0.48	1.20	66	0.30	1.13
Homicide	38	-0.50	1.60	19	-0.42	1.31	9	0.25	1.18	95	0.06	1.12

the homicide group shows a larger height for age Z-score than the accident group never reach significant t-values.

None of the Levene's tests were significant, indicating that t-tests were appropriate for use in all cases to assess whether differences existed between height for age Z-scores of accident and homicide victims across the age groups and samples. In the case of New Zealand juveniles, a Mann-Whitney was used due to small sample sizes, and the standardized Z-test statistic is reported (Table 4.6). Positive t-values indicate that the mean height for age Z-score of the homicide group is smaller than that of the accident group, and negative t-values indicate that the mean height for age Z-score of the homicide group is larger than that of the accident group. Two categories, children in Cuyahoga County and in New Mexico, yielded significantly smaller height for age Z-scores in the homicide group at a p level of 0.10 (Table 4.6). Differences for infants in New York City as well as adolescents in Cuyahoga County also yielded smaller height for age Z-scores in the homicide group but p-values were borderline significant at the 0.10 level.

Table 4.6: F and p-values from independent sample t-tests of height for age Z-scores between accidental deaths and homicides for each age group within each collection

	Infant		Child		Juvenile		Adolescent	
	t	p	t	p	t	p	t	p
Australia	0.91	0.37	-0.44	0.66	0.60	0.56	-0.76	0.45
New Zealand	-0.73	0.47	-0.06	0.96	-0.65	0.52	1.23	0.22
New Mexico	-0.38	0.71	1.85	0.07**	0.58	0.57	-	-
New York	1.57	0.12*	0.24	0.81	0.30	0.77	0.33	0.74
Cuyahoga County	-1.05	0.30	2.51	0.02***	0.51	0.61	1.39	0.17*

*** p<0.05; ** p<0.10; * p<0.20.
New Zealand juveniles were compared using a Mann-Whitney U, the standardized Z-test statistic is shown

While there is no patterning of t-values in Australia and New Zealand, in the other three coronial samples (New Mexico, New York and Cuyahoga County), the accident victims were consistently longer/taller than the homicide victims. This difference seemed largest in the child category, where the two largest t-values of the entire sample were obtained. Adolescent may also show a larger difference in Z-scores, however no adolescents were available for assessment from New Mexico.

Of the cases where the individuals in the accident group are smaller/shorter than the homicide group, New Zealand children had practically identical height for age Z-score means for the manner of death groups (accident=0.70, homicide=0.71), and the t-test was the least significant (p=0.96). The juvenile group in New Zealand also had smaller/shorter accidents victims than homicides, but the sample size for the homicide group was quite small (N=3). In two more cases, infants in New Mexico and infants in Cuyahoga County, the victims in the accident group ware quite small/short (mean Z-scores=-0.33 and -0.83, respectively).

Chapter 5. Discussion

5.1. Using accidental death as a comparative group

In this study, accident victims were used as a proxy for the population on which age estimation methods are developed. Selecting a local group of children for comparison is advantageous because it creates a control group that is subject to some of the same environmental pressures as the homicide group, and controls for potential secular effects and any problems of representativeness with the reference sample. The accidental death group is therefore assumed to represent a healthy cross-section of the population. This is a key assumption which if valid would allow for more straightforward conclusions about the difference in growth between the reference and target populations of forensic age estimation methods. Fatal accidents, however, are not distributed evenly across the population. The risk for accidents is higher in the disadvantaged end of the socioeconomic gradient (Dougherty et al. 1990, Engström et al. 2002, UCL Institute for Health Equity 2015). Thus, the sample of accidental deaths used in this analysis may not accurately represent the demographic group of children on which age estimation methods tend to be developed, although it may reflect the average population more closely than the homicide group.

Because the accident group used in this analysis is likely composed of a larger proportion of disadvantaged children than the general population, their growth may also be somewhat compromised. Thus, differences in growth found between the accidental death and homicide groups in this study should be taken as conservative estimates of real differences between average children and the forensic sample.

5.2. Conversion from cadaver length to stature

In order to better understand how the how well the accidental death group represented the average population, cadaver length was converted into standing height (stature) and compared to the CDC-2000 growth reference. Only one method for converting between cadaver length and stature in children was available (Krishan and Sharma 2002), which is a simple linear regression formula purported to be applicable through the adolescent growth spurt. To better understand the performance of this method, it was compared to two other methods meant for use in adult males, using all 17 year-old male accident victims in the sample.

In the 17 year-old male group, Krishan and Sharma (2002) yielded the smallest average correction of the three tested methods. This suggests that in the adolescents, and possibly throughout the rest of the age range, Krishan and Sharma (2002) provided estimates of stature which were taller than the true stature. This impacts the interpretation of the growth status of the accident victims. In the Cuyahoga County 17 year-old male group, a correction using Krishan and Sharma (2002) yielded a negative deviation from the reference sample significant at the 0.10 level (p=0.09), while correcting using the Trotter and Gleser (1952) factor yielded a more significant negative deviation (p=0.05). This suggests that cases of borderline negative deviation from the sample should be considered carefully, as using another conversion method may yield significant deviation from the sample.

Using Krishan and Sharma (2002) was necessary because no other method for converting between cadaver length and stature was available for children. However, the correction obtained may be smaller than the actual difference between the two measurements. Thus, accident victims may actually be considerably shorter than the reference, but results were not significant due to problems with stature correction. Results from this analysis should again be considered a conservative estimate of deviation from the reference. When deviations from the reference are significant this indicates that growth of children the accident sample is compromised.

5.3. Findings by age groups

The age groups were assigned based on changes in growth velocity, but they also hold meaning in terms of changing injury and victimization patterns. For example, children and infants are more dependent on their families for food and care than adolescents, and are accordingly more likely to become victims of caretakers rather than acquaintances and strangers (Christoffel 1990, Finklehor 1997, Alder and Polk 2001). Thus, exploring trends in differences in height for age Z-scores between manner of death groups in conjunction with knowledge about shifting mortality patterns can yield a deeper understanding of why differences between the groups arise.

The ANCOVA analysis for the infants indicated differences in length between the coronial samples rather than the manner of death groups, however difference between the manner of death groups was not uniformly absent from all coronial samples. T-test results from both the Australian and New Zealand samples suggested that there was no difference in growth between the accident and homicide victims, but results from infants in New York City showed that infants in the homicide group were smaller for age than in the accidental death group. This makes sense considering the high proportion of deaths related to child abuse and neglect (31.82%) compared to other samples in the analysis. Results from New Mexico showed that infants in the homicide group were longer for age than in the accidental death group. However in this case the accidental death group may not be representative of the normal population, as they are small for age compared to the growth reference. In Cuyahoga County, no significant differences were found in growth status between the homicide and accident victims, but the accidental death group also did not reflect the average population.

The preliminary ANCOVA test for the coronial samples combined, for all age groups besides the infants, suggested that the victims in the homicide group were shorter for age than those in accidental death group. As with the infants, results for the child age group from Australia and New Zealand indicated that there was no difference in growth status between the manner of death groups. On the other hand, the New Mexico and

Cuyahoga County both showed that the homicide victims were shorter than the accident victims, although in both cases children in the accidental death group is slightly taller than the reference sample. In the New York sample, the homicide victims were slightly smaller for age than the accident victims.

Interpretation of results in the juvenile age category is complicated by the small sample size within these groups. In Australia and New Zealand, no difference in growth status was found between the manner of death groups. Although no difference was found in the samples from the United States, there is a continuation of the tendency for homicide victims to be shorter than accident victims.

Adolescents in Australia did not differ in growth status between the manner of death groups. New Zealand homicides victims were shorter than the accidental death victims, but not significantly. No difference in growth between the manner of death groups was found in New York City. In this case the homicide victims were shorter for age than the accident victims but the accidental victims were also short for age compared to the reference. Although this difference between manner of death groups was not quite significant, it might be stronger if another method was used to convert between cadaver length and stature. In Cuyahoga County, the homicide victims were also shorter for age than the accident victims, albeit not statistically significantly so.

5.4. Trends across the age range

In the infant age category, more variation in length for age was seen between the samples rather than between the manners of death. This may be due to the interaction of several factors besides manner of death, which include a link between socioeconomic inequality and infant mortality rates, infant vulnerability and associated difficulties with classifying manner of death, and the process through which stressors impact growth that may be dependent on the sample.

Social inequality is positively correlated with infant mortality (Babones 2008). Areas characterized by lower socioeconomic status tend to have higher rates of negative health and lifestyle behaviours both during and after gestation, which leads to high rates of negative birth outcomes (UCL Institute of Health Equity 2015). The United States has high inequality and high infant mortality rates compared to other high income countries such as Australia and New Zealand (Kim and Saada 2013). This effect may be reflected in the results from New Mexico and Cuyahoga County, where accident victims are smaller for age than the reference sample, suggesting that there is a higher risk of accidental death for infants from lower socioeconomic groups in these samples. This effect does not seem to be at play or is less strong in the Australian and New Zealand samples, where there is less socioeconomic inequality. Reduced inequality in Australia and New Zealand means that homicide victims in these samples are less impacted by negative health environments associated with low socioeconomic status than homicide victims in the United States.

Infancy is a particularly delicate period of life, during which infants are dependent on caretakers. They are not yet able to escape or hide from threatening situations, and the way in which they communicate their needs, through crying, can be distressing and irritating for caretakers. Homicide in these situations can stem from a single violent disciplinary response from a caretaker (Cavanagh et al. 2007). On the other hand, because infants are so vulnerable, precise manner of death can be difficult to determine. Determining between an accidental fall and a homicidal push or drop, or between accidental asphyxia in sleep and homicidal suffocation using a soft object for example, can be nearly impossible (Goetting 1995). In some cases, a full autopsy may not be conducted in order to spare the family further trauma, which could mean missing indications of homicide (Alder and Polk 2001). The difficulty of assessing manner of death may be affecting the results in this age category, and concealing any significant differences between the manner of death groups.

Lastly, environmental stressors affect growth if they are acting on a child over the long term (Gilmour and Skuse 1999, Ehouxou et al. 2009). On the short term, child neglect and abuse has been linked with weight loss (Karp et al. 1989). Perhaps differences between the manner of death groups would be better explored in weight for age in this group as a precursor to differences in length for age. This difference in weight for age, however, would not impact skeletal age estimation.

The child age category captures an age when children are more mobile and independent, but daily routine is still largely supervised by caretakers (Christoffel 1990, Finklehor 1997, Alder and Polk 2001). Because many are not yet attending school, they are not regularly interacting with individuals outside of their family group, and remain at risk for abuse and neglect (Christoffel 1990, Finklehor 1997). Increased mobility compared to infancy reduces the risks associated with accidental drowning (often in bathtubs), asphyxia in sleep, and other causes which infants cannot protect themselves against. Reduction in fatal accidents due to these causes, combined with an increased certainty in assigning cause of death, may work to reduce socioeconomic bias in the child accidental death groups. However, this was not supported by the results: this is also the age group in which the highest proportion of coronial samples show accident groups which negatively deviate from the reference sample.

Children from lower socioeconomic contexts have been exposed to adverse growth environments over a longer period of time than the infants. The accumulation of environmental insults over a lifetime means that child homicide victims may show more growth deficits than the infants did. This was supported by the results, as this is the age group in which the highest proportion of coronial samples have homicide groups which are significantly shorter for age than the homicide victims.

The juvenile age category is somewhat difficult to interpret, because it is a transitional time. Children are increasingly insulated against accidents and homicides perpetuated by caretakers, but they are too young to engage in risky behaviour alongside their peers (Christoffel 1990). Mortality is thus lowest in this group, and cause of death is expected to be a mix of the patterns found in the child and adolescent categories. Indeed, sample sizes in this group are smaller than all other groups: in the case of New Zealand only 12 individuals are included in this group. None of the results from this age group were significant, potentially because of the reduced risk of mortality or lack of power due to sample size.

The adolescent age group represents a time when children begin to pull away from the family and engage in potentially risky behaviour away from the home (Christoffel 1990, Finklehor 1997, Muftić and Moreno 2010), perhaps changing the demographic of victims in both the accident and the homicide groups, as the homicide group becomes dominated by male victims, who are often killed in connection with another crime (Muftić and Moreno 2010). In the accident group, we can expect to see more traffic accidents, which are more common in lower socioeconomic status groups (UCL Institute of Health Equity 2015). Australia is the only sample in which adolescents in the homicide group are not shorter for age than those in the accident group, however New York City is the only sample in which accident victims are shorter for age than expected, potentially reflecting a greater proportion of low socioeconomic status individuals than in other samples.

Differences in growth between the manner of death groups in the infants may be affected by external factors such as cause of death classification and socioeconomic variation between the manner of death groups in infant mortality rates. Children show the most differences in growth between the manner of death groups. Juveniles show the least, potentially due to small sample sizes in this age group. Adolescents show more differences in height for age between the groups than the juveniles, but less than the children. Trends in the differences between the manner of death groups across the age range are potentially due to shifting mortality patterns, which may reflect shifting socioeconomic composition of the manner of death groups. In infancy, accidents occur in greater rates in low socioeconomic groups, which may blur the distinction in growth between accident and homicide victims. The gradient may persist in the child and adolescent age categories.

5.5. Findings by coronial samples

Differences in height for age between the manner of death groups are not consistent across the samples: all cases where homicide victims are significantly shorter than the accident victims occur in U.S. coronial samples. For this reason, trends in differences in height for age may be clarified by examining patterns within each coronial sample.

In the infant group, Australia showed no significant difference in growth between the groups, and the homicide victims were slightly shorter for age than the accident victims. No consistent trends were seen for the rest of the age range. In no age category was the difference in growth significantly different between the groups. In addition, in two groups (children and adolescents) the homicide group was slightly taller for age than the accident group, while in the third one (juveniles) the opposite was true. Overall, the homicide and accident groups do not differ significantly in Australia and whatever small differences are detected are inconsistent: in some cases the homicides are shorter for age than the accidents, and in others they are taller.

New Zealand infants did not show a significant difference between the manner of death groups, although the homicide victims were slightly smaller for age than the accident victims were. In the other groups, no consistent trend was obvious. In the child and juvenile age categories, the groups were very similar, although the homicide victims were slightly taller for age than the accident victims. The opposite was true for the adolescent group, where homicide victims were slightly shorter than the accident victims. Overall, there was no obvious trend in difference in growth between the manner of death groups throughout the age range in New Zealand.

In New Mexico, the infants showed no significant difference between the manner of death groups, but the accident victims were somewhat smaller than average children included in the CDC-2000 reference. In both the child and the juvenile groups, the homicide victims were shorter for age than the accident victims, although this difference was only approaching significance in the children. These results show that in New Mexico, there is a consistent tendency for the homicide victims to be shorter for age than the accident victims, although not always statistically significantly so.

In New York City, homicide victims were consistently smaller/shorter for age than accident victims in all age groups, and only in infants did these differences seem to be significant. The difference in growth status between the homicide and accident group was likely underestimated in the adolescents, where the accident victims are shorter than the reference.

Infants in Cuyahoga County had the same finding as those in the New Mexico sample, where the accident victims were smaller for age than average infants as indicated by the CDC-2000 reference. Throughout the rest of the age range, homicide victims were consistently smaller for age than accident victims at varying levels of significance. In Cuyahoga County, there is a consistent tendency for the homicide group to be shorter for age than the accident group.

5.6. Trends in the samples: Australia and New Zealand versus the United States

Examination of trends in the coronial samples excluding infants, reveal two distinct trends emerging between the samples. Australia and New Zealand show inconsistent, never significant, differences between the groups. In contrast, coronial collections from the United States (New Mexico, New York City, and Ohio) seem to be showing a consistent, although variably strong, difference in height for age between the accident and homicide groups.

Differences between the Australian and New Zealand samples and those from the United States are apparent from analysis of difference in growth status between the manner of death groups. In Australia and New Zealand, the strongest difference between the manner of death groups is seen in the New Zealand adolescents, but the United States collections have three categories which show more significant differences: New Mexico children, and Cuyahoga County children and adolescents. The magnitudes of the differences seen between the manner of death groups in the United States are not present in Australia and New Zealand. For the U.S. coronial samples, homicide victims are smaller/shorter for age than accident victims. This is not the case in Australia and New Zealand, where the differences between the groups are inexistent or inconsistent in direction.

This divide between the two groups of coronial samples is further supported by the difference in mean length/height for age Z-score between the samples in each manner of death group. Homicide victims in the United States are shorter for age than those in Australia and New Zealand. This is not simply a case of children from Australia and New Zealand being taller for age: children in the accident groups from all coronial samples are quite similar in mean height for age Z-scores throughout the age range. Overall, homicide victims are shorter for age in the United States than they are in New Zealand and Australia, and have potentially been exposed to more growth insults over their lives.

Further support for differing trends across the sample is found when accident victims are compared to the reference sample for deviation from zero in order to assess the validity of the accidental death sample. In Australia and New Zealand, the accidental death sample never deviated significantly from the reference sample. In coronial samples from the United States, deviation was more frequent. This occurred both significantly at the 0.10 level (Cuyahoga County infants, New Mexico children), as well as borderline significantly at this same level (Cuyahoga County children). Deviation from the reference sample may be even more frequent in the United States samples, considering that the method used to convert cadaver length to standing height may have underestimated it significantly.

When growth of children in the accidental death group deviates significantly from the reference sample, this suggests a relatively higher proportion of lower socioeconomic status individuals in the accident group. Consequently, this group does not represent an accurate cross-sectional sample of the overall population. This is either because of the socioeconomic context of the sample, yielding a greater proportion of children who are shorter for age than healthy normal children, or because fatal accidents are disproportionately affecting children of lower socioeconomic status. In both cases, the increased frequency of cases in which individuals in the accidental death sample are smaller/shorter for age than the reference in the United States compared to Australia and New Zealand indicates a further impact of social inequality on the samples.

Differences in growth status between the accident and homicide victims in the United States samples are potentially conservative estimates of real differences, because the differences found between the growth of children in the accident group and the reference sample in the U.S. samples suggest that the accident group is less representative of the general population. Children in the accident group are likely smaller/shorter for age than the general population, and thus the difference in growth status between the homicide and the accident victims presented in this study may be an underrepresentation or conservative estimation of the difference in growth between the general and forensic population in the United States. This is not necessarily the case in Australia and New Zealand, where the comparison is more likely to be accurate.

This difference in results between coronial samples from the United States and those from New Zealand and Australia may result from the fact that the United States is a much more socioeconomically unequal country. Socioeconomic inequality has been linked to poorer health, higher rates of child maltreatment, higher homicide rates, and higher rates of infant mortality (Lee and Bankston 1999, Daly et al. 2001, Wilkinson and Pickett 2006, Nadanovsky and Cunha-Cruz 2009, Kim and Saada 2013, United Nations Children's Fund 2014). Several measures are available to measure socioeconomic inequality (e.g.: Theil index, Atkinson index), but the Gini coefficient is the most widely calculated indicator, and has been correlated to negative health outcomes by various researchers (Babones 2008, Roberts and Willits 2015). The Gini ranges from 0 to 1, with higher numbers indicating higher income inequality.

A robust examination of the correlation of economic inequality with differences in growth is beyond the scope of this study and would require data from many more coronial bodies in more countries. Additionally, each organization calculates the Gini coefficient slightly differently. This makes direct comparison between the samples in this study difficult: the Gini is calculated by the Organization for Economic Co-operation and Development (OECD) for Australia, New Zealand, and the United States, and is given for geographies within the United States (New Mexico, New York, and Cuyahoga County) by the U.S. Census Bureau. These are not reconcilable.

Even with the small sample size, there seems to be a relationship between the national Gini coefficients (as calculated by averaging OECD-provided values for available years between 2008 and 2010) and the differences in growth seen in this study. The U.S., which has the highest Gini coefficient (0.379), showed more consistent differences between the homicide and accident groups than either Australia or New Zealand, which had smaller Gini coefficients (0.335 and 0.327 respectively). Although these observations are based on a very limited sample of cases, it follows that in areas with higher inequality, differences in growth between the general and forensic populations may be more pronounced. Growth differences between the accident and homicide victims seem to vary across contexts, with higher levels of socioeconomic inequality yielding greater differences between the groups. This means that in countries or settings with greater socioeconomic inequality, the difference in growth between the general and forensic populations will be greater.

Differences in growth between the general and forensic populations seem to be linked to social inequality. Consequently, the difference in growth between these groups is likely to be limited in relatively equal countries such as Australia, but is more pronounced in unequal countries like the United States. This may be even more marked in developing countries with higher inequality, where the highest rates of child homicide occur (United Nations Children’s Fund 2014). Here, the impact of the difference in growth between the population on which methods are developed and used is even more important, as methods are developed on healthy children living in developed countries and yet are used on children growing in disadvantaged settings within developing nations. Given that children are a more significant component of forensic anthropology casework in developing nations, these differences in growth are particularly important to recognize when choosing, applying and interpreting the results of juvenile age estimation methods in these countries.

5.7. Importance of findings for forensic age estimation

Throughout the analysis, conservative methodological choices, such as using the accident group as a comparison, were made to raise confidence in the findings. Even the small differences in linear growth found in this study, however, result in measurable differences in height, which can be translated in differences in long bone length, and therefore in age between children in the accident and homicide groups or between the general and forensic population.

To gauge the impact on age estimation that these difference might have, the mean difference between the length/height for age Z-scores of the accident and homicide in New York infants and in New Mexico and Cuyahoga County children were converted back to age using the reference samples (Kuczmarski et al. 2000). The Z-score calculation formula was re-arranged to solve for height. Height at the beginning and end of the age range in infants and children was calculated using the LMS parameters for the age and sex group and the difference in mean Z-score for that comparison. For the infant group, the "beginning" of the age range was set at 1 year, because beginning at birth would yield a negative age. Obtained height or length at the beginning and end of the age category was then compared to the reference sample to find the age in months at which it most closely matched the 50^{th} percentile. The difference between that age and the original age (either the beginning or end of each age category) was obtained as an estimate of the overall difference in skeletal ages at the same chronological age between the accident and homicide groups. This estimate of the difference in skeletal ages assumes that a measurable difference in mean height for age is a good proxy for difference in mean long bone length for age.

The difference between skeletal ages at the same chronological age ranged from 1 to 3 months in infants and 6 to 10 months in children. In a forensic investigation, where age is estimated from skeletal remains, children who are several months behind the average general population may be included in the 95% prediction interval of a well-crafted method, but the portion of cases whose real age is missed in the 95% prediction interval would be larger than 5%. For example, Stull et al. (2014) found a 95% prediction interval for age estimation from diaphyseal length of the femur that ranged from 2 months at birth to 5 years at 12 years of age, so between 1 month and 2.5 years on either side of the point estimate at various points in the age range. The difference in skeletal age of 10 months at the chronological age of 7 years between the means of the general and forensic groups found in this study implies that even the average forensic child will be in the youngest end of the prediction interval provided by Stull et al. (2014). Children who are even shorter than the forensic group mean may be excluded.

This growth delay can be expected to extend to other age and maturity indicators. While long bone length is a popular way of estimating juvenile age (Stull et al. 2014), methods based on growth indicators such as epiphyseal union and dental development can also be used. Other researchers have found an effect of socioeconomic status on hand-wrist age skeletal age estimation (Schmeling et al. 2000). Because dental development is less affected by environmental factors, dental age estimates would be less impacted by the differences in linear growth found. However, dental development is not completely insulated against these pressures (Cardoso 2007), and differences in maturation between a healthy population and the forensic sample still have the potential to influence age estimates.

The results of this study did not find consistently significant differences between the forensic and general population on which skeletal age estimation methods are developed. However, the consistent trend for the forensic population to be shorter than the general population means that there is potential for methods to yield inaccurate or at least biased results when they are developed on the a health population and used on a forensic population. The notion that the reference sample should match the target population is not new to physical anthropology (Lampl and Johnston 1996, Saunders 2000, Franklin 2010), however it is particularly relevant in forensic contexts. The growth deficit that might exist between the general and forensic populations may delay or prevent correct victim identification by means of an inaccurate age estimate.

5.8. Limitations of the study

This study presents several limitations, which result from the type of data used: cadaver lengths taken at autopsy. First, as cadaver lengths were recorded by various technicians working within several organizations, there is a legitimate concern with the consistency of cadaver length measurement methods. Using a group of accident victims to compare to homicide victims within each coronial sample controls for these inconsistencies in measurements, as the same cadaver measurement methods are applied to both the accidental death and homicide groups.

Additionally, cadaver length is used in this analysis as measurement of linear growth. In a forensic investigation, long bone length, not cadaver length, would be the measurement taken to estimate age. However, differences in growth will also exist in the long bones of the body, which contribute to making up stature or cadaver length.

Lastly, using the homicide group as a representation of the forensic anthropology population may not be perfect comparison. Little is known about the makeup of the forensic anthropology population, but it is likely to be largely composed of homicide victims. The homicide group is therefore most appropriate proxy for the forensic anthropology population from which data can be readily obtained, although some difference in the magnitude of the growth delay between the forensic and general populations and the homicide and accident victims can be expected.

Chapter 6. Conclusion

This study drew upon cadaver lengths collected during autopsies of children performed between 2002 and 2015 in Australia, New Zealand, and the United States to compare the growth status of the general and forensic populations. Two research questions were posed.

First, do differences in growth occur between accident victims, which represent the normal population, and homicide victims, which represent the forensic population? Differences in growth between the homicide and accident victims did not occur uniformly throughout the samples and age groups. When differences were statistically significant, the homicides were always shorter for age than the accident victims. In Australia and New Zealand, homicide and accident victims did not differ in growth status. In samples from the United States, homicide victims were always shorter for age than accident victims. The lack of difference between the manner of death groups in Australia and New Zealand compared to the differences found in the United States may be due to greater socioeconomic inequality in the United States than in either Australia or New Zealand.

Second, does the accidental death sample accurately represent the normal population? The accident victims rarely deviated significantly from the reference sample, indicating that in most cases the accident group was an appropriate comparison for the average population. However, the method used to obtain stature from cadaver length may have overestimated stature, thus underestimating the difference between the accident victims and average, healthy children. When differences did occur between the accident victims and the reference sample, they were in samples from the United States. This indicates that the growth deficit of homicide victims in samples from the United States may be underestimated.

Homicide victims tend to be smaller/shorter for age than accident victims, although the strength of that difference seems to be linked to socioeconomic inequality. Forensic anthropologists should carefully consider the samples from which they derive skeletal age estimation methods intended for use in a forensic context, as the growth differences between the groups could slow or halt victim identification.

References

Abubakar A, Van de Vijver F, Van Baar A, Mbonani L, Kalu R, Newton C, Holding P.
2006. Socioeconomic status, anthropometric status, and psychomotor development of Kenyan children from resource-limited settings: a path-analytic study. Early Human Development 84:613-621.

Alder C and Polk K.
2001. Child victims of homicide. Cambridge, UK: Cambridge University Press.

Armstrong JS.
2012. Illusions in regression analysis. International Journal of Forecasting 28:689-694.

Anderson M, Messner MB, Green WT.
1964. Distribution of lengths of the normal femur and tibia in children from one to eighteen years of age. The Journal of Bone and Joint Surgery 46:1197-1202.

Australian Institute of Criminology.
2008. Homicide in Australia: 2006-07 National Homicide Monitoring Program annual report. Canberra, Australia: Australian Institute of Criminology.

Aykroyd RG, Lucy D, Pollard AM, Solheim T.
1997. Technical note: regression analysis in adult age estimation. American Journal of Physical Anthropology 104:259-265.

Babones SJ.
2008. Income inequality and population health: correlation and causality. Social Science and Medicine 66:1614-1626

Behrman JR, Calderon MC, Preston SH, Hoddinott J, Martorell R, Stein AD.
2009. Nutritional supplementation in girls influences growth of their children: prospective study in Guatemala. American Journal of Clinical Nutrition 90:1372-1379.

Bidmos MA.
2005. On the non-equivalence of documented cadaver lengths to living stature estimates based on Fully's methods on bones in the Raymond A. Dart collection. Journal of Forensic Sciences 50:1-6.

Bielicki T, Szczotka H, Charzewski J.
1981. The influence of three socioeconomic factors on body height in Polish military conscripts. Human Biology 53:543-555.

Block RW, Krebs NF.
2005. Failure to thrive as a manifestation of child neglect. Pediatrics 116:1234-1237.

Boccone S, Micheletti Cremasco M, Bortoluzzi S, Moggi-Cecchi J, Rabino Massa E.
2010. Age estimation in subadult Egyptian remains. HOMO – Journal of Comparative Human Biology 61:337-358.

Bogin B.
1999. Patterns of human growth, 2nd edition. Cambridge, UK: Cambridge University Press.

Bor W, Najman JM, Andersen M, Morrison J, Williams G.
1993. Socioeconomic disadvantage and child morbidity: an Australian longitudinal study. Social Science & Medicine 36:1053-1061.

Bradley RH, Corwyn RF.
2002. Socioeconomic status and child development. Annual Review of Psychology 53:371-399.

Brooks-Gunn J, Duncan GJ.
1997. The effects of poverty on children. The Future of Children 7:55-71.

Buyken AE, Hahn S, Kroke A.
2005. Differences between recumbent length and stature measurement in groups of 2- and 3-y-old children and its relevance for the use of European body mass index references. International Journal of Obesity 29:24-28.

Cavanagh K, Dobash RE, Dobash RP.
2007. The murder of children by fathers in the context of child abuse. Child Abuse and Neglect 31:731-746.

Cardoso HFV.
2007. Environmental effects on skeletal versus dental development; using a documented subadult skeletal sample to test a basic assumption in human osteological research. American Journal of Physical Anthropology 133:223-233.

Cardoso HFV, Abrantes J, Humphrey LT.
2014. Age estimation of immature human skeletal remains from the diaphyseal length of the long bones in the postnatal period. International Journal of Legal Medicine 128:809-824.

Cardoso HFV, Marinho L, Albanese J.
2016. The relationship between cadaver, living and forensic stature: a review of current knowledge and a test using a sample of adult Portuguese males. Forensic Science International 258:55-63

Cattaneo C.
2007. The skeletal remains of a child, victim of organized crime: the study of post-mortem interval, personal identification and manner of death. In: Brickley MB, Ferllini R, eds. Forensic anthropology: case studies from Europe. Springfield, Illinois: Charles C Thomas, 137-150.

Chilton M, Chyatte M, Breaux J.
2007. The negative effects of poverty & food insecurity on child development. Indian Journal of Medical Research 126:262-272.

Christiansen N, Mora JO, Herrera MG.
1975. Family social characteristics related to physical growth of young children. Journal of Epidemiology & Community Health 29:121-130.

Christoffel KK.
1990. Violent death and injury in US children and adolescents. American Journal of Diseases of Children 144:697-706.

Clegg EJ.
1982. The influence of social, geographical and demographic factors on the size of 11-13 year old children from the Isle of Lewis, Scotland. Human Biology 54:93.

Cole T.
1990. The LMS method for constructing normalized growth standards. European Journal of Clinical Nutrition 44:45-60.

Cole TJ.
2003. The secular trend in human physical growth: a biological view. Economics and Human Biology 1:161-168.

Conceição ELN, Cardoso HFV.
2011. Environmental effects on skeletal versus dental development II: further testing of a basic assumption in human osteological research. American Journal of Physical Anthropology 144:462-470.

Cunha E, Baccino E, Martrille L, Ramsthaler F, Prieto J, Schuliar Y, Lynnerup N, Cattaneo C.
2009. The problem of aging human remains and living individuals: a review. Forensic Science International 193:1-13

Currie E.
2007. Pulling apart: notes on the widening gap in the risks of violence. Criminal Justice Matters 70:37-38.

Crooks D.
1999. Child growth and nutritional status in a high-poverty community in Eastern Kentucky. American Journal of Physical Anthropology 109:129-142.

Daly M, Wilson M, Vasdev S.
2001. Income inequality and mortality in Canada and the United States. Canadian Journal of Criminology 43:219-236.

Danforth ME, Wrobel GD, Armstrong CW, Swanson D.
2009. Juvenile age estimation using diaphyseal long bone lengths among ancient Maya populations. Latin American Antiquity 20:3-13.

De Onis M, Garza C.
2006. Foreword. Acta Paediatrica Supplement 450:5-6.

De Onis M, Onyango AW, Boghi E, Siyam A, Nishida C, Siekmann J.
2007a. Development of a WHO growth reference for school-aged children and adolescents. Bulletin of the World Health Organization 85:660-667.

De Onis M, Garza C, Onyango AW, Borghi E.
2007b. Comparison of the WHO child growth standards and the CDC 2000 growth charts. The Journal of Nutrition 137:144-148.

Dettmeyer RB, Verhoff MA, Schütz HF.
2014. Forensic medicine: fundamentals and perspectives. Heidelberg, Germany: Springer.

Dirkmaat DC, Cabo LL, Ousley SD, Symes SA.
2008. New perspectives in forensic anthropology. Yearbook of Physical Anthropology 51:33-52.

Doolan M.
2006. Child homicide in New Zealand: 2001-2005. Te Awatea Review 4:25.

Dougherty G, Pless IB, Wilkins R.
1990. Social class and the occurrence of traffic injuries and death in urban children. Canadian Journal of Public Health 81:204-209.

Douglas EM, Finkelhor D.
2005. Child maltreatment fatalities factsheet. Durham, NH: Crimes against Children Research Center.

Douglas E, Vanderminden J.
2014. A longitudinal, multilevel analysis of homicide against children aged 0-9 years using state-level characteristics: 1979-2007. Violence and Victims 29:757-770.

Duren DL, Nahhas RW, Sherwood RJ.
2015. Do secular trends in skeletal maturity occur equally in both sexes? Clinical Orthopaedics and Related Research 473:2559-2567.

Đurić M, Rakočević Z, Đonić D.
2005. The reliability of sex determination of skeletons from forensic context in the Balkans. Forensic Science International 147:159-164.

Eckhardt CL, Adair LS.
2002. Differences in stunting prevalences calculated from two similar growth references may be large and inconsistent in undernourished children. Annals of Human Biology 29:566-578.

Ehouxou NZ, Zuzunegui MV, Séguin L, Nikiema B, Gauvin L.
2009. Duration of lack of money for basic needs and growth delay in the Quebec Longitudinal Study of Child Development birth cohort. Journal of Epidemiology and Community Health 63:45-49.

Elmer E, Gregg GS.
1967. Developmental characteristics of abused children. Pediatrics 40:596-602.

Engström K, Diderichsen F, Laflamme L.
2002. Socioeconomic differences in injury risks in childhood and adolescence: a nation-wide study of intentional and unintentional injury in Sweden. Injury Prevention 8:137-142.

Ericksen MF.
1982. How “representative” is the Terry collection? Evidence from the proximal femus. American Journal of Physical Anthropology 59:345-350.

Facchini F, Veschi S.
2004. Age determination on long bones in a skeletal subadults sample (b-12 years). Collegium Antropologicum 1:89-98.

Ferrante L, Skrami E, Gesuita R, Cameriere R.
2015. Bayesian calibration for forensic age estimation. Statistics in Medicine 34:1779-1790.

Feldesman MR.
1992. Femur/stature ratio and estimates of stature in children. American Journal of Physical Anthropology 87:447-459.

Fernald LC, Grantham-McGregor SM.
2002. Nutrition growth retardation is associated with changes in the stress response system and behavior in school-aged Jamaican children. The Journal of Nutrition 132:3674-3679.

Fieguth A, Günther D, Kleeman WJ, Tröger HD.
2002. Lethal child neglect. Forensic Science International 130: 8-12.

Finkelhor D.
1997. The homicides of children and youth: a developmental perspective. In: Kaufman Kantor G, Jasinski JL, eds. Out of the darkness: contemporary perspectives on family violence. London, UK: Sage Publications, 17-34.

Finkelhor D and Ormrod R.
2001. Homicides of children and youth. Juvenile Justice Bulletin Oct 2001:1-12.

Franklin D.
2010. Forensic age estimation in human skeletal remains: current concepts and future directions. Legal Medicine 12:1-7.

Galić I, Vodanović M, Cameriere R, Nakaš E, Galić E, Selimović E, Brkić H.
2011. Accuracy of Cameriere, Haavikko, and Willems radiographic methods on age estimation on Bosnian-Herzegovian children age groups 6-13. International Journal of Legal Medicine 125:315-321.

Garn SM, Clark DC.
1975. Nutrition, growth, development, and maturation: findings from the Ten-State Nutrition Survey of 1968-1970. Pediatrics 56:300-319.

Garn SM, Pesick SD, Pilkington JJ.
1984. The interaction between prenatal and socioeconomic effects on growth and development in childhood. In Borms J, Hauspie R, Sand C, Susanne C, Hebbelinck M, eds. Human growth and development. New York City, NY: Plenum. 59-70.

Ghantus MK.
1951. Growth of the shaft of the human radius and ulna during the first two years of life. American Journal of Roentgenology and Radium Therapy 65:784-786.

Gilbert R, Widom CS, Browne K, Fergusson D, Webb E, Janson S.
2009. Burden and consequences of child maltreatment in high-income countries. Lancet 373:68-81.

Gilmour J, Skuse D.
1999. A case-comparison study of the characteristics of children with a short stature syndrome induced by stress (hyperphagic short stature) and a consecutive series of unaffected "stressed" children. Journal of Child Psychology and Psychiatry 40:969-978.

Gindhart PS.
1973. Growth standards for the tibia and radius in children aged one month through eighteen years. American Journal of Physical Anthropology 39:41-48.

Goetting A.
1995. Homicide in families and other special populations. New York City, NY: Springer.

Greves Grow HM, Cook AJ, Arterburn DE, Saelens BE, Drewnowski A, Lozano P.
2010. Child obesity associated with social disadvantage of children's neighborhoods. Social Science & Medicine 71:584-591.

Grimberg A, Ramos M, Grundmeier R, Feemster KA, Pati S, Chucciara AJ, Stallings VA.
2009. Sex-based prevalence of growth faltering in an urban pediatric population. Journal of Pediatrics 154:567-572.

Hamill PVV, Johnson CL, Reed RB, Roche, AF.
1977. NCHS growth curves for children. Vital and Health Statistics: Series 11, Data from the National Health Survey 165:1-75.

Hermanussen M, Godina E, Rühli FJ, Blaha P, Boldsen JL, van Buuren S, MacIntyre M, Aßmann C, Ghosh A, de Stefano GF, Sonkin VD, Tresguerres JAF, Meigen C, Scheffler C, Geiger C, Lierberman LS.
2010. Growth variation, final height and secular trend. Proceedings of the 17th Aschauer Soiree, 7th November 2009. HOMO – Journal of Comparative Human Biology 61:277-284.

Holobinko A.
2012. Forensic human identification in the United States and Canada: a review of the law, admissible techniques, and the legal implications of their application in forensic cases. Forensic Science International 222:394.e1-394.e13

Hunt D.
1990. Sex determination in the subadult ilia: an indirect test of Weaver's non-metric sexing method. Journal of Forensic Sciences 35:881-885.

Hunt DR, Albanese J.
2005. History and demographic composition of the Robert J. Terry anatomical collection. American Journal of Physical Anthropology 127:406-417.

Jackes M.
2000. Building the bases for paleodemographic analysis: adult age estimation. In: Katzenberg MA, Saunders SR, eds. Biological anthropology of the human skeleton. New York City, NY: Wiley-Liss, 417-466.

Johnson W, Choh AC, Soloway LE, Czerwinski SA, Towne B, Demerath EW.
2012. Eighty year trends in infant weight and length growth: the Fels longitudinal study. Journal of Pediatrics 160:762-768.

Jorge CB, Doris CP.
2014. Z-score anthropometric indicators derived from NCHS-1977, CDC-2000, and WHO-2006 in children under 5 years in central area of Peru. Universal Journal of Public Health 2:73-81.

Julia M, Van Wissenruch MM, Delemarre-van de Waal HA, Surjono A.
2004. Influence of socioeconomic status on the prevalence of stunted growth and obesity in prepubertal Indonesian children. Food and Nutrition Bulletin 25:354-360.

Karp RJ, Scholl TO, Decker E, Ebert E.
1989. Growth of abused children contrasted with the non-abused in an urban poor community. Clinical Pediatrics 28:317-320.

Kellogg ND, Lukefahr JL.
2005. Criminally prosecuted cases of child starvation. Pediatrics 116:1309-1316.

Kemkes-Grottenthaler A.
2002. Aging through the ages: historical perspectives on age indicator methods. In: Hoppa RH, Vaupel JW, eds. Paleodemography: age distributions from skeletal samples. Cambridge, UK: Cambridge University Press, 48-72.

Kim D, Saada A.
2013. Social determinants of infant mortality and birth outcomes in modern developed nations: a cross-country systematic review. International Journal of Environmental Research and Public Health 10:2296-2335.

Kimmerle EH, Jantz RL, Koningsberg LW, Baraybar JP.
2008. Skeletal estimation and identification in American and East European populations. Journal of Forensic Sciences 53:524-532.

Klebanoff MA, Schulsinger C, Mednick BR, Secher NJ.
1997. Preterm and small-for-gestational-age birth across generations. American Journal of Obstetrics and Gynecology 176:521-526.

Komar D.
2003. Lessons from Srebrenica: the contributions and limitations of physical anthropology in identifying war crimes. Journal of Forensic Science 48:713-6.

Komar DA, Grivas C.
2008. Manufactured populations: what do contemporary reference skeletal collections represent? A comparative study using the Maxwell Museum documented collection. American Journal of Physical Anthropology 137:224-233.

Koningsberg LW, Frankenberg SR.
1994. Paleodemography: "Not quite dead." Evolutionary Anthropology 3:92-105.

Koningsberg LW, Frankenberg SR, Walker RB.
1994. Regress what on what?: paleodemographic age estimation as a calibration problem. In: Paine RR, ed. Integrating archaeological demography: multidisciplinary approaches to prehistoric population. Carbondale, IL: Southern Illinois University Press, 64-88.

Koningsberg LW, Hens SM, Jantz LM, Jungers WL.
1998. Stature estimation and calibration: Bayesian and maximum likelihood perspectives in physical anthropology. Yearbook of Physical Anthropology 41:65-92.

Kuczmarski RJ, Ogden CL, Grummer-Strawn LM, Felgal KM, Guo SS, Wei R, Mei Z, Roche AF, Johnson CL.
2000. CDC Growth Charts: United States. Advance Data 314:1-28.

Krishan K, Sharma JC.
2002. Intra-individual difference between recumbent length and stature among growing children. The Indian Journal of Pediatrics 69:565-569.

Lampl M, Johnston FE.
1996. Problems in the aging of skeletal juveniles: perspectives from maturation assessments of living children. American Journal of Physical Anthropology 101:345-355.

Langley-Shirley N, Jantz RL.
2010. A Bayesian approach to age estimation in modern Americans from the clavicle. Journal of Forensic Sciences 55:571-583.

Lasker GW, Mascie-Taylor CGN.
1989. Effects of social class differences and social mobility on growth in height, weight, and body mass index in a British cohort. Annals of Human Biology 16:1-8.

Lasker GW, Mascie-Taylor CGN.
2005. Biosocial correlates of stature in a British national cohort. Journal of Biosocial Science 37:245-251.

Lee MR, Bankston WB.
1999. Political structure, economic inequality, and homicide: a cross-national analysis. Deviant Behavior 20:27-55.

Lewis ME, Rutty GN.
2003. The endangered child: the personal identification of children in forensic anthropology. Science & Justice 43:201-209.

Liversidge HM, Molleson TI.
1999. Developing permanent tooth length as an estimate of age. Journal of Forensic Sciences 44:917-920.

López-Costas O, Rissech C, Trancho G, Turbón D.
2012. Postnatal ontogenesis of the tibia. Implications for age and sex estimation. Forensic Science International 214:207.e1-207.e11.

Madea B.
2005. Death as a result of starvation: diagnostic criteria. In: Tsokos M, ed. Forensic pathology reviews, vol. 2. Totowa, NJ: Humana Press, 3-23.

Maresh MM.
1943. Growth of major long bones in healthy children. A preliminary report on successive roentgenograms of the extremities from early infancy to twelve years of age. American Journal of Diseases of Children 66:227-257.

Maresh MM.
1955. Linear growth of long bones of extremities from infancy through adolescence. American Journal of Diseases of Children 89:725-742.

Maresh MM
1970. Measurements from roentgenograms. In: McCannon, RW, ed. Human growth and development. Springfield, IL: Charles C Thomas, 157-188.

Maresh MM, Beal VA.
1970. A longitudinal survey of nutrition intake, body size, and tissue measurements in healthy subjects during growth. Monographs of the Society for Research in Child Development 35:33-39.

Martorell R and Zongrone A.
2012. Intergenerational influences on child growth and undernutrition. Paediatric and Perinatal Epidemiology 26S1:302-314.

Meindl RS, Russell KF, Lovejoy CO.
1990. Reliability of age at death in the Hamann-Todd collection: validity of subselection procedures used in blind tests of the summary age technique. American Journal of Physical Anthropology 83:349-357.

Mimasaka S, Funayama M, Adachi N, Nata M, Morita M.
2000. A fatal case of infantile scurvy. International Journal of Legal Medicine 114: 122-124.

Mittler D, Sheridan S.
1992. Sex determination in subadults using auricular surface morphology: a forensic science perspective. Journal of Forensic Sciences 37:1068-1075.

Moffat T, Galloway T, Latham J.
2005. Stature and adiposity among children in contrasting neighborhoods in the city of Hamilton, Ontario, Canada. American Journal of Human Biology 17:355-367.

Moffat T, Galloway T.
2007. Adverse environments: investigating local variation in child growth. American Journal of Human Biology 19:676-683.

Muftić LR and Moreno RD.
2010. Juvenile homicide victimization: differences and similarities by gender. Youth Violence and Juvenile Justice 8:386-398.

Nadanovsky P, Cunha-Cruz J.
2009. Relative contribution of income inequality and imprisonment to the variation in homicide rates among developed (OECD), South and Central American countries. Social Science & Medicine 69:1343-1350.

Nelson S, Hans MG, Broadbent BH, Dean D.
2000. The Brush Inquiry: an opportunity to investigate health outcomes in a well-characterized cohort. American Journal of Human Biology 12:1-9.

Norgan NG, Bogin B, Cameron N.
2012. Nutrition and growth. In: Cameron N and Bogin B, eds. Human growth and development, 2nd edition. London, UK: Academic Press, 123-152.

Palmer PB, O'Connell DG.
2009. Regression analysis for prediction: understanding the process. Cardiopulmonary Physical Therapy Journal 20:23-26.

Pfau RO, Sciulli PW.
1994. A method for establishing the age of subadults. Journal of Forensic Sciences 39:165-176.

Primeau C, Friis L, Sejrsen B, Lynnerup N.
2012. A method for estimating age of Danish medieval sub-adults based on long bone length. Anthropologischer Anzeiger 69:317-333.

Primeau C Friis L, Sejrsen B, Lynnerup N.
2016. A method for estimating age of medieval sub-adults from infancy to adulthood based on long bone length. American Journal of Physical Anthropology 159:135-145.

Prince DA, Koningsberg LW.
2008. New formulae for estimating age-at-death in the Balkans utilizing Lamendin's dental technique and Bayesian analysis. Journal of Forensic Sciences 53:524-532.

Rissech C, Schaefer M, Malgosa A.
2008. Development of the femur – implications for age and sex determination. Forensic Science International 180:1-9.

Rissech C, López-Costas O, Turbón D.
2013. Humeral development from neonatal period to skeletal maturity – application in age and sex assessment. International Journal of Legal Medicine 127:201-212.

Roche AF.
1992. Growth, maturation and body composition: the Fels longitudinal study 1929-1991. Cambridge, UK: Cambridge University Press.

Roberts A, Willits D.
2015. Income inequality and homicide in the United States: consistency across different income inequality measured and disaggregated homicide types. Homicide Studies 19:28-57.

Robbins Schug G, Gupta S, Cowgill LW, Sciulli PW, Blatt SH.
2013. Panel regression formulas for estimating stature and body mass from immature human skeletons: a statistical approach without reference to specific age estimates. Journal of Archaeological Science 40:3076-3086.

Ross AH.
2011. Fatal starvation/malnutrition: medicolegal investigation from the juvenile skeleton. In Ross AH, Abel SM, eds. The juvenile skeleton in forensic abuse investigations. New York City, NY: Humana Press, 151-166.

Saunders SR.
2000. Subadult skeletons and growth-related studies. In Katzenberg MA, Saunders SR, eds. Biological anthropology of the human skeleton. New York City, NY: Wiley-Liss, 135-162.

Schell LM, Gallo MV, Ravenscroft J.
2009. Environmental influences on human growth and development: historical review and case study of contemporary influences. Annals of Human Biology 36:459-477.

Schaefer M, Black S, Scheuer L.
2008. Juvenile osteology: a laboratory and field manual. London, UK: Academic Press.

Schmeling A, Reisinger W, Loreck D, Vendura K, Markus W, Geserick G.
2000. Effects of ethnicity on skeletal maturation: consequences for forensic age estimations. International Journal of Legal Medicine 113:253-258.

Schutkowski H.
1993. Sex determination of infant and juvenile skeletons: 1 - morphological features. American Journal of Physical Anthropology 90:199-205.

Sidebotham J, Fraser J, Covington T, Freemantle J, Petrou S, Pulikottil-Jacob R, Cutler T, Ellis C.
2014. Understanding why children die in high-income countries. Lancet 384:915-927.

Simmons T.
2007. Fleeing Srebrinica: a surprise attack on a group of men and boys in the moutains. In Brickley MB, Ferllini R, eds. Forensic anthropology: case studies from Europe. Springfield, IL: Charles C. Thomas. 165-182.

Singh GK, Kogan MD.
2007. Widening socioeconomic disparities in US childhood mortality. American Journal of Public Health 97:1658-1665.

Skuse DH.
1989. Emotional abuse and delay in growth. ABC of Child Abuse 299:113-115.

Smith SL.
2007. Stature estimation of 3-10-year-old children from long bone lengths. Journal of Forensic Sciences 52:538-546.

Smith SL, Buschang PH.
2009. Growth in root length of the mandibular canine and pre-molars in a mixed-longitudinal orthodontic sample. American Journal of Human Biology 21:623-634.

Snow CC, Luke JL.
1970. The Oklahoma City child disappearances of 1967: forensic anthropology in the identification of skeletal remains. Journal of Forensic Sciences 15:125-153.

Solarino B, Grattagliano I, Catanesi R, Tsokos M.
2012. Child starvation and neglect: a report of two fatal cases. Journal of Forensic and Legal Medicine 19:171-174.

Soyer E, Hogarth RM.
2012. The illusion of predictability: How regression statistics mislead experts. International Journal of Forecasting 28:695-711.

Steckel RH.
2012. Social and economic effects on growth. In: Cameron N and Bogin B, eds. Human growth and development, 2nd edition. London, UK: Academic Press, 225-244.

Stuart HC, Reed RB, and associates.
1959. Description of Project. Pediatrics 24: 875-885.

Stull KE, L'Abbé EN, Ousley SD.
2014. Using Multivariate Adaptive Regression Splines to estimate subadult age from diaphyseal dimensions. American Journal of Physical Anthropology 154: 376-386.

Taitz LS, King JM.
1988. Growth patterns in child abuse. Acta Paediatrica Scandinavica 343:62-72.

Taleb NN, Goldstein DG.
2012. The problem is beyond psychology: the real world is more random than regression analyses. International Journal of Forecasting 28:715-716.

Towne B, Demerath EW, Czerwinski SA.
2012. The genetic epidemiology of growth and development. In: Cameron N and Bogin B, eds. Human growth and development, 2nd edition. London, UK: Academic Press, 173-224.

Trotter M, Gleser GC.
1952. Estimation of stature from long bones of American Whites and Negroes. American Journal of Physical Anthropology 10:463-514

Tsai A, Stamoulis C, Dixby SD, Breen MA, Connolly SA, Kleinman PK.
2016. Infant bone age estimation based on fibular shaft length: model development and clinical validation. Pediatric Radiology 46:342-356.

Tyrell AR, Reilley T, Troup JDG.
1985. Circadian variation in stature and the effects of spinal loading. Spine 10:161-164.

UCL Institute of Health Equity
2015. Social inequalities in the leading causes of early death: a life course approach. London, UK: Department of Health.

United Nations Children's Fund.
2014. Hidden in plain sight: a statistical analysis of violence against children. New York City, NY: UNICEF.

United Nations Development Program.
2015. Human development report 2015: work for human development. United Nations: New York, NY.

U.S. Census Bureau.
2010. 2006-2010 American Community Survey.

Wales JKH, Herber SM, Taitz LS.
1992. Height and body proportions in child abuse. Archives of Disease in Childhood 67:632-635.

Weaver D.
1980. Sex differences in the ilia of a known sex and age sample of fetal and infant skeletons. American Journal of Physical Anthropology 52:191-195.

Wehner F, Schieffer MC, Wehner HD.
1999. Percentile charts to determine the duration of child abuse by chronic malnutrition. Forensic Science International 102:173-180.

World Health Organization.
1983. Measuring changes in nutritional status: guidelines for assessing the nutritional impact of supplementary feeding programs on vulnerable groups. Geneva, Switzerland: World Health Organization.

World Health Organization.
2016. International classification of diseases and health related problems. 10th revision, 5th edition. Geneva, Switzerland: World Health Organization

WHO Multicenter Growth Reference Study Group.
2006a. WHO child growth standards based on length/height for age. Acta Paediatrica Supplement 450:76-85.

WHO Multicenter Growth Reference Study Group.
2006b. Enrolment and baseline characteristics in the WHO Multicenter Growth Reference Study Group. Acta Paediatrica Supplement 450:7-17

Wilkinson RG, Pickett KE.
2006. Income inequality and population health: a review and explanation of the evidence. Social Science & Medicine 62:1768-1784.

Yamaoka Y, Tamaiya N, Fujiwara T, Yamasaki Y, Matsuzawa A, Miyaishi S.
2015. Child deaths with persistent neglected experiences from medico-legal documents in Japan. Pediatrics International 57:373-380.