Panel regression formulas for stature and body mass estimation in immature human skeletons

By: Gwen Robbins Schug, Sat Gupta, Libby W. Cowgill, Paul W. Sciulli, and Samantha H. Blatt

Robbins Schug, G., Gupta, S., Cowgill, L.W., Sciulli, P.W., & Blatt, S. (2013). Panel regression formulas for stature and body mass estimation in immature human skeletons. *Journal of Archaeological Science, 40(7),* 3076–3086. https://doi.org/10.1016/j.jas.2013.02.025

Abstract:

Anthropologists require methods for accurately estimating stature and body mass from the human skeleton. Age-structured, generalized Least Squares (LS) regression formulas have been developed to predict stature from femoral length and to predict body mass in immature human remains using the width of the distal metaphysis, midshaft femoral geometry (*J*), and femoral head diameter. This paper tests the hypothesis that panel regression is an appropriate statistical method for regression modeling of longitudinal growth data, with longitudinal and cross-sectional effects on variance. Reference data were derived from the Denver Growth Study; panel regression was used to create one formula for estimating stature (for individuals 0.5–11.5 years old); two formulas for estimating body mass from the femur in infants and children (0.5–12.5 years old); and one formula for estimating body mass from the femoral head in older subadults (7–17.5 years old). The formulas were applied to an independent target sample of cadavers from Franklin County, Ohio and a large sample of immature individuals from diverse global populations. Results indicate panel regression formulas accurately estimate stature and body mass in immature skeletons, without reference to an independent estimate for age at death. Thus, using panel regression formulas to estimate stature and body mass in forensic and archaeological specimens may reduce second stage errors associated with inaccurate age estimates.

Keywords: Stature | Body mass | Femur | Method | Bioarchaeology | Childhood | Forensic anthropology

Article:

1. Introduction

Bioarchaeologists and forensic anthropologists require methods to estimate body mass and stature from immature human skeletons. There is no doubt that femoral length can be used to estimate stature accurately in immature archaeological skeletons (Feldesman, 1992; Ruff, 2007; Telkka et al., 1962), but body mass estimation is more difficult for both adult (Smith, 1993) and immature remains (Ruff, 2007). Measures of the lower appendicular skeleton are an obvious choice for body mass estimation as supporting body weight is a principle mechanical strain in young infants and children (van der Meulen et al., 1993) and strong correlations exist among body mass, diaphyseal robusticity, and articular breadths (Ruff et al., 2013; Pearson and Lieberman, 2000; Robbins et al., 2010; Robbins Schug, 2011; Ruff, 2007). However, lower limb morphology is also shaped by a variety of other forces, including genetics, sex, activity, nutritional status, and hormones (Cowgill, 2008; Moro et al., 1996; Robbins, 2007; Robbins et al., 2010; Robbins Schug, 2011; Ruff, 2000, 2003a, 2003b, 2005a; Ruff et al., 1993; Sumner and Andriacchi, 1996; van der Meulen et al., 1993; van der Meulen et al., 1996; Wallace et al., 2012).

Mechanical and morphometric methods currently exist to estimate body mass from the immature femur. Generalized Least Squares (LS) regression formulas were developed to estimate body mass in individuals 1–12 years of age using either the distal femoral metaphyseal breadth (Ruff, 2007) or a measure of femoral midshaft cross-sectional geometry (polar second moments of area, or *J*) (Robbins et al., 2010). For adolescents (12–17 years), the femur metaphyseal breadth and the midshaft are both poor estimators of body mass but the diameter of the femoral head is a significant predictor (Ruff, 2007). However, all of these methods were created using reference data from the Denver Growth Study, data that are affected by an autoregressive trend and an increase in variance between the dependent and independent variable with age. Thus, a technique in regression called “panel regression” represents a clearer alternative for developing prediction formulas from these data.

The goal of this paper is to examine the utility of a panel regression approach for developing methods from age structured data for use in forensics and bioarchaeology. When Generalized LS regression techniques are used for estimation from a reference sample with longitudinal and cross-sectional variation, the autocorrelation is dealt with by breaking the sample up into age categories. The problem with this approach is that it requires independent age estimates. Techniques for age estimation in archaeology must be chosen based on which skeletal elements are preserved, not necessarily which technique is most accurate. Each age estimation technique has a unique set of errors and commonly, the variance increases with age. Similarly, formulas for stature and body mass also have their own error term. When age-structured formulas are used, the second stage error of prediction is compounded by the first stage error of age estimation.

Panel regression theoretically represents an appropriate alternative to Generalized LS regression for developing prediction models from age-structured data in anthropology because it was developed specifically for data with multiple subjects and a time series aspect (i.e. repeated measurements on each subject at regular time intervals) (Gatignon and Hanssens, 1987; Mela et al., 1998). Potentially, there are several advantages of using a panel regression approach as compared to running the models using separate time series or cross-section data approaches (Baltagi, 2001). When both the cross-sectional and the longitudinal aspects are accounted for by the model, one has the benefit of using the larger data set (more than one measure per year), thereby increasing the degrees of freedom; the parameter estimates are thus more efficient. Also the scope of inference is broader since it allows prediction of stature and body mass without having to first estimate specific age. Because prediction is based on one formula applied to immature remains without regard to age, the error term remains consistent across the age pyramid.

In this paper, we provide a single formula for stature estimation based on femur diaphyseal length, developed for individuals 0.5–11.5 years of age (prior to epiphyseal fusion). Three formulas are provided for estimating body mass for individuals 0.5–17.5 years, based on three different femoral dimensions: midshaft *J*, the breadth of the distal end, and head diameter. The panel regression formulas were developed from the Denver Growth Study data and applied to a target sample of known stature and body mass from Franklin County, Ohio. Estimates from age structured generalized LS regression formulas were compared to estimates from panel regression formulas to evaluate the correlation among the two sets of estimates and the significance of any differences. We then examined the utility of the panel regression approach for global populations that represent a wide range of human variation in body shape and size by comparing estimates from the two statistical methods for seven archaeological populations from diverse temporal and geographic contexts.

2. Materials and methods

2.1. Materials and methods

The panel regression model was applied to the Denver Growth Study data, a longitudinal sample of measurements from 20 well-nourished immature individuals, selected from a database compiled by the Denver Child Research Council from 1941 to 1967 and used in several previous studies (Robbins et al., 2010; Robbins Schug, 2011; Ruff, 2003a, 2003b, 2005a, 2007). Radiographs were made for the Denver Study at 2, 4, 6, and 12 months for the first year of life and at 6 month intervals through the age of 17.5 years. Ruff measured femoral lengths, femoral head diameters, and cortical bone thicknesses (at 45.5% of diaphyseal length) from the Denver sample anteroposterior radiographs (Ruff, 2003a, 2003b). Medullary diameter (M) was calculated as diaphyseal external diameter (T) minus combined cortical thickness, and torsional rigidity, J, as $\pi/32 \times (T^4 - M^4)$, assuming a concentric elliptical model (O'neill and Ruff, 2004). Magnification error was corrected as described previously (Ruff, 2007). An intra-observer measurement error of 3.1% for J was reported (Ruff, 2007).

A panel regression formula was created to estimate stature from the length of the femoral diaphysis. The analysis of stature was limited to younger individuals because after age 11 years, bone diaphyseal length data are affected by epiphyseal fusion; thus, diaphyseal length measurements are only available in younger individuals whose secondary centers are unfused. For this analysis, all of the data points were used for each individual, when available, between the ages of 0.5 and 11.5 years. Three formulas were created from these data to estimate body mass. For younger individuals (0.5–12.5 years), body mass can be estimated from the breadth of the distal end of the femur (Ruff, 2007) and femur midshaft torsional rigidity, or J (Robbins et al., 2010). For older individuals (6.5–17.5 years old), body mass is traditionally estimated from the femoral head diameter (Ruff, 2007). Thus, the reference population was divided into two samples – younger individuals in age categories 1–12 and older individuals in age categories 7–17. This division was based on the strength of the relationship between body mass and three morphological measures demonstrated in previous studies (Robbins et al., 2010; Ruff, 2007).

Outliers, defined as individuals with values for BMI (body mass index, body mass in kilograms/stature in meters^2) outside the 95% confidence limits for age according to national standards (Must et al., 1991), were eliminated in previous studies because they dramatically reduced the accuracy and precision of the regression models (Robbins et al., 2010; Ruff, 2007). The outliers consisted of one female (with high BMI from ages 3.5–8 years) and one male (with high BMI from ages 5.5–8 years). In the present study, these outliers were not eliminated from the sample used to create a formula to predict stature from bone length; the outliers were eliminated in the sample used to create formulas for body mass estimation, reducing that sample size to 9 males and 9 females (18 individuals). The sexes were pooled in this analysis because sex determination is difficult, inaccurate and not commonly attempted in studies involving archaeological immature skeletal samples.

The dependent and/or independent variables were missing for 3.75% of the sample 12.5 years and under. Linear interpolation was used to estimate these measurements. After interpolation, the total sample size of diaphyseal length and stature measurements was 440 (22 measurement events on 20 individuals 0.5–12.5 years of age); 432 measurements of *J*, distal end width, and body mass were used (24 measurement events on 18 individuals 0.5–12.5 years of age); and, 378 measurements of the femoral head and body mass were used (21 measurement events for 18 individuals from 7 to 17 years of age). Panel regression is not tolerant of empty cells in some panels so measurements for individuals taken prior to 0.5 years of age were excluded from the analysis because missing data points could not be interpolated. It is also true that femoral dimensions do not have a strong correlation with body mass in infants less than 6 months of age (Robbins et al., 2010; Ruff, 2007).

2.2. Statistical modeling

Panel regression analysis was performed in SAS (Version 9.2). Panel regression utilizes both cross-sectional and longitudinal aspects of variation in the data. The cross-sectional aspect pertains to the set of individuals in the sample on which observations are taken. The longitudinal aspect pertains to the time series component, that periodic observations were made using the same sample of individuals over a particular time span. Each time series is called a panel. The model can be developed using fixed effect or random effect models. A fixed effect model investigates how the intercepts in the regression model vary across group and/or time periods while the random effect model investigates how the error variance structures are affected by group and/or time. A one-way model uses dummy variables in the regression model only for one factor (group or time but not both) while a two-way model uses two sets of dummy variables corresponding to both group and time. The fixed effect model used here was based on a specification test (Hausman, 1978) that compares the utility of fixed effect and random effect models.

In this analysis, the type of model to be fitted and the error structure must be specified in the model statement. During preliminary analysis it was observed that in each of the time series corresponding to different subjects there was a strong positive *lag*(1) auto-correlation. This was confirmed by the Durbin–Watson (DW) Test (Durbin and Watson, 1950, 1951) which measures serial independence of errors against the hypothesis of first order autoregressive behavior. Typically a DW statistic value less than 2.0 indicates significant positive *lag*(1) auto-correlation.

For both stature and body mass, all DW values were less than 2.0. For stature, 85% DW values were below 1.5; and for body mass, 95% DW values were below 1.5 Because of this auto-correlation, we used the Parks (1967) option to describe the error structure in our panel regression model. Our fitted model is given by

$$Y_{it} = \propto + \beta X_{it} + u_{it}, i = 1,2,\ldots N; T = 1,2,\ldots t \tag{1}$$

where N = number of subjects, T = number of time points, Y_{it} = Observation on the response variable corresponding to the ith subject at time t, X_{it} = Observation on the predictor variable corresponding to the ith subject at time t, α is the intercept term, β is the slope and u_{it} is given by

$$u_{it} = \rho_i u_{it-1} + \varepsilon_{it} \tag{2}$$

where ε_{it} is the zero-mean error term with no correlation within the same panel but possibly non-zero correlations across different panels and ρ_i is the $AR(1)$ (auto-regressive model of order 1) parameter for the ith panel. The procedure first estimates $\rho_i's$. The covariance matrix of the error term u_{it} is then estimated using Ordinary Least Square regression on the transformed data given by

$$Y_{it} - \hat{p}_{it} Y_{it-1}, X_{it} - \hat{p}_{it-1} \tag{3}$$

Finally (α,β) were estimated using the Generalized LS method. The terms R^2 (coefficient of determination), *SSE* (Standard Error of the Estimate), and *MSE* (Mean Square Error) are based on the fitted Generalized LS estimates. These quantities and standard errors of estimated regression coefficients are all reported in the SAS output (for details, refer to Parks (1967) or SAS Version 9.2).

2.3. Applying the formulas to the target samples

In this research, four formulas were developed – one to predict stature from femoral diaphyseal length; one to predict body mass from the width of the distal end of the femur; one to predict body mass from midshaft *J*; and, one to predict body mass from femoral head diameters. The panel regression formulas were applied to an independent target sample from the Franklin County, Ohio Coroner's office (Pfau and Sciulli, 1994; Sciulli, 1994; Sciulli and Blatt, 2008); the formula for estimating body mass from the femoral head diameter was excluded from this analysis because those data were unavailable. If panel regression represents an appropriate method for developing prediction equations for use in anthropology, then the estimates from age structured formulas and panel regression formulas should not be significantly different when chronological age is known for the target sample.

The Ohio sample consists in total of 186 immature individuals, 0.04–20 years of age, who died between July 1, 1990 and June 30, 1991. The present analysis was limited to 36 individuals between one and 12 years of age. The sample includes European-American and African-American males and females. Dates of birth, death, sex, ancestry, body mass, and stature were obtained from medical records. Long bones were radiographed shortly after death (Pfau and Sciulli, 1994; Sciulli and Blatt, 2008). Blatt collected the following measurements from the radiographs: femoral distal metaphyseal breadth and external diaphyseal and medullary breadths (at 50% of diaphyseal length). Blatt calculated polar second moments of area (*J*) using the method described above for the calculation in the Denver sample. Intra-observer error was evaluated on measures of twenty individuals (17.8%) made on two separate occasions. The mean standard deviation was 0.12 mm for the midshaft diameter and 0.47 mm for the medulla. Stature and body mass formulas were applied to this sample; bias was measured and compared with age-structured formulas. Bias was defined as the signed difference between observed and predicted values (Sciulli and Blatt, 2008).

The panel regression formulas were also applied to a large sample of immature remains from seven global populations (Table 1). Our goal in this comparison was to examine the utility of using a panel regression approach for geographically and temporally diverse populations, representing a variety of lifestyle and subsistence behaviors. Theoretically, the main advantage to using panel regression is that independent age estimates are not required to apply the method and thus second stage error is reduced. However, this is only true if there are no significant interaction effects between the statistical method and population level differences in mean stature and body mass. Estimates of stature and body mass were compared for four different measures of the skeleton. Plots of the estimated marginal means and repeated measures ANOVA were used to examine the within and between subjects effects.

Table 1. Provenience and sample sizes of archaeological samples used in this study.

Sample	Original location	Approximate time period	Sample size in this study	Sample location
California Amerindian	Northern California	500–4600 BP	63	Phoebe A. Hearst Museum at the University of California, Berkeley (Berkeley, CA)
Dart	Johannesburg, South Africa	20th century	41	School of Medicine, University of Witwatersrand (Johannesburg, South Africa)
Indian Knoll	Green River, Kentucky	4143–6415 BP	70	University of Kentucky, Lexington (Lexington, KY)
Kulubnarti	Batn el Hajar, Upper Nubia	Medieval	80	University of Colorado, Boulder (Boulder, CO)
Luis Lopes	Lisbon, Portugal	20th century	37	Bocage Museum (Lisbon, Portugal)
Mistihalj	Bosnia–Herzegovina	Medieval (15th century)	35	Peabody Museum at Harvard University (Cambridge, MA)
Point Hope	Point Hope, Alaska	300–2100 BP	50	American Museum of Natural History (New York, NY)

3. Results

3.1. Formula to estimate stature from femur length

For the stature data, 440 measurements were used (20 subjects with measurements at 22 regularly spaced time intervals for each subject 0.5–11.5 years of age). A scatter plot suggests no need for any log transformation (Fig. 1). Femoral measurements (x) were regressed on body size (y). For femur diaphyseal length and stature, most of the time series show a significant first order autoregressive structure with all DW values less than 2.0 and 85% of them falling below 1.5. Hence, the panel regression is modeled using the Parks (1967) option. The fitted model is

$$\text{Stature} = 31.0390 + 0.3221 * \text{femur diaphyseal length},$$

The R^2 value is 0.9995, with SSE = 397.0700 and MSE = 0.9065. The high R^2 value suggests that the model will predict stature from femur length with a high degree of precision. Since no log transformation was used, the final predicted values will be mean stature and no detransformation is required.

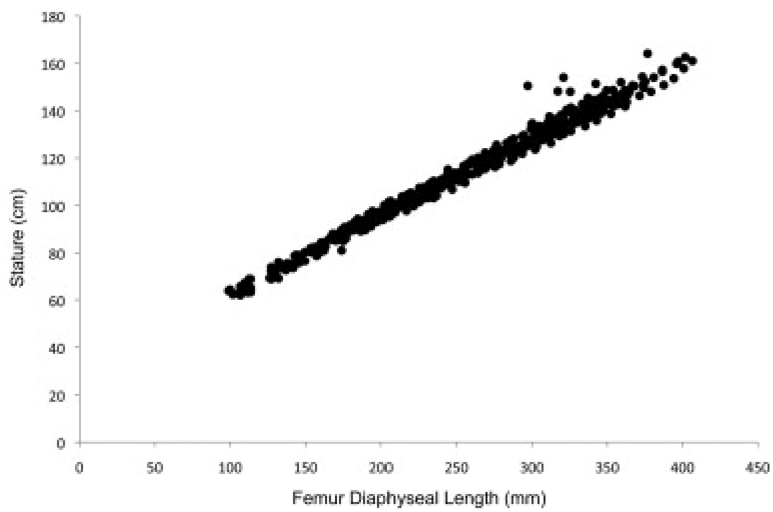

Fig. 1. Scatter plot of stature (y) versus femur diaphyseal length (x) demonstrates that the data do not require log transformation.

A simple scatter plot of body mass versus the width of the distal end of the femur, midshaft J, and the diameter of the femoral head suggests a clear need for log transformation in these data. Log transformation has two main objectives – it attempts to normalize the data and tries to stabilize the variance. We used a natural log transformation (ln) and it may be noted that variance was significantly stabilized by this transformation (Fig. 2). Also, for the analyses involving the midshaft and distal end, body weight (y) was not normal without log transformation (Kolmogorov–Smirnov test, *p-value* = .024), but normality was achieved after log transformation (Kolmogorov–Smirnov test, *p-value* = .254). Despite transformation, the femoral head diameter data did not achieve normality. It may however be noted that with sample size this large, normality is not very critical. Again, a strong first order auto-regressive behavior in each of the time series was confirmed using the Durbin–Watson test. For all of the panels, the DW value was less than 2 with 95% of the panels showing a DW value of less than 1.5.

3.2. Formula to estimate body mass from femur midshaft geometry

Using the Parks (1967) option within SAS, the following model was generated for the 432 measurements of femur midshaft J and body mass (24 measurement events on 18 individuals 0.5–12.5 years of age):

$$\ln\{\text{bodymass}\} = 2.0683 - 0.3126 * \ln\{J\} + 0.0477 * \ln\{J\}^2,$$

It was observed that the second order model provided better fit as compared to the first order model for this as well as the other two predictors (in the models given below).

The R^2 value is 0.9719 with SSE = 373.0541 and MSE = 0.8696 Again, the high R^2 value suggests that the model will predict body mass from femur midshaft J with a high degree of precision.

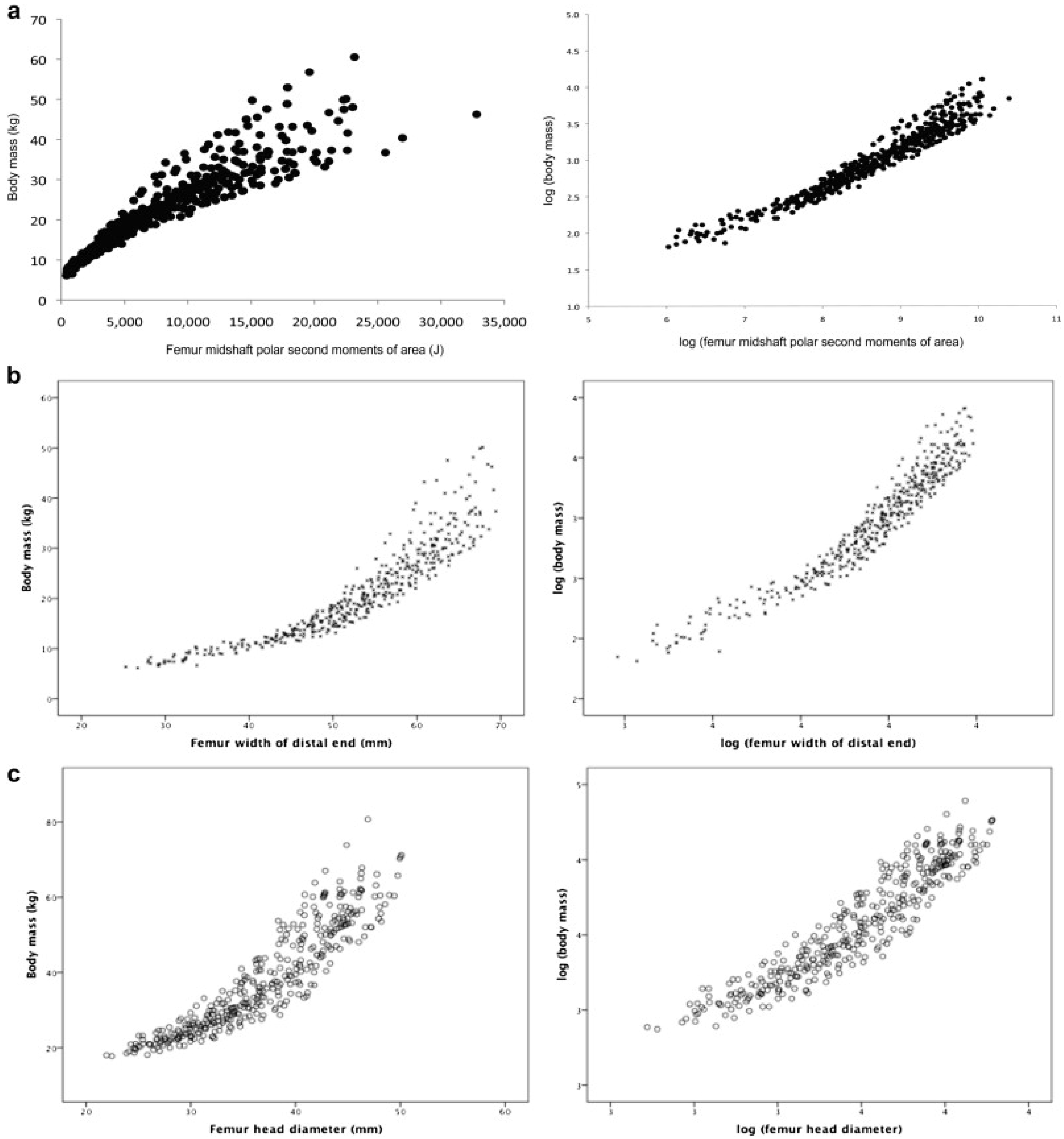

Fig. 2. Scatter plots of body mass (*y*) versus femur dimensions (*x*) demonstrate that the data clearly require log transformation (left) and that transformation achieved variance stabilization of the variance (right) in the midshaft polar second moments of area (*J*) (a, top row) and width of the distal end of the femur (b, middle row). Variance was not stabilized in the femoral head diameters (c, bottom row).

3.3. Formula to estimate body mass from the breadth of the distal metaphysis of the femur

Another model was generated using 432 measurements of the breadth of the distal end of the femur (24 measurement events on 18 individuals 0.5–12.5 years of age):

$$\ln\{\text{body mass}\} = 13.0615 - 7.3338 * \ln\{\text{breadth of the distal end of the femur}\} + 1.2058 * \ln\{\text{breadth of the distal end of the femur}\}^2,$$

The R^2 value is 0.9010 with *SSE* = 372.7655 and *MSE* = 0.8689. The high R^2 value suggests that the model will predict body mass from the breadth of the distal end of the femur with precision, although this measure is slightly less effective as a predictor than the midshaft.

3.4. Formula to estimate body mass from the diameter of the femoral head

Finally, using 378 measurements of the femoral head and body mass (21 measurement events for 18 individuals from 7 to 17 years of age) the following model was generated to predict body mass for older subadult skeletons:

$$\ln\{\text{body mass}\} = 11.5770 - 6.2969 * \ln\{\text{diameter of the femoral head}\} + 1.1297 * \ln\{\text{diameter of the femoral head}\}^2,$$

The R^2 value is 0.9880 with *SSE* = 323.3666 and *MSE* = 0.8623. The high R^2 value suggests that the model will also predict body mass from the diameter of the femoral head with a high degree of precision for older juvenile and adolescent skeletons.

3.5. Issues in detransformation of estimates

Caution needs to be exercised in detransforming the log values for body mass obtained using these fitted models. We cannot simply take the anti-log of the predicted ln{body mass} value since this will give us predicted median value and not the mean value. Either predictions must be made on the log scale or it should be recognized that the predicted value is median body mass, not mean body mass. To illustrate the problem, suppose log transformation is applied to a skewed variable X to get a normalized variable Y. Then we have,

$$\begin{aligned} \text{Mean}(Y) &= \text{Median}(Y), && \text{since } Y \text{ has a normal distribution} \\ &= \text{Median}(\text{Log } X) && \text{since } Y = \text{Log}(X) \\ &= \text{Log}(\text{Median}(X)) && \text{since log preserves ordering} \end{aligned}$$

Hence Antilog{Mean(Y)} = Median(X).

Using a bias correction approach when taking the anti-log of the predicted mean value on log scale (Smith, 1993) is not sound mathematically since it will give predicted median value and not the mean value, as explained above. When detransformed values are used in the results of this study, the estimates are reported as medians, not means.

3.6. Application to the target samples

The panel regression formula for stature made accurate predictions from femur diaphyseal length when applied to an independent target sample of children from Franklin County, Ohio (Table 2). Unfortunately, the target sample included only 12 measurements for femur length (from 12 individuals in age categories one and two), which limits the robusticity of this analysis. The bias, or signed difference between the two estimation techniques, was 1.25–2.61 cm. A comparison of observed versus predicted values for stature in the Ohio cadaver sample demonstrates that the age structured formulas ($R^2 = 0.922$) and the panel regression formulas ($R^2 = 0.912$) both predict stature with a relatively high degree of accuracy when age is known (Fig. 3a). The two statistical approaches also produced estimates that are strongly correlated with one another (Fig. 3b). For samples of known age, the age structured formulas are slightly more accurate however, if age is unknown, the panel model should be preferred for stature estimation to limit second stage error from age estimation.

Table 2. Bias[a] of two statistical techniques to estimate stature in the Ohio cadaver sample.

			Bias (cm)	
Age category	***n***	**Median stature (cm)**	**Age structured[b]**	**Panel regression**
1	10	126.9	1.25	1.68
2	3	165.5	2.19	2.61

[a] Bias (observed stature − estimated stature).
[b] Equations in Ruff (2007).

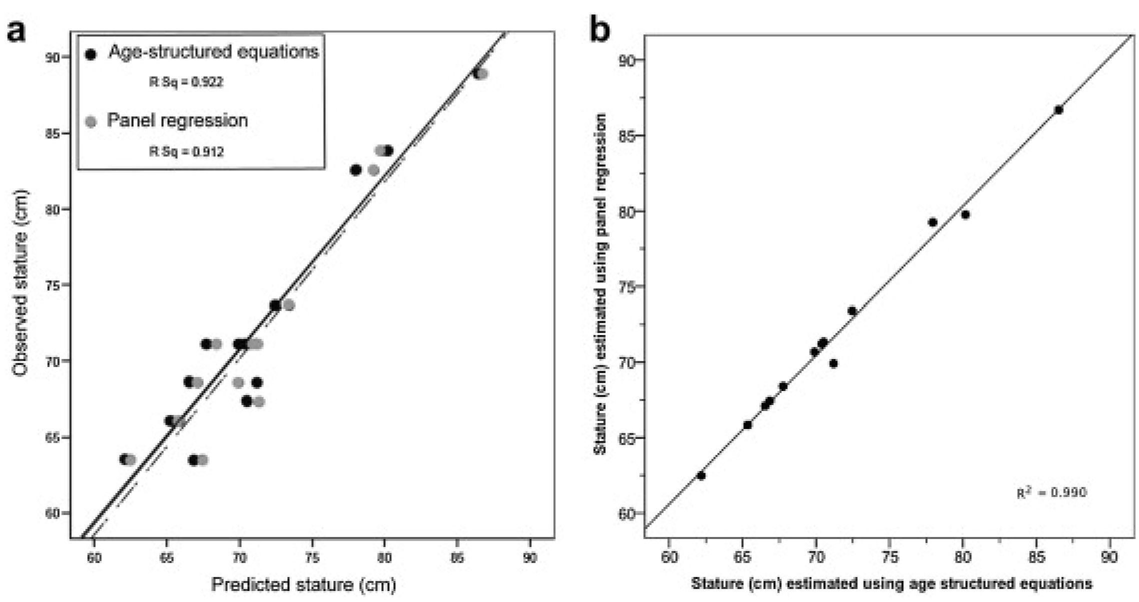

Fig. 3. Scatter plot of observed versus predicted stature (a) demonstrates that the panel regression formula predicts stature with accuracy ($R^2 = 0.912$) comparable to previous age structured formulas ($R^2 = 0.922$) from Ruff (2007). The two statistical approaches produced estimates that are strongly correlated (b).

The age structured and panel regression formulas to predict body mass were also applied to 36 individuals from the Ohio target sample, age 0.5–12.5 years (Table 3). Femoral head diameters were not available for the Ohio sample so this analysis is limited to two skeletal measures: the breadth of the distal end of the femur and the midshaft (*J*). The body mass estimates from both statistical approaches are strongly correlated with one another, for both the distal end of the femur (R^2 = 0.965) and the midshaft (*J*) (R^2 = 0.997) (Fig. 4).

Table 3. Bias[a] and %SEE[b] for body mass estimates by age from the femur midshaft *J* in the Ohio cadaver sample.

			Bias (kg) Midshaft *J*		Bias (kg) metaphysis		%SEE for age structured LS regression formulas	
Age category	*n*	Median weight (kg)	Age structured	Panel regression	Age structured	Panel regression	Metaphysis	Midshaft
1	11	9.29	0.43	−1.06	0.36	−0.32	6.7	7.1
2	9	13.12	−0.91	−1.55	1.41	−0.12	5.9	4.8
3	3	15.74	−0.52	−2.72	1.30	−1.76	6.8	4.8
4	2	15.89	2.59	−4.73	0.83	−0.44	6.5	6.5
5	1	19.52	−5.26	−0.02	1.74	0.50	6.1	6.2
6	1	25.88	1.85	−6.99	6.37	5.62	6.4	6.6
7	2	32.01	2.42	−8.34	7.60	3.55	6.4	6.3
8	1	31.78	9.29	−14.13	7.51	8.78	7.2	9.2
9	1	30.87	0.21	−6.30	−2.73	−3.49	14.3	14.4
10	1	39.95	−25.95	25.26	1.14	2.44	15.8	15.8
11	1	44.49	7.72	−14.13	3.28	5.39	16.9	18.0
12	3	41.62	−4.90	−0.71	−9.51	−6.51	16.4	17.6
Pooled sample	36	18.90	−0.61	−2.18	0.73	−0.11	9.6	9.8

[a] Bias (observed body mass − estimated body mass).
[b] %SEE = SEE/mean body mass (kg) for age.

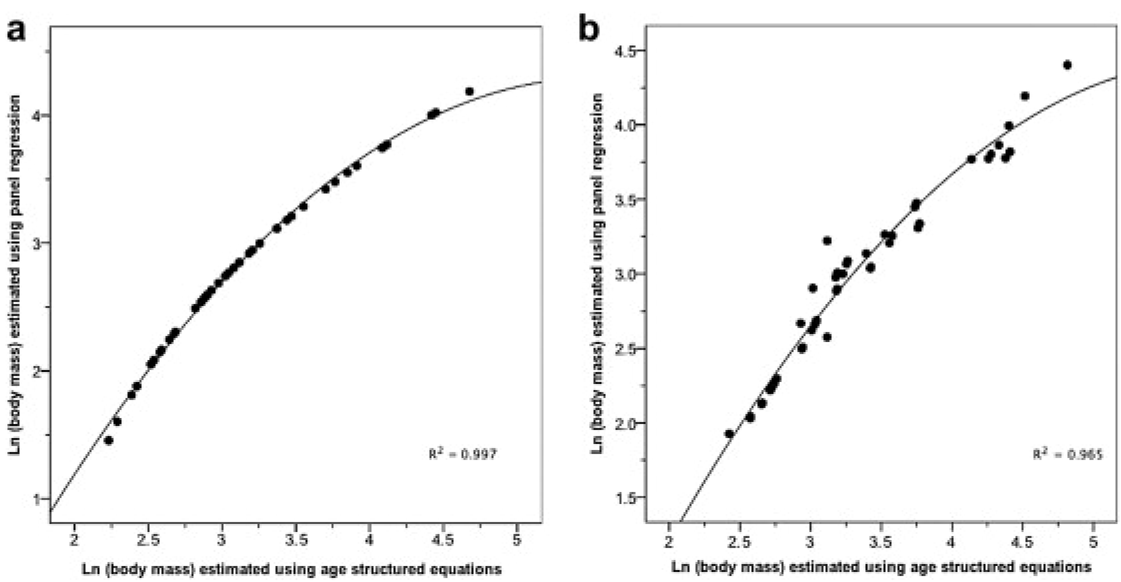

Fig. 4. Scatter plots of body mass estimated using panel regression (y) versus age structured formulas (x) demonstrate that the two statistical approaches predict body mass with similar accuracy using the midshaft (a) and the distal end of the femur (b) on the Ohio cadaver sample. Accuracy of both skeletal measures was also similar, as measured by the correlation between observed and predicted estimates (R = 0.866 and 0.864, respectively).

If accuracy is measured by %SEE, then the panel model should be preferred for both skeletal measures. The %SEE for both panel models was 5.1%, compared to 7.6% and 7.7 %SEE for the midshaft and the breadth of the metaphysis. However, an examination of the level of bias (the signed difference between the two sets of estimates) indicates that there are significant differences in the level of accuracy between the two statistical approaches and the two skeletal measures.

The panel model is more accurate when body mass was estimated from the breadth of the distal metaphysis in children 1–12 years. When observed and expected values are compared, bias was −0.11 kg for panel regression and 0.73 kg for the age structured formulas. However, the age structured formulas were more accurate when body mass was estimated using the midshaft (*J*). In this comparison, bias was −2.18 kg for panel regression and −0.61 kg for the age structured formulas. In a previous publication, it was argued that body mass can be estimated with similar levels of accuracy using either skeletal measure for children 1–7 years of age; accuracy declines for estimates made using the midshaft for children 8–12 years of age (Robbins et al., 2010). The present analysis demonstrates that age structured formulas should be preferred when body mass is estimated for children 1–7 years using the midshaft femur.

Although it has been used for this purpose in previous publications, the Ohio cadaver sample is small and limited in its power to test formulas for stature and body mass. Stature and body mass were also estimated for a large sample of immature remains from seven global populations to compare the results for the panel regression versus age structured regression formulas published previously (Robbins et al., 2010; Ruff, 2007).[1] Estimates of stature from the two statistical approaches were not significantly different ($T = 1.065$, $df = 337$, $p\text{-value} = 0.288$) (Table 4). There was a strong correlation ($r = 0.957$, CI = 0.947–0.965, $n = 331$) between estimates made using both age structured and panel regression formulas (Fig. 5). Age was estimated using only dental development for 284 individuals (85.8% of the archaeological sample). The two sets of stature estimates were also strongly correlated in this sub-sample, with confidence intervals (CI) that overlap with the total sample ($r = 0.951$, CI = 0.938–0.961), indicating that age estimation did not contribute significantly to second stage error in this analysis.

Repeated measures ANOVA method was used to examine the within and between subject effects for stature estimates made using the age-structured and panel regression formulas (Fig. 6). Population membership had no significant effect on the accuracy of stature estimates. Although there is variation in mean stature among these different prehistoric populations, there was no significant effect related to use of the different statistical methods for estimation ($p\text{-value} = 0.677$) and there was no significant interaction between population membership and estimation method ($p\text{-value} = 0.131$). Thus, the panel regression formula for stature performs well when applied to global populations, despite differences in mean stature and temporal, geographic, and biological variation.

[1] Caution is warranted when applying these formulas here because the Denver sample has a limited range of variation compared with the geographically and temporally diverse archaeological population. The ranges of variation for femur length, midshaft *J*, and the distal end of the femur are narrower in the Denver sample than in the archaeological samples (see tables in Supplemental materials).

Table 4. Median and range of estimates for stature and body mass (untransformed data) in a sample of subadults from seven archaeological sites.

	Stature (cm)[a] femoral length				Body mass (kg) midshaft *J*				Body mass (kg) distal end				Body mass (kg) femoral head			
Age	*n*	Age structured formulas	Panel regression formula	Median bias[b] (cm)	*n*	Age structured formulas	Panel regression formula	Median bias (cm)	*n*	Age structured formulas	Panel regression formula	Median bias (kg)	*n*	Age structured formulas	Panel regression formula	Median bias (kg)
1	61	70.47	71.29	0.82	61	8.68	8.21	−0.47	52	8.54	8.33	−0.17	.	.	.	.
2	40	80.54	80.15	−0.38	40	11.16	10.38	−0.78	37	10.48	10.59	0.11	.	.	.	.
3	27	87.11	86.10	−1.00	27	12.49	11.56	−0.93	24	11.42	10.86	−0.57	.	.	.	.
4	29	92.87	91.26	−1.60	29	13.67	12.21	−1.47	28	13.6	12.93	−0.66	.	.	.	.
5	29	99.41	98.66	−0.74	29	15.45	13.03	−2.42	25	14.88	13.74	−1.16	.	.	.	.
6	29	104.00	102.21	−1.72	29	18.18	15.73	−2.46	25	17.29	15.64	−1.65	.	.	.	.
7	30	110.54	109.30	−1.25	30	21.00	17.95	−3.05	28	19.49	17.46	−2.03	25	22.19	23.68	1.49
8	24	118.20	116.54	−1.66	24	22.55	20.16	−2.39	25	22.65	19.49	−3.12	24	23.86	24.56	0.69
9	25	119.68	118.96	−0.73	25	24.87	22.13	−2.74	23	26.05	20.09	−5.88	21	26.45	25.84	−0.57
10	20	123.82	122.82	−1.00	20	24.93	23.44	−1.49	20	26.44	22.43	−4.12	23	27.70	26.83	−0.86
11	24	118.57	125.88	7.31	24	28.40	25.41	−2.96	23	29.96	23.57	−6.41	22	30.18	28.69	−1.47
12					37	33.70	31.54	−2.15	19	31.19	25.53	−5.56	27	32.40	31.02	−1.31
13													21	36.70	34.44	−2.10
14													15	43.10	36.40	−6.70
15[c]													20	.[c]	39.27	.
16[d]													14	9.52[d]	44.15	34.63
17													5	58.35	49.41	−8.94
Total	375	98.99	98.19	−0.77	375	17.12	14.45	−1.44	329	15.80	14.30	−1.09	217	27.70	29.39	−0.77

[a] Stature was not estimated for individuals over 11.5 years of age.

[b] Median bias = (median estimate from panel regression formula) − (median estimate from age structured formulas).

[c] Age structured formula not provided for body mass estimation in 14.5–15.49 year olds (Ruff, 2007).

[d] Age structured formula for body mass estimation in 15.5–16.49 year olds not statistically significant (Ruff, 2007).

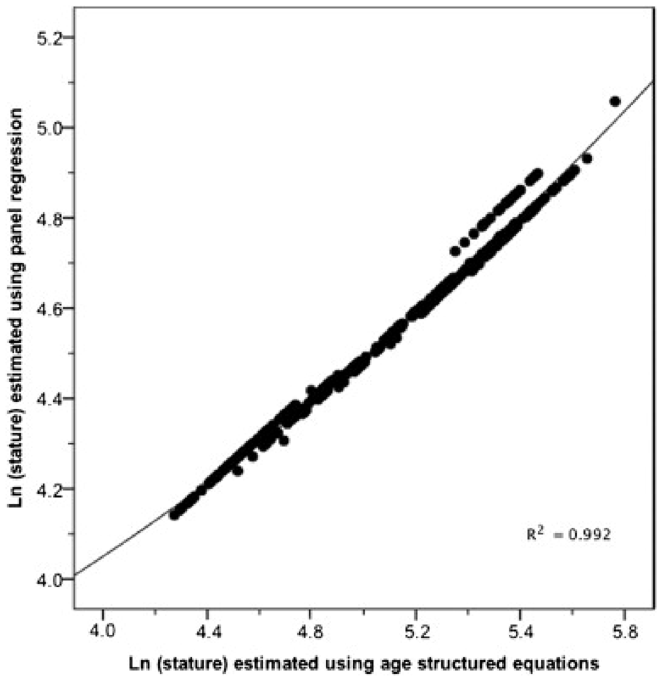

Fig. 5. Scatter plot of stature predicted using panel regression (*y*) versus age structured formulas (*x*) demonstrates that both methods predict stature with similar accuracy (R^2 = 0.992) in a large, global archaeological sample.

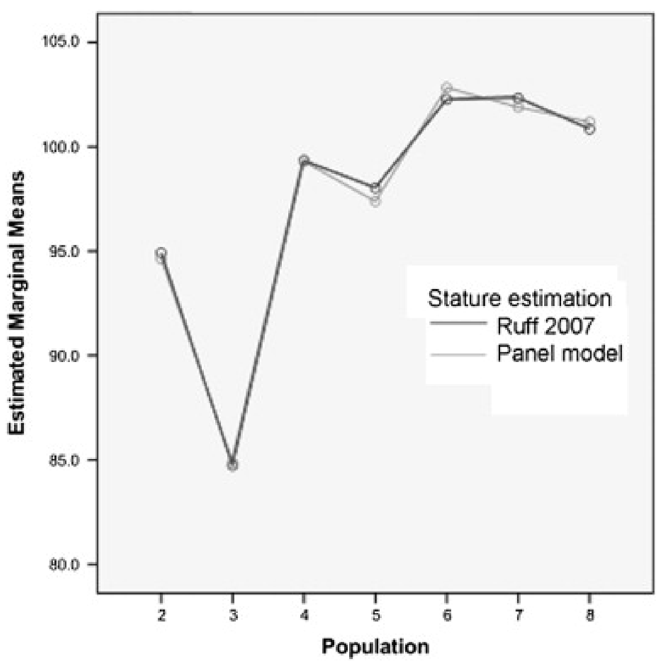

Fig. 6. Plot of estimated marginal means for stature estimates in seven archaeological populations using two estimation methods, panel regression and age structured LS regression formulas published previously (Ruff, 2007). The graph demonstrates there are no statistically significant effects from population membership, statistical technique, or the interaction of the two.

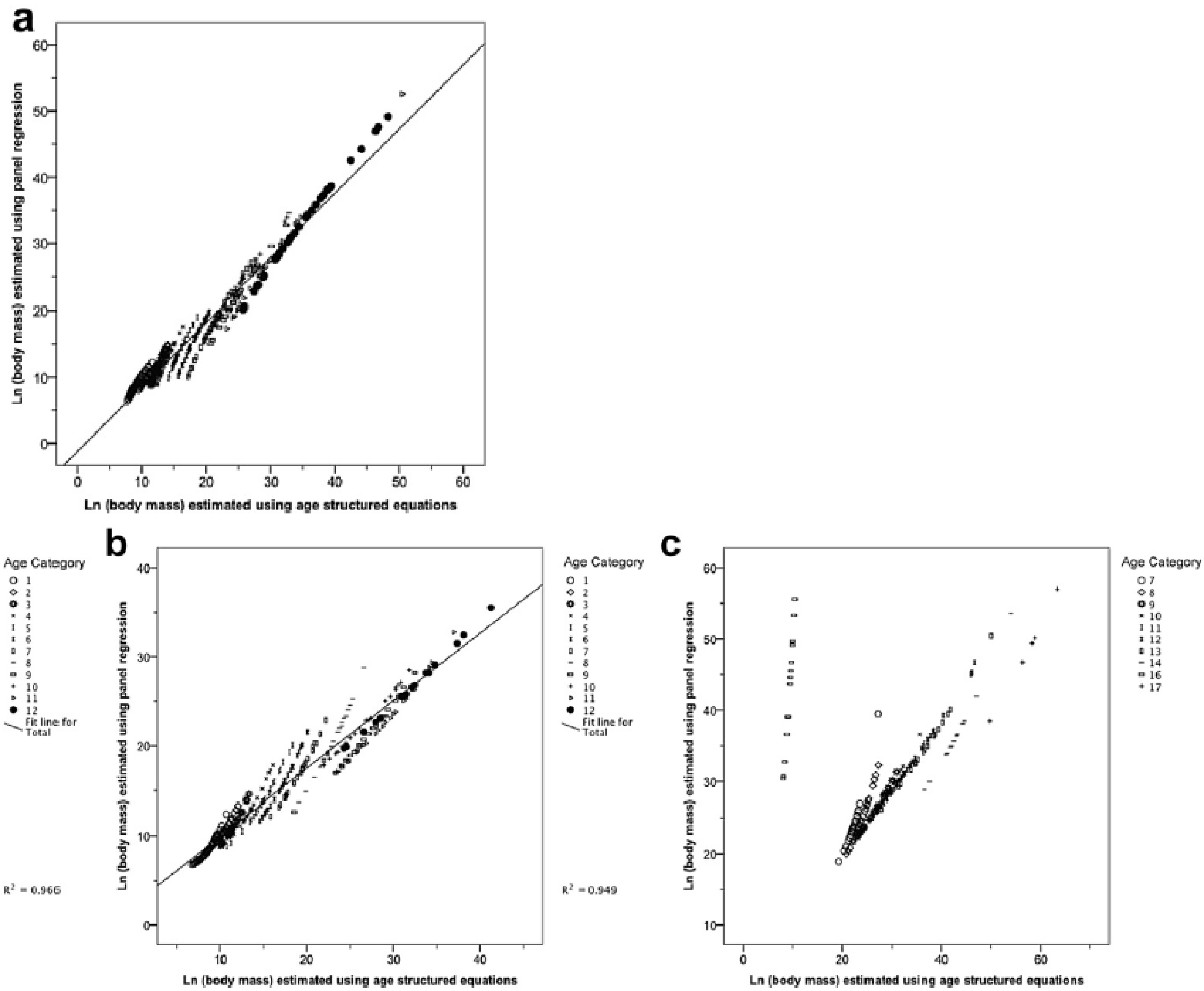

Fig. 7. Scatter plot of body mass predicted using panel regression (y) versus age structured formulas (x) demonstrates that there are few significant differences among the estimates from the two statistical approaches to estimation using the midshaft (a) and the metaphysis (b) in a large, global archaeological sample. Estimates made using the femoral head (c) differed significantly for individuals 15.5–16.49 years of age, probably because the predictors for the age structured formula for age 16 were not statistically significant (Ruff, 2007).

Body mass was also estimated for the archaeological samples using three skeletal measures: midshaft J, breadth of the distal end, and the diameter of the femoral head (Table 4). Estimates made using the two statistical approaches were compared. There was a strong correlation between body mass estimates made using the two statistical approaches for the midshaft femur (Fig. 7a) and the distal metaphysis (Fig. 7b). Despite this strong correlation, estimates made using the two statistical approaches were significantly different for the midshaft femur ($T = 21.813$, $df = 374$, p-$value < 0.001$) and the metaphysis ($T = 14.913$, $df = 328$, p-$value < 0.001$). When body mass estimates based on the femoral head were compared (Fig. 7c), the two statistical techniques also produced estimates that were significantly different ($T = -2.110$, $df = 196$, p-$value = 0.036$). This result is not surprising given that the skeletal measures used are predictive of body mass, but the relationship is not as strong as the correlation between stature and long bone length.

Plots of the estimated marginal means and repeated measures ANOVA were used to examine the significance of effects from population membership, statistical technique, and the interaction between these two variables. A significant effect from population membership indicates that mean body mass differs significantly among the different archaeological populations. We expect to see a significant effect from population membership because the archaeological samples included in this analysis are derived from seven highly diverse, global populations, which include temporal, geographical, genetic, social, and lifestyle variation. More important for this research are the ANOVA's that examine significant differences among statistical techniques and significant interaction terms for population membership and statistical technique because these demonstrate whether the different statistical techniques contribute to significant differences in body mass estimates, or whether the techniques perform similarly for the seven different samples (with different mean body mass).

A plot of the estimated marginal means for body mass in the 1–12 year olds demonstrates significant differences related to population membership (p-$value < 0.005$), confirming expectations about the diversity of body shape and size in these samples. More importantly, there was a significant effect from statistical method for both the midshaft (p-$value < 0.001$) and the breadth of the distal end of the femur (p-$value < 0.001$). The age structured generalized LS regression formulas systematically produced estimates for body mass that were higher than estimates made using panel regression for both the midshaft (Fig. 8a) and the breadth of the distal end (Fig. 8b). This result conforms to the results from the known age target sample from Ohio. Statistical method and population membership were not interacting significantly, supporting the use of the panel regression model in global populations with diverse body shapes and sizes.

Body mass was estimated for the older immature individuals (7–17 years) using the diameter of the femoral head (Ruff, 2007). Caution should be used in interpreting the results of this analysis because neither the age structured formulas published previously, nor the panel model formula for the femoral head provided here, have been verified using an independent target sample of known body mass. A plot of the estimated marginal means and a repeated measures ANOVA were again used to examine within and between subjects effects. When the femoral head is used to estimate body mass, there is no significant effect from population membership (p-$value = 0.122$). This result contradicts expectations based on temporal, geographical, and biocultural differences among these populations. There was a significant effect from statistical technique (p-$value = 0.001$) but the results did not demonstrate a consistent trend. Age structured and panel regression formulas produced significantly different results in four of the seven samples; estimates from the age structured formulas were lower than estimates made using the panel model (Fig. 9).

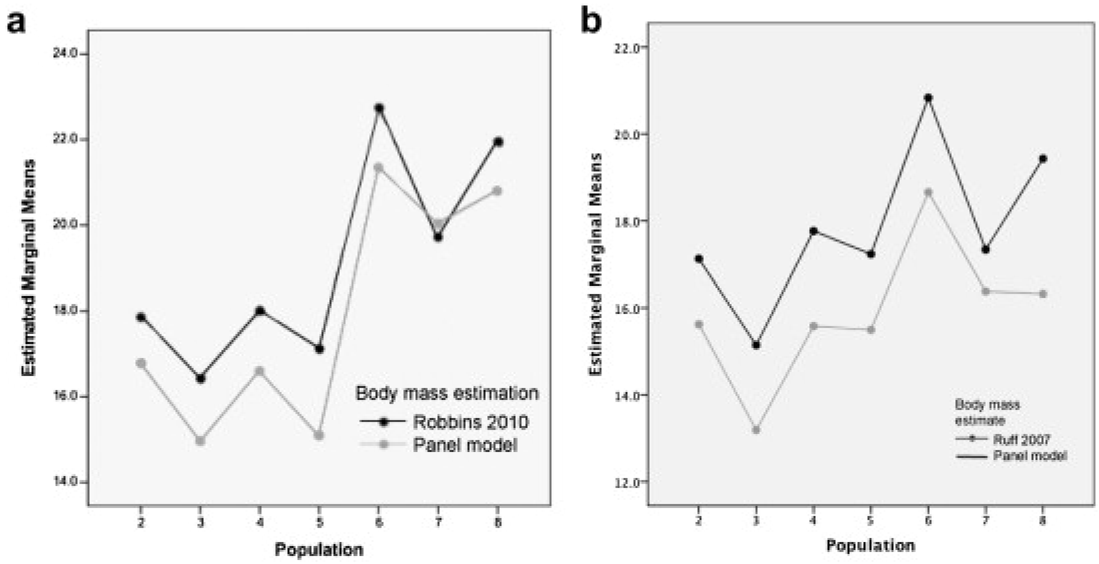

Fig. 8. Plot of estimated marginal means for body mass estimates in seven archaeological populations. Body mass estimates were made using both age structured and panel regression formulas based on the femoral midshaft (a) and the distal end of the femur (b) in immature individuals 1–12 years of age. Plots demonstrate no significant interaction among statistical technique and populations membership for this global sample of diverse body shape.

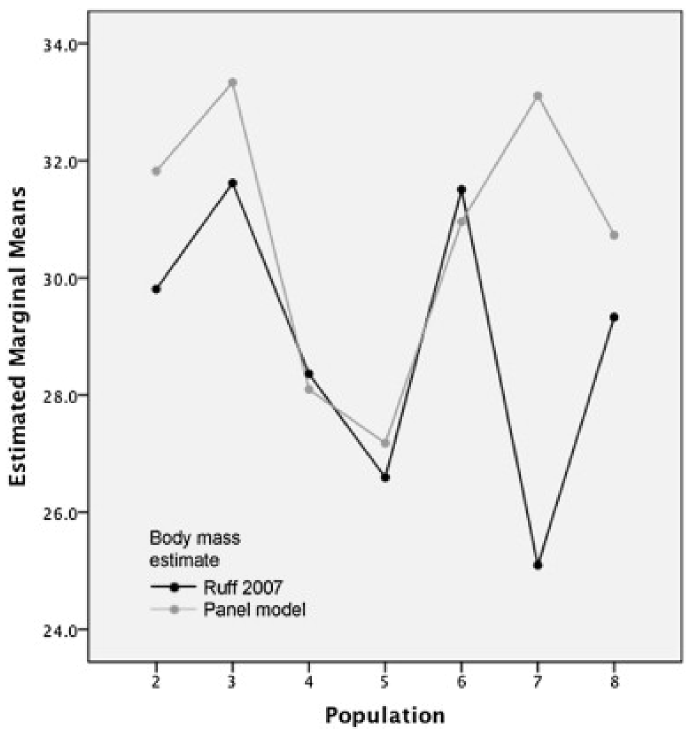

Fig. 9. Plot of estimated marginal means for body mass estimates in seven archaeological populations. Body mass estimates were made using both age structured and panel regression formulas based on the diameter of the femoral head in immature individuals 7–17 years of age. The plot shows significant interaction among statistical technique and population membership for this method.

The results suggest significant issues are present when estimating body mass in adolescent specimens using either the age structured or the panel formulas for the femoral head. For example, the means diverged most significantly for the sample from Mistihalj, for which there was a mean difference of 8 kg between the two techniques. Fourteen percent of this sample was in age categories 15 and 16. The age structured formulas for these two age categories were not statistically valid (Ruff, 2007) and an examination of Table 4 indicates the body mass estimates from the age structured formula for age category 16 are unlikely to be correct. There was one positive result from this analysis, for those who need to estimate body mass using the femoral head for adolescent skeletons; there was no significant effect from the interaction of method and population membership (*p-value* = 0.157). This result indicates that although this skeletal measure is relatively inaccurate for estimating body mass in adolescents, it performs with equivalent accuracy (or inaccuracy) in diverse populations (with different mean body mass).

4. Conclusion and discussion

This paper provides panel regression formulas for estimating stature and body mass in immature human skeletons. Formulas were developed using Denver Growth Study data, the same reference population used to create and validate age structured, generalized LS regression formulas published previously (Robbins et al., 2010; Ruff, 2007). Accuracy of the panel regression formulas was examined by applying these formulas to a cadaver sample from Franklin County, Ohio and a global sample of archaeological skeletons. Estimates from age-structured and panel regression formulas were compared for significant and/or systematic differences in the target sample.

Our results support the hypothesis that panel regression is an appropriate model for describing statistical relationships in or creating prediction equations from age structured reference data. Advantages to using panel regression instead of age structured generalized LS models are, 1) repeated measures in the reference sample can be included in the model, thereby increasing the degrees of freedom, 2) specific independent age estimates are not required to make predictions using the panel models, and 3) panel regression approaches are appropriate for anthropological applications to specimens from diverse temporal, geographic, social, and behavioral contexts.

Our results confirm the specific utility of the panel models for estimating stature and body mass in immature human skeletons. We compared stature and body mass estimates made using the panel model and the generalized LS regression model in a large sample of immature archaeological skeletons from global populations that represent a significant range of variation in stature and body mass (see supplemental materials). The stature estimates from the panel model were not significantly different from the generalized LS regression estimates for stature. Femur diaphyseal length is a very strong predictor of stature; the %SEE for age structured LS regression formulas ranges from 1.9 to 2.4% (Ruff, 2007). The slope of the line describing this relationship in the age structured formulas ranges from 0.269 to 0.320 across the entire subadult age range (1–17 years). Stature can be estimated with a similar level of precision using either panel regression or age structured generalized LS regression. The panel model should be preferred for ease of use in anthropological specimens, particularly when age is unknown, or dental material is not available for independent age estimates.

Body mass is more difficult to estimate in immature skeletons because the skeletal measures used are influenced by variables other than body mass. Some researchers have argued that the breadth of the distal femoral metaphysis should be preferred for body mass estimation in immature remains because midshaft geometry is sensitive to behavioral and environmental influences (Ruff et al., 1991, 1993, 2013; Trinkaus et al., 1994). While midshaft geometry responds more strongly to environmental cues, these include changes in body mass (Wallace et al., 2012). Supporting body weight is a principal mechanical strain for young infants and children and thus both skeletal measures demonstrate a strong correlation with body mass in 1–7 year old children from the Denver reference population and the Ohio sample. Anthropologists should choose the estimation method that provides the most accurate results for the skeletal element available—the panel regression formula for the breadth of the distal metaphysis in individuals 1–12 years of age and the age structured formulas for the midshaft in children 1–7 years. These body mass prediction equations for young children are appropriate for anthropological samples from diverse global populations.

Our results suggest a caution for the application of body mass formulas based on the breadth of the femoral head to adolescent skeletons. The diameter of the femoral head was used to estimate body mass in older subadults (7–17 years) from our global archaeological sample. Our concerns derive from the following: 1) neither the age structured formulas published previously (Ruff, 2007), nor the panel formula for the femoral head provided here, have been verified using an independent target sample of known body mass; 2) variation in mean body mass among diverse global populations was not detected using this method (I.e. these seven populations were expected to have different mean body mass but this skeletal measure did not detect that variation); 3) estimates for body mass from the panel and the generalized LS models were significantly different; 4) unlike the other predictors, there were significant interaction effects between population membership and statistical approach for this skeletal measure.

An examination of the original publication for the generalized LS formulas demonstrates the diameter of the femoral head is not a significant predictor of body mass for individuals in age categories 15 and 16 (Ruff, 2007). Indeed Table 3 in this study shows that the age structured formulas produced highly unlikely body mass estimates for the individuals in age category 16. Other skeletal measures (I.e. bi-iliac breadth) have yet to be tested on target samples of known age or archaeological samples of significant size. Our results suggest doing so might reveal significant issues with these predictors. Additional research on body mass estimation should be undertaken to specifically address the problems associated with prediction in adolescent skeletons.

Variation among the archaeological populations was demonstrated in this study and probably reflects extrinsic environmental variables (like latitude) and ontogenetic differences in activity levels, diet, disease status, as well as body mass. Because both the articulations and the midshaft are subject to different constraints, using both measures to estimate body mass in immature skeletons may elucidate interesting patterns of variation. For example, body mass, bone mass, and activity levels are all expected to be reduced in cases of severe biocultural stress, such as emaciation. Body mass and stature have already been used to examine the presence of stunting and wasting in archaeological populations (Robbins Schug, 2011). A comparison of body mass estimates from the articular end and the midshaft might yield interesting insights about skeletal growth and emaciation in immature skeletal remains (Robbins, 2007; Cowgill, 2008, 2010).

We hypothesized that the panel regression approach would theoretically have several advantages over age-structured LS regression formulas for applications in forensics and archaeology. First, because stature and body mass can be estimated without reference to age at death, this technique minimizes second stage errors; inaccurate estimates for age at death will not lead to inaccurate estimates for stature and body mass. Second, panel regression provides one formula for estimation in immature individuals without reference to specific ages, making application simpler. The results of this analysis support our hypothesis. However, the target sample is fairly small and additional research should test this idea further. Unfortunately, samples of immature individuals, with known age, stature, body mass, and the necessary measurements are relatively uncommon. The methods should also be further tested by application to additional archaeological samples, representing a greater share of the range of human variation in body shape and size. Certainly, when applying any of these methods to a target population, the range of variation in the target sample should be compared to the Denver reference sample before any of these formulas are applied (see supplemental materials). Future research should also examine the potential for using panel regression for other anthropological purposes. Any anthropological method developed from a reference population with both longitudinal and cross-sectional aspects of variation would potentially benefit from this approach.

Acknowledgments

This research was sponsored by the American Institute of Indian Studies, the George Franklin Dales Foundation, Fulbright IIE, the University of Oregon Department of Anthropology, the Leakey Foundation, the Wenner Gren Foundation, and the National Science Foundation (BCS-0549925). Data from the Denver Growth Study were provided by Chris Ruff. We would like to thank him for providing these data, comments on the manuscript, and for many productive discussions about the theoretical and practical aspects of body mass estimation in immature remains over the years. We received permission to use these data from the Lifespan Health Research Center at Wright State University and we thank them for allowing these data to be used here. The authors wish to thank the following institutions for granting access to the archaeological samples included in this analysis: Phoebe A. Hearst Museum at UC Berkeley School of Medicine; University of Witwatersrand; University of Kentucky, Lexington; University of Colorado, Boulder; Bocage Museum, Lisbon; Peabody Museum at Harvard university; and American Museum of Natural History.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.jas.2013.02.025.

References

Baltagi, B.H., 2001. Econometric Analysis of Panel Data. Wiley, John & Sons, New York.

Cowgill, L.W. 2008. The Ontogeny of Late Pleistocene Postcranial Robusticity [Dissertation]. Washington University, St. Louis. 319 pp.

Cowgill, L.W., 2010. The ontogeny of Holocene and late Pleistocene human postcranial strength. Am. J. Phys. Anthropol. 141, 16-37.

Durbin, J., Watson, G.S., 1950. Testing for serial correlation in Least Squares regression, I. Biometrika 37, 409-428.

Durbin, J., Watson, G.S., 1951. Testing for serial correlation in Least Squares regression, II. Biometrika 38, 159-179.

Feldesman, M.R., 1992. Femur stature ratio and estimates of stature in children. Am. J. Phys. Anthropol. 87, 447-459.

Gatignon, H., Hanssens, D.M., 1987. Modeling marketing interactions with application to salesforce effectiveness. J. Market Res. 24, 247-257.

Hausman, J.A., 1978. Specification tests in econometrics. Econometrica 46, 1251-1271.

Mela, C.F., Jedidi, K., Bowman, D., 1998. The long-term impact of promotions on consumer stockpiling behavior. J. Market Res. 35, 250-262.

Moro, M., van der Meulen, M., Kiratli, B., Marcus, R., Bachrach, L.K., Carter, D.R., 1996. Body mass is the primary determinant of midfemoral bone acquisition during adolescent growth. Bone 19, 519-526.

Must, A., Dalal, G.E., Dietz, W.H., 1991. Reference data for obesity: 85th and 95th percentiles of body mass index (Wt/Ht2) and triceps fold thickness. Am. J. Clin. Nutr. 53, 839-846.

O'neill, M.C., Ruff, C.B., 2004. Estimating human long bone cross-sectional geometric properties: a comparison of noninvasive methods. J. Hum. Evol. 47, 221-235.

Parks, R.W., 1967. Efficient estimation of a system of regression equations when disturbances are both serially and contemporaneously correlated. J. Am. Stat. Assoc. 62, 500-509.

Pearson, O.M., Lieberman, D.E., 2000. The ageing of Wolff's "Law": ontogeny and responses to mechanical loading in cortical bone. Yearbook Phys. Anthropol. 47, 63-99.

Pfau, R.O., Sciulli, P.W., 1994. A method for establishing the age of subadults. J. Forensic Sci. 39, 165-176.

Robbins, G., 2007. Population Dynamics, Growth and Development in Chalcolithic Sites of the Deccan Plateau, India [Dissertation]. University of Oregon, Eugene. 350 pp.

Robbins, G., Sciulli, P.W., Blatt, S., 2010. Estimating body mass in subadult human skeletons. Am. J. Phys. Anthropol. 143, 146-150.

Robbins Schug, G., 2011. Bioarchaeology and Climate Change: a View from South Asian Prehistory. University Press of Florida, Gainesville.

Ruff, C.B., 2000. Body size, body shape, and long bone strength in modern humans. J. Hum. Evol. 38, 269-290.

Ruff, C.B., 2003a. Growth in bone strength, body size, and muscle size in a juvenile longitudinal sample. Bone 33, 317-329.

Ruff, C.B., 2003b. Ontogenetic adaptation to bipedalism: age changes in femoral to humeral length and strength proportions in humans, with a comparison to baboons. J. Hum. Evol. 45, 317-349.

Ruff, C.B., 2005. Growth tracking of femoral and humeral strength from infancy through late adolescence. Acta Paediatr. 94, 1030-1037.

Ruff, C.B., 2007. Body size prediction from juvenile skeletal remains. Am. J. Phys. Anthropol. 133, 698-716.

Ruff, C.B., Scott, W.W., Liu, A.Y.C., 1991. Articular and diaphyseal remodeling of the proximal femur with changes in body mass in adults. Am. J. Phys. Anthropol. 86, 397-413.

Ruff, C.B., Trinkaus, E., Walker, A., Larsen, C., 1993. Postcranial robusticity in Homo I: temporal trends and biomechanical interpretation. Am. J. Phys.Anthropol. 91, 21-53.

Ruff, C.B., Garofalo, E., Holmes, M.A., 2013. Interpreting human growth in the past from a functional and physiological perspective. Am. J. Phys. Anthropol. 150, 29-37.

SAS 9.2, 2010. SAS Institute Inc., Cary.

Sciulli, P.W., 1994. Standardization of long bone growth in children. Int. J. Osteoarchaeol. 4, 257-259.

Sciulli, P.W., Blatt, S., 2008. Evaluation of stature and body mass prediction. Am. J. Phys. Anthropol. 136, 387-393.

Smith, R.J., 1993. Logarithmic transformation bias in allometry. Am. J. Phys. Anthropol. 90, 215-228.

Sumner, D.R., Andriacchi, T.P., 1996. Adaptation to differential loading: comparison of growth-related changes in cross-sectional properties of the human femur and humerus. Bone 19, 121-126.

Telkka, A., Palkama, A., Virtama, P., 1962. Prediction of stature from radiographs of long bones in infants and children. J. Forensic Sci. 7, 474-479.

Trinkaus, E., Churchill, S.E., Ruff, C.B., 1994. Postcranial robusticity in Homo II: humeral bilateral asymmetry and bone plasticity. Am. J. Phys. Anthropol. 93, 1-34.

van der Meulen, M.C.H., Ashford, M.W., Kiratli, B.J., Bachrach, L.K., Carter, D.R., 1996. Determinants of femoral geometry and structure during adolescent growth. J. Orthop. Res. 14, 22-29.

van der Meulen, M.C.H., Beaupré, G.S., Carter, D.R., 1993. Mechanobiologic influences in long bone cross-sectional growth. Bone 14, 635-642.

Wallace, I.J., Tommasini, S.M., Judex, S., Garland Jr., T., Demes, B., 2012. Genetic variations and physical activity as determinants of limb bone morphology: an experimental approach using a mouse model. Am. J. Phys. Anthropol. 148, 24-35.