CrossMark

Introduction to special issue: Body mass estimation — Methodological issues and fossil applications

Christopher B. Ruff [a, *], Markku Niskanen [b]

[a] Center for Functional Anatomy and Evolution, Johns Hopkins University School of Medicine, 1830 E. Monument St., Baltimore, MD 21205, USA
[b] Department of Archeology, University of Oulu, Oulu 90014, Finland

ARTICLE INFO

1. Introduction

The significance of body mass — as a biological parameter itself, as a predictor of other biological, behavioral, and ecological variables, and as the "denominator" in many other analyses, such as relative brain size (encephalization) and relative skeletal strength (robusticity) — is so well known as to be almost a truism (Calder, 1984; Schmidt-Nielson, 1984; Jungers, 1985; Damuth and Mac-Fadden, 1990a). The measurement and interpretation of body mass is not without its own issues, even in living or recently deceased animals — regional, seasonal, and individual variation in body composition can significantly affect body mass and obscure relationships to other variables (e.g., Van Valkenburgh, 1990; also see below). Problems are even more pronounced in analyses involving past specimens and taxa, where body mass must be estimated from preserved skeletal or dental remains. Which dimensions provide the most accurate and consistent estimates and which extant reference groups are most appropriate for particular target specimens are two major issues that must be addressed in any such study.

The present collection of papers originated in a 2015 American Association of Physical Anthropologists symposium that brought together researchers from several disciplines — primatology, paleontology, bioarchaeology, and forensic anthropology — who could provide different perspectives on common issues related to body mass estimation. The specific goals and challenges in estimating body mass are somewhat different across disciplines. For example, in forensic anthropology, predicting actual body mass at the time of death might be the objective, while in paleontology, an

"ideal" or healthy young adult value might be more appropriate, e.g., for cross-species comparisons. The two issues may be intertwined, however, as forensic (autopsy) samples have often been used as reference samples in human paleontological research. Several of the studies in this special issue investigate the effects of variation in body composition and body mass relative to other variables (such as stature) in developing and applying estimation formulae, in both living and autopsy samples, with implications for their use in earlier specimens. Age-related variation in the relationship between body mass and skeletal dimensions is another issue that crosses interdisciplinary lines and is addressed in several of the studies using living, forensic, and archaeological samples. Fundamentally different approaches to body mass estimation, based on either mechanical principles or different methods of general morphological reconstruction, are also explored, again in a number of contexts ranging from broad cross-species to within-population analyses. It is hoped that these different levels and types of comparisons will stimulate researchers in all fields, and provide insight into factors that should be considered in such analyses, regardless of the specific aims of the study.

In this spirit, rather than providing a paper-by-paper summary of findings, the following review is organized thematically, focusing on the overlap between studies and the different ways in which these common issues can be addressed.

2. Which predictor variables?

Although other approaches are possible, as discussed further below, the most common method for estimating body mass from skeletal/dental remains is through predictive equations based on single, or less frequently, multiple, metric dimension(s). Almost all of the studies in this special issue employed postcranial dimensions as estimators, although dental estimates were also considered in comparisons (Dagosto et al., 2018; Perry et al., 2018) and employed in a methodological study (Yapuncich, 2018). Dental and cranial dimensions have certain advantages when estimating body mass, particularly in paleontological contexts, including the greater availability of material (especially dental) and certainty of taxonomic assignment (which is usually based on craniodental

* Corresponding author.
  E-mail addresses: cbruff@jhmi.edu (C.B. Ruff), markku.niskanen@oulu.fi (M. Niskanen).

characteristics). However, it is generally acknowledged that postcranial dimensions, when available, provide more accurate estimates because of their more direct physiological relationship with body mass (Damuth and MacFadden, 1990b). The mechanical role of weight-bearing elements, in particular limb bones (lower limb bones in humans), makes them logical candidates for generating predictor variables. Articulations have some potential advantages in this regard, as they are less developmentally plastic and so should not reflect individual differences in behavior to the same extent as, for example, diaphyseal breadth dimensions (Trinkaus et al., 1994; Lieberman et al., 2001; Ruff et al., 2018). This issue is addressed by Cowgill (2018) in comparisons between body mass estimates calculated from femoral head versus diaphyseal dimensions in juveniles from a number of archaeological samples, using equations derived from a modern living sample (Denver Growth Study: McCammon, 1970). The diaphyseal-based equations produce greater body mass estimates than the articular-based equations in the archaeological samples, with the greatest differences in two samples that were likely very active, consistent with the expected generally higher activity levels of past populations and greater developmental plasticity of diaphyses. This calls for caution in applying diaphyseal-based equations developed in one reference sample to other samples with potentially different lifestyles. This issue is of less concern where the reference and target samples are expected to be more similar in overall behavior, here for example in applications to modern forensic cases.

Articular or peri-articular breadths or areas of weight-bearing limb bones are used as predictor variables in several studies in the special issue, ranging from broader inter-species analyses (Dagosto et al., 2018; Perry et al., 2018; Burgess et al., 2018) to studies focused on hominins (Walker et al., 2018; Ruff et al., 2018; Niskanen et al., 2018; Young et al., 2018). In the two taxonomically broadest studies — of euarchontans, non-catarrhine primates, and small mammals (Dagosto et al., 2018), and catarrhine and platyrrhine primates (Perry et al., 2018) — mediolateral breadth of the knee joint (articular or periarticular) was found to be the best predictor of body mass when compared with several other articular dimensions, as judged by statistical precision (percent standard error of estimate, percent prediction error), and consistency of results across different taxonomic groupings, due to similar scaling relationships. This is similar to results reported earlier (Ruff, 2003) for catarrhine primates, where knee transverse breadth relative to body mass was found to be "locomotor blind". This may result from the generally conservative nature of knee morphology and function (Dye, 1987), and the minor effect of knee mediolateral breadth on joint excursion, compared to other articular dimensions (Jenkins and Camazine, 1977; MacLatchy and Bossert, 1996; Hammond, 2014). Thus, in analyses involving broad taxonomic/functional reference groups, i.e., where the target specimen is not closely affiliated with a particular modern group, knee transverse breadth may be the safest, least biased body mass estimator.

Interestingly, however, in a study more narrowly focused on hominoids (Burgess et al., 2018) knee mediolateral breadth dimensions are not the best (smallest %SEE, most consistent) body mass estimators among a number of limb bone articular breadths. As discussed in that paper, more subtle differences in knee joint structure and function among these closely related taxa, and possibly fewer differences in more proximal joints such as the femoral head, may account for this result. This highlights the importance of the scope of an investigation in choosing the most appropriate predictor variable(s) — if a target specimen is considered to be closely affiliated with modern hominoids, then a femoral head equation might be preferable, while a less certain taxonomic attribution (or indications that it differed functionally) could suggest using a knee breadth equation. Similarity in inter- and intra-

limb proportions between reference samples and target specimens is one criterion that can be used to help identify the most appropriate formulae to apply (Perry et al., 2018).

Variation in lower limb articular proportions and its effect on body mass estimation in hominins is explored in Ruff et al. (2018). Femoral head breadth is often used as an estimator or surrogate for body mass in hominins (McHenry, 1992; Ruff et al., 1997; Steudel-Numbers and Tilkens, 2004; Churchill et al., 2012; Holliday, 2012; Plavcan, 2012; Trinkaus and Ruff, 2012; Arsuaga et al., 2015; Grabowski et al., 2015; Will and Stock, 2015; Cowgill, 2018; Niskanen et al., 2018; Young et al., 2018), although whether its scaling relative to body mass is similar in modern humans and early hominins has been a matter of continuing debate (Napier, 1964; Lovejoy et al., 1973; McHenry, 1975; Jungers, 1991; Ruff, 1998; Richmond and Jungers, 2008; Ruff et al., 2016). Ruff et al. (2018) argue that based on comparisons between hip and knee breadth proportions, australopiths are more ape-like, with a relatively smaller femoral head, while early *Homo* (1.5–2.0 Ma) specimens are similar to modern humans. The implied relative reduction in hip joint loading is compensated for by including a correction factor for australopiths. New estimation formulae based on femoral head and knee breadths are developed using morphometrically-estimated body masses (see below) for a diverse sample of modern humans.

The effects on body mass estimation of incorporating a bone length (or estimated stature) dimension in addition to femoral head breadth are investigated in Niskanen et al. (2018). When compared against morphometrically-estimated body mass in several skeletal samples or reported living body mass, some improvement in estimates is achieved, indicating that linear body proportions and variation in stature have some effect on estimation error. In some ways, then, this combines aspects of both "mechanical" and "morphometric" (see below) methods of body mass estimation. Young et al. (2018) test three previous femoral head estimation equations on a living female British sample and find that all three perform relatively well, as long as the living sample is restricted to normal-weight (non-under-weight or obese) subjects, an issue that is returned to below. A novel estimation method based on pubis length is introduced, with promising results. More work with male samples and samples of varying body shape will help to establish the general efficacy of this method.

3. Morphometric methods

A different approach to body mass estimation involves attempts to reconstruct body size and shape more directly from preserved or estimated skeletal dimensions. This approach has been termed "morphometric", to distinguish it from more "mechanically" based methods that rely on mechanical loading of a skeletal feature, such as an articulation (Auerbach and Ruff, 2004). Stature and bi-iliac (maximum pelvic) breadth are two dimensions that can be estimated or measured from skeletal remains, and have been shown to predict body mass in modern humans of diverse body size and shape with relatively high accuracy, i.e., little bias, again as long as the modern sample is restricted to normal-weight individuals (Ruff et al., 2018; also see below). Although this method has been applied (with modifications) to *Australopithecus afarensis* A.L. 288-1 ("Lucy") (Ruff, 1998), it has most commonly been used with material assigned to *Homo*, because of the greater similarity in general body form with modern humans and thus fewer assumptions necessary (Arsuaga et al., 1999; Rosenberg et al., 2006; Ruff et al., 2006).

Two studies in this special issue (Walker et al., 2018; Junno et al., 2018) test the stature/bi-iliac method in samples of modern humans. Walker et al. (2018) use anthropometric data for a worldwide sample of 6–12 year-olds to examine how well the equations

work in systematically small-bodied humans (i.e., children), which is particularly relevant to paleoanthropology given the small body size of many early hominins. Prediction errors vary by age, which can be attributed to changes in relative limb length and possibly muscularity during growth. Overall, except for the youngest age group, the adult equations work relatively well, particularly the equation based on adult females, who are smaller on average than adult males (using the female formula, 75% or more of predicted values fall within ±20% of true values). In addition to extending the testing of the method into a lower body mass range, this study demonstrates the impact of body proportions, especially relative lower limb length and shoulder breadth, on prediction accuracy using this method, issues that were also apparent in earlier studies among adults (Ruff, 2000; Schaffer, 2016). Estimating such body proportions from skeletal material remains a challenge. Junno et al. (2018) examine the effects of age (among adults) and body composition on body mass estimation errors using the stature/bi-iliac method. All relevant skeletal and anthropometric variables (including relative lean and fat mass) change throughout adulthood, with systematic effects on body mass estimation errors. However, average estimation errors remain low (within about 7%) in younger, relatively lean adults (until the sixth decade), who are likely most representative of earlier human populations. Application of the formulae to modern forensic samples, which tend to be older and relatively heavier, would be more problematic.

Brassey et al. (2018) also apply a morphometric approach to body mass estimation in hominins, but use a completely different methodology based on convex hull modeling. In this method, the outer bounds of a reconstructed complete skeleton are used to estimate the total "shrink-wrapped" volume of the specimen, which is then converted to actual body mass using relationships between the two in extant specimens. The method is applied to A.L. 288-1, with a resulting best estimate of 20.4 kg (95% prediction interval 13.5–30.9 kg), which is lower than almost all previous estimates. Various reasons for this difference are discussed, including limitations of the modern extant sample used to develop the conversion factor, difficulties in precisely reconstructing the complete skeleton of A.L. 288-1, and possible overestimation of body mass in previous studies. More work is needed to further validate the approach, ideally on modern humans and nonhuman specimens of diverse size, shape, and body composition. It is also applicable only to relatively complete skeletons. It would be interesting to try the method on other, non-hominin fossil primate specimens that fulfill this criterion.

4. Which reference groups?

Since body mass can never be directly observed in past specimens, its estimation always involves use of a modern reference sample of known body mass. Choice of the most appropriate reference sample can be influenced by a number of factors, several of which were addressed in special issue papers.

4.1. Matching of size range between reference sample and target specimens

Extrapolation beyond the size range of a reference sample is potentially dangerous on both statistical and biological grounds. This problem becomes particularly acute when target specimens fall close to or outside the size range of the most taxonomically or functionally relevant extant groups. Dagosto et al. (2018) address this issue when estimating the body mass of two very small early Eocene primates (*Teilhardina* and *Archicebus*) by including a variety of modern mammalian reference samples, encompassing both a wide range of body masses and ranges restricted to small-bodied

taxa. Their results indicate that in this case the choice of reference sample, including size range, has less of an effect on estimated body masses of the fossils than the choice of skeletal predictor variable. However, the two factors are interrelated — the choice of reference sample is less important with the best performing predictor variables (distal femoral mediolateral width and calcaneo-cuboid facet area of the calcaneus) and has more of an effect with the more poorly performing variables. This implies that the functional relationship between the best performing variables and body mass is more constant across taxa and across size ranges.

Body mass estimation in early hominins is another example where extrapolation beyond the size range of the (arguably) most appropriate modern reference sample — humans — is an issue. Walker et al.'s (2018) creative approach of using modern children to test the applicability of the stature/bi-iliac method was designed with this in mind, with average body masses in their sample extending down into the range of estimates for even the smallest-bodied early hominins. The potential concerns of using a juvenile test sample, including systematic differences in body proportions from those of adults, are also illustrated in this study (see below on juvenile body mass estimation in general).

The very large (by modern standards) body size of two Pleistocene platyrrhines, *Caipora bambuiorum* and *Cartelles coimbrafilhoi*, well outside the range of living New World monkeys, prompted the inclusion of living hominoids in some analyses of these specimens (Perry et al., 2018). This had little effect on estimates derived from the distal femur, for reasons discussed earlier (i.e., similarity in scaling of this dimension across broad taxonomic/functional groupings). Thus, the performance of some dimensions as estimators may be less affected by extrapolation than others. The general similarity in allometric scaling of several articular breadths to body mass in older juvenile and adult nonhuman hominoids reported by Burgess et al. (2018) also suggests that moderate extrapolation may be warranted in certain situations — in this case, when based on articular proportions within similar taxonomic/functional groupings.

4.2. Taxonomic/functional similarity

The issue of phylogenetic effects on body mass estimation is also addressed by Dagosto et al. (2018). They find little effect of phylogenetic correction on their estimates, although again this varied by the predictor variable, with the better predictors showing less effect. These variables may thus carry less "phylogenetic signal", which could be a result of a more constant functional relationship to body mass across diverse taxa. No other study in this special issue explicitly incorporates phylogenetic models in analyses, although both Perry et al. (2018) and Burgess et al. (2018) compare scaling of dimensions to body mass in different taxonomic groupings. Since these largely correspond to groups that also differ functionally (e.g., greater proximal limb joint excursion in hominoids versus cercopithecoids), it is difficult to distinguish between the two effects. The relatively simple dimensions included here — maximum articular breadths in a single plane — may also limit the ability to pick up either phylogenetic or functional signals that may be apparent in more complex morphological assessments, e.g., of full articular surface area distributions (MacLatchy and Bossert, 1996; Ruff, 2002). This could also be a benefit, however, in the context of body mass estimation, by limiting the influence of potentially confounding variables.

Whether only modern human or a combination of human and nonhuman hominoid reference samples should be used in estimating the body mass of early hominins is a long-standing issue that depends in large part on whether early hominins are viewed as fully modern or intermediate in locomotor behavior between

humans and apes (Jungers, 1988; McHenry, 1991, 1992; Nakatsukasa et al., 2007). Applying a femoral head breadth equation derived from African apes (Burgess et al., 2018) to the *Ardipithecus ramidus* specimen ARA-VP-6/50 results in a body mass estimate very similar to one estimated using a modern human reference sample, if the fossil's femoral head size is adjusted as recommended based on observed lower limb articular proportions (Ruff et al., 2018). This does not necessarily mean that the fossil mechanically loaded its femoral head in a manner similar to that of modern African apes, however. Other functional explanations for relatively reduced hip joint size in early hominins have been proposed (Lovejoy et al., 1973; Ruff, 1998). The functional context of any body mass predictor — in both reference and target samples — is critical in evaluating its applicability and probable accuracy (Damuth and MacFadden, 1990b). Both Perry et al. (2018) and Burgess et al. (2018) discuss the effects of locomotor differences between broader taxonomic groups on the scaling of skeletal dimensions to body mass, with implications for body mass prediction.

4.3. Body composition

Taxonomic or functional (e.g., locomotor) differences are not issues when applying body mass estimation equations within the same taxon, of course. However, issues regarding the most appropriate reference samples can still be important even in this context. In particular, variation among recent humans in body composition can have significant effects on the development and evaluation of body mass estimation equations.

As reviewed in Ruff et al. (2018), a number of studies have used human cadaveric specimens to develop estimation equations, although it has long been known that body weights associated with cadaveric material can be problematic, due to the effects of illness prior to death and variation in post-mortem conditions. Ruff et al. (2018) show that typical specimens derived from such collections have quite variable and overall very low body mass indices (BMI, weight for height^2), and that preselecting against the lowest BMI (i.e., most emaciated) individuals does not correct for this effect. Recent equations developed from such samples therefore produce much lower body mass estimates. A different problem arises when using modern Western living populations to develop or test equations. Such populations tend to be overweight relative to earlier or non-Western populations, part of a recent "epidemic" of obesity (Hill and Peters, 1998). Thus, equations based on samples from these populations produce body mass estimates that are likely to be too high when applied to earlier, leaner populations (e.g., Ruff et al., 1991). Controlling for BMI or other measures of relative fatness is therefore critical when selecting living reference samples (Lorkiewicz-Muszynska et al., 2013; Squyres and Ruff, 2015; Schaffer, 2016; Ruff et al., 2018). Such equations could be appropriate, however, for actually estimating body mass in similar populations, e.g., for forensic purposes.

Two other studies included in the special issue also examine the effects of body composition on body mass estimation in humans. As noted above, Junno et al. (2018) compare body masses estimated using the morphometric stature/bi-iliac technique (based on a world-wide sample of generally lean individuals — see Ruff et al., 2018) with actual and theoretically derived body masses of individuals with different relative lean and fat masses. The technique produces reasonable results when applied to leaner populations, but not to living Western populations, where body mass is underestimated. This is similar to results obtained in other applications to similar populations (Elliott et al., 2016). Thus, the proposed target sample must be considered when evaluating such equations. Young et al. (2018) emphasize the same point in their study, noting that

researchers should select equations that are most likely to be accurate for a particular population or specimen. Previously developed estimation equations based on the femoral head generally show the greatest accuracy in living individuals with "normal" BMIs (18.5–24.9), except for the unmodified equation from Ruff et al. (1991), which the original authors argued should be adjusted downwards because of the greater adiposity of the sample. This unmodified equation worked well for "overweight" individuals, however (BMI 25–29.9), as would be predicted.

The issue of body composition is also relevant to the study by Brassey et al. (2018). They test their equation for converting convex hull estimated body mass to true body mass on a series of nonhuman primate carcasses and living human specimens. The humans vary in BMI from 21.4 to 31.7, i.e., from "normal" to "obese". The body masses of individuals with higher BMIs are progressively more underestimated using their technique, as would be expected. However, even individuals in the "normal" range are somewhat underestimated (by an average of 9%). Two *Macaca fuscata* are underestimated by 27–29%. The authors point out that the macaques, which were captive, may have been obese, although their body masses (4.5 and 10.2 kg, both females) bracket the average given in the literature (8 kg; Smith and Jungers, 1997). As noted earlier, more work clearly needs to be carried out to further verify this method on individuals of varying body composition.

5. Juvenile body mass estimation

Although a number of previous studies have examined ontogenetic scaling of primate limb bones (e.g., Jungers and Susman, 1984), very few have attempted to develop body mass prediction equations for juveniles (Hartwig-Scherer and Martin, 1992; Ruff, 2007). Such equations are useful in a variety of contexts: many important paleontological specimens are juveniles (Brown et al., 1985), studies of growth and development in past populations rely on accurate body size prediction (Ruff et al., 2013), and forensic cases often involve juveniles (Sciulli and Blatt, 2008).

Three studies in the special issue included estimation of body mass in juveniles — in nonhuman hominoids (Burgess et al., 2018), human archaeological samples (Cowgill, 2018), and a living human anthropometric sample (Walker et al., 2018). The generally similar allometric scaling of long bone articulations, within taxon, to body mass in adults and juveniles in Burgess et al.'s (2018) study suggests that adult estimation formulae can be applied to older juveniles (with well-formed epiphyses), and that older juveniles and adults can also be combined to increase the size of reference samples. The similar scaling is interesting in that a previous study of humans (Ruff, 2007) showed articulations to be relatively large in juveniles, and thus that different formulae were needed until late adolescence. Whether this indicates a different growth pattern in humans or is a product of small sample size in some of the nonhuman primate juvenile samples requires further testing. Body mass estimation equations based on long bone metaphyseal rather than epiphyseal breadths work well in both nonhuman hominoids (Burgess et al., 2018) and humans (Ruff, 2007). Metaphyseal breadths avoid some of the potential problems in measuring partially formed epiphyses and are applicable to very young juveniles without epiphyses; they are therefore generally preferable as estimators.

Walker et al.'s (2018) study shows that adult stature/bi-iliac formulae for predicting body mass can be applied to children, although systematic differences in body proportions and probably composition between adults and children increase estimation error, especially in young children. In any event, the technique is likely to be applicable to skeletal remains of only older adolescents with fused hip bones. Equations for individuals in this age range

have been presented, but using femoral length rather than stature (to avoid estimation of stature in juveniles; Ruff, 2007).

Cowgill's (2018) study was not designed to test the accuracy of human body mass prediction equations, since true body mass was not known in her juvenile archaeological samples. However, her results indicate that caution should be exercised in applying equations to target samples that may differ substantially from reference samples in environmental conditions during growth. This effect is likely to be more severe for skeletal parameters that exhibit more growth plasticity, e.g., diaphyseal breadths.

6. Statistical issues

Body masses of fossil specimens are often estimated using prediction equations based on extant species means, in part because individually associated body masses for extant skeletal specimens can be relatively rare for some taxa, and because taxonomic breadth in the reference sample may be desired, particularly for fossils without clear phylogenetic affinities to modern taxa. How estimation error should be assessed in such situations has been a matter of discussion (Radinsky, 1982; Van Valkenburgh, 1990; Smith, 2002; Ruff, 2003). Yapuncich (2018) compares the use of two kinds of percent prediction errors (%PE) generated from application of an equation to a novel (test) sample: that based on the mean error of multiple individual predictions of body mass, and that based on the error of mean body mass calculated from the mean of the predictor variable. Molar length or area is used as the predictor variable in several previously available equations (derived from species means), and tested using a cercopithecoid sample. Based on the results of this analysis, Yapuncich (2018) concludes that more accurate estimates of the mean body mass of target samples can be derived using the mean of predictor variables, rather than averaging the estimates for individuals. Conversely, equations based on individuals can be used on individual target specimens. Trade-offs between increasing sample size for target specimens and increasing precision by choosing better (but more rarely preserved) dental/skeletal predictors are also discussed.

Among the studies in the special issue that included more than one species, Dagosto et al. (2018) used species means, whereas Perry et al. (2018) and Burgess et al. (2018) used individually associated body masses to generate prediction equations. Perry et al. (2018) applied these equations to individual fossil specimens, and then averaged results for taxa, as recommended given the above results. Dagosto et al. (2018) applied their equations to mean predictor values of the fossil taxa, again as recommended (although fossil sample sizes were unavoidably very low).

Methods of line-fitting for regression analyses is another issue that has been discussed at length in the literature with reference to body size prediction (Smith, 1994, 2009; Konigsberg et al., 1998; Hens et al., 2000; Ruff, 2007; Squyres and Ruff, 2015). Least Squares (LS) (or Ordinary Least Squares, or Inverse Calibration) has usually been employed for predictive purposes when the reference sample and target sample are similar in size distribution, because under these circumstances it minimizes estimation error (Hens et al., 2000). Conversely, Reduced Major Axis (RMA) regression is often used to examine scaling relationships, because it incorporates reasonable assumptions regarding error (biological and statistical) in both variables (Aiello, 1992) and does not assume a causal relationship (i.e., it is "symmetric"; Smith, 2009). Difficulties arise in using LS regression for prediction near or outside the data limits of the reference sample, however (Sjøvold, 1990; Konigsberg et al., 1998). In this case, Model II regression techniques such as RMA are useful, as long as extrapolation is not too great (Konigsberg et al., 1998).

Among the papers in the special issue that develop or apply regression equations for predicting body mass, most use LS for this purpose, although two studies employ RMA (Ruff et al., 2018; Niskanen et al., 2018), and RMA regression is also used in several studies to study scaling relationships between body mass and skeletal variables (Perry et al., 2018; Burgess et al., 2018; Ruff et al., 2018). In Ruff et al. (2018) and Niskanen et al. (2018), the RMA equations are derived for relatively large samples of individual humans spanning a limited size range (when compared with broader interspecies analyses). In this situation, correlations tend to be relatively low and LS line fits are more likely to significantly overestimate body mass for smaller individuals and underestimate body mass in larger individuals. In Ruff et al. (2018), target samples include very small-bodied early hominins with predictor variables below the size range of the modern reference sample, as well as some more recent fossil specimens near the upper end of the modern sample distribution; thus, RMA is argued to be the better choice for prediction. A special circumstance arises when multiple regression is used to predict body mass, as in the case of the stature/bi-iliac technique. Here, Model II methods are not as readily available, so all studies utilizing this technique employ LS regressions. However, body mass predictions using the technique do not appear to suffer from statistical biases near the ends of the reference sample distribution (Ruff et al., 2018), or even well below it (Walker et al., 2018).

7. Conclusions

Body mass estimation from skeletal or dental remains is subject to a number of limitations, both technical and interpretive (Smith, 1996): statistical rather than functionally validated relationships may be used to generate estimation equations, confidence limits for individual specimens are often wide, and general scaling relationships may obscure biologically meaningful variation. The papers in this special issue take some steps to address these issues, by examining the effects of choice of body mass estimator, including both "mechanical" and "morphometric" approaches, exploring the effects of using different reference samples, examining phylogenetic and developmental influences on estimation, investigating allometric (i.e., "size"-related) effects, and studying the effects of variation in body composition, a component of body mass estimation not often considered in such studies. Some of these analyses require living (or recently deceased) samples, so forensic or anthropometric data are relevant, even in paleoanthropological contexts. Broader interspecies analyses demonstrate relationships between taxonomic or locomotor differences and the scaling of predictor variables. All levels of analysis contribute to a better understanding of factors influencing the accuracy of body mass estimation and the choice of methodology to employ.

One other contribution that these papers make to the field is to provide large data sets included as supplementary tables. Species mean body masses for 98 euarchontan taxa are given by Dagosto et al. (2018). Burgess et al. (2018) provide body mass and articular breadth data on 157 wild-collected hominoid individuals, both adult and immature. Perry et al. (2018) give individual body masses and articular dimensions for 204 wild-collected catarrhines and platyrrhines; the platyrrhine data are unique to this study. It is hoped that these data will be useful in a variety of future investigations.

Body mass estimation will always be associated with significant error, given the lability of body mass itself in living animals and its imperfect association with skeletal or dental size. Thus, there are limits to estimation precision for individual specimens, with mean percent prediction errors of 10–15% for even the best estimators (Burgess et al., 2018; Perry et al., 2018; Ruff et al., 2018; also see

Ruff, 2003; Dagosto et al., 2018). This is why target sample size is so important in such analyses: the probability of deriving a true estimate of mean taxon body mass is in part dependent on the number of individuals included in the analysis (Yapuncich, 2018). Ensuring accuracy, i.e., non-biased estimates, is perhaps even more important. In this context, it is instructive to review the comments of Damuth and MacFadden (1990b: 4, italics original) in their summary of a conference held 30 years ago:

> "The general theme that emerged at the Gainesville workshop, and that runs throughout the book, is that *body mass and functional morphological interpretation are not separable ….* The different ways the same anatomical element may be related to overall mass in different species can to some extent be predicted from functional considerations."

The fallacy of using statistical criteria alone to evaluate the accuracy and usefulness of body mass predictive equations, without understanding functional relationships between predictor variables and body mass, was also highlighted by Smith (1996). This involves consideration of functional morphology in both the reference and target samples, which can be difficult with fragmentary and rare fossil material. Perry et al. (2018) make one attempt to match reference and target samples functionally by comparing intra- and inter-limb proportions, which are related to locomotor behavior among extant taxa. A similar approach is used to adjust estimation equations for early hominins by comparing hip to knee breadth proportions among hominins and African apes (Ruff et al., 2018). Functional considerations may explain why knee transverse breadth shows a relatively invariant relationship with body mass in broad inter-taxon comparisons (Ruff, 2003; Dagosto et al., 2018; Perry et al., 2018). Consideration of developmental constraints is also important in evaluating how accurate prediction equations based on different skeletal parameters are likely to be when applied to other samples (Cowgill, 2018). Continued investigation of the physiological relationship between body mass and skeletal dimensions under a variety of conditions is necessary to optimize development and use of predictive equations.

References

Aiello, L.C., 1992. Allometry and the analysis of size and shape in human evolution. J. Hum. Evol. 22, 127—148.

Arsuaga, J.-L., Lorenzo, C., Carretero, J.-M., Gracia, A., Martínez, I., Garcia, N., Bermúdez de Castro, J.-M., Carbonell, E., 1999. A complete human pelvis from the Middle Pleistocene of Spain. Nature 399, 255—258.

Arsuaga, J.L., Carretero, J.M., Lorenzo, C., Gomez-Olivencia, A., Pablos, A., Rodríguez, L., Garcia-Gonzalez, R., Bonmatí, A., Quam, R.M., Pantoja-Pérez, A., Martínez, I., Aranburu, A., Gracia-Tellez, A., Poza-Rey, E., Sala, N., Garcia, N., Alcazar de Velasco, A., Cuenca-Bescos, 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.

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.

Brassey, C.A., O'Mahoney, T.G., Chamberlain, A.T., Sellers, W.I., 2018. A volumetric technique for fossil body mass estimation applied to *Australopithecus afarensis*. J. Hum. Evol. 115, 47—64.

Brown, F., Harris, J., Leakey, R., Walker, A., 1985. Early *Homo erectus* skeleton from West Lake Turkana, Kenya. Nature 316, 788—792.

Burgess, M.L., McFarlin, S.C., Mudakikwa, A., Cranfield, M.R., Ruff, C.B., 2018. Body mass estimation in hominoids: age and locomotor effects. J. Hum. Evol. 115, 36—46.

Calder III, W.A., 1984. Size, Function, and Life History. Cambridge Univ. Press, Cambridge.

Churchill, S.E., Berger, L.R., Hartstone-Rose, A., Zondo, B.H., 2012. Body size in African Middle Pleistocene *Homo*. In: Reynolds, S.C., Gallagher, A. (Eds.), African Genesis: Perspectives on Hominin Evolution. Cambridge Univ. Press, pp. 319—346.

Cowgill, L., 2018. Juvenile body mass estimation: A methodological evaluation. J. Hum. Evol. 115, 78—84.

Dagosto, M., Gebo, D., Ni, X., Smith, T., 2018. Estimating body size in early primates: the case of *Archicebus* and *Teilhardina*. J. Hum. Evol. 115, 8—19.

Damuth, J., MacFadden, B.J., 1990a. Body Size in Mammalian Paleobiology: Estimation and Biological Implications. Cambridge Univ. Press, Cambridge.

Damuth, J., MacFadden, B.J., 1990b. Introduction: Body size and its estimation. In: Damuth, J., MacFadden, B.J. (Eds.), Body Size in Mammalian Paleobiology: Estimation and Biological Implications. Cambridge Univ. Press, Cambridge, pp. 1—10.

Dye, S.F., 1987. An evolutionary perspective of the knee. J. Bone Jt. Surg. 69A, 976—983.

Elliott, M., Kurki, H., Weston, D.A., Collard, M., 2016. 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.

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.

Hammond, A.S., 2014. In vivo baseline measurements of hip joint range of motion in suspensory and nonsuspensory anthropoids. Am. J. Phys. Anthropol. 153, 417—434.

Hartwig-Scherer, S., Martin, R.D., 1992. Allometry and prediction in hominoids: A solution to the problem of intervening variables. Am. J. Phys. Anthropol. 88, 37—58.

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.

Hill, J.O., Peters, J.C., 1998. Environmental contributions to the obesity epidemic. Science 280, 1371—1374.

Holliday, T.W., 2012. Body size, body shape, and the circumscription of the genus *Homo*. Curr. Anthropol. 53(Supp. 6), S330—S345.

Jenkins, F.A., Camazine, S.M., 1977. Hip structure and locomotion in ambulatory and cursorial carnivores. J. Zool. Lond. 181, 351—370.

Jungers, W.L., 1985. Size and Scaling in Primate Biology. Plenum Press, New York.

Jungers, W.L., 1988. New estimates of body size in australopithecines. In: Grine, F.E. (Ed.), Evolutionary History of the "Robust" Australopithecines. Aldine de Gruyter, New York, pp. 115—125.

Jungers, W.L., 1991. A pygmy perspective on body size and shape in *Australopithcus afarensis* (AL 288-1, "Lucy"). In: Coppens, Y., Senut, B. (Eds.), Origine(s) de la bipédie chez les hominidés. Centre National de la Recherche Scientifique, Paris, pp. 215—224.

Jungers, W.L., Susman, R.L., 1984. Body size and skeletal allometry in African apes. In: Susman, R.L. (Ed.), The Pygmy Chimpanzee: Evolutionary Morphology and Behavior. Plenum Press, New York, pp. 131—177.

Junno, J.A., Niskanen, M., Maijanen, H., Holt, B., Sládek, V., Niinimäki, S., Berner, M., 2018. The effect of age and body composition on body mass estimation of males using the stature/bi-iliac method. J. Hum. Evol. 115, 122—129.

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. Yearbook Phys. Anthopol. 41, 65—92.

Lieberman, D.E., Devlin, M.J., Pearson, O.M., 2001. Articular area responses to mechanical loading: Effects of exercise, age, and skeletal location. Am. J. Phys. Anthropol. 116, 266—277.

Lorkiewicz-Muszynska, D., Przystanska, A., Kociemba, W., Sroka, A., Rewekant, A., Zaba, C., Paprzycki, W., 2013. Body mass estimation in modern population using anthropometric measurements from computed tomography. Forensic Sci. Int. 231(405), e401—406.

Lovejoy, C.O., Heiple, K.G., Burstein, A.H., 1973. The gait of *Australopithecus*. Am. J. Phys. Anthropol. 38, 757—780.

MacLatchy, L.M., Bossert, W.H., 1996. An analysis of the articular surface distribution of the femoral head and acetabulum in anthropoids, with implications for hip function in Miocene hominoids. J. Hum. Evol. 31, 425—453.

McCammon, R.W., 1970. Human Growth and Development. Charles C. Thomas, Springfield, IL.

McHenry, H.M., 1975. Biomechanical interpretation of the early hominid hip. J. Hum. Evol. 4, 343—355.

McHenry, H.M., 1991. Petite bodies of the "robust" australopithecines. Am. J. Phys. Anthropol. 86, 445—454.

McHenry, H.M., 1992. Body size and proportions in early hominids. Am. J. Phys. Anthropol. 87, 407—431.

Nakatsukasa, M., Pickford, M., Egi, N., Senut, B., 2007. Femur length, body mass, and stature estimates of Orrorin tugenensis, a 6 Ma hominid from Kenya. Primates 48, 171—178.

Napier, J.R., 1964. The evolution of bipedal walking in the hominids. Arch. Biol. (Leige) 75(Suppl), 673—708.

Niskanen, M., Junno, J.A., Maijanen, H., Holt, B., Sládek, V., Berner, M., 2018. Can we refine body mass estimations based on femoral head breadth? J. Hum. Evol. 115, 112—121.

Perry, J.M.G., Cooke, S.B., Runestad Connour, J.A., Burgess, L., Ruff, C.B., 2018. Articular scaling and body mass estimation in platyrrhines and catarrhines: modern variation and application to fossil anthropoids. J. Hum. Evol. 115, 20—35.

Plavcan, J.M., 2012. Body size, size variation, and sexual size dimorphism in early *Homo*. Curr. Anthropol. 53, S409—S423.

Radinsky, L., 1982. Some cautionary notes on making inferences about relative brain size. In: Armstrong, E., Falk, D. (Eds.), Primate Brain Evolution: Methods and Concepts. Plenum Press, Chicago, pp. 29—37.

Richmond, B.G., Jungers, W.L., 2008. *Orrorin tugenensis* femoral morphology and the evolution of hominin bipedalism. Science 319, 1662–1665.

Rosenberg, K.R., Lü, Z., 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.

Ruff, C.B., Burgess, M.L., Squyres, N., Junno, J.-A., Trinkaus, E., 2018. Lower limb articular scaling and body mass estimation in Pliocene and Pleistocene hominins. J. Hum. Evol. 115, 85–111.

Ruff, C.B., 1998. Evolution of the hominid hip. In: Strasser, E., Rosenberger, A., McHenry, H., Fleagle, J. (Eds.), Primate Locomotion: Recent Advances. Plenum Press, Davis, CA, pp. 449–469.

Ruff, C.B., 2000. Body mass prediction from skeletal frame size in elite athletes. Am. J. Phys. Anthropol. 113, 507–517.

Ruff, C.B., 2002. Long bone articular and diaphyseal structure in Old World monkeys and apes, I: Locomotor effects. Am. J. Phys. Anthropol. 119, 305–342.

Ruff, C.B., 2003. Long bone articular and diaphyseal structure in Old World monkeys and apes, II: Estimation of body mass. Am. J. Phys. Anthropol. 120, 16–37.

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

Ruff, C.B., Burgess, M.L., Ketcham, R.A., Kappelman, J., 2016. Limb bone structural proportions and locomotor behavior in A.L. 288-1 ("Lucy"). PLoS One 11, e0166095.

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

Ruff, C.B., Holt, B.M., Sladek, V., Berner, M., Murphy Jr., W.A., zur Nedden, D., Seidler, H., Recheis, W., 2006. Body size, body proportions, and mobility in the Tyrolean "Iceman". J. Hum. Evol. 51, 91–101.

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., Holliday, T.W., 1997. Body mass and encephalization in Pleistocene *Homo*. Nature 387, 173–176.

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.

Schmidt-Nielson, K., 1984. Scaling: Why is Animal Size So Important? Cambridge Univ. Press, Cambridge.

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

Sjøvold, T., 1990. Estimation of stature from long bones utilizing the line of organic correlation. Hum. Evol. 5, 431–447.

Smith, R.J., 1994. Regression models for prediction equations. J. Hum. Evol. 26, 239–244.

Smith, R.J., 1996. Biology and body size in human evolution. Statistical inference misapplied. Current Anthropology 37, 451–481.

Smith, R.J., 2002. Estimation of body mass in paleontology. J. Hum. Evol. 43, 271–287.

Smith, R.J., 2009. Use and misuse of the reduced major axis for line-fitting. Am. J. Phys. Anthropol. 140, 476–486.

Smith, R.J., Jungers, W.L., 1997. Body mass in comparative primatology. J. Hum. Evol. 32, 523–559.

Squyres, N., Ruff, C.B., 2015. Body mass estimation from knee breadth, with application to early hominins. Am. J. Phys. Anthropol. 158, 198–208.

Steudel-Numbers, K.L., Tilkens, M.J., 2004. The effect of lower limb length on the energetic cost of locomotion: implications for fossil hominins. J. Hum. Evol. 47, 95–109.

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.

Trinkaus, E., Ruff, C.B., 2012. Femoral and tibial diaphyseal cross-sectional geometry in Pleistocene *Homo*. PaleoAnthropology 2012, 13–62.

Van Valkenburgh, B., 1990. Skeletal and dental predictors of body mass in carnivores. In: Damuth, J., McFadden, B.J. (Eds.), Body Size in Mammalian Paleobiology. Cambridge Univ. Press, Cambridge, pp. 181–205.

Walker, C.S., Yapuncich, G.S., Sridhar, S., Cameron, N., Churchill, S.E., 2018. Evaluating morphometric body mass prediction equations with a juvenile human test sample: accuracy and applicability to small-bodied hominins. J. Hum. Evol. 115, 65–77.

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.

Young, M., Johannesdottir, F., Poole, K., Shaw, C., 2018. Assessing the accuracy of body mass estimation equations from pelvic and femoral variables among modern British women of known mass. J. Hum. Evol. 115, 130–139.