Prominent methods and theories in the estimation of body mass from skeletal remains

by Kyra O'Neill

Estimating body mass from skeletal remains is considered a gap in the creation of a biological profile. Over the last few decades, there have been attempts to fill this gap using different elements from the skeleton. Using various academic databases, a study was done to investigate the prominent methods and theories in body mass estimation. These methods include the use of the femur, the articular surfaces, and the interpretation of musculoskeletal stress markers at the entheses. Calculations using the femur found success in adults most prominently when the cortical area is used. The cortical area provided a percent error margin of 14–22%, with the error decreasing when sex and ancestry-specific equations were used. Musculoskeletal stress markers correlated with heavier body mass in various regions when looking at robusticity. However, these results could not be distinguished between higher body mass individuals and athletic individuals. The articular surface area exhibited no change when body mass is considered, although other features such as osteoarthritis can potentially provide insight into body mass. In addition, subadult femurs were investigated and provided error percentages of 5–7% for juveniles 7 years and younger, and the bi-iliac breadth with long bones can be used for those 15–17 years old with an error margin of 5–8%. These methods exhibit limitations in the demographics of the study, the lack of weight extremely investigated, and various confounding factors. However, these methods and theories in body mass estimations from skeletal remains provide a promising start.

One of the primary goals of the forensic anthropologist is to provide a biological profile of skeletonized remains to law enforcement. This is important for the positive identification of unknown individuals. The traditional aspects of this biological profile tend to include estimations of sex, age-at-death, stature, and ancestry (St. George 2015). These techniques have been tested in numerous studies and have been used in the field for decades. Yet, it has been suggested that there is a missing element in these profiles and forensic death investigations: body mass (Moore 2008). As a response to this hypothetical gap, many researchers have been investigating how to estimate body mass from skeletal remains. This paper aims to explore the extent to which body mass can be inferred from skeletal remains, through a literature review exploring the major methods and theories of body mass estimation in existence to date. These methods are based on the femur, entheses, and bony articular surfaces. It will also present some methods for assessing subadult body mass, and will conclude with a discussion of the overarching critiques and problems in estimating body mass. Future avenues of study will be presented as potential solutions.

Kyra O'Neill is a BA alum of the University of Alberta (Department of Anthropology, 13–15 HM Tory Building, Edmonton, Alberta, T6G 2H4; Campus Saint-Jean, 8406 Rue Marie-Anne Gaboury, Edmonton, Alberta, T6C 4G9 [keoneill@ualberta.ca]).

Methodology

For this paper, a literature search for medical, forensic and biological anthropology journal publications was carried out using online databases, including the Wiley Online Library, the University of Alberta online library database, ScienceDirect and JSTOR. The keywords to search the databases included a combination of terms such as: "body mass," "skeletal remains," "physical activity," "skeletal stress markers," "weight versus activity," "musculoskeletal stress markers," and "weight estimation." The results were then reviewed for pertinence in the current field of forensic anthropology, hinging on factors such as efficacy or impact of the method in the field, and a modern study group. This aims to provide a comprehensive, though not exhaustive, overview of the prominent theories and methods around body mass estimation from skeletal remains.

Femur

The femur is one of the most important bones in constructing a biological profile and is arguably the human skeleton's most analyzed long bone (İşcan 2005). It has been used for sex estimation and stature reconstruction methods with reasonable success (Byers 2016). It has also been considered as an indicator of body mass. The logic behind this is straightforward: just as it contributes to stature, the femur makes up a portion of the body weight as well as anchoring several strong and potentially heavy muscles (such as the gluteus muscles). Furthermore, as one of the limbs of the lower body, it is responsible for a significant amount of load-bearing. As such, it is reasonable to assume that if body mass has any impact on the skeleton, the femur would be a likely element to show a response to weight.

Ruff (1988) conducted a study on both the hindlimb articular surface dimensions and diaphyseal scaling, in relation to body mass for Hominoidea and *Macaca*. It was found that modern humans showed a positive allometric relationship with body mass and femoral head diameter. This led to another study by Ruff, Scott, and Liu (1991) using proximal femoral dimensions to estimate body weight. This study had a few hypotheses: first, the femoral diaphyseal cross-sectional size would be highly correlated with current body weight. Second, the femoral head size would be highly correlated with body weight at 18 years old, as the femoral head is essentially fixed around that time. Third, the femoral neck would show an intermediate correlation.

Methods

Ruff, Scott, and Liu (1991) conducted a study with the use of radiographs on living subjects whose current weight and weight at 18 years of age was known. To conduct this study, 80 outpatients from John Hopkins Hospital were chosen. The ages of the group were from 24 to 81 years, with roughly half females and half males, and two-thirds White, and one-third Black. Measurements were taken from the radiographs for the superior-inferior head (A) and neck breadths (B), and mediolateral subperiosteal and cortical breadths (C) (used to calculate two cross-sectional geometric properties: cortical area and the second moment of area in the mediolateral plane) of the proximal diaphysis. These measurements can be seen in Figure 1.

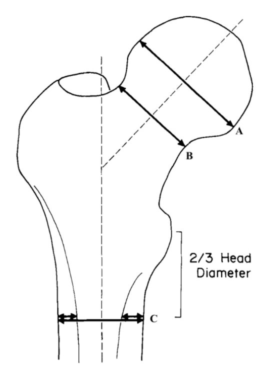

Figure 1. Outline of proximal femur measurements. Reproduced with permission from Ruff et al. (1991).

This was done using Helios dial calipers to 0.1mm. If both hips could be measured, the average was taken for the analysis. If both hips could not be measured, whichever one was available would be used. The current and past body weights of the subjects were determined by patient recall through questioning by the attending physician. The analysis was done by comparing correlations between the femoral dimensions and the current and 18-year-old weights, and creating least square regression equations to estimate body mass. The analysis of the correlations was divided into groups: total, White male, White female, White, and Black. There was no sex division among the Black individuals as the category was too small to divide and be statistically significant.

Results

The results of Ruff, Scott and Liu (1991) were concurrent with their hypothesis. The shaft dimensions were correlated with current body weight, while the femoral head breadth was more highly correlated with body weight at 18 years of age, and the femoral neck breadth fell between the two. The only deviation from this involved the female subgroup which showed low correlations of head breadth with former body weight. The equations were found to be more accurate with raw data than log-transformed data and presented a percent standard error range of roughly 14–25% depending on the chosen area of measurement and subgroup. The error was found to be lowest when using the femoral shaft breadth of cortical area index, lower for White, and lower for males.

These equations were then tested for accuracy on an independent sample of eight individuals not included in the base study. The body weight predictions were the most accurate using the femoral cortical area, less accurate using the femoral shaft breadth, and the least accurate using the femoral head breadth. It was generally found that using race- and sex-specific equations slightly improves prediction, but never by more than 3% for the mean absolute percent error. The mean absolute percent errors showed a 10–13% error when using the femoral cortical area, a 16% error for shaft breadth, and a 17–20% error for head breadth. These decreased slightly upon the removal of one subject, a female with a very high body mass. However, the estimations for this particular individual noted interesting findings. The femoral head and shaft breadth equations consistently underestimated her weight by 50% or more, but the femoral cortical area equations did fall within 12%, and with the sex-specific equation, within 1%. The authors remarked that this individual had more than doubled her weight since the age of 18 and it was the shaft cortical thickness that increased in response. They concluded by stating that the proximal femoral dimensions and body masses among adult humans have a scaling relationship (Ruff, Scott, and Liu 1991). This study demonstrated that the body weight of chronologically recent U.S. adults can be predicted with an average percent of error of 10–16% based on the independent test sample, which the authors deemed to be reasonably accurate.

This article has inspired further investigation of the bones of the lower limbs, such as using the first metatarsal to calculate the femoral head size in instances of a missing or damaged femur. It has also been used independently in equations of body mass with a percent error of just over 7%; however, this result did not account for the cumulative error that exists as this test was done using previously published equations (De Groote and Humphrey 2011).

Critiques

Ruff, Scott, and Liu (1991) have had a large impact on the study of body mass estimation from skeletal remains. Despite that, there are still areas of improvement to consider in this study. Studies such as Lacoste Jeanson et al. (2017) and Niskanen et al. (2018) have evaluated the equations based on the femur for accuracy and have found them to have a higher percent error than originally suggested. However, both studies investigated the femoral head breath. It is unclear why this particular measurement was investigated as it was proven to be a less effective method of calculating body mass in the original study.

In addition, the body weights provided in the study were given by self-reporting through patient recall. Research has found that weight reports tend to have a gender-related bias with men overestimating their weights, and women underestimating them in the U.S. (Villaneuva 2001). Furthermore, some of these individuals were being asked to recall their weight from as much as 63 years later, which likely added to some error.

Entheses

Another avenue of inquiry in body mass estimation is the use of entheses, the location where muscles or ligaments attach to bone (Lieverse et al. 2009). Entheses are involved in the creation of musculoskeletal stress markers (MSMs) also referred to as occupational stress markers or entheseal changes (Lieverse et al. 2009; Byers 2016). Entheses leave imprints on bone, with more severe or complex imprints being associated with heavier or more frequent labour or activity (Byers 2016). Currently, researchers are investigating how heavier body masses might also produce MSMs. The theory suggests that individuals with larger body masses will exhibit greater MSMs due to the increased weight they carry. While analyzing the entheses, researchers tend to look for two things at muscle origins and insertions: robusticity and stress. Robusticity is the deposition of bone at the attachment site. This produces a rough and raised rugged marking on the bone that at the most extreme can present as a sharp ridge or crest. Stress is the resorption of bone at the attachment site, which can manifest as a depression or furrow in the cortex of the bone (Hawkey and Merbs 1995).

Based on this knowledge, Godde and Taylor (2011, 2013) conducted two studies. The former investigated the MSMs of the upper limbs and pectoral girdle (Godde and Taylor 2011), and the latter the lower limbs (Godde and Taylor 2013). These studies were primarily done to determine if obesity could be identified in unknown skeletal material. Both investigated the differences in MSMs across different activity levels and body mass indexes. At the time, very little research had been done regarding obesity and the skeleton.

Upper Limb and Pectoral Girdle Methods

Godde and Taylor (2011) hypothesized that there would be a difference in the patterning of MSMs based on weight and activity. They analyzed the remains of 108 modern White males from the William M. Bass Donated Skeletal Collection for their study. Each skeleton had documented information, including activity level, height, and weight. The study consisted of average-weight individuals (n = 50, BMI = 17–24), obese individuals (n = 50, BMI > 30), and individuals documented as being athletic or very active in their occupation (n = 8). The entheses chosen were those associated with the sit-to-stand transition to ensure they were used by all groups. They scored the MSMs of 25 entheses on a scale of 0–3 and input the results into a logistic regression analysis with separate analyses for robusticity markers, stress markers, and combined markers. These logistic regression equations were used to calculate the significance of variables and the percentage of correctly classified cases.

Upper Limb and Pectoral Girdle Results

The results indicated that 13 muscles had robusticity scores significantly associated with activity level and body mass, and two muscles had significant stress scores. When robusticity and stress were considered together, only nine muscles were statistically significant. The researchers investigated the accuracy of predicting BMI/activity level through classification percentages. Of the robusticity classification percentages, none of the muscles exceeded a 60% classification rate. Of the stress classification percentages, none of the muscles were classified above chance, or above 50%. However, when the researchers assessed them together, two muscle groups, the flexors and extensors of the humerus, exceeded a 60% classification rate. The researchers concluded that there is a significant correlation in the MSMs studied, with some muscles being better indicators of activity and body mass than others. They noted that this technique should only be looked at as a group along with other indicators of obesity in estimating an appropriate body mass. Nevertheless, the pattern was observed and showed promise in body mass estimation (Godde and Taylor, 2011).

Lower Limb Methods

Godde and Taylor (2013) hypothesized that both obese and athletic individuals would show hypertrophy of muscle attachment compared to average weight, but that the MSMs affected would be different. To conduct their study, 105 White males from the William M. Bass Donated Skeletal Collection from the ages of 31–81 years were used. The researchers looked at 18 muscle attachment sites of the lower limb. Within this group, they included individuals of average weight (n = 50, BMI = 18–29), obese (n = 50, BMI > 30), and athletic individuals (n = 5). Each set of remains had documented height, weight, age, sex and activity level. The height and weight were provided by the family or self-reported. The muscle attachment sites (16 insertions, two origins) for the study were used in the sit-to-stand transition. They were scored to assess robusticity and stress on a 0–3 scale. These scores were used in logistic regression analyses and linear log models to determine their results, and the researchers used an analysis of variance (ANOVA) to detect age bias.

Lower Limb Results

The results showed that robusticity markers in the lower limb were generally reliable predictors of both BMI and activity level, whereas the stress markers were not. In a normal vs obese analysis, the classifications ranged from 54–75%; in the normal vs. active they ranged from 51-88% (Godde and Taylor 2013). The gluteal muscles were classified very strongly with both groups, and the semimembranous was the strongest predictor in the normal vs. active subjects. There was no significant difference between obese and active individuals, and they noted that age differences were significant in robusticity scores for most entheses. Godde and Taylor (2013) conclude that entheses may be useful in identifying normal BMI individuals, but this technique should not be used by itself—rather, it could be beneficial with other indicators of obesity in estimating an appropriate body mass.

Additional research on this topic has been explored, such as by Myszka and Piontek (2011), who investigated the development of musculoskeletal stress markers and the length and circumference measurements of long bones on male and female skeletal remains from a medieval burial ground in Cedynia, Poland. Similar techniques as the previous two studies were used, and the researchers found that while individual muscle attachment sites may develop in relation to body form, a greater robusticity of the musculoskeletal system is seen in those with a more robust skeleton (larger circumference bones, and potentially larger weight) compared to those with a gracile skeleton.

Critiques

When conducting the second study, Godde and Taylor (2013) provided clarification on how the documented information about height, weight, age and activity was known (i.e., self-reported or family-provided data) and added the age demographic of their sample.

The use of the William M. Bass Donated Skeletal Collection for this study is an appropriate collection. The range of individuals in this collection currently includes the birth years up to 2016, with most being after the 1940s (University of Tennessee Knoxville Forensic Anthropology Center n.d.) This group would be reasonably chronologically representative of the population these techniques may be used on today. However, it is worth noting that this study was conducted only on White men. It is known that different ancestries and different sexes present varying degrees of general robusticity regardless of any other factors (Byers 2016). The study's application on only White men is not representative of the population today and these MSMs should be investigated with different reference groups before it is used with any certainty.

While the second study notes where documentation of height and weight was found (which are crucial for BMI calculations), it is well known that self-reported weight and height are not always accurate (Villaneuva 2001; Gorber et al. 2007). It is not known if this data was updated close to the time of death, or how accurate the individual was in their self-reporting. The reliance on this data, therefore, has a potentially negative impact on the study. In addition, while the second study did include the age of the remains at the time of death, and tested for age bias, there was no further detail on how age may impact the MSMs interpretations, as older individuals have greater muscle markers than younger individuals, which has been related to the accumulation stress of activity patterns over time (Weiss 2003). Furthermore, Godde and Taylor (2011, 2013) never explain what defines an "athletic individual," or what the weight of the athletic individuals were. These individuals could fall into the average or obese category as being athletic is not a weight. It would be worthwhile to know what category they would have been a part of if they were not classified as athletic. These individuals also make up less than 5% of the overall study and are therefore not quantifiably representative of the greater population.

There is further research that needs to be done to account different demographics, age, and the use of a more clearly defined and quantifiably representative group of "active individuals." Nevertheless, Godde and Taylor (2011, 2013) do provide another method of assessing body mass that offers promising results.

Articular surfaces

The final method for this review includes the articular surfaces. The logic behind this theory follows much of the same as the femur: the articular surfaces, especially in the lower body, are responsible for the majority of load-bearing in the body. This considerable load bearing has been proven to have an impact on the joints in various forms of deterioration and accessory bone growth (Lieberman, Devlin, and Pearson 2001). This relationship led researchers to investigate these articular surfaces for evidence of body mass, notably with Lieberman, Devlin, and Pearson, (2001), who examined the phenotypic plasticity of the subchondral articular surface area (ASA) of the following joints: the distal scapula, distal humerus, proximal femur, proximal tibia, anterior astragalus, posterior astragalus, proximal metacarpal, distal metacarpal, proximal metatarsal, and distal metatarsal (Lieberman, Devlin, and Pearson 2001:269).

The researchers compared subchondral ASA responses to loading, and midshaft diaphyseal responses to loading. They hypothesised that animals that experience high levels of load-bearing activity would have larger ASAs relative to body mass. This was tested against the null hypothesis that stated ASAs do not display a substantial degree of phenotypic plasticity. Lastly, the authors assessed if there was variation in static or dynamic loading.

Methods

To test their hypothesis, 40 male sheep of different ages were used: juvenile (40 days old at the start of the experiment), subadult (265–275 days old at the start of the experiment), and mature (400–430 days old at the start of the experiment). For each age group, an exercised and a control group was created. Over 90 days, the exercised animal ran every day for 60 minutes at a speed of approximately 4 kph on a treadmill. Loading activity was restricted to minor locomotor activity and sedentary weight support by housing the animals in 1 m2 cages. At the end of the experiment, the animals were euthanized and the bones were isolated. The articulating surfaces were measured using a latex cast method with a correction factor to compensate for latex shrinkage. The researchers also observed the cross-sectional geometry of the midshaft diaphyses of the hindlimb bones. All of the statistical analyses were done using Statview 4.5.

Results

Lieberman, Devlin, and Pearson (2001) found that the ASA did not produce statistically significant differences between exercises and non-exercised animals for any age group. There were some significant differences in body mass and ASA between the different age categories, but what was most interesting is how little the ASA dimensions changed within each age group in response to exercise. On the other hand, the cross-sectional geometry showed some significant differences between exercised and control sheep, mostly in the distal hindlimb of the juvenile animal. The authors conclude that the mechanical loading experienced by sheep had no significant effect on the animals, and the ASAs showed no substantial degree of phenotypic plasticity. They state that "in order to make inferences about the behaviour and body mass of individual for a given species, it may be more appropriate to use the diaphyseal area, which is sensitive to changes in both body mass and activity level throughout life" (Lieberman, Devlin, and Pearson 2001:276).

Critiques

While this study did not provide anything directly helpful to forensic anthropology, it does offer different avenues of inquiry to biological anthropology, such as the potential to get an average body mass of a species or animal. However, there are still areas of improvement that this study would have benefitted from. For example, female sheep are not represented in the study, and the number of male sheep is small.

Addressing these factors could have potentially provided more information to the overall result, like sex-based differences in ASAs. In addition, it is not specified what the weight of any of the sheep was or if exercise has any impact on pre/post-running weight. If the exercise the sheep were doing was not having any impact on body mass, then it would seem obvious there would not be any significant differences between runners and controls in any age group standardized by body mass. Finally, sheep and humans are very different species in terms of mobility and movement; thus, the information provided in this study might not be representative of changes in humans. Although there are different avenues open to refining this study, such as the use of a different species, it is unlikely it would be relevant in a forensic context to estimate the body mass of an unidentified human individual.

Osteoarthritis

While the articular surface area is currently an unlikely indicator of body mass on an individual level, there is another feature of the articular surfaces that research has found to show a relationship to body mass: osteoarthritis. Osteoarthritis is the destruction of articular cartilage at the joints of the body. It also includes the formation of bony lipping and osteophytes around the margins of the bone (St. George 2015). This has been seen both clinically and in anthropological studies to have a relationship with individuals of a higher body mass (Weiss 2006; Moore 2008; St. George 2015). At the same time, it is also widely known that osteoarthritis is highly correlated with age (Weiss 2006; Lieverse et al. 2016; Calce et al. 2018), nutrition, physical activity, mechanical stress, and habitual activities (Lieverse et al. 2016), and there have also been studies that show differences in osteoarthritis between males and females, as well as different activity levels (St. George 2015). With all of these compounding factors, most notably the high correlation with age, osteoarthritis itself is not a suitable determinant of heavy body mass. At present, this degeneration should be interpreted with caution and with a suite of other indicators.

Juveniles

The previous studies have all mentioned possible methods of estimating body mass from skeletal remains based predominantly (if not exclusively) on adult remains. This is not surprising due to the different growth and morphological changes of the skeleton during the subadult period. The adult equations for estimating body mass have been found to be unsuitable in application on juvenile remains for the following reasons. First, the formulae have been found to generally overestimate body mass. Second, the measurements that the equations rely on are often difficult or impossible to measure on juveniles. Third, general body size to bone length approximations are difficult, as the relationship between those factors changes throughout development (Cowgill 2018). The estimation of body mass for subadults nonetheless could be extremely beneficial in adding more individualized information since sex is not reliably known until puberty (Byers 2016). Fortunately, a novel study was conducted by Ruff (2007) to attempt to create formulae for juvenile stature and body mass estimations. At the time, there had been very little research on stature and no research conducted on body mass estimations for subadults. Only body mass will be reviewed for the purpose of this paper.

Methods

To create his formula, Ruff (2007) examined the femoral head breadth, the distal metaphysis of the femur, and bi-iliac breadth in relation to long bone length. He observed radiographs of children from 1–17, with 10 girls and 10 boys for each age, using the Denver Growth Study to conduct his research. Participants had standardized radiographs taken between 1941 and 1967 at either 6-month or 1-year intervals. The femoral head breadth was measured using digital calipers and taken perpendicular to the femoral head-neck axis. The age of this appearance varied but all were able to be measured by age 7. The mediolateral breadth of the distal metaphysis was also measured with digital calipers and this measurement was possible before the age of 1. This became progressively less visible in all adolescents after 14 years of age as the distal femoral epiphysis blocked the view on the radiograph. Bi-iliac breadth and long bone length were already available as a measurement from the original studies archives. All of the radiographic measurements were corrected for magnification and the equations were generated using ordinary least squares regression. The equations were created with discrete intervals, and sex was not subdivided due to the potential of creating a questionable statistical validity, as well as the lesser pronounced sex-related differences up until adolescence. The exception to this involves the bi-iliac breadth and long bone length due to the significant changes in the male and female pelvis with the onset of puberty. The study considered the percent standard error of estimate (%SEE) to determine if the equations created effective predictions. This percentile also allowed comparisons between different ages and methods. Standard error of estimate (SEE) was calculated for each equation and the statistics were done using SYSTAT (1990).

Results

The results showed body mass estimation error as the smallest in years 2–7 (%SEE 5-6%), slightly larger in year 1 (%SEE 7%) and the error increased from 8 years onwards (Ruff 2007). The femoral head breadth overlapped but did have smaller errors than the distal metaphyseal breadth in estimating body mass, suggesting that once the femoral head becomes usable, it is the preferable method. Once the individual reaches adolescence, however, neither measurement provided very precise estimates with the percent standard error estimates from 10 years onward being above 13%. Ruff (2007:704) suggested that this was "likely due to volatile fluctuations in body mass that were observed in several individuals in this general age range." Once the individual reached 17 years, the %SEEs were much lower. The bi-iliac breadth and long bone lengths for ages 15–17, provided estimation errors of approximately 5–8%. The femoral head breadth and distal metaphysis breadth both provide reasonable body mass estimates in juvenile skeletons up to 8 years of age, and the bi-iliac breadth with long bone length provides reasonable estimates in later adolescence. These were found to be comparable to what has been seen in similar equations for adults.

Further Developments

This ground-breaking study was continued by Robbins, Sciulli, and Blatt (2010) who furthered the data acquired by Ruff (2007) by considering if other areas of the femur could be used in the event of missing or damaged bones. They applied the previously described techniques to the femoral midshaft to determine if it was an appropriate measurement for body mass estimations.

Methods

To conduct this study, Robbins, Sciulli, and Blatt (2010) used a set of age-structured least squares regression formulas based on the measurements from radiographs from the Denver Child Research Council of children from 2 months to 17 years, with 20 subjects in each category. The measurements were acquired from previous studies. The formulas were then tested on an independent sample from the Franklin County, Ohio Coroner's office consisting of 112 subadults of European-American and African-American males and females from 1–15 years old who died in the years of 1990 and 1991. Documentation of sex, ancestry, weight, and stature were available from medical records. They calculated the estimated body mass based on 1) femoral midshaft, 2) the width of the distal metaphysis for ages 1–12, and 3) the femoral head for ages 7–17 as constructed by Ruff (2007).

Results

The results of the Denver sample showed the midshaft femur as being a reliable predictor of body mass for ages 1–8 with a %SEE of under 8% (Robbins, Sciulli, and Blatt 2010). The use of this measurement was less accurate for ages 9–17 with a %SEE of roughly 12% and above. The results from the independent test sample showed a similar trend with the results for 1–8 years of age being comparably accurate. The researchers found that when they removed 4 individuals who were of a higher BMI outside the 95% confidence range the mean bias decreased, and when the sample was analyzed by sex the formulas were found to underestimate males slightly (-0.5kg) and overestimate females (+0.8kg). Despite this, the formulas are still within a reasonable margin of error. They conclude that estimations of body mass of individuals from 1–8 years old using the equations they created based on the femoral midshaft are just as reliable as other methods. Therefore, in cases where the bone ends are damaged, this technique can be used with confidence to estimate body mass. In the older categories of 9–17 years, the midshaft is generally less accurate and precise so measurement from the femoral head would be preferable. They also suggest limiting both the femoral midshaft and distal metaphysis equations from individuals with a high BMI.

Critiques

Both of these studies are based on the Denver Child Growth Study, a longitudinal study of well-nourished, European-American children (Robbins, Sciulli, and Blatt 2010). As mentioned in other studies, this is not a good representation of the demographics these techniques might be used on (i.e., impoverished individuals) and could introduce some bias in the estimation of body mass. In addition, studies have found that repeated exposure to radiation poses health risks such as cancer (Kleinerman 2006). This creates a possibility that the original test group is not a representation of the average healthy child due to this continuous exposure to radiation from the radiographs taken throughout their childhood.

On a separate note, all of the equations given are associated with ages. While this likely provides a more accurate estimation based on that age, this depends on the age of a child being known. Within the forensic and biological context as a whole, it is not always possible to create an accurate and precise age estimation. Elements may be missing or damaged, and individuals may fall between or into multiple age ranges. This specificity will likely not be possible in many situations, and the authors of both studies offer no solution to this problem.

Despite these critiques, the methods presented do offer a range of techniques with reasonable error estimates. They can no doubt be useful in many situations and adapted to fit different situations. These techniques should be tested using different reference populations before any widespread use.

Discussion

This literature review presented several methods for estimating body mass from skeletal remains and specific critiques for each study were highlighted. However, there are still some overarching issues to address within many of the methods of estimating body mass: population demographics, BMI, reported weight, and weight extremes.

All of the studies showed issues with the reference groups included to conduct their studies. While some did include Black populations and females, the vast majority of these studies are based on data from White men. This is not a representative demographic, particularly in North America. For example, a study conducted by the Canadian Femicide Observatory for Justice and Accountability (2018) found that while Indigenous women make up approximately 5% of the population of Canada, they account for 36% of women and girls who are killed by violence. Within this context, it is very possible for unidentified remains to belong to Indigenous women, thus Indigenous data is important to be included when assessing these body mass estimations methods. While no sample can be fully representative, basing a study on only one ancestral group or one sex limits the application of this technique in the real world. Without assessing the potential variation that exists across groups, or having the knowledge of if these calculations are appropriate for everyone, a forensic anthropologist could run the risk of derailing an investigation in providing a body mass by using a calculation that is only applicable for a White man. Before these techniques can ever be responsibly incorporated into forensic anthropology, they need to be re-evaluated with different populations. Stature estimate equations are separated by sex and ancestry to allow for better accuracy (Byers 2016); thus, it seems completely reasonable to produce different equations or evaluate different features in accordance with specific groups for body mass.

Another issue is the use of the BMI or body mass index. The formula for the BMI is simply kg/m2 and does not calculate for any sort of tissue. This means that a relatively light and extremely tall person could be classified as underweight, or an extremely muscular person could be classified as obese (St. George 2015). This index does not indicate actual health, nor does it reflect what a person might visibly look like as people can carry weight in very different ways due to body shape, sex, or muscle size. This categorization can bias researchers into only considering one perceived type of individual when in reality, they might look completely different. This information will likely be shared with law enforcement, and potentially the general public in hope of identifying unidentified skeletal remains. The bias these classifications inherently have could pose larger issues, such as a group being ignored from possible ID-ing.

There are further problems regarding reported weight. As mentioned previously, self-reported data is often inaccurate (Villaneuva 2001; Gorber et al. 2007). Many methods of confirming the body mass of an individual would be based on self-reported data, such as the weight and height on a driver's license. In addition to this problem, unlike stature—which is arguably consistent once adulthood is reached—body mass will fluctuate. A study in the Health Reports of Statistics Canada showed that for both men and women, the average change in body weight over two years was + 0.5–1.0 kg (Orpana, Tremblay, and Finès 2007). This is merely an average and the accumulation years after could make a dramatic change. This means that unless a person was completely accurate in their reporting and had not changed at all since the time of doing so, documented body mass is potentially not accurate to what a person weighs at their time of death. This presents a possible limiting factor in what you can compare the body mass estimation to.

A final problem is that of weight extremes. Ruff (1991) and Robbins, Sciulli, and Blatt (2010) both had a few individuals of high body mass that were removed from the study to improve the accuracy. This suggests that these equations may not be appropriate for everyone. It is possible these discrepancies were due to sample size, and as such they should be evaluated on different samples of body weight. The point of using these methods is to provide an estimation of the unknown body mass. If there is a subset of the population these calculations do not work for, they might not be practical methods at present.

Fortunately, there are ways to navigate many of these issues. As Godde and Taylor (2011, 2013) suggested, multiple methods should be used in tandem. As with sex and stature, it would likely be best to consider many measurements and many variables when trying to assess body mass. This is the logic Moore and Schaefer (2011) have followed. They created a regression tree based on many variables to estimate body weight. Using numerous factors will provide a broader understanding, as different methods can potentially cover the gaps in others.

Furthermore, the argument could be made that, at least in forensic anthropology, the need for an exact body mass might not be necessary. It has already been discussed that provided records of body mass themselves may not be accurate due to reporting and weight fluctuations. In addition, the difference between an individual who is 170 cm and 55 kg versus another who is 170 cm and 65 kg, may be relatively negligible in terms of identification. Persons of diverse build carry weight differently, making appearances highly variable. On the other hand, if the remains are able to show that this person is 170 cm and under 45 kg or over 136 kg, this provides more distinct groups of individuals for investigators to compare and could narrow the pool of candidates in ongoing investigations. Some of the studies excluded the extremes of weight, but it could be worthwhile to find a way to confirm these extremes. These are probably more distinctive than any "average weight" and would have been noticeable to those that knew the deceased.

Conclusion

This literature review provided an overview of many of the notable theories and studies in the estimation of body mass from skeletal remains. In conducting this review, several conclusions were found: the femur can provide body mass estimates of reasonable error margins for both adults and juveniles most prominently within White demographics. The entheses are shown to have a relationship with heavier body mass, but there is no way as of yet to sort out athleticism from this grouping. The articular surface area shows no impact from body mass, but other features, such as osteoarthritis, are shown to be impacted in some ways by heavier body mass (though age seems to be a more predominant factor). In summary, there are ways in which body mass can be inferred from the skeleton, some with smaller error margins than others. All of these methods require further investigation with different populations, body weight extremes, and association with other confounding factors (such as age, ancestry and sex) before they can be used with a high level of certainty. Despite the gaps in the literature, there have been some promising starts in estimating body mass from skeletal remains that can be researched in the future in hopes of one day being possible to apply to forensic investigations.

References Cited

Byers, Steven N. (2016). *Introduction to forensic anthropology*. 5th edition. London: Routledge.

Calce, Stephanie E., Helen K. Kurki, Darlene A. Weston, and Lisa Gould. 2018. The relationship of age, activity, and body size on osteoarthritis in weight-bearing skeletal regions. *International Journal of Paleopathology* 22:45–53, https://doi.org/10.1016/j.ijpp.2018.04.001.

Canadian Femicide Observatory for Justice and Accountability. 2018. Call it femicide: understanding gender-related killings of women and girls in Canada 2018. https://femicideincanada.ca/callitfemicide.pdf.

Cowgill, Libby. 2018. Juvenile body mass estimation: a methodological evaluation. *Journal of Human Evolution* 115:78–84, https://doi.org/10.1016/j.jhevol.2017.07.007.

De Groote, Isabelle and Louise T. Humphrey. 2011. Body mass and stature estimation based on the first metatarsal in humans. *American Journal of Physical Anthropology* 144:625–632, https://doi.org/10.1002/ajpa.21458.

Godde, Kanya and Rebecca W. Taylor. 2011. Musculoskeletal stress marker (MSM) differences in the modern American upper limb and pectoral girdle in relation to activity level and body mass index (BMI). *Forensic Science International* 210(1–3):237–242, https://doi.org/10.1016/j.forsciint.2011.03.014.

———. 2013. Distinguishing body mass and activity level from the lower limb: can entheses diagnose obesity? *Forensic Science International* 226(1–3): 303.e1–303.e3037, https://doi.org/10.1016/j.forsciint.2013.01.027.

Gorber, Connor, Tremblay, Mark Moher, David, and B. Gorber. 2007. A comparison of direct vs. self-report measures for assessing height, weight and body mass index: a systematic review. *Obesity Reviews* 8(4):307–326, https://doi.org/10.1111/j.1467-789X.2007.00347.x.

Hawkey, Diane E. and Merbs F. Charles. 1995. Activity-induces musculoskeletal stress markers (MSM) and subsistence strategy changes among ancient Hudson Bay Eskimos. *International Journal of Osteoarchaeology* 5(4):324–338, https://doi.org/10.1002/oa.1390050403.

İşcan, Mehmet Y. 2005. Forensic anthropology of sex and body size. *Forensic Science International* 147*(*2–3):107–112, https://doi.org/10.1016/j.forsciint.2004.09.069.

Kleinerman, Ruth A. 2006. Cancer risks following diagnostic and therapeutic radiation exposure in children. *Pediatric Radiology* 36:121–12, http://doi.org/10.1007/s00247-006-0191-5.

Lacoste, Jeanson, A., Frédéric Santos, Chiara Villa, Ján Dupej, Niels Lynnerup, and Jaroslav Brůžek. 2017. Body mass estimation from the skeleton: an evaluation of 11 methods. *Forensic Science international* 281:183.e1–183.e8, https://doi.org/10.1016/j.forsciint.2017.10.026.

Lieberman, Daniel E., Maureen J. Devlin, and Osbjorn M. Pearson. 2001. Articular area responses to mechanical loading: effects of exercise, age, and skeletal location. *American Journal of Physical Anthropology* 116:266–277, https://doi.org/10.1002/ajpa.1123.

Lieverse Angela R., Vladimir I. Bazaliiskii, Olga I. Goriunova, and Andrzej W. Weber. 2009. Upper limb musculoskeletal stress markers among middle Holocene forager of Siberia's Cis-Baikal region. *American Journal of Physical Anthropology* 138:458-472, https://doi-org.login.ezproxy.library.ualberta.ca/10.1002/ajpa.20964.

Lieverse, Angela R., Brianna Mack, Vladimir I. Bazaliiskii, and Andrzej W. Weber. 2016. Revisiting osteoarthritis in the Cis-Baikal: understanding behavioral variability and adaptation among middle Holocene foragers. *Quaternary International* 405:160–171, https://doi.org/10.1016/j.quaint.2015.03.019.

Moore, M. K. 2008. Body mass estimation from the human skeleton. PhD dissertation, University of Tennessee, Knoxville. https://trace.tennessee.edu/utk_graddiss/365.

Moore, Megan K., and Eric M.S. Schaefer. 2011. A comprehensive regression tree to estimate body weight from the skeleton. *Journal of Forensic Sciences* 56:1115–1122, https://doi.org/10.1111/j.1556-4029.2011.01819.x.

Myszka, Anna, and Janusz Piontek. 2011. Shape and size of the body vs. musculoskeletal stress markers. *Anthropologischer Anzeiger* 68(2):139–152. http://www.jstor.org/stable/41262720.

Niskanen, Markku, Juho-Antti Junno, Heli Maijanen, Brigitte Holt, Valdimir Sladék, and Margit Berner. 2018. Can we refine body mass estimations based on femoral head breadth? *Journal of Human Evolution* 115:112–121, https://doi.org/10.1016/j.jhevol.2017.10.015.

Orpana, Heather M., Mark S. Tremblay, and Phillipe Finès. 2007. Trends in weight change among Canadian adults. *Health reports* 18(2):9–16. https://www150.statcan.gc.ca/n1/en/pub/82-003-x/2006005/article/trends-tendances/9633-eng.pdf?st=yuPOtiGL.

Robbins, Gwen., Paul W. Sciulli, and Samantha H. Blatt. 2010. Estimating body mass in subadult human skeletons. *American Journal of Physical Anthropology* 143:146–150, https://doi.org/10.1002/ajpa.21320.

Ruff, Christopher B., and Wilson C. Hayes. 1988. Sex differences in age-related remodeling of the femur and tibia. *Journal of Orthopaedic Research* 6:886–896, https://doi.org/10.1002/jor.1100060613.

Ruff, Christopher B., William W. Scott, and Allie Y.-C. Liu. 1991. Articular and diaphyseal remodeling of the proximal femur with changes in body mass in adults. *American Journal of Physical Anthropology* 86:397–413, https://doi.org/10.1002/ajpa.1330860306.

Ruff, Christopher B. 2007. Body size prediction from juvenile skeletal remains. *American Journal of Physical Anthropology* 133:698–716, https://doi.org/10.1002/ajpa.20568.

St. George, Karen R. B. 2015. The estimation of body mass from human skeletal remains. PhD Dissertation, Boston University, Boston.

University of Tennessee Knoxville Forensic Anthropology Center. n.d. UTK Donated Skeletal Collection. https://fac.utk.edu/utk-donated-skeletal-collection/ (accessed April 3, 2023)

Weiss, Elizabeth. 2003. Understanding muscle markers: aggregation and construct validity. *American Journal of Physical Anthropology* 121(3):230–240. https://doi.org/10.1002/ajpa.10226

———. 2006. Osteoarthritis and body mass. *Journal of Archaeological Science* 33:690–695, https://doi.org/10.1016/j.jas.2005.10.003.

Villanueva, Elmer V. 2001. The validity of self-reported weight in US adults: a population based cross-sectional study. *BMC Public Health* 1:11 https://doi.org/10.1186/1471-2458-1-11