



Evaluating morphometric body mass prediction equations with a juvenile human test sample: accuracy and applicability to small-bodied hominins



Christopher S. Walker^{a, b, c, *, 1}, Gabriel S. Yapuncich^{b, 1}, Shilpa Sridhar^b, Noël Cameron^{d, c}, Steven E. Churchill^{b, c}

^a Department of Molecular Biomedical Sciences, College of Veterinary Medicine, North Carolina State University, 1060 William Moore Drive, Raleigh, NC, 27607, USA

^b Department of Evolutionary Anthropology, Box 90383, Duke University, Durham, NC, 27708, USA

^c Evolutionary Studies Institute, University of the Witwatersrand, Private Bag 3, Wits 2050, Johannesburg, South Africa

^d School of Sport, Exercise and Health Sciences, Loughborough University, Loughborough, LE11 3TU, United Kingdom

ARTICLE INFO

Article history:

Received 30 June 2016

Accepted 14 March 2017

Available online 3 May 2017

Keywords:

Children

Australopithecus

Homo

Body size

Stature

Bi-iliac breadth

ABSTRACT

Body mass is an ecologically and biomechanically important variable in the study of hominin biology. Regression equations derived from recent human samples allow for the reasonable prediction of body mass of later, more human-like, and generally larger hominins from hip joint dimensions, but potential differences in hip biomechanics across hominin taxa render their use questionable with some earlier taxa (i.e., *Australopithecus* spp.). Morphometric prediction equations using stature and bi-iliac breadth avoid this problem, but their applicability to early hominins, some of which differ in both size and proportions from modern adult humans, has not been demonstrated. Here we use mean stature, bi-iliac breadth, and body mass from a global sample of human juveniles ranging in age from 6 to 12 years ($n = 530$ age- and sex-specific group annual means from 33 countries/regions) to evaluate the accuracy of several published morphometric prediction equations when applied to small humans. Though the body proportions of modern human juveniles likely differ from those of small-bodied early hominins, human juveniles (like fossil hominins) often differ in size and proportions from adult human reference samples and, accordingly, serve as a useful model for assessing the robustness of morphometric prediction equations. Morphometric equations based on adults systematically underpredict body mass in the youngest age groups and moderately overpredict body mass in the older groups, which fall in the body size range of adult *Australopithecus* (~26–46 kg). Differences in body proportions, notably the ratio of lower limb length to stature, influence predictive accuracy. Ontogenetic changes in these body proportions likely influence the shift in prediction error (from under- to overprediction). However, because morphometric equations are reasonably accurate when applied to this juvenile test sample, we argue these equations may be used to predict body mass in small-bodied hominins, despite the potential for some error induced by differing body proportions and/or extrapolation beyond the original reference sample range.

© 2017 Elsevier Ltd. All rights reserved.

1. Introduction

As a frequent proxy for body size, body mass has broad implications for the study of allometry and can inform reconstructions of

a variety of associated ecological and physiological characteristics, including diet (Aiello and Wheeler, 1995), locomotion (Cant, 1992; Rubenson et al., 2007), predation risk (Isbell, 1994), life history (Robson and Wood, 2008), and energetic demands (Aiello and Wells, 2002; Churchill, 2006; Snodgrass and Leonard, 2009). Because body mass covaries with many ecological and biomechanical attributes, it has immense importance in the study of early hominin biology. However, body mass cannot be observed directly in fossil taxa, thus researchers frequently resort to predicting its

* Corresponding author.

E-mail address: christopher_walker@ncsu.edu (C.S. Walker).

¹ These authors contributed equally to this work.

value from morphological variables that are preserved in the paleontological or archeological records. As a result, the reliability of many assessments of hominin paleoecology are partially dependent on the accurate prediction of body mass.

For fossil hominins, two predominant approaches to body mass prediction have been employed: a “mechanical” method that uses dimensions of load-bearing skeletal elements such as femoral head diameter or knee breadth (Ruff et al., 1991, 2012; McHenry, 1992; Ruff, 1994, 2000a; Grine et al., 1995; Squyres and Ruff, 2015; Elliott et al., 2016a,b), and a “morphometric” method that models the body as a cylinder, under the assumption that mass is proportional to the volume of the cylinder (Ruff, 1991, 1994, 2000b; Ruff et al., 1997, 2005). With the morphometric method, the diameter of the cylinder is represented by bi-iliac breadth, while the height of the cylinder is set equal to stature. In both approaches, predictor variables are regressed against known body masses of a reference sample using ordinary least squares (for bivariate regressions) or multiple regression (for multivariate regressions); this procedure is also known as “inverse calibration” (Konigsberg et al., 1998).

Mechanical prediction equations derived from recent human samples allow for the reasonable prediction of body mass in more recent, and generally larger, hominins (including many members of the genus *Homo*) from hip joint dimensions (e.g., Ruff et al., 1997; Ruff, 2010). Potential differences in hip abductor biomechanics across hominin taxa, however, render their use questionable with earlier, generally smaller, hominins such as *Australopithecus* spp. (Jungers, 1988a; McHenry, 1992; Ruff, 1995; Ruff et al., 1999; Auerbach and Ruff, 2004). Morphometric prediction equations do not require the assumption of similar lower limb biomechanics (though they do require the assumption that body proportions are broadly similar). Though multiple variables may be included in morphometric prediction equations (e.g., Schaffer, 2016), stature/bi-iliac breadth equations have been most widely applied to archeological and paleontological samples (e.g., Ruff and Walker, 1993; Ruff et al., 1997, 2006; Arsuaga et al., 1999; Trinkaus et al., 1999; Holt, 2003; Rosenberg et al., 2006; Vercellotti et al., 2008; Siegmund and Papageorgopoulou, 2011; Pomeroy and Stock, 2012) because the incorporated variables can often be gathered from skeletonized material. Accordingly, while not necessarily the most accurate possible morphometric equations (see Schaffer, 2016), stature/bi-iliac breadth equations (henceforth used synonymously with “morphometric equations” or “morphometric method”) are the most widely applicable and have been recommended over mechanical methods when the variables can be reconstructed with confidence (Auerbach and Ruff, 2004).

Even when measurements of bi-iliac breadth and stature are available, some methodological difficulties remain for researchers who wish to apply the morphometric method to fossil hominins. First, osteological measurements of some fossil taxa fall outside the range of extant human samples, so that predicted body masses for these taxa require extrapolation. Extrapolation greatly increases uncertainty around predicted values (Aiello, 1992; Hens et al., 1998; Konigsberg et al., 1998; Ruff, 2007) and may violate a fundamental assumption of the prediction process – that predicted values belong to the same population as the reference population (Smith, 2009). Second, many fossil hominins have different proportions than those observed in recent human samples (Johanson et al., 1982; McHenry and Berger, 1998; Asfaw et al., 1999; Richmond et al., 2002; Morwood et al., 2005; Lovejoy et al., 2009; Berger et al., 2010; Churchill et al., 2013; Schmid et al., 2013; Berger et al., 2015), which may reduce accuracy when prediction equations are applied to fossil taxa. Ruff et al. (2005) examined the effect of certain proportional indices on the predictive accuracy of the morphometric method, including relative sitting height (= sitting height/stature \times 100) and biacromial/bi-iliac breadth. They found

that “only biacromial/bi-iliac breadth has a significant effect on prediction bias ($p < 0.01$)” (Ruff et al., 2005: 386), and that this effect was restricted to males. Additionally, Ruff (2000b) evaluated the reliability of the morphometric method by predicting the body masses of Olympic athletes, a sample that also deviates from the body proportions of non-athlete populations in a variety of ways. Ruff (2000b) demonstrated that the morphometric method produced fairly accurate body mass predictions for the athletic sample, particularly for individuals in events that emphasized a general combination of agility, endurance, strength, and speed (e.g., decathletes). Continued evaluation of morphometric methods with samples that exhibit a range of proportions may reveal additional correlations between prediction error and body proportions.

Several previous studies have addressed issues of body mass prediction in small-bodied hominins by evaluating mechanical and morphometric equations in small-bodied archeological samples (Kurki et al., 2010) or by developing and testing prediction equations for application to juveniles (Ruff, 2007; Sciulli and Blatt, 2008; Robbins et al., 2010; Robbins Schug et al., 2013). However, most juvenile prediction equations follow the mechanical method, utilizing femoral metaphyseal breadth (Ruff, 2007), femoral head breadth (Ruff, 2007), and femoral midshaft second polar moments of area (Robbins et al., 2010; Robbins Schug et al., 2013). Ruff (2007) published juvenile morphometric equations that used bi-iliac breadth and long bone lengths (rather than stature); the accuracy of these equations has yet to be evaluated with a novel test sample.

While published juvenile mechanical equations have proven to be reasonably accurate when predicting body masses of small-bodied hominins, no previous study has tested the accuracy of morphometric prediction equations with a small-bodied human sample. This line of investigation is particularly important since predictions from morphometric equations have been used as a comparative baseline to assess mechanical equations (Auerbach and Ruff, 2004; Kurki et al., 2010). Here we use mean stature, bi-iliac breadth, and body mass from a global sample of human juveniles to evaluate the accuracy of morphometric body mass prediction equations when applied to small-bodied fossil hominins and examine the effect of juvenile body proportions (specifically lower limb length/stature and biacromial/bi-iliac breadth) on predictive accuracy. Treating human juveniles as a small-bodied fossil hominin proxy, we focus on two potential sources of error with possible repercussions for the application of morphometric prediction equations to the hominin fossil record: 1) extrapolation error due to differing body sizes between test and reference samples and 2) differing body proportions between test and reference samples. With respect to body size, human juveniles largely overlap with the estimated statures and masses of *Australopithecus* (and some specimens attributed to *Homo*), allowing for a direct assessment of extrapolation error over a range of sizes (e.g., Eveleth and Tanner, 1976; Grabowski et al., 2015). Body proportions in human juveniles and small-bodied fossil hominins, however, are particularly variable (Eveleth and Tanner, 1976; Johanson et al., 1982; Berge, 1998; McHenry and Berger, 1998; Asfaw et al., 1999; Bogin, 1999; Richmond et al., 2002; Ruff, 2002; Bogin et al., 2002; Morwood et al., 2005; Green et al., 2007; Lovejoy et al., 2009; Berger et al., 2010; Kibii et al., 2011; Schmid et al., 2013; Churchill et al., 2013; DeSilva et al., 2013; Berger et al., 2015; Marchi et al., 2017; Feuerriegel et al., 2017). Body proportions change during growth and development, differ across human juvenile populations, and even diverge within juvenile populations according to environmental factors (Eveleth and Tanner, 1976; Berge, 1998; Bogin, 1999; Bogin et al., 2002; Ruff, 2002; Temple et al., 2011). Small-bodied fossil hominin body proportions are similarly diverse, particularly with respect to measures central to morphometric prediction equations – limb length and hip breadth (Jungers, 1982;

Aiello, 1992; Richmond et al., 2002; Berger et al., 2010; Kibii et al., 2011; Churchill et al., 2013; Berger et al., 2015; Walker et al., 2015; Feuerriegel et al., 2017; Marchi et al., 2017). Accordingly, generalized comparisons of juvenile human and small-bodied fossil hominin body proportions are not practical. Importantly, however, each of these groups possesses body proportions that often differ from those of adult humans. As such, with respect to body proportions, human juveniles serve not as a direct proxy for all small-bodied fossil hominins, but rather a test sample to examine the general effect of varying body proportions in a highly variable group.

2. Materials and methods

To evaluate the accuracy of morphometric prediction equations, population mean values of stature (ST: cm), bi-iliac breadth (BIB: cm), and body mass (BM: kg) were collected from the literature (Eveleth and Tanner, 1976). This global sample includes 530 age- and sex-specific group annual means of human juveniles, ranging in age from 6 to 12 years, aggregated from populations (following descriptions by Eveleth and Tanner, 1976) of European, African, Asian, New Guinean, and Indo-Mediterranean descent, residing in 33 distinct countries or regions (Table 1). Summary statistics for the sample are provided in Table 2. Mean body

masses of the older ages partially overlap with the smallest mean body mass values in the adult reference samples (Ruff, 1994), as well as the inferred body mass range of australopiths. Body masses at younger ages are smaller than the reference sample and thus permit evaluation of the effects of extrapolation outside the reference sample range.

Prediction equations from three studies, the male and combined sex equations from Ruff (1994), the female equation from Ruff et al. (1997), and the male and female equations from Ruff et al. (2005) (Table 3), were appraised with this juvenile test sample. These equations are derived from similar reference samples (Ruff et al. [1997] updated the female equation of Ruff [1994] by correcting the value of Aleut female mean body mass and Ruff et al. [2005] incorporated mean body masses for male and female Finns), so differences in predicted body masses are expected to be modest. However, because Finns have both larger stature and broader bi-iliac breadth compared to other populations in the reference sample, their inclusion may have a strong impact on the regression equation (and, thus, have a strong effect on prediction at the opposite end of the body mass spectrum). For brevity, we refer to the Ruff (1994) and Ruff et al. (1997) equations as the “initial” set and the Ruff et al. (2005) equations as the “revised” set. Since the juvenile morphometric equations published by Ruff (2007) use long bone lengths rather than stature, our test sample could not be used to test the predictive accuracy of these equations.

2.1. Accuracy in extrapolation

Our analytical approach first evaluated which set of ST-BIB prediction equations (initial or revised) was most accurate. Within that set of equations, we then assessed which particular equations (female, male, or combined sex) were most accurate for both females and males. Given the age range of our test sample, sexual dimorphism is likely minimal (Feldesman, 1992; Humphrey, 1998; Hauspie, 2002), so an equation developed for one sex in adults might produce accurate predictions for both sexes in juveniles. Additionally, since the reference samples of the female equations overlap more substantially with the juvenile test sample, negative effects of extrapolation may be reduced relative to the male equations.

Accuracy was determined by calculating prediction error (PE: kg) as the difference between observed (BM_{obs}) and predicted body masses (BM_{pred}). Negative values indicate an overprediction of body mass, while positive values indicate underprediction. Following Smith (1980, 1984), relative percentage prediction error (%PE) was calculated as $(BM_{obs} - BM_{pred})/BM_{pred} * 100$, while absolute percentage prediction error was calculated as $|BM_{obs} - BM_{pred}|/BM_{pred} * 100$. Mean percentage prediction error (%MPE) was calculated by averaging relative or absolute %PEs across all populations by age and sex. Since both the reference and test samples comprise population mean values, there is no change in the units of analysis and prediction errors should not be inflated (Yapuncich, 2018). For both sexes and all ages, observed and predicted body masses were compared with non-parametric Mann–Whitney U-tests since observed body masses were often not normally distributed and predicted body masses had greater variance than observed body masses.

While it is useful to determine which of the morphometric equations is most accurate relative to one another, these comparisons do not indicate whether or not the most accurate morphometric equation generates error at an acceptably low level. In this study, we consider both a mean %PE less than $\pm 20\%$ and the majority of test cases with %PE less than $\pm 20\%$ to indicate an acceptable level of prediction error. These performance criteria have frequently been used in other studies of prediction error (e.g.,

Table 1
Sample origins (from Eveleth and Tanner, 1976).

Ancestry	Residence
Africa	Cuba
	Nigeria
	Rwanda
	South Africa
	Surinam
	Tanzania
	USA–Philadelphia
	USA–Washington, DC
Asia	West Indies
	Brazil
	Guatemala
	Hong Kong
	Mexico
	Singapore
Europe	USA–Wainwright
	Argentina
	Belgium
	Brazil
	Canada
	Cuba
	France
	Guatemala
	Hungary
	Italy
	Poland
	Romania
	South Africa
	United Kingdom
	USA–Berkeley
	USA–Cleveland
USA–Denver	
USA–Iowa	
USA–Philadelphia	
Indo-Mediterranean	Yugoslavia
	Egypt
	India
New Guinea	South Africa
	Manus
	Mt. Hagen

Table 2
Descriptive statistics (mean, standard deviation [SD] and range) for samples.

Sex	Age (Years)	N	Stature (cm)			Bi-iliac breadth (cm)			Body mass (kg)		
			Mean	SD	Range	Mean	SD	Range	Mean	SD	Range
Female	6	24	112.5	3.2	106.7–117.6	18.0	0.8	16.0–19.7	19.4	1.7	16.2–21.9
	7	39	117.8	3.9	106.6–123.8	18.5	1.1	15.2–20.0	21.4	2.2	17.5–24.8
	8	41	122.2	4.5	111.0–130.1	19.0	1.2	15.4–21.0	23.3	2.6	19.0–28.1
	9	42	127.4	4.9	115.5–135.7	19.7	1.3	16.0–21.9	25.9	3.2	20.5–31.6
	10	42	132.9	5.6	118.3–141.5	20.7	1.4	17.0–23.0	29.1	3.8	23.7–35.9
	11	41	138.5	6.1	123.5–148.6	21.7	1.6	17.4–24.4	32.6	4.6	26.0–40.7
Male	12	43	144.5	6.1	130.8–155.2	22.9	1.7	18.0–25.5	36.8	5.2	27.9–46.1
	6	24	113.0	3.3	107.2–117.7	18.1	0.8	16.3–19.7	20.0	1.5	16.7–22.4
	7	38	118.3	4.2	108.0–124.8	18.7	1.0	16.8–20.4	22.1	2.2	18.0–25.9
	8	39	123.2	4.6	110.8–130.0	19.1	1.1	17.4–21.2	24.1	2.5	20.0–28.6
	9	39	128.4	4.9	116.1–136.0	19.9	1.2	17.9–21.8	26.7	3.0	22.0–32.2
	10	38	133.0	5.4	120.7–141.3	20.7	1.2	18.7–22.8	29.3	3.5	23.9–35.1
	11	39	137.7	5.8	124.3–146.5	21.3	1.3	19.0–23.9	32.2	4.3	25.3–39.4
	12	41	142.4	6.0	129.4–152.2	22.1	1.3	19.3–24.3	35.5	4.7	27.2–43.3

Table 3
Body mass prediction equations^a.

Group	ST coefficient	BIB coefficient	Intercept	Source
Male	0.373	3.033	–82.5	Ruff, 1994
Female	0.522	1.809	–75.5	Ruff et al., 1997
Combined	0.414	2.782	–83.0	Ruff, 1994
Male	0.422	3.126	–92.9	Ruff et al., 2005
Female	0.504	1.804	–72.6	Ruff et al., 2005

^aST = stature; BIB = bi-iliac breadth.

Dagosto and Terranova, 1992; Aiello and Wood, 1994; Delson et al., 2000; Yapuncich et al., 2015; Elliott et al., 2016a,b).

2.2. Effects of differences in proportions

To evaluate the effects of differing proportions on predictive accuracy, additional mean somatic variables were collected from the literature (Eveleth and Tanner, 1976). The ratio of mean lower limb length (LL: cm) to mean ST was taken as a measure of lower limb to trunk proportions (higher LL/ST values reflect relatively longer lower limbs). The ratio of mean biacromial breadth (BAB: cm) to mean BIB was taken as a measure of trunk shape (higher BAB/BIB values reflect relatively broader shoulders). Finally, mean triceps skinfold thickness (TST: mm) was taken as a measure of subcutaneous fatness. The distributions of LL/ST and BAB/BIB in juveniles were compared to distributions of adult population means collected from the literature (LL/ST from Eveleth and Tanner, 1976 and Ruff et al., 2005; BAB/BIB from Hiernaux, 1985 and Ruff et al., 2005).

We used generalized linear mixed models (GLMM) with repeated measures in JMP Pro 12.0 (SAS Institute, 2015) to evaluate the relationship between prediction error and body mass, as well as the additional somatic variables. Because our sample comprises pseudo-longitudinal observations (that is, age series of juveniles from the same population), there may be a certain amount of autocorrelation within each population (Cnaan et al., 1997; Dobson and Barnett, 2008). To account for this autocorrelation, we employed a first-order, autoregressive covariance structure. The first GLMM examined the relationship between prediction error and body mass, with body mass modeled as a fixed effect. The second GLMM examined the relationship between prediction error and the additional somatic variables, LL/ST, BAB/BIB, TST. All possible interactions were modeled as fixed effects, while body mass was modeled as a random effect. Males and females were analyzed separately in both GLMMs.

3. Results

3.1. Accuracy in extrapolation

The relative and absolute %PE for the initial and revised ST-BIB morphometric equations are presented in Table 4. Figure 1 shows boxplots of relative %PE for both sets of equations. While prediction error varied by sex and age, there were similar trends in both sets of equations. In terms of mean %PE and range of %PE, female equations were better predictors of female juvenile body mass than male equations were of predicting male juvenile body mass. Compared to male equations, female equations generated lower mean %PEs and had reduced variance in %PEs. At younger ages (6–8 years), all equations tended to underpredict body mass, while at older ages (10–12 years), all equations tended to overpredict body mass. In the initial equations, predictions for females were the most accurate among 9 year olds (lowest relative and absolute %PE), and were reasonably precise for 8 and 10 year olds (absolute %PE <10%). Among males, predictions for 10 and 11 year olds were most accurate (gauged by relative and absolute %PE, respectively). Predictions for 10 and 12 year old males also had an absolute %PE less than 10%.

The revised equations generated less accurate predictions overall (Table 4). These inaccuracies were usually pronounced underpredictions, particularly in males at age 6 years (Fig. 1). However, the revised equations did have slightly lower %PE at specific ages. Predictions for females were most accurate at 8 years old by relative %PE and 9 years old by absolute %PE. Among males, the revised equations were most accurate for 12 year olds.

Since the initial set of equations had greater accuracy than the revised equations, we predicted body masses for female and male juveniles with all three equations of the initial set (females, males, and combined). Table 5 shows the relative %PE and the percentage of cases under 20% PE for each equation by age and sex; Figure 2 shows boxplots of relative %PE for all three equations. Again, body masses tended to be underpredicted at younger ages (6–8 years) and slightly overpredicted at older ages (10–12 years). However, only at younger ages (6–8 years) did the majority of cases have greater than 20% prediction error; for the female equation, only 6 year olds failed to meet this performance criterion (Table 4). In fact, the initial female equation generates prediction errors that are less than 10% in the majority of test cases for all age groups, except 6 year olds, in both females and males. The female equation had the least variance in relative %PE, particularly at younger ages. The most accurate predictions (lowest relative %PE) were achieved in 9 year olds for the female equation, and in 10 year olds for the male and combined equations.

Table 4
Relative and absolute percentage prediction error for stature and bi-iliac breadth equations.

	Sex	Age (Years)	Initial Equations ^a			Revised Equations ^b		
			Mean	SD	Range	Mean	SD	Range
Relative %MPE	Female	6	25.06	15.19	8.10 to 77.54	18.76	12.83	4.55 to 62.07
		7	11.68	13.85	−5.23 to 50.09	7.57	12.16	−7.53 to 41.06
		8	4.41	11.47	−14.30 to 40.37	1.47	10.32	−16.85 to 32.90
		9	−2.14	9.07	−24.50 to 26.78	−4.12	8.35	−26.07 to 21.65
		10	−6.53	8.03	−28.17 to 23.12	−7.83	7.55	−29.18 to 19.06
		11	−9.32	7.11	−32.96 to 11.37	−10.14	6.85	−33.52 to 8.79
	Male	6	41.64	27.02	9.83 to 133.17	90.34	66.76	28.06 to 355.22
		7	25.88	24.91	−2.51 to 107.93	55.13	47.88	8.38 to 241.48
		8	15.52	18.36	−9.23 to 63.03	33.96	28.86	2.36 to 122.03
		9	5.85	14.27	−18.89 to 37.99	17.47	20.41	−11.58 to 66.70
		10	−0.08	12.02	−23.63 to 37.22	7.83	15.85	−18.37 to 56.23
		11	−3.10	11.14	−27.16 to 36.49	2.59	13.73	−23.28 to 51.31
Absolute %MPE	Female	6	25.06	15.19	8.10 to 77.54	18.76	12.83	4.55 to 62.07
		7	13.06	13.49	0.41 to 50.09	10.04	10.91	0.32 to 41.06
		8	8.60	8.70	0.44 to 40.37	7.59	7.05	0.62 to 32.90
		9	7.25	5.76	0.50 to 26.78	7.45	5.50	0.90 to 26.07
		10	8.45	5.92	1.12 to 28.17	9.15	5.83	0.48 to 29.18
		11	10.17	5.79	0.91 to 32.96	10.72	5.87	0.41 to 33.52
	Male	6	10.57	6.51	0.28 to 34.49	11.00	6.45	0.33 to 34.81
		7	41.64	27.02	9.83 to 133.17	90.34	66.76	28.06 to 355.22
		8	26.07	24.70	0.17 to 107.93	55.13	47.88	8.38 to 241.48
		9	17.33	16.61	0.03 to 63.03	33.96	28.86	2.36 to 122.03
		10	11.29	10.39	0.15 to 37.99	18.98	18.98	0.32 to 66.71
		11	9.30	7.47	0.54 to 37.22	11.79	13.09	0.21 to 56.23
		12	9.23	6.81	0.14 to 36.49	10.05	9.57	0.06 to 51.31
		12	9.46	7.35	0.76 to 40.08	9.23	8.72	0.55 to 51.09

^a Male equation from Ruff (1994) and female equation from Ruff et al. (1997). SD = standard deviation.

^b Male and female equations from Ruff et al. (2005). SD = standard deviation.

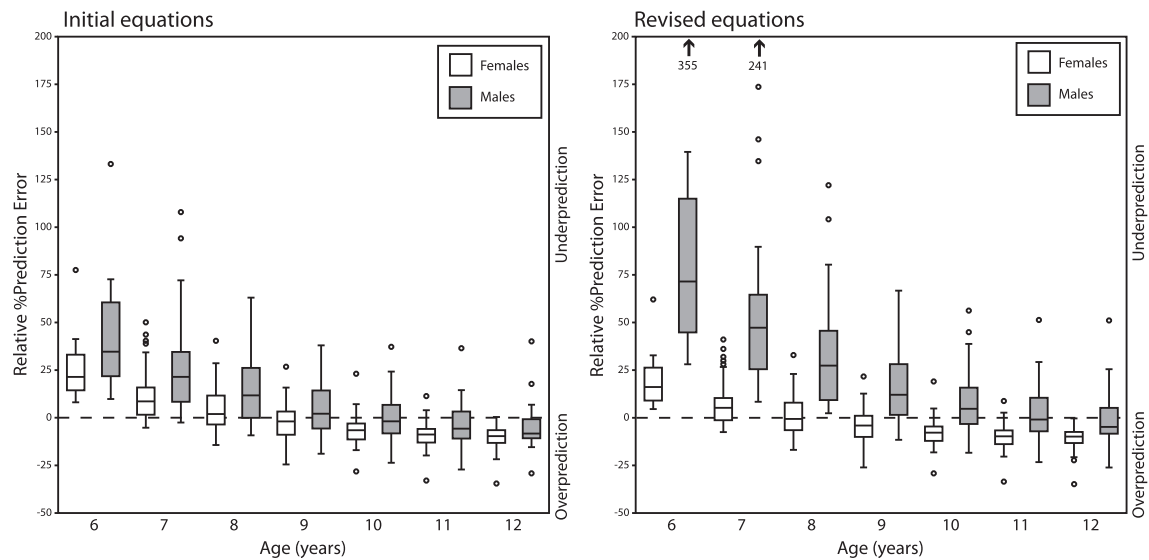


Figure 1. Relative %prediction error boxplots by age and sex for “initial” equations (male equation of Ruff, 1994; female equation of Ruff et al., 1997) and “revised” equations (both from Ruff et al., 2005). Female body masses were predicted with female equations, and male body masses were predicted with male equations. Boxes include 25–75% quartiles; whiskers extend to farthest points less than 1.5 times interquartile range; circles indicate outliers, with two extreme outliers indicated by arrows; dashed line demarcates over- and underprediction.

Overall, the initial female equation generated the most accurate predictions for both female and male juveniles. Averaged across all age groups, female masses were overpredicted by 1.04% (standard deviation [sd] 3.45%) and male masses were overpredicted by 0.63% (sd 3.32%). When body masses predicted by the initial female equation were compared to observed body masses (Table 6), there were no significant differences for either females or males at ages 8 and 9 years. Predicted body masses are most distinct from observed

mean body masses in the youngest or oldest examined ages in both sexes.

3.2. Effects of differences in proportions

Significant differences in both LL/ST and BAB/BIB were observed between multiple juvenile age groups (Supplementary Online Material [SOM] Table S1), as well as between certain juvenile

Table 5
Relative percentage prediction error for initial stature and bi-iliac breadth equations^a.

Sex	Age (Years)	Female Equation				Male Equation				Combined Equation			
		Mean	SD	Range	%PE <20%	Mean	SD	Range	%PE <20%	Mean	SD	Range	%PE <20%
Female	6	25.06	15.19	8.10 to 77.54	36	44.52	33.57	12.04 to 177.76	9	48.65	33.94	17.22 to 183.69	9
	7	12.47	14.05	-5.23 to 50.09	77	28.11	33.55	0.10 to 159.53	38	29.86	31.31	1.87 to 143.39	31
	8	4.41	11.47	-14.30 to 40.37	86	16.96	23.98	-14.02 to 111.03	61	18.08	22.42	-11.25 to 98.05	54
	9	-2.14	9.07	-24.50 to 26.78	97	6.97	17.62	-26.82 to 70.78	72	7.69	16.53	-24.61 to 62.55	72
	10	-6.53	8.03	-28.17 to 23.12	97	-0.72	13.75	-32.29 to 37.56	90	-0.07	13.05	-30.29 to 38.25	90
	11	-9.32	7.11	-32.96 to 11.37	100	-5.56	11.70	-37.14 to 26.95	93	-4.98	11.03	-35.53 to 23.04	93
	12	-10.55	6.54	-34.49 to 0.39	100	-8.53	10.09	-38.53 to 18.38	100	-7.96	9.47	-37.14 to 15.25	100
Male	6	25.01	14.30	5.02 to 62.56	27	41.64	27.02	9.83 to 133.17	9	45.81	27.46	12.77 to 138.07	9
	7	12.50	15.13	-4.94 to 64.22	76	25.88	24.91	-2.51 to 107.93	28	28.21	25.40	-0.42 to 115.4	20
	8	4.50	11.47	-8.71 to 42.09	92	15.52	18.36	-9.23 to 63.03	46	16.80	18.27	6.26 to 68.42	42
	9	-1.99	9.06	-17.15 to 21.37	92	5.85	14.27	-18.89 to 37.99	73	6.72	14.03	-16.64 to 38.98	73
	10	-5.62	8.20	-21.65 to 21.30	96	-0.08	12.02	-23.63 to 37.22	88	0.66	11.80	-21.78 to 37.68	88
	11	-7.58	8.12	-24.91 to 22.77	96	-3.10	11.14	-27.16 to 36.49	96	-2.54	10.89	-25.55 to 36.79	96
	12	-8.37	8.30	-26.27 to 29.21	96	-5.12	10.90	-29.17 to 40.08	96	-4.63	10.62	-27.61 to 40.42	96

^aSD = standard deviation; %PE = relative percentage prediction error.

ages and adults (Fig. 3). Younger juveniles had significantly lower LL/ST ratios than adults, indicating shorter lower limbs relative to stature. Older juveniles tended to have higher LL/ST ratios than adults (longer lower limbs relative to stature), but these differences were not significant. For BAB/BIB ratios, female adults had significantly lower ratios (greater hip breadth relative to shoulder breadth) than many female juvenile age groups (ages 7, 9–11 years), while the only significant difference observed in males was a higher ratio in adults (greater shoulder breadth relative to hip breadth) compared to 6 year olds.

Parameter estimates for both general linear mixed models are presented in Table 7. The first model reveals a significant negative relationship between body mass and prediction error for both females ($p < 0.0001$) and males ($p < 0.0001$). The second model shows there are few significant relationships between prediction error and the additional somatic variables. Among females, there is a significant positive relationship between prediction error and the interaction of LL/ST and BAB/BIB ($p = 0.0285$). Among males, there is a significant negative relationship between prediction error and LL/ST ($p = 0.0034$) and significant positive relationships between prediction error and BAB/BIB ($p = 0.0178$) and the interaction between LL/ST*BAB/BIB ($p = 0.0105$). In females, LL/ST has the strongest total effect (0.357); in males, BAB/BIB and LL/ST have similarly strong effects (0.295 and 0.251, respectively). Tricep skinfold thickness did not have a significant relationship with prediction error in either males or females. SOM Table S2 shows the main and total effects for each parameter, and marginal model plots for the second model are shown in SOM Figure S1.

4. Discussion

Based on our results, juvenile male and female body masses are most accurately predicted (as gauged by prediction error) using the female ST-BIB equation of Ruff et al. (1997). Mean %PEs for this equation are considerably lower than those observed for mechanical equations evaluated by Ruff with a juvenile test sample (Ruff, 2007; his Fig. 2). Since the majority of cases have %PE less than 10%, the initial female equation would even meet the more rigorous performance criterion proposed by Ruff et al. (2005), exceeding the more lenient standard utilized in this study and other studies evaluating prediction equations (e.g., Dagosto and Terranova, 1992; Aiello and Wood, 1994; Elliott et al., 2016a,b). Relative to the other prediction equations (Table 3), this equation has the highest ST

coefficient, the second lowest BIB coefficient, and the second highest intercept. Thus, in the Ruff et al. (1997) female equation, the impact of stature is maximized, the impact of bi-iliac breadth is minimized, and the intercept makes a relatively small adjustment. The much lower intercepts of the male equations likely drive the observed underprediction of juvenile body masses.

The female equation of Ruff et al. (1997) also differs from the other equations in its reference sample composition: it has the smallest grand mean and the most overlap with the juvenile test sample. Mbuti ($\bar{x} = 38.2$ kg), West Aka ($\bar{x} = 42.7$ kg), and Australian Aboriginal ($\bar{x} = 45.4$ kg) female samples used in Ruff (1994) and Ruff et al. (1997) are all smaller than the most massive juvenile sample (12 year old females from Berkeley, CA [$\bar{x} = 46.1$ kg]). Due to this overlap, predicting juvenile body masses requires less extrapolation with the initial female equation than the other equations tested. Still, even the most accurate morphometric equation consistently overpredicts body mass in the inferred size range of australopiths (Fig. 2, Table 6), as well as in two of the three aforementioned small-bodied adult female populations (West Aka and Australian Aboriginal female populations are overpredicted by 1.1 kg and 5.6 kg, respectively).

The choice of regression model may be one factor contributing to the consistent overprediction of body mass in older juveniles. All ST-BIB prediction equations examined in this study are multiple regression equations that employ ordinary least squares (OLS; Model I) to optimize regression parameters. There is substantial debate about the most appropriate regression method for prediction (Konigsberg et al., 1998; Warton et al., 2006; Smith, 2009), but OLS may be particularly unsuitable for extrapolation, as this regression model biases predicted values toward the mean of the reference sample, particularly in instances of extrapolation. Still, Grabowski et al. (2015) argue that the reduced variance in body masses predicted through OLS may be desirable if the objective of the study is a comparison of body mass variation among several taxa. In other contexts, many authors have suggested that reduced major axis (Model II) regression is more appropriate than OLS in extrapolation (Ricker, 1973; Jungers, 1988b; Aiello, 1992; Ruff, 1998). "Classical" calibration (regressing the predictor variable on body mass and then solving for body mass) has also been proposed as the best option to minimize the effects of extrapolation (Konigsberg et al., 1998; Hens et al., 2000). Although there is no real consensus on the most appropriate regression method in extrapolation, multivariate OLS regression (and, therefore, ST-BIB

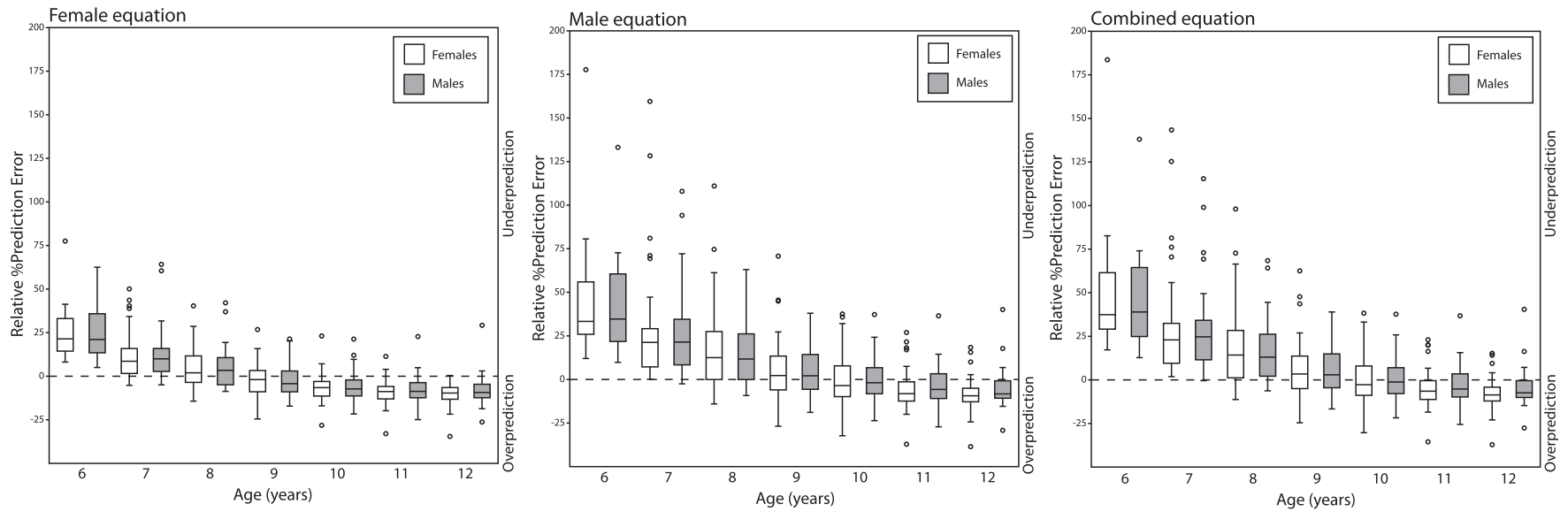


Figure 2. Relative %prediction error boxplots by age and sex for all “initial” equations (male and combined equations of Ruff, 1994; female equation of Ruff et al., 1997). Each sex was predicted with all three equations. Boxes include 25–75% quartiles; whiskers extend to farthest points less than 1.5 times interquartile range; circles indicate outliers; dashed line demarcates over- and underprediction.

Table 6
Body masses (BM) predicted with initial stature and bi-iliac breadth equations^a.

Sex	Age (Years)	Mean Observed BM	SD	Mean Predicted BM	SD	Range	Mann–Whitney U	p-value
Female	6	19.4	1.7	15.8	2.6	9.4–20.3	65.0	0.0000
	7	21.4	2.2	19.4	3.5	12.3–25.1	518.0	0.0156
	8	23.3	2.6	22.7	4.0	14.5–30.2	772.0	0.5281
	9	25.9	3.2	26.7	4.5	17.4–34.7	769.5	0.3163
	10	29.1	3.8	31.3	4.9	19.9–39.6	627.0	0.0228
	11	32.6	4.6	36.1	5.6	24.4–45.3	517.0	0.0027
Male	12	36.8	5.2	41.3	5.7	29.0–50.9	515.0	0.0004
	6	20.0	1.5	16.3	2.7	10.5–20.9	63.5	0.0000
	7	22.1	2.2	20.0	3.5	11.4–25.5	466.0	0.0079
	8	24.1	2.5	23.4	3.9	14.4–30.1	687.0	0.4655
	9	26.7	3.0	27.6	4.3	18.2–34.7	655.0	0.2940
	10	29.3	3.5	31.3	4.6	21.3–39.0	526.0	0.0422
11	32.2	4.3	35.0	5.0	24.1–42.9	503.0	0.0102	
12	35.5	4.7	38.9	5.2	28.0–47.4	526.0	0.0036	

^aSD = standard deviation.

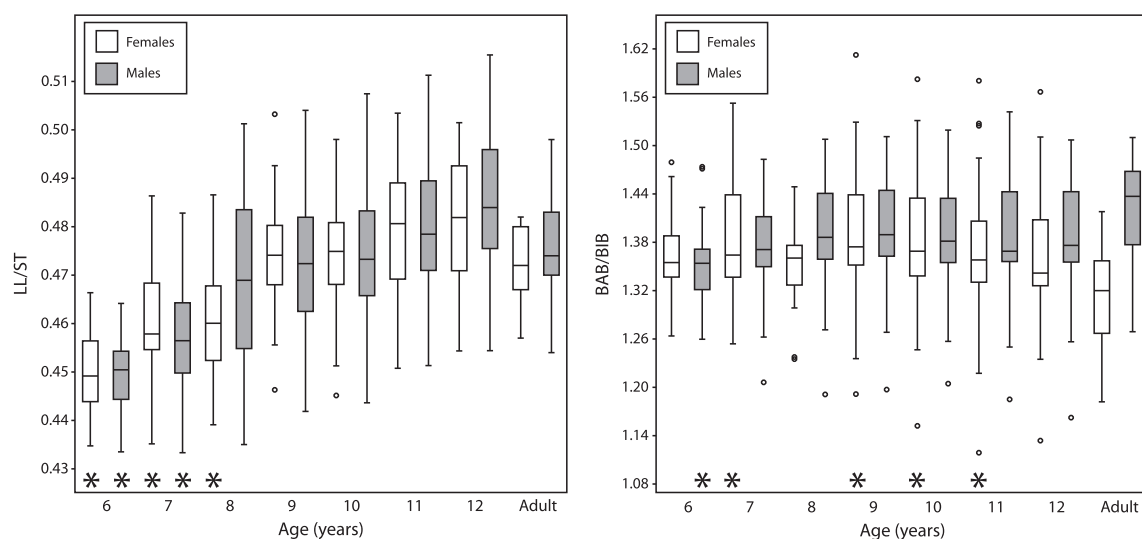


Figure 3. Boxplots of lower limb length/stature (LL/ST) and biacromial/bi-iliac breadth (BAB/BIB) by age and sex. Adult LL/ST ratios were derived from data published by [Eveleth and Tanner \(1976\)](#) and [Ruff et al. \(2005\)](#). Adult BAB/BIB ratios were derived from data published by [Hiernaux \(1985\)](#) and [Ruff et al. \(2005\)](#). Boxes include 25–75% quartiles; whiskers extend to farthest points less than 1.5 times interquartile range; circles indicate outliers. Asterisks indicate significant differences between juvenile and adult sex-specific means.

morphometric equations) will tend to overpredict values beyond the lower bound and underpredict values beyond the upper bound. It is possible that ST-BIB equations reformulated with other, non-OLS multivariate regression methods (see [Konigsberg et al., 1998](#); [Hens et al., 2000](#); [Uhl et al., 2013](#)) would not generate overpredictions beyond the lower bound of the reference sample, but this reformulation of morphometric equations is beyond the scope of the present study.

Investigating the predictive accuracy of morphometric equations is particularly important, as they have been used as a comparative baseline for evaluating mechanical prediction equations. [Auerbach and Ruff \(2004\)](#) demonstrate consistent overprediction of mass by mechanical (femoral head dimensions) relative to morphometric (ST-BIB) methods. [Jungers et al. \(2016\)](#) highlight major body mass overprediction by [Will and Stock \(2015\)](#) due to their reliance on femoral head-based mechanical methods. Using a small-bodied sample, [Kurki et al. \(2010\)](#) found that, with the exception of [McHenry's \(1992\)](#) femoral head equation, mechanical equations tend to generate greater predicted body masses than morphometric equations. For this reason, both [Auerbach and Ruff \(2004\)](#) and [Kurki et al. \(2010\)](#) recommend that [McHenry's \(1992\)](#) femoral head equation be used for small-bodied samples, particularly when stature and bi-iliac breadth cannot be

reconstructed with confidence (although between-species differences in hip abductor mechanics may still require the use of correction coefficients when applying femoral head mechanical prediction equations to australopiths; see [Ruff, 2015](#)). Since morphometric methods also overpredict body mass in small-bodied fossil hominins, our results suggest that the degree of overprediction by mechanical methods (including [McHenry's \[1992\]](#) equation) may be greater than previously recognized. It should be noted that the congruence of predictions from alternative methods does not sufficiently address the accuracy of the prediction equations (i.e., whenever possible, predictions should be compared to observed values rather than other predictions).

It is possible to visualize the consistent overprediction for most of our sample by plotting prediction error against body mass and using a locally weighted smoothing regression ([Fig. 4](#)). At the smallest sizes, body mass is underpredicted. Error decreases until ~25 kg, when body mass is well predicted. From 25 kg to 30 kg, prediction error again increases (now as overpredicted masses) and then stabilizes at about 4 kg over the remaining range of the test sample. In the estimated range of australopith body mass, body masses are consistently overpredicted.

The strong effects of LL/ST and BAB/BIB in the linear mixed models demonstrate that body proportions are important

Table 7

Parameter estimates, standard errors (SE), denominator degrees of freedom (Den df), t ratios, and probabilities ($\text{Pr} > |t|$) for GLMM models^a.

Sex	Model term	Estimate	SE	Den df	t ratio	$\text{Pr} > t $
Model 1						
Female	Intercept	8.77	0.71	184.7	12.35	<0.0001
	BM	-0.34	0.02	262.1	-15.38	<0.0001
Male	Intercept	9.20	0.81	192.9	11.29	<0.0001
	BM	-0.34	0.03	253.5	-13.15	<0.0001
Model 2						
Female	Intercept	5.63	10.74	57.3	0.52	0.6023
	TST	0.10	0.18	58.8	0.57	0.5697
	LL/ST	-35.65	19.38	56.9	-1.84	0.0711
	TST*LL/ST	0.99	10.02	56.5	0.10	0.9217
	BAB/BIB	6.66	4.51	58.0	1.48	0.1448
	TST*BAB/BIB	3.09	2.39	56.5	1.29	0.2007
	LL/ST*BAB/BIB	946.93	422.31	62.1	2.24	0.0285
	TST*LL/ST*BAB/BIB	215.82	165.24	58.0	1.31	0.1967
Male	Intercept	4.97	10.21	55.7	0.49	0.6285
	TST	0.20	0.15	57.3	1.29	0.2037
	LL/ST	-58.18	19.12	65.3	-3.04	0.0034
	TST*LL/ST	-18.47	9.88	59.2	-1.87	0.066
	BAB/BIB	14.68	6.01	54.3	2.44	0.0178
	TST*BAB/BIB	2.54	1.87	57.8	1.36	0.1796
	LL/ST*BAB/BIB	691.67	262.01	61.5	2.64	0.0105
	TST*LL/ST*BAB/BIB	227.77	128.31	58.0	1.78	0.0811

^aBM = body mass; TST = triceps skinfold thickness; LL = lower limb length; ST = stature; BAB = biacromial breadth; BIB = bi-iliac breadth.

considerations for morphometric prediction methods, as they represent another mode by which a test sample may differ from the reference population. The importance of these variables corresponds well with recent model-fitting analyses performed by Schaffer (2016), who found that regression models that included some combination of BAB, BIB, and sitting height had more explanatory power than the ST-BIB equation in six modern populations. Unfortunately, while it may be possible to collect these

additional measurements in some forensic contexts, it is not possible to measure BAB or sitting height from fully skeletonized remains (as discussed further below), and they are, therefore, not practical for archeological or paleontological applications.

In our test sample, significant differences from adults in LL/ST ratios likely drive high prediction errors in the youngest age groups (6–8 years) of our sample (Fig. 3, SOM Table S1). All equations examined in this study have the greatest prediction error in these groups, which can also be seen in Figure 4. These errors could be caused solely by extrapolation beyond the observed range of adult body masses, as the body masses at these ages are much smaller than any of the adult reference samples. However, body masses in the youngest juveniles are strongly underpredicted, which is opposite of what we would expect when extrapolating beyond the observed range of an OLS regression equation. Rather than being driven by extrapolation, it seems more likely that prediction errors for the youngest ages are caused by significant differences in body proportions. In addition to significantly lower LL/ST ratios than adults, 6–8 year olds also have high sitting height/stature ratios (Martorell et al., 1988). When applied to a sample with proportionally shorter lower limbs and longer trunks (such as 6–8 year olds), prediction equations derived from adult reference samples would tend to underpredict the mass of juveniles.

At 9 years of age, LL/ST ratios resemble those of adults (Fig. 3, SOM Table S1). Beyond 9 years old, pubertal acceleration of growth, with attendant changes in body proportions, may also increase prediction error (manifested as overprediction). Although there is considerable variation within and between human populations in the timing, intensity, and duration of developmental events, growth data on American children (who are presumably well-fed and in generally good health) provide some idea of the early age range of these events. American girls may enter puberty (and the adolescent growth spurt) as early as 8 years of age (Lee, 1980), and reach peak growth velocity (in stature), on average, at 11 years of

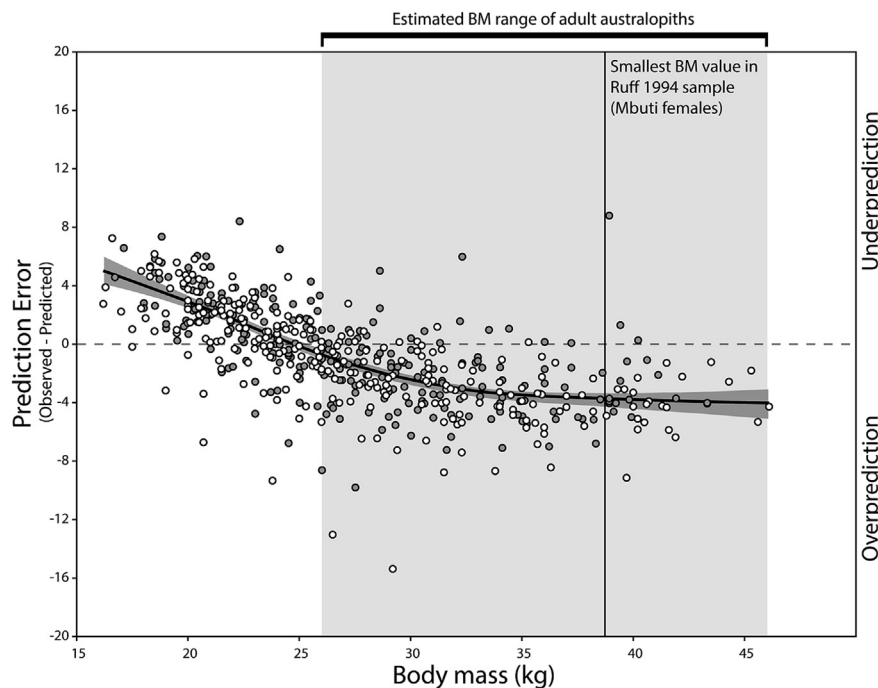


Figure 4. Local regression (LOESS) of prediction error (smoothing factor = 0.5) of the “initial” female equation (Ruff et al., 1997) to observed body mass. 95% confidence interval of regression indicated by dark gray region. White circles indicate females; gray circles indicate males. With the juvenile test sample, the ST-BIB prediction equation consistently overpredicts body masses across the estimated body mass range of australopiths.

age (Bock et al., 1973). American boys may begin the adolescent growth spurt by 9.7 years of age (Lee, 1980), and reach peak growth velocity (in stature) around age 13 (Bock et al., 1973). Our oldest-aged samples (10–12 years) capture the early stages of puberty for both sexes, during which lower limb length growth velocity surpasses that of trunk length (Buckler, 2012). Additionally, this peri- and post-pubescent period is associated with marked increases in absolute and relative muscle mass (McCannon, 1970; Ruff, 2003; Malina et al., 2004), suggesting that even the children in our oldest age groups possess relatively much lower muscle masses than the adults in the reference sample. Accordingly, the adult-based prediction equations may overpredict mass in older, relatively long-legged, and relatively smaller-muscled juveniles. Though children in our youngest age groups (for whom mass is underestimated) are also expected to possess relatively low muscle masses (compared to adults and older children), we suspect – if relative muscle mass is indeed a factor in body mass estimation error – that the effect of limb proportions predominates. The effect of limb proportions in older children is hinted at in the LL/ST ratios of Figure 3: the oldest-aged samples have higher mean LL/ST ratios than adults (though the differences are not statistically significant). In later adolescence (beyond the range of our samples), trunk length growth velocity is greater than leg length growth velocity (Buckler, 2012), and late adolescents increasingly assume adult body proportions and muscle mass. Thus, prediction errors would likely trend toward zero if 13–18 year old groups were analyzed.

Based on results of the linear mixed models, LL/ST has the strongest influence on prediction error of the examined somatic variables. This is promising for the potential application of morphometric equations to fossil hominins, since LL/ST ratios can be reasonably estimated from relatively complete fossil specimens. It is possible that novel prediction equations could be generated using bi-iliac breadth, sitting height, and lower limb length; these equations may be more reliable in populations with different lower limb or trunk proportions.

Biacromial breadth/bi-iliac breadth is also a significant factor (either alone or in interaction with LL/ST) in both linear mixed models. However, in contrast to LL/ST, anthropometric biacromial breadth is not possible to calculate from skeletal elements alone (Ruff, 1991; Ruff et al., 2005), though some studies have calculated biacromial breadth solely from osteological measurements (e.g., Eyre, 2014). If BAB cannot be estimated for fossil hominins, then differences between modern humans and fossil taxa in shoulder-to-hip ratios could potentially generate large over- or under-predictions of body mass. This issue might be addressed by examining the relationship between osteological measurements (clavicle length, manubrium width) and anthropometric biacromial breadth to develop a correction factor, similar to one used to convert skeletal bi-iliac breadth to living bi-iliac breadth (Ruff et al., 1997). Applying morphometric prediction equations to samples with different BAB/BIB ratios (such as non-human hominids) would also be informative.

When considering application to fossil taxa, differing BAB/BIB proportions appear to be more problematic than differing LL/ST proportions. However, the direction of significant differences between BAB/BIB ratios in juveniles and adults (Fig. 3) gives some indication that this ratio may be less problematic for the adult female prediction equation. In our sample, when female juvenile BAB/BIB ratios are significantly different from female adults (at ages 7, and 9–11), the juveniles have higher BAB/BIB ratios than the adults. Since many fossil hominins have very broad and laterally flaring ilia (Ruff, 1991; Simpson et al., 2008; Kibii et al., 2011; Arsuaga et al., 2015; Berger et al., 2015), they may have broader hips than shoulders, and thus BAB/BIB ratios that are more similar

to adult females than adult males (or any of the juvenile age groups observed here). This similarity may reduce error caused by proportional differences, and supports applying an adult female equation to fossil hominins (although it is certainly possible that some fossil specimens have significantly lower BAB/BIB ratios than adult females).

As a soft tissue metric, triceps skinfold thickness (TST) cannot be estimated for fossil specimens. Although TST does not have a significant effect on prediction error for this sample, we are cautious to interpret this as evidence that subcutaneous fat is not a potential source of error in body mass prediction, as other studies have found that fat mass has a strong effect on predictive accuracy (Lorkiewicz-Muszyńska et al., 2013; Schaffer, 2016). Other studies have noted the tendency for increasing accumulation of adipose tissue throughout adulthood (Hruschka et al., 2014), so it is also possible that juveniles likely do not have as much subcutaneous fat as adults (particularly older adults). If so, juveniles are not likely to be the most representative sample for evaluating the relationship between prediction error and subcutaneous fat.

Finally, we offer some cautionary comments concerning the construction and use of morphometric body mass prediction equations. For fossil specimens, both predictor variables (stature and bi-iliac breadth) need to be derived from prediction equations generated from modern reference samples. There are many published prediction equations for stature (e.g., Jungers, 1988b; Konigsberg et al., 1998; Hens et al., 2000; Raxter et al., 2006, 2007). Because bi-iliac breadth in the ST-BIB equation is effectively living bi-iliac breadth (rather than osteological bi-iliac breadth), it also needs to be predicted for skeletonized remains using an equation derived from a modern reference sample (Ruff et al., 1997; Ruff, 2010). Though mean predicted values might be very accurate, the statistical uncertainty surrounding these predictions grows substantially as predicted values (stature and living bi-iliac breadth) are incorporated into subsequent predictions (body mass). Martin (1990) has warned against the practice of layering predicted values on top of one another, and Elliott et al. (2016b) recently expressed similar concern. Future work should attempt to reduce the reliance on predicted values for morphometric prediction equations, similar to equations published by Ruff (2007) that use bi-iliac breadth and long bone lengths (rather than stature) to predict body mass. We believe these developments are particularly important for predicting body mass in archeological and paleontological contexts. Additionally, these results suggest that morphometric prediction methods may have some utility in forensic or bioarcheological contexts in which juvenile body mass is of interest, though not without limitation. With respect to forensic applications, it is important to remember that while mass prediction techniques (whether mechanical or morphometric) may do a reasonable job of predicting the average mass of individuals with certain skeletal dimensions, their prediction errors tend to make them relatively poor estimators of the mass of a given individual. When combined with uncertainty about the body composition (that is, fat mass) of an unknown individual, predicting the mass of individuals in forensic cases – especially juvenile individuals – should be done with extreme caution. In bioarcheological contexts, in which juvenile body size may be of interest, we encourage the direct comparison of skeletal dimensions (e.g., femoral head diameters) between samples as a means of avoiding the unwanted addition of prediction error into comparative analyses.

5. Conclusions

Determining the reliability of morphometric equations when applied to small-bodied fossil hominins hinges largely on resolving

sources of prediction error. We have argued that the systematic overprediction of body mass in juveniles aged 9–12 years (20–46 kg) is in part caused by extrapolation beyond the range of observed body masses in the adult reference sample. As morphometric equations are OLS multiple regression equations, extrapolation beyond the lower bound of the reference sample will generally result in overpredicted values (Smith, 1984). Overprediction in these age groups may also be related to relatively longer lower limbs and lower muscle mass than the adult modern human reference samples.

Body masses in younger age groups (6–8 years) are generally underpredicted; we believe this is most likely a result of significant differences in lower limb length/stature proportions. Ratios of BAB/BIB and TST do not have as obvious (or significant, in the case of TST) effects on prediction error as LL/ST, which is encouraging, as these somatic variables are difficult or impossible to measure in fossil taxa.

The results of this study demonstrate that morphometric body mass prediction equations derived from adult human samples (and especially adult female samples) generate reasonably accurate predictions of body mass when applied to juvenile human remains. Morphometric prediction equations may be acceptable tools for predicting body mass of small-bodied fossil hominins, provided the fossil specimens are complete enough to permit assessment of body proportions (such as LL/ST) as well as reliable predictions of stature. Future work should focus on generating equations that do not rely on predicted values (living bi-iliac breadth and stature) as input. For example, replacing stature with long bone lengths (as in Ruff, 2007) would help reduce the statistical uncertainty around predicted body masses. Evaluating morphometric equations with data from non-human great apes (an extreme case of extrapolation) would help elucidate the effects of differences in body proportions on prediction error. Finally, efforts to generate prediction equations from juvenile samples would likely help alleviate prediction error induced by extrapolation beyond the lower bound of body masses in the reference sample.

Acknowledgments

We thank Chris Ruff for the invitation to participate in the AAPA symposium “Body Mass Estimation in Biological Anthropology,” and to contribute to this special issue of the Journal of Human Evolution. Chris Ruff, Bill Jungers, and an anonymous reviewer provided helpful comments that improved this paper. This work was supported by the National Science Foundation (NSF BCS 1540421 to GSY and Doug M. Boyer) and Duke University.

Supplementary Online Material

Supplementary online material related to this article can be found at <http://dx.doi.org/10.1016/j.jhevol.2017.03.009>.

References

- Aiello, L.C., 1992. Allometry and the analysis of size and shape in human evolution. *J. Hum. Evol.* 22, 127–147.
- Aiello, L.C., Wells, J., 2002. Energetics and the evolution of the genus *Homo*. *A. Rev. Anthropol.* 31, 323–338.
- Aiello, L.C., Wheeler, P., 1995. The expensive-tissue hypothesis: the brain and the digestive system in human and primate evolution. *Curr. Anthropol.* 36, 199–221.
- Aiello, L.C., Wood, B.A., 1994. Cranial variables as predictors of hominine body mass. *Am. J. Phys. Anthropol.* 95, 409–426.
- Arsuaga, J., Lorenzo, C., Gracia, A., Martí, I., 1999. A complete human pelvis from the Middle Pleistocene of Spain. *Nature* 399, 255–258.
- Arsuaga, J., Carretero, J.M., Lorenzo, C., Gómez-Olivencia, A., Pablos, A., Rodríguez, L., García-González, R., Bonmati, A., Quam, R.M., Pantoja-Pérez, A., Martínez, I., Aranburu, A., Gracia-Téllez, A., Poza-Rey, E., Sala, N., García, N., Alcázar de Velasco, A., Cuenca-Bescós, G., Bermúdez de Castro, J.M., Carbonell, E., 2015. Postcranial morphology of the middle Pleistocene humans from Sima de los Huesos, Spain. *Proc. Natl. Acad. Sci.* 112, 11524–11529.
- Asfaw, B., White, T.D., Lovejoy, C.O., Latimer, B.M., Simpson, S., Suwa, G., 1999. *Australopithecus garhi*: a new species of early hominid from Ethiopia. *Science* 284, 629–635.
- Auerbach, B.M., Ruff, C.B., 2004. Human body mass estimation: a comparison of “morphometric” and “mechanical” methods. *Am. J. Phys. Anthropol.* 125, 331–342.
- Berge, C., 1998. Heterochronic processes in human evolution: an ontogenetic analysis of the hominid pelvis. *Am. J. Phys. Anthropol.* 105, 441–459.
- Berger, L.R., de Ruiter, D.J., Churchill, S.E., Schmid, P., Carlson, K.J., Dirks, P.H.G.M., Kibii, J.M., 2010. *Australopithecus sediba*: a new species of *Homo*-like australopithecine from South Africa. *Science* 328, 195–204.
- Berger, L.R., Hawks, J., de Ruiter, D.J., Churchill, S.E., Schmid, P., Deleuzene, L.K., Kivell, T.L., Garvin, H.M., Williams, S.A., DeSilva, J.M., Skinner, M.M., Musiba, C.M., Cameron, N., Holliday, T.W., Harcourt-Smith, W., Ackermann, R.R., Bastir, M., Bogin, B., Bolter, D., Brophy, J., Cofran, Z.D., Congdon, K.A., Deane, A.S., Dembo, M., Drapeau, M., Elliott, M.C., Feuerriegel, E.M., Garcia-Martinez, D., Green, D.J., Gurtov, A., Irish, J.D., Kruger, A., Laird, M.F., Marchi, D., Meyer, M.R., Nalla, S., Negash, E.W., Orr, C.M., Radovic, D., Schroeder, L., Scott, J.E., Throckmorton, Z., Tocheri, M.W., VanSickle, C., Walker, C.S., Wei, P., Zipfel, B., 2015. *Homo naledi*, a new species of the genus *Homo* from the Dinaledi Chamber, South Africa. *eLife* 4, e09560.
- Bock, R.D., Wainer, H., Petersen, A., Thissen, D., Murray, J., Roche, A., 1973. A parameterization for individual human growth curves. *Hum. Biol.* 45, 63–80.
- Bogin, B., 1999. *Patterns of Human Growth*. Cambridge University Press, Cambridge.
- Bogin, B., Smith, P., Orden, A.B., Varela Silva, M.I., Loucky, J., 2002. Rapid change in height and body proportions of Maya American children. *Am. J. Hum. Biol.* 14, 753–761.
- Buckler, J., 2012. *A Longitudinal Study of Adolescent Growth*. Springer-Verlag, London.
- Cant, J.G.H., 1992. Positional behavior and body size of arboreal primates: a theoretical framework for field studies and an illustration of its application. *Am. J. Phys. Anthropol.* 88, 273–283.
- Churchill, S.E., 2006. Bioenergetic perspectives on Neanderthal thermoregulatory and activity budgets. In: Hublin, J.J., Harvati, K., Harrison, T. (Eds.), *Neanderthals Revisited: New Approaches and Perspectives*. Springer, Netherlands, pp. 113–134.
- Churchill, S.E., Holliday, T.W., Carlson, K.J., Jashashvili, T., Macias, M.E., Mathews, S., Sparling, T.L., Schmid, P., de Ruiter, D.J., Berger, L.R., 2013. The upper limb of *Australopithecus sediba*. *Science* 340, 1233477.
- Cnaan, A., Laird, N.M., Slator, P., 1997. Tutorial in biostatistics: using the general linear mixed model to analyse unbalanced repeated measures and longitudinal data. *Stat. Med.* 16, 2349–2380.
- Dagosto, M., Terranova, C.J., 1992. Estimating the body size of Eocene primates: a comparison of results from dental and postcranial variables. *Int. J. Primatol.* 13, 307–344.
- Delson, E., Terranova, C.J., Jungers, W.L., Sargis, E.J., Jablonski, N.G., Dechow, P., 2000. Body mass in Cercopithecidae (Primates, Mammalia): estimation and scaling in extinct and extant taxa. *Anthropol. Pap. Am. Mus. Nat. Hist.* 83, 1–159.
- DeSilva, J.M., Holt, K.G., Churchill, S.E., Carlson, K.J., Walker, C.S., Zipfel, B., Berger, L.R., 2013. The lower limb and mechanics of walking in *Australopithecus sediba*. *Science* 340, 1232999.
- Dobson, A., Barnett, A., 2008. *An Introduction to Generalized Linear Models*, 3rd ed. Chapman and Hall/CRC Press, Boca Raton.
- Elliott, M., Kurki, H., Weston, D.A., Collard, M., 2016a. Estimating body mass from skeletal material: new predictive equations and methodological insights from analyses of a known-mass sample of humans. *Archaeol. Anthropol. Sci.* 8, 731–750.
- Elliott, M., Kurki, H., Weston, D.A., Collard, M., 2016b. Estimating body mass from postcranial variables: an evaluation of current equations using a large known-mass sample of modern humans. *Archaeol. Anthropol. Sci.* 8, 689–704.
- Eveleigh, P., Tanner, J., 1976. *Worldwide Variation in Human Growth*. Cambridge University Press, Cambridge.
- Eyre, J., 2014. Bi-acromial and clavicular scaling in hominoids: implications for locomotor behavior. *Am. J. Phys. Anthropol.* 153, 115.
- Feldesman, M.R., 1992. Femur/stature ratio and estimates of stature in children. *Am. J. Phys. Anthropol.* 87, 447–459.
- Feuerriegel, E.M., Green, D.J., Walker, C.S., Schmid, P., Hawks, J., Berger, L.R., Churchill, S.E., 2017. The upper limb of *Homo naledi*. *J. Hum. Evol.* 104, 155–173.
- Grabowski, M., Hatala, K.G., Jungers, W.L., Richmond, B.G., 2015. Body mass estimates of hominin fossils and the evolution of human body size. *J. Hum. Evol.* 85, 75–93.
- Green, D.J., Gordon, A.D., Richmond, B.G., 2007. Limb-size proportions in *Australopithecus afarensis* and *Australopithecus africanus*. *J. Hum. Evol.* 52, 187–200.
- Grine, F.E., Jungers, W.L., Tobias, P.V., Pearson, O.M., 1995. Fossil *Homo* femur from Berg Aukas, Northern Namibia. *Am. J. Phys. Anthropol.* 97, 151–185.
- Hauspie, R., 2002. Adolescence: somatic growth and sex differences. In: Cameron, N. (Ed.), *Human Growth and Development*. Academic Press, San Diego, pp. 45–64.
- Hens, S.M., Konigsberg, L.W., Jungers, W.L., 1998. Estimation of African ape body length from femur length. *J. Hum. Evol.* 34, 401–411.

- Hens, S.M., Konigsberg, L.W., Jungers, W.L., 2000. Estimating stature in fossil hominids: which regression model and reference sample to use? *J. Hum. Evol.* 38, 767–784.
- Hiernaux, J., 1985. A comparison of the shoulder-hip-width sexual dimorphism in sub-Saharan Africa and Europe. In: Ghesquiere, J., Martin, R.D., Newcombe, F. (Eds.), *Human Sexual Dimorphism*. Taylor and Francis, Philadelphia, pp. 191–206.
- Holt, B.M., 2003. Mobility in Upper Paleolithic and Mesolithic Europe: evidence from the lower limb. *Am. J. Phys. Anthropol.* 122, 200–215.
- Hruschka, D.J., Hadley, C., Brewis, A., 2014. Disentangling basal and accumulated body mass for cross-population comparisons. *Am. J. Phys. Anthropol.* 153, 542–550.
- Humphrey, L.T., 1998. Growth patterns in the modern human skeleton. *Am. J. Phys. Anthropol.* 105, 57–72.
- Isbell, L.A., 1994. Predation on primates: ecological patterns and evolutionary consequences. *Evol. Anthropol.* 3, 61–71.
- Johanson, D.C., Lovejoy, C.O., Kimbel, W.H., White, T.D., Ward, S.C., Bush, M.E., Latimer, B.M., Coppens, Y., 1982. Morphology of the Pliocene partial hominid skeleton (A.L. 288-1) from the Hadar formation, Ethiopia. *Am. J. Phys. Anthropol.* 57, 403–451.
- Jungers, W.L., 1982. Lucy's limbs: skeletal allometry and locomotion in *Australopithecus afarensis*. *Nature* 297, 676–678.
- Jungers, W.L., 1988a. Relative joint size and hominoid locomotor adaptations with implications for the evolution of hominid bipedalism. *J. Hum. Evol.* 17, 247–265.
- Jungers, W.L., 1988b. Lucy's length: stature reconstruction in *Australopithecus afarensis* (A.L. 288-1) with implications for other small-bodied hominids. *Am. J. Phys. Anthropol.* 76, 227–231.
- Jungers, W.L., Grabowski, M., Hatala, K.G., Richmond, B.G., 2016. The evolution of body size and shape in the human career. *Philos. T. R. Soc. B* 371, 125–129.
- Kibii, J.M., Churchill, S.E., Schmid, P., Carlson, K.J., Reed, N.D., de Ruiter, D.J., Berger, L.R., 2011. A partial pelvis of *Australopithecus sediba*. *Science* 333, 1407–1411.
- Konigsberg, L.W., Hens, S.M., Jantz, L.M., Jungers, W.L., 1998. Stature estimation and calibration: Bayesian and maximum likelihood perspectives in physical anthropology. *Am. J. Phys. Anthropol.* 107, 65–92.
- Kurki, H.K., Ginter, J.K., Stock, J.T., Pfeiffer, S., 2010. Body size estimation of small-bodied humans: applicability of current methods. *Am. J. Phys. Anthropol.* 141, 169–180.
- Lee, P.A., 1980. Normal ages of pubertal events among American males and females. *J. Adolescent Health* 1, 26–29.
- Lorkiewicz-Muszyńska, D., Przystańska, A., Kociemba, W., Sroka, A., Żaba, C., Paprzycki, W., 2013. Body mass estimation in modern population using anthropometric measurements from computed tomography. *Forensic Sci. Int.* 231, 405 e1–405.e6.
- Lovejoy, C.O., Suwa, G., Simpson, S.W., Matternes, J.H., White, T.D., 2009. The great divides: *Ardipithecus ramidus* reveals the postcrania of our last common ancestors with African apes. *Science* 326, 100–106.
- Malina, R.M., Bouchard, C., Bar-Or, O., 2004. Growth, Maturation, and Physical Activity, 2nd Edition. Human Kinetics, Champaign, IL.
- Marchi, D., Walker, C.S., Wei, P., Holliday, T.W., Churchill, S.E., Berger, L.R., DeSilva, J.M., 2017. The thigh and leg of *Homo naledi*. *J. Hum. Evol.* 104, 174–204.
- Martin, R.A., 1990. Estimating body mass and correlated variables in extinct mammals: travels in the fourth dimensions. In: Damuth, J., MacFadden, B.J. (Eds.), *Body Size in Mammalian Paleobiology: Estimation and Biological Implications*. Cambridge University Press, Cambridge, pp. 49–68.
- Martorell, R., Malina, R.M., Castillo, R.O., Mendoza, F.S., Pawson, I.G., 1988. Body proportions in three ethnic groups: children and youths 2–17 years in NHANES II and HHANES. *Hum. Biol.* 60, 205–222.
- McCammon, R.W., 1970. *Human Growth and Development*. Charles C. Thomas, Springfield, Ill.
- McHenry, H.M., 1992. Body size and proportions in early hominids. *Am. J. Phys. Anthropol.* 87, 407–431.
- McHenry, H.M., Berger, L.R., 1998. Body proportions of *Australopithecus afarensis* and *A. africanus* and the origin of the genus *Homo*. *J. Hum. Evol.* 35, 1–22.
- Morwood, M.J., Brown, P., Jatmiko, Sutikna, T., Saptomo, E.W., Westaway, K.E., Due, R.A., Roberts, R.G., Maeda, T., Wasisto, S., Djubiantono, T., 2005. Further evidence for small-bodied hominins from the Late Pleistocene of Flores, Indonesia. *Nature* 437, 1012–1017.
- Pomeroy, E., Stock, J.T., 2012. Estimation of stature and body mass from the skeleton among coastal and mid-altitude Andean populations. *Am. J. Phys. Anthropol.* 147, 264–279.
- Raxter, M.H., Auerbach, B.M., Ruff, C.B., 2006. Revision of the Fully technique for estimating statures. *Am. J. Phys. Anthropol.* 130, 374–384.
- Raxter, M.H., Ruff, C.B., Auerbach, B.M., 2007. Technical note: revised Fully stature estimation technique. *Am. J. Phys. Anthropol.* 133, 817–818.
- Richmond, B.G., Aiello, L.C., Wood, B., 2002. Early hominin limb proportions. *J. Hum. Evol.* 43, 529–548.
- Ricker, W.E., 1973. Linear regressions in fishery research. *J. Fish. Res. Board Can.* 30, 409–434.
- Robbins, G., Sciuili, P.W., Blatt, S.H., 2010. Estimating body mass in subadult human skeletons. *Am. J. Phys. Anthropol.* 143, 146–150.
- Robbins Schug, G., Gupta, S., Cowgill, L.W., Sciuili, P.W., Blatt, S.H., 2013. Panel regression formulas for estimating stature and body mass from immature human skeletons: a statistical approach without reference to specific age estimates. *J. Archaeol. Sci.* 40, 1–11.
- Robson, S.L., Wood, B., 2008. Hominin life history: reconstruction and evolution. *J. Anat.* 212, 394–425.
- Rosenberg, K.R., Zünig, L., Ruff, C.B., 2006. Body size, body proportions, and encephalization in a Middle Pleistocene archaic human from northern China. *Proc. Natl. Acad. Sci.* 103, 3552–3556.
- Rubenson, J., Heliams, D.B., Maloney, S.K., Withers, P.C., Lloyd, D.G., Fournier, P.A., 2007. Reappraisal of the comparative cost of human locomotion using gait-specific allometric analyses. *J. Exp. Biol.* 210, 3513–3524.
- Ruff, C.B., 1991. Climate and body shape in hominid evolution. *J. Hum. Evol.* 21, 81–105.
- Ruff, C.B., 1994. Morphological adaptation to climate in modern and fossil hominids. *Am. J. Phys. Anthropol.* 37, 65–107.
- Ruff, C.B., 1995. Biomechanics of the hip and birth in early *Homo*. *Am. J. Phys. Anthropol.* 98, 527–574.
- Ruff, C.B., 1998. Evolution of the hominid hip. In: Strasser, E., Fleagle, J., McHenry, H., Rosenberger, A. (Eds.), *Primate Locomotion: Recent Advances*. Plenum Press, New York, pp. 449–469.
- Ruff, C.B., 2000a. Body size, body shape, and long bone strength in modern humans. *J. Hum. Evol.* 38, 269–290.
- Ruff, C.B., 2000b. Body mass prediction from skeletal frame size in elite athletes. *Am. J. Phys. Anthropol.* 113, 507–517.
- Ruff, C.B., 2002. Variation in human body size and shape. *Annu. Rev. Anthropol.* 31, 211–232.
- Ruff, C.B., 2003. Growth in bone strength, body size, and muscle size in a juvenile longitudinal sample. *Bone* 33, 317–329.
- Ruff, C.B., 2007. Body size prediction from juvenile skeletal remains. *Am. J. Phys. Anthropol.* 133, 698–716.
- Ruff, C.B., 2010. Body size and body shape in early hominins – implications of the Gona pelvis. *J. Hum. Evol.* 58, 166–178.
- Ruff, C.B., 2015. The effects of hip joint loading on body mass estimation in early hominins. *Am. J. Phys. Anthropol.* 156 (Suppl. 60), 273.
- Ruff, C.B., Walker, A., 1993. Body size and body shape. In: Walker, A., Leakey, R. (Eds.), *The Nariokotome *Homo erectus* Skeleton*. Harvard University Press, Cambridge, MA, pp. 234–265.
- Ruff, C.B., Scott, W.W., Liu, A.Y., 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., Holliday, T.W., 1997. Body mass and encephalization in Pleistocene *Homo*. *Nature* 387, 173–176.
- Ruff, C.B., McHenry, H.M., Thackeray, J.F., 1999. Cross-sectional morphology of the SK 82 and 97 proximal femora. *Am. J. Phys. Anthropol.* 109, 509–521.
- Ruff, C.B., Niskanen, M., Junno, J.A., Jamison, P., 2005. Body mass prediction from stature and bi-iliac breadth in two high latitude populations, with application to earlier higher latitude humans. *J. Hum. Evol.* 48, 381–392.
- Ruff, C.B., Holt, B.M., Sládek, V., Berner, M., Murphy, W.A., zur Nedden, D., Seidler, H., Recheis, W., 2006. Body size, body proportions, and mobility in the Tyrolean “Ice Man”. *J. Hum. Evol.* 51, 91–101.
- Ruff, C.B., Holt, B.M., Niskanen, M., Sládek, V., Berner, M., Garofalo, E., Garvin, H.M., Hora, M., Majajani, H., Niinimäki, S., Salo, K., Schuplerová, E., Tompkins, D., 2012. Stature and body mass estimation from skeletal remains in the European Holocene. *Am. J. Phys. Anthropol.* 148, 601–617.
- Schaffer, W.C., 2016. Total body mass estimation from anthropometric measurements in modern young adult U.S. populations with healthy body fat percentages (NHANES III). *J. Forensic Sci.* 61, 1431–1439.
- Schmid, P., Churchill, S.E., Nalla, S., Weissen, E., Carlson, K.J., de Ruiter, D.J., Berger, L.R., 2013. Mosaic morphology in the thorax of *Australopithecus sediba*. *Science* 340, 1234598.
- Sciuili, P.W., Blatt, S.H., 2008. Evaluation of juvenile stature and body mass prediction. *Am. J. Phys. Anthropol.* 136, 387–393.
- Siegmund, F., Papageorgopoulou, C., 2011. Body mass and body mass index estimation in medieval Switzerland. *Bull. Schweizerischen Gesellschaft Anthropol.* 17, 35–44.
- Simpson, S.W., Quade, J., Levin, N.E., Butler, R., Dupont-Nivet, G., Everett, M., Semaw, S., 2008. A female *Homo erectus* pelvis from Gona, Ethiopia. *Science* 322, 1089–1092.
- Smith, R.J., 1980. Rethinking allometry. *J. Theoret. Biol.* 87, 97–111.
- Smith, R.J., 1984. Allometric scaling in comparative biology: problems of concept and method. *Am. J. Physiol.* 246, R152–R160.
- Smith, R.J., 2009. Use and misuse of the reduced major axis for line-fitting. *Am. J. Phys. Anthropol.* 140, 476–486.
- Snodgrass, J., Leonard, W., 2009. Neandertal energetics revisited: insights into population dynamics and life history evolution. *PaleoAnthropology* 2009, 220–237.
- Squyres, N., Ruff, C.B., 2015. Body mass estimation from knee breadth, with application to early hominins. *Am. J. Phys. Anthropol.* 158, 198–208.
- Temple, D.H., Okazaki, K., Cowgill, L.W., 2011. Ontogeny of limb proportions in late through final Jomon period foragers. *Am. J. Phys. Anthropol.* 145, 415–425.
- Trinkaus, E., Churchill, S.E., Ruff, C.B., Vandermeersch, B., 1999. Long bone shaft robusticity and body proportions of the Saint-Césaire 1 Châtelperronian Neanderthal. *J. Archaeol. Sci.* 26, 753–773.
- Uhl, N.M., Rainwater, C.W., Konigsberg, L.W., 2013. Testing for size and allometric differences in fossil hominin body mass estimation. *Am. J. Phys. Anthropol.* 151, 215–229.
- Vercellotti, G., Alciati, G., Richards, M.P., Formicola, V., 2008. The Late Upper Paleolithic skeleton Villabruna 1 (Italy): a source of data on biology and behavior of a 14000 year-old hunter. *J. Anthropol. Sci.* 86, 143–163.

- Walker, C.S., Desilva, J.M., Holliday, T.W., Marchi, D., Garvin, H.M., Cofran, Z., Hawks, J., Berger, L.R., Churchill, S.E., 2015. Relative length of the immature *Homo naledi* tibia U.W. 101-1070: evidence for elongation of the leg. *Am. J. Phys. Anthropol.* 159, 326.
- Warton, D.I., Wright, I.J., Falster, D.S., Westoby, M., 2006. Bivariate line-fitting methods for allometry. *Biol. Rev.* 81, 259–291.
- Will, M., Stock, J.T., 2015. Spatial and temporal variation of body size among early *Homo*. *J. Hum. Evol.* 82, 15–33.
- Yapuncich, G.S., 2018. Alternative methods for calculating percentage prediction error and their implications for predicting body mass in fossil taxa. *J. Hum. Evol.* 115, 140–145.
- Yapuncich, G.S., Gladman, J.T., Boyer, D.M., 2015. Predicting euarchontan body mass: a comparison of tarsal and dental variables. *Am. J. Phys. Anthropol.* 157, 472–506.