7-20-2021

Reconstruction of Femur Length Using the Epiphysial and Diaphysial Diameters in Contemporary Egyptian Sample, with Application to Ancient Egyptians

MennattAllah Hassan Attia
Alexandria University, Alexandria, Egypt

Fatma Mohamed Magdy Badr El-Dine
Alexandria University, Alexandria, Egypt

Nancy Mohamed Aly El-Sekily
Alexandria University, Alexandria, Egypt

Reconstruction of Femur Length Using the Epiphysial and Diaphysial Diameters in Contemporary Egyptian Sample, with Application to Ancient Egyptians

MennattAllah Hassan Attia,[1] Fatma Mohamed Magdy Badr El-Dine,[1] and Nancy Mohamed Aly El-Sekily[2]*

[1]Forensic Medicine and Clinical Toxicology, Faculty of Medicine, Alexandria University, Alexandria, Egypt.

[2]Anatomy and Embryology, Faculty of Medicine, Alexandria University, Alexandria, Egypt.

*Correspondence to: Nancy Mohamed Aly El-Sekily, Alexandria University, Mowasat Medical Campus, Anatomy and Embryology, Faculty of Medicine, Alexandria Governorate, Egypt. E-mail: nancyelsekily@yahoo.com.

Short Title: Femur Length Reconstruction by Nutrient Foraminal Models

KEY WORDS: FEMUR RECONSTRUCTION, DOMINANT NUTRIENT FORAMINA, MODERN UNDOCUMENTED OSTEOLOGICAL COLLECTIONS, ANCIENT EGYPTIANS, FRAGMENTARY FEMORA, INDIRECT STATURE ESTIMATION.

Abstract

Inferences in bioarchaeology and forensic contexts require mathematical stature estimation using long bone lengths. This study is in hand to identify predictors of femur length (FL) from epiphyseal and diaphysial width measurements that are not bound to assumptions of sex or laterality. Both standard and new measurements around dominant foramen nutricium (NF) were collected on modern femora (n=64) from Alexandria university unidentified skeletal Collection to compute linear regression models. Four equations were then validated on Ancient Egyptian sample (n=73) from Goldman's Osteometric dataset to evaluate effect of sex subdivision on the prediction accuracy of FL and indirect stature estimation using Raxter's formulae. Most of models reflected significant positive association r>0.60) between width variables and FL. Oddly, the distance from proximal end to NF correlated weakly with FL (r=0.34). The stepwise selected equations preferred measurements around NF to midshaft where the anteroposterior diameter was included in proximal fragment model (r=0.77) and circumference in diaphyseal fragment model (r=0.62). Tested equations performed consistently on the ancient Egyptian sample. Measurements from femoral proximal fragment are more reliable predictors than distal fragment with the exception of femur neck diameter. However, distal epicondylar breadth is a better predictor of FL in females than in males. Indirect stature estimation showed a reasonable degree of accuracy in both sexes. These models can be applied successfully in Contemporary and Ancient Egyptians fragmentary remains however, due to larger size of femora from Old Kingdom sample, they would be most applicable to individuals from the following dynasties.

Several procedures in paleo- and forensic anthropology involve estimation of the biological profile parameters which are considered as unobservable quantity (such as stature) using observable quantities (long bone lengths) (Auerbach, 2011; Konigsberg, Hens, Jantz, & Jungers, 1998). The body height is an important datum for inferring the growth status of contemporary or ancient populations (Shuler, Danforth, & Auerbach, 2011; Mays, 2016). Notwithstanding, the practice of estimating the stature remains plagued by technical problems and the accuracy of the estimates depends on the method employed to calculate them. The anatomical method entails that lengths of individual skeletal elements are summed to provide a direct stature estimate. A correction factor for the spinal curvatures, pelvic inclination and missing non-bone body parts can be applied to modify skeletal length to living body length (Fully and Pineau, 1960; Porter, 2002). Alternatively, a mathematical technique can be employed in which regression formulae (or ratios) based on the proportionality of long bone lengths to stature. These regression equations provide stature estimates with a certain margin of error which can be limited through a careful selection of the regression method used (Raxter et al., 2006; Giannecchini and Moggi-Cecchi, 2008). A large number of these equations employ measurements from the femur due to its resilience; being the strongest weight-bearing bone of the appendicular skeleton (Mays, 2016), and it generally shows the closest linear correlation with stature (Mays, 2016; Ruff, Holt, Niskanen, Sladék & Berner, 2012; Feldesman & Fountain, 1996). The femur/stature ratio method is considered as “a special case of classical calibration” in which the intercept equals zero (Hens, Konigsberg & Jungers, 2000). The femur length is multiplied by the stature/femur ratio to obtain the stature estimate. The femoral length (FL) is on average 26.74% of the stature across different populations (Feldesman & Fountain, 1996).

While the anatomical method is the most reliable because there is no dependence on a correlation in a modern reference sample, it can hardly be applied in palaeoanthropology.

Unfortunately, it requires nearly complete skeletons which is rarely fulfilled in archeological samples (Porter, 2002) in addition to inapplicability in commingled contexts (Anzellini & Toyne, 2019). On the other hand, the proportionality of femur to stature (allometry) is variable across populations as well as present and past populations and it follows certain spatiotemporal pattern due to eco-geographic and growth plasticity factors (Béguelin, 2011; Raxter, Ruff, Azab, Erfan, et al., 2008; Ruff, 2002) leading to regional and temporal biases in stature estimates (Hens et al. 2000). When the FL is compared among 3 modern population groups from the same continent, discerned differences in their mean values are noted. For example, the Thai population (Mahakkanukrauh, Khanpetch, Prasitwattanseree, & Vichairat et al., 2011), the mean values of FL were 402.7 mm in females and 435.5 mm in males; the Indian population (Prasad, Vettivel, Jeyaseelan, Isaac, & Chandi, 1996), FL= 417.7 mm in females, 448.6 mm in males, and in the pooled sample equals 434.7 mm; the Sri Lankan population (Nanayakkara, Vadysinghe, & Nawarathna, 2018), FL= 428.6 mm in the pooled sample. Subsequently, a multitude of stature and/or FL estimation formulae have been created for the diverse populations around the world. Applying population-specific equations to individuals with similar proportions should guarantee the same level of accuracy in stature estimation because they account or control for variation in body proportions (Holliday and Ruff, 1997; Auerbach & Ruff, 2004).

Raxter et al. (2008) attempted to address the problem of applying non-population specific stature estimation equations to archaeological Egyptian specimens by developing stature estimation equations directly from a diverse skeletal remains dating to old kingdoms. If a statistically sufficient subsample of individuals with the key skeletal elements is available, a "hybrid" approach can be applied in which the anatomical stature is reconstructed in those individuals, and then used to compute sample-specific mathematical formulae for estimating stature from long bone lengths in less well preserved skeletons (Ruff, Niskanen, Maijanen, Mays, 2019). Therefore, establishing population data for archaeological samples requires a mathematical method in addition to the anatomical method (Sciulli, Schneider & Mahaney, 1990; Giannecchini and Moggi-Cecchi, 2008).

The aforementioned equations are available to researchers for estimating body size/height among the Egyptian skeletal remains, however, it is unclear how widely applicable they are in cases of fragmentation. Archaeological and forensic materials are sometimes damaged and fragmented due to taphonomic alterations naturally or during the process of excavation which result in loss of contextual information and restrict their use to the estimation of a minimum number of individuals within the sample (Hoppa & Gruspier, 1996; Meyer, Frater, Seiler, Bickel & Böni ,2020; Tomsová & Schierová, 2016). Bones can be destroyed either by loss of the organic (collagen) phase or chemical dissolution of bone mineral which exposes the protein to microbial attack in cases of long-term burial depending on the conditions in the burial environment (Collins, Nielsen–Marsh, Hiller, & Smith, et al., 2002).

Nevertheless, sufficient skeletal material must be present in any assemblage for reliable reconstruction of the biological profile at the individual level which is based on the intact element portions for measurements collection (Komar & Potter, 2007). Methods accounting for fragmented and/or commingled human remains shared several key aspects of their approach which include (1) dividing the femur into several linear segments using certain landmarks as defined by (Steele & McKern, 1969 ; Simmons, Jantz & Bass, 1990) and the proportion of each segment to the maximum length of the femur is then calculated, or (2) collection of standard width measurements from the epiphyses and/or the diaphysis. An estimate of the living stature can be obtained using a two-step approach by plugging the reconstructed maximum femur length into the appropriate stature reconstruction formulae.

The direct method allows for the estimation of stature from individual or combinations of measurements of fragments of long bones (Bidmos, 2008; Bidmos, 2009).

Previous studies provided several critiques of the longitudinal measurements methods for FL reconstruction with regards the difficulty in locating some anatomical landmarks that define the bone segments (Shuler et al., 2011; Simmons et al., 1990; Bidmos, 2008; Bidmos, 2009). Although Wright and Vasquiz, 2003 found that landmarks based on the articular surfaces and secondary ossification centers are more easily to identify than the variable muscular attachment sites, they stated that longer femoral fragments are required to apply their models such that the entire diaphysis as well as some marginal articular bone should be present in order to apply some equations. Nonetheless, Gidna and Domínguez-Rodrigo, 2013 introduced a method to estimate the length of femora from incomplete diaphyseal fragments using the length of linea aspera. However, their methods did not account for small fragments of the diaphysis.

On the other hand, the methods employing the transverse measurements were proposed by several authors due to their reproducibility and ease of identification (Simmons et al., 1990; Fongkete et al., 2016; Prasad et al., 1996). As a result, several authors suggested the use of standard width measurements in the estimation of stature and maximum length of long bones in adults (Simmons et al., 1990; Bidmos, 2008; Fongkete, Singsuwan, Prasitwattanaseree & Riengrojpitak, 2016; Prasad et al, 1996; Bidmos, 2009; Timonov, & Fusova, 2016; Nanayakkara, Vadysinghe & Nawarathna, 2018; Abledu, Offei, & Osabutey, 2016) and sub-adults (Hoppa & Gruspier, 1996).

Traditionally, the external dimensions of diaphyses were frequently used as morphometric measures for quantifying sexual dimphorphism, robusticity and diaphyseal shape (Attia, Badr El-Dine, Attia & El-Sekily, 2020; Stock & Shaw, 2007). Although femoral Midshaft measurements are routine in these standardized anthropological measurements, the diaphyseal nutrient foramen is considered as a recognizable landmark that may be of value in cases where the midshaft diameter can't be located accurately (Attia et al, 2020; Feldesman & Fountain, 1996; Buck, 2010). Steele and Mckern, 1969 showed that the variability of the position of the nutrient foramen (NF) prohibited the inclusion of the longitudinal distance from proximal end to NF in their study. In the present study, the transverse measurements around the NF instead of the linear segment were employed as an alternative to midshaft counterparts in order to be used when only the diaphyseal region is recovered.

Up to our knowledge, there are no formulae for reconstruction of the maximal femur length (FL) from its fragments available for use in the Egyptian population. In this context, we aimed to (1) establish the correlation between the FL and the new measurements around NF, and (2) derive linear regression models for the reconstruction of FL which account for various recovery scenarios of fragmentary remains. The models were tested on an ancient Egyptian sample as a completely independent test of models performance using samples from a different time period.

Materials and Methods

Reference Sample

A modern Egyptian sample was assembled according to strict selection characteristics that prevent the occurrence of morphological and topographical factors affecting the femoral proportions and/or measurements acquisition as well as absence of nutrient foramina. This sample is composed of a balanced sex ratio (Attia et al., 2020). All the specimens included were skeletally adult of 18 years or more, defined as having united epiphyses however, their exact ages were unknown. The skeletal materials were prepared from cadavers utilized for routine dissections by the undergraduate academic program. They represent individuals from the middle to low socioeconomic strata who grew and lived during the second half of the past century. Of the 64 individuals in this analysis, 23 were presented by only a right femur and 41 by only a left femur; considering that directional bilateral asymmetry in femora is very small (Steele & Mckern, 1969; Auerbach & Ruff, 2006). Approval from the Alexandria faculty of medicine ethical committee was obtained prior to data acquisition. The serial number is 0304406/9/2019.

Measurements Acquisition and Data Analysis

The various measurements taken are standardized for use in bioanthropology following the definitions described in (Auerbach & Ruff, 2006; Bass, 2005; Curate, Coelho, Gonçalves, Coelho & Ferreira, 2016) to represent different fragments of femur such as proximal end: the transverse head diameter (THD), vertical head diameter (VHD), femur neck axis length (FNAL), femur neck width (FNW), mid-diaphysis: mediolateral diameter (MLD), anteroposterior diameter (APD), femoral circumference (CF), and distal end: distal epicondylar breadth (DEB). In addition to these, four new measurements around the dominant nutrient foramena namely proximal end to nutrient foramen (PENF), anteroposterior diameter (APNF), mediolateral diameter (MLNF), and circumference (CNF) were collected due to the ease of identifying NF as a landmark (Table 1). The dominant nutrient foramina were initially identified in the unsexed (pooled) sample by the elevated margins and distinct groove leading to the external orifice and by being capable of admitting at least the tip of a 24-gauge hypodermic needle (0.56 mm in outer diameter) while those smaller were excluded from analysis (Johnson, Beckett, & Márquez-Grant, 2017; Attia et al., 2020). Seven femora were remeasured 4 weeks after the original analysis to test for intra-observer variation using intra-class correlation coefficient analysis (ICC) based on single rater/measurement, absolute agreement, 2-way mixed-effects model. Descriptive statistics, including minimum, maximum, mean and standard deviation were calculated. Pearson’s r was used to assess the association between FL and the other epiphysial and diaphyseal anthropometric measurements. Additionally, the mean differences between the new measurements at the level of NF and their counterparts at the mid-shaft level were compared and Pearson's r correlation analysis were performed.

Test Sample

The equations were tested on an independent ancient Egyptian sample (n=73, F=28, M=45) obtained from the Goldman Osteometric dataset which was collected by Dr. Benjamin Auerbach from various museum collections around the world. The sample comprised of Old Kingdom period from Gizeh and later dynasties from El-Hesa materials to represent different Dynastic periods (for more details see Auerbach & Ruff, 2006; Goldman's Osteometric dataset). Five left femoral measurements were selected namely FL,THD, DEB, APD, and MLD because this side exhibits less missing values. An independent samples t-test was used to analyze the mean differences between the femoral measurements in the combined sex samples of modern and ancient Egyptians. Pearson's r was used to establish the association between FL and the other epiphysial and diaphyseal anthropometric measurements.

Reconstruction of Femur Length Using Linear Regression Method and Goodness of Fit Measures

Data were pooled in a single dataset, without distinguishing sides and sex (Steele & McKern, 1969) because it may not always be possible to sex the bone fragments confidently and it is preferable to have both sex represented in a reference population (Albanese, Tuck, Gomes & Cardoso, 2016). The normality assumption was violated in 4 variables (APD, CF, APNF, and CNF), however, simple linear regression is robust against this violation (Gidna & Domínguez-Rodrigo, 2013). The Ordinary Least Squares (OLS) Regression was implemented to establish the strength of relationship between the maximal femur length and the standard and/or new measurements of femur. The simple linear regression model is in the form of:

$$Y = b_0 + b_1X_1$$

where "Y" represents FL (dependent variable), "X" the femoral width measurement (independent variable), "b_0" the intercept and "b_1" the slope. As more independent variables are added "X_n", their respective new coefficients of the slope "b_n" are calculated and the model is called multiple regression. Multiple regression equations were derived using the stepwise method which is a combination of forward selection and backward elimination (Mc Henry, 1974).

After modeling, the estimated coefficients and the distribution of errors were checked using the residual plots and were found to follow a normal distribution (Hoppa & Gruspier, 1996; Gidna & Domínguez-Rodrigo, 2013). Part of the regression output including a plot of residuals versus predicted values, normal (P-P) probability plot, and two statistical tests for normality (the Kolmogorov-Smirnov test, the Shapiro-Wilk test) for the four non normally distributed variables are present in the online supplemental materials. A variance inflation factor (VIF) < 5.00 indicated low multicollinearity (Chiba et al., 2018). The VIF ranged from 1.319 to 2.450 in the multiple regression models. ANOVA test was performed to check the significance of the fitted model with the null hypothesis that the model explains zero variance in the dependent variable (FL). Selected models showed positive correlation >0.60 with the FL, with a p-value <0.05. The best model was the one with the highest coefficient of determination (R^2), adjusted R^2 values and the minimum standard error of the estimate (SEE). All statistical analyses were performed using Microsoft Excel (2013) and Statistical Package for Social Sciences (SPSS), Version 21.0.

Assessment of Model Performance in the Training and Test Samples

The bias and inaccuracies of models were assessed using the mean of raw residuals (MD), and the mean of the absolute value of the residuals (MAD) calculated for each sample. The MD (bias indicator) is calculated as measured FL minus estimated FL. A positive and negative MD indicates a tendency to underestimate and overestimate the measured FL, respectively. MAD (measure of overall error) is the mean of the magnitude of the absolute individual errors. The measurement of utility is the percentages of cases whether the estimated range bracketed the measured FL or not within plus or minus one and two SEEs around the estimated femur length (Albanese et al., 2016).

A two-step validation was adopted to test the models on the ancient Egyptians. Sex specific analyses were performed to assess the effects of subdivision by sex on prediction accuracy of FL (Sjøvold, 1990). Analyses were conducted by entering the THD, DEB, APD, and APD+MLD into the appropriate equation and calculating an estimated femur length. A straightforward application of stature estimation formula can be then implemented using Raxter's formula (Raxter, Ruff, Azab, Erfan & Soliman, 2008) after conversion of the predicted FL to centimeters. To obtain the range of standard deviation for the predicted stature, the standard deviation calculated for the long bone was multiplied by the first constant in the stature formulae, plus the standard deviation of stature following Steele and Mckern, 1969.

Results

The general characteristics of the study calibration (contemporary) and validation (ancient Egyptians) samples are shown in Tables 2 and 3. All the measurements show an acceptable level of consistency between observational series i.e., ICC >0.90 indicating excellent correlation with the exception of FNW (ICC=0.88) indicating good correlation. Therefore, the measurements error bias should have negligible impact on the results.

The descriptive statistics in Table 2 indicated minimal decrease in the antero-posterior diameter from the midshaft level to NF level by only 0.2 mm whereas the MLNF and CNF were slightly increased by 0.66 and 1.4 mm, respectively. Pearson correlation coefficients were calculated for three new nutrient foramina related measurements and their counterparts at the midshaft. The R values were high varied from 0.886 to 0.929 at P=0.000. The highest correlation coefficient was observed for CNF vs. CF and the lowest value was observed for MLNF vs. MLD.

All measurements show a statistically significant moderate to strong positive correlation with femur length, except PENF was weak (see Table 2) according to the arbitrary limits for the absolute values of r (Swinscow & Campbell, 1997). The measurements of epiphysial ends of femur showed higher correlation with FL than diaphysial measurements in our sample (Table 2). However, the APD is correlated the best with FL in the ancient Egyptian and the correlation coefficient was higher than the correlation coefficient in the contemporary Egyptian sample (Table 3).

Reconstruction of Femur Length Using Linear Regression Method

Table 4 shows the linear regression models and goodness of fit statistics for univariable and stepwise selected multivariable models. Observing the univariate linear regression models, it can be seen that THD, DEB, VHD, and FNAL are the most reliable measurements for predicting FL in the training dataset. THD provides better fit than VHD and DEB in the combined sex equations. The diaphyseal measurements including CNF, APD, and APNF were next to the epiphyseal measurements in the rank and the highest correlation with FL was obtained by CNF. In general, MLD and MLNF were poorer predictor of FL than either APD or APNF. The inclusion of THD in the stepwise selected bivariate regression models provides a better fit of the line to the data than the univariable equation with smaller SEE which span 19.688-20.384 mm. APNF and FNAL is involved in the most probable models with THD, but DEB and APD were excluded. The NF stepwise model selected the CNF over the other 2 variables.

Assessment of Performance of Best Models in the Training Sample

Table 5 indicates that the generic formulae including measurements of epiphyseal ends provided an overall lower MAD than diaphyseal measurements with the percentage of bracketed FL in range within 2 SEE spanned 95.3% to 98.4% using FNAL and DEB, respectively. The APNF and CNF (NF) equations have only slightly higher MAD values than APD (midshaft) equation. The APD model provided the best results within 1SEE but performed slightly lower than NF models within 2 SEE. Broadly, the diaphyseal models have slight tendency to overestimate FL. The best bivariate stepwise selected model included THD+APNF with highest percentage of correctly bracketed FL in range 98.4% and the lowest MAD 15.92 mm with slight tendency to overestimate FL. On the other hand, the second best bivariate model included THD and FNAL with slightly higher MAD and tendency to underestimate the FL.

Testing the Models on a Sample of Ancient Egyptians: A Two-Step Validation

The results for the 2-step tests of the generic models for the ancient Egyptians for the sex-specific sample are presented in Tables 6 and 7, respectively. In step 1, the results for four generic equations of femur length reconstruction tested using Goldman's Osteometric dataset of ancient Egyptians are presented in Table 6. In general, the generic equations performed consistently well for both sexes with comparable results to the contemporary reference sample. The estimated range bracketed the estimated FL between 68.2% and 73.3% in males, 71.4% and 76.9% in females within 1 SEE. The overall accuracy increased between 86.67% and 88.89% in males and 100% of the time for females within 2 SEE. The best performance and accurate FL reconstruction in females was achieved by DEB model while in males the THD model provided the best estimates. The univariate and bivariate midshaft models performed only marginally better on females than in males as measured by the MAD and the percent correctly bracketed within 1 and 2 SEE. The MAD was similar for males and females for all equations albeit males had a slightly higher MAD than females for all equations with the exception of DEB model. The differences between the male MAD and female MAD for any given equation was between 1.6 and 5.05 mm.

In step 2, the formulae tend to underestimate stature in the Ancient Egyptian males (MD values from 1.42 to 2.74 cm), whereas in female sample tend to overestimate stature as indicated with negative MD values of -0.74 and -1.53 cm. The males had a slightly higher MAD than females for all equations with a small difference in MAD values that ranged between 0.27 to 0.66 cm. The APD model and the model employing both diameters of the mid shaft performed equally and were better at bracketing the calculated stature, but using the bivariate model gave more precision in the estimates. Despite the slight differences in the direction of the error by sex (MD) for equations and moderate average error (MAD), the 4 equations tested using the ancient Egyptians, correctly bracketed the estimated stature by Raxter's formula (the best case scenario), and would have provided useful stature information in actual archeological contexts using detached femoral pieces (the worst case scenario) (see Table 7). Figure 1 depicts the scatterplots of all individuals in both datasets and shows that while results cluster around the fit line, there is still a noteworthy diverging individuals who were 4 males from the pyramidien Giza using all models and 1 small male from El Hesa sample using DEB model only.

Discussion

Stature is a keystone measure of body size that is especially important for human evolutionary studies (Hens et al., 2000).

The present work introduces a technique to estimate FL from incomplete elements, including short diaphyseal fragments bearing only single nutrient foramen as well as the standard width measurements from proximal and distal ends and the mid-diaphysis. In the current study, the accuracy varies according to strength of linear association between these width measurements and FL. Additionally, the variation in SEE was according to which part of the femur is represented and the number of variables employed (De Groote & Humphrey, 2011). In comparison to other studies, SEE obtained for FL reconstruction in our study is comparable to other populations of both sexes (for example Thai (Fongkete et al., 2016), SEE= 15.6-19.1; Indian (Prasad et al ., 1996), SEE= 20.1-26.9; Sri Lankan (Nanayakkara et al., 2018), SEE= 18.85-21.91). It should be noted that SEE are generally larger than for regression formulae based on longitudinal measurements of femur, which is not surprising considering that longer zones contribute to femur length (Wright & Vásquez, 2003).

As regards the range of correlation between femoral measurements and FL in the Egyptian population, there was higher correlation coefficient of THD than VHD with FL (r= 0.725 and 0.674) which is in agreement with Abledu et al., 2016 in the Ghanaian sample (r=0.714 and 0.704) but in contrast with Nanayakkara et al., 2018 in the Sri Lankan sample (r=0.569 and 0.670). We also pointed out that the r values of FNW, FNAL, and DEB spanned from 0.548 to 0.694 with the lowest and highest values obtained for FNW and DEB, respectively. This range is higher than those obtained by Simmons et al., 1990 for the American African (r=0.315–0.592) and American White samples (r=0.384–0.606) where the lowest and highest values for FNW and FNAL, respectively.

In general, measurements from the proximal end of femur -with the exception of FNW- displayed the highest correlation with FL which is consistent with the observation made by several authors (Timonov & Fusova, 2016; Abledu et al., 2016; Meeusen, Christensen & Hefner, 2015) where the upper epicondylar length, a proximal femoral measurement, consistently showed the best correlation with FL in both sexes. Recent studies mentioned the presence of high positive association of FNAL with femur length and body height (Meeusen et al., 2015; Nissen, Hauge, Abrahamsen, Jensen & Mosekilde, 2005). The combinations of THD + APNF or THD + FNAL increased the accuracy of FL prediction. These two stepwise multiple regression models were computed as a potential indicator of FL when the proximal end is well preserved or when the distal end is missing. They provided the best fit of the data, resulting in better estimates (higher R^2 and lower SEE) when compared to the simple linear regression equations (Torimitsu, Makino, Saitoh, Sakuma & Ishii et al., 2015; Albanese et al., 2016).

Further, Attia et al., 2020 referred to the presence of similarities in the distribution of NF location among different populations which spanned from ca. 30% to 65% and the mean values of foraminal index (a proxy of NF location calculated as percentage of maximal femur length i.e., dividing PENF by FL then multiplication by 100) were more or less related to midshaft position. Davies and Stock, 2014 examined the correlation between the femoral cross-sectional geometric properties and the relative body breadth (as a proxy of body shape) throughout the mid-diaphyseal region (within 60–30% of femur shaft length taken from proximal part of the bone towards its distal end). In general, the authors noted the presence of (relative) mediolateral strengthening of the femoral shaft among both males and females with increased correlation coefficients towards both ends of the femur diaphysis (please see figure 5, p. 828). However, the statistical significance is dependent on which diaphyseal property is being examined and sex. For every cross-sectional geometric property (Ix, Iy, Imax, Imin, and J), between 60-30% of FL, the absolute values for males are lower by ca. 0.2 than females and non-significant (except at section 60% in the Ix, Imax, and J) whereas in females, the correlations with relative body breadth are significant throughout the femur mid diaphysis at either $p<0.05$ or <0.001 but it is still property- and level- dependent. Moreover, the correlation coefficients for Iy and Imin begin to exceed those observed for Imax and Ix in the region of the femur midshaft, suggesting that mediolateral dimensions retain a greater relationship with relative body breadth in both males and females. In contrast, the total area (TA) which is an external quantification of the combined cortical bone and medullary areas, showed statistically insignificant low magnitude of correlation coefficients in both sexes (with the exception of the section 30% of FL in females). Furthermore, in another study, the TA correlation coefficients according to stature diminish toward the epiphysial ends of the diaphysis, suggesting that stature may be a factor that influence the femoral diaphyseal architecture more than the epiphysis (Santos, Lacoste Jeanson, 2019). Notwithstanding, Santos et al., 2019 demonstrated that the femoral diaphyseal cortical thickness displays a moderate correlation coefficient values (r) with stature that reach 0.4 (please see Figure 3 in p. 5), between the anterior and lateral surfaces as well as the whole length of the posterior surface of the femoral diaphysis. Obviously, the range of the NF distribution and the correlations with cross sectional properties and diaphyseal thickness are comparable explaining why the CNF and APNF measurements retained significant relation to femur length regardless the NF location along the diaphysis.

The general trend in this study as well as similar studies is the low association between mediolateral diameters at both levels (midshaft and NF) and FL in comparison with other femoral measurements in both sexes (Simmons et al., 1990; Nanayakkara et al., 2018). A possible explanation is that the mediolateral dimensions have more correlation with relative body breadth and body mass even at the level of midshaft as it was stated above

(Davies & Stock, 2014). Ruff (1995) stated that the femoral shaft increases in mediolateral breadth in all land vertebrates towards the epiphyses because it is important to increase the mediolateral breadth of the knee for weight transfer across the joint. Another study by Agostini and Ross (2011) showed that the body mass index (BMI) has a significant effect on the mediolateral diameters rather than anteroposterior diameters at different sections of bone length with the exception of the 65% of femur shaft length location measured from the distal end proximally. The difference between normal-weight and obese individuals in midshaft diaphyseal dimensions is great enough that Agostini and Ross (2011) were able to correctly classify 88% and 77% of normal-weight and overweight individuals, respectively, using the femur midshaft ML dimension. Eliott et al (2016) noted that a series of medio-lateral shaft breadths at different levels from 20% to 80% of FL consistently performed better than VHD in body mass estimation. Moreover, Ruff (1991) and Elliot et al (2016) utilized the femoral neck width (FNW) in body mass estimation because it may exhibit a pattern of correlation with body mass between that of the head and shaft due to its intermediate location. In accordance with Steele and Mckern, 1969, the PENF was the least reliable measurement for estimating FL.

For the archaeological remains, mathematical techniques are often employed for stature estimation because the anatomical methods require the summation of the measurements from the cranium through foot bones (Raxter et al., 2008). From the methodological standpoint, the selection of a particular linear regression model for body size/height estimation should be based on the most significant variables biologically and statistically, the ease of measurements collection, presence of a suitable reference sample (Giannecchini and Moggi-Cecchi, 2008) as well as robust line fitting approach (Sjøvold,1990; Holliday and Ruff, 1997; Konigsberg et al., 1998; Auerbach and Ruff, 2004; Raxter et al., 2006). There is a smidgen of studies that investigate the accuracy of FL and stature reconstruction in archeological samples and accounts for the discovery of fragmentary femora through the mathematical (regression) approach (Shuler et al., 2011; Mays, 2016; Meyer et al., 2020).

In the current study, we observed comparable correlation between the articular breadths and diaphyseal measurements and the measured FL in both samples which explain the considerable accuracy obtained by our models. In fact, we note that the most evident differences were in the correlation between APD and FL being higher in ancient Egyptians (Figure 1). There are a myriad of different factors that can affect the femoral morphology, robusticity, and rigidity such as physical activity, muscle strength and postadulthood weight changes (Wescott & Zephero, 2016). However, the extent to which diaphyseal dimensions are influenced by one, or any combination of these factors is not fully understood (Pearson and Lieberman 2004). The mechanical loading of a long bone is not only a function of physical activity and muscle strength, but also of its linear dimensions and body weight (Ruff et al., 1991). Moreover, the mechanical loading and activity related effects may not be the same in past and present populations (Wright, & Vásquez, 2003; Ruff 1994; Eliott, 2016).

Although data regarding activity patterns and muscle mass are not available in the present research, Modern Egyptians of both sexes are taller and heavier than ancient Egyptians and these changes are statistically significant for both sexes ($p < 0.001$) (please see Table 20 in Raxter's thesis). The differences are for males only 1.5 cm and 11.9 kg, while the differences for females are 4.8 cm and 11.2 kg in mean height and body mass, respectively. Therefore, mechanical loading may not be the same in the past and present groups (Elliott et al., 2016). While the pooled sex sample of ancient Egyptians have femoral epiphysial widths and diaphyseal length similar to those of modern Egyptians, the mid-diaphyseal width measurements are significantly different. A relatively more pronounced increase of MLD measurements in the modern sample rather than APD was noted which may suggest that the contemporary Egyptian populations in our sample are experiencing favorable living conditions allowing for increasing the body weight as compared to those of the ancient period. It is also not surprising that the APD has increased over time in the pooled modern sample due to increased stature which increase the anteroposterior bending stress at the femoral mid-shaft (Ruff et al., 2006b, Premory and Zakrzewski, 2009). While increased mechanical loading induces increased apposition and/or decreased resorption rates of diaphyses during life, the epiphyses do not have these structural changes which may be attributed to the physiological constraints on joint remodeling (Ruff, Scott, & Liu, 1991; Ruff, Trinkaus, Walker, Larsen, 1993; Lieberman, Devlin, Pearson, 2001). Therefore, articular external dimensions appear to be less confounded by mechanical loading changes than the diaphyseal morphology (Lieberman et al., 2001). This was evident in Figure 1 where the coefficients of determination R^2 were comparable in the epiphysial ends models while differences were noted in the APD model.

In ancient Egyptians, the APD alone performed slightly better than the two midshaft measurements model and achieving more or less balanced accuracy of FL reconstruction in males and females, probably due to the higher coefficient of determination of APD in the ancient Egyptians sample. In stature estimation, however, both models achieved similar accuracies albeit the APD+MLD model provided more precise estimates in males (i.e., least MAD values). Previous studies established that prediction of height from FL can be improved by the addition of APD as a width measurement from the same bone or the calculation of height directly (Porter, 2002; Porter, 1999; Reynolds, MacGregor, Alston-Knox, Meredith & Barry, 2018). Similarly, the inclusion of midshaft width measurement (APD) decrease SEE and improve FL prediction due to the strong correlation with FL (Nanayakkara et al., 2018; Wescott & Zephro, 2012).

Another observation is the higher prediction accuracy of FL in females using DEB model than males in the ancient Egyptians sample. There is an obvious relationship between DEB and sex because females having much smaller measurements overall; as FL and DEB increase and subsequently the stature increase. Reynolds et al., 2018 showed that DEB had a higher inclusion probability than sex in their regression models. This supports their claims that distal bicondylar breadth should replace sex when estimating stature because DEB has higher inclusion propability in their models than sex. Moreover, a multitude of studies recorded higher correlation coefficients of DEB with FL in females than in males for example in the indigenous South Africans (Bidmos, 2008a), r=0.746 vs 0.560, respectively and in the European descent South African (Bidmos, 2008b), r= 0.722 vs 0.400, respectively. In Simmons et al, 1990, the same pattern was preserved being r=0.537 in females vs 0.521 in males.

Despite sex-based drifts are seen in the stature estimations being slightly underestimated in males and over-estimated in females by the four pooled sex FL models, these drifts were still within the errors seen in other sex-dependent stature estimation formulae. The tendency to underestimate stature in the ancient Egyptians males is largely attributable to the relatively smaller epiphyseal and diaphyseal measurements in this sample compared to the contemporary Egyptian sample and subsequently underestimation of FL. In contrast, the predicted FL was more profoundly underestimated in four male individuals from OK with large femora and another one was overestimated in short individual from El-Hesa. Similar biased results were reported in other studies employed OLS regressions near the extremes of the size distribution of the calibration sample (Ruff et al., 2012; De Groote & Humphrey, 2011). Male individuals from the OK have larger FL than the following dynasties (Raxter, 2011). Nevertheless, MAD of all the FL reconstruction models was small (<2 cm) for the luxury of making no assumptions about sex. Considering the more pronounced secular changes in Egyptian females' stature in comparison to males (as mentioned above), the non-isometric FL changes in relation to stature among females over this period might lead to higher MAD of the estimated stature in females than males in the 2-step approach. In fact, the MAD values in females were consistently lower than males according to Table 7. Therefore, the generic femur length reconstruction formulae do not appear to significantly affect the accuracy of stature estimates when separating the sample by sex.

OLS regression equations are useful for predicting the values of dependent variable (FL) from the independent variables of femoral width measurements within the observed range of the calibration population (Konigsberg et al., 1998). Figure 1 showed that FL estimation using equations devised from the Ancient Egyptians as caliberation sample to the same target archeological population did not perform much better than those based on the contemporary population. These results also emphasize that proximal lower limb bone (femur) is less susceptible to environmental stressors as compared to the distal parts of lower limb (tibia and fibula), leading to greater consistency of femoral metrics regardless of reference population (Pomeroy et al., 2012; Anzellini and Toyne, 2019; Mahakkanukrauh et al., 2011; Albanese et al., 2016). Considering the same ecogeographic zone as a criterion to select the representative population and the similar moderate-to-high correlations between femoral measurements and FL obtained in the present study, these findings confirm the usefulness of our models and reduce the potential errors (Mays, 2016; Béguelin, 2011).

We presented multiple sets of equations available for use in fragmented femora contexts, some of them based solely on measurements from the epiphyseal ends of femur, and the others depend upon the presence of a portion of the diaphysis with or without the proximal epiphysis. In summary, then, the following procedures are recommended when reconstructing the femur length in Egyptian assemblages or fragmented femora: (1) There is generally a dominant foramen or multiple nutrient foramina in the middle third of the diaphysis however, this is an unknown variable that should be accounted for. Previous clinical studies demonstrate that femora may have as many as nine NF, and that the majority of femora across populations have more than one NF (Murlimanju, Prashanth, Prabhu, Chettiar et al., 2011; Kawasaki, Kinose, Kato, Sakai, et al. 2019; Mazengenya & Fasemore 2015). Nevertheless, the dominant nutrient foramina can be identified by the elevated margins and distinct groove leading to the external orifice and by being capable of admitting at least the tip of a 24-gauge hypodermic needle while those smaller should be excluded (Johnson, Beckett, & Márquez-Grant, 2017). Although there are advantages in the diaphysial measurements for FL reconstruction of poorly preserved skeletons, some problems in the applicability of the proposed method should be carefully considered such as absence of NF. Fortunately, this condition is rarely reported in different populations/ancestries (please see Table 1 in Murlimanju et al. 2011; Kizilkanat, Boyan, Ozsahin, Soames, et al., 2007; Bridgeman & Brookes, 1996 and Table 3 in Mazengenya & Fasemore, 2015), (2) The CNF model allows estimates based on extremely fragmentary femoral evidence when typically less than 40% of bones are preserved in a non-diagnostic region like the diaphysis (Feldesman & Lundy, 1988) and/or the midshaft point can not be accurately determined (Jerković, Bašić, Kružić & Anđelinović, 2016), and (3) When both epiphyses are not present, we may apply this method without knowing the status of epiphyseal union because age at death can be determined from other parts of the excavated skeletons or inferred from the roughness of linea aspera (Sołtysiak, 2015) which is most frequently preserved in in extremely shattered archeological materials (Gidna & Domínguez-Rodrigo, 2013).

The *limitations* of the study should be acknowledged. *First*, we proposed a series of generic equations for estimating the FL which are best suited for bioarchaeological studies. Generic equations are bet-hedging strategies that minimize the potential wrongful selection of the model or loss of information due to inapplicability in unknown/ambiguously sexed specimens and technical difficulties in estimating the age from fragmentary remains or commingled contexts in addition to the use of few skeletal metric predictors (Feldesman & Fountain, 1996; Meyer et al., 2020; Albanese et al., 2016; Reynolds et al., 2018; Nikita & Chovalopoulou, 2017; Anzellini & Toyne, 2019). Moreover, the combined-sex equations are derived from larger sample size than each sex alone, and include wide spectrum of variation to provide the best fit of the line to the data and statistically more robust (Albanese et al., 2016).

Second, we could not test the new subset of measurements in the ancient Egyptians sample because only 5 measurements were common between both datasets. However, the regression models diagnostics indicated statistical significance at p=0.000, R^2 >15% and small standard error of the regression coefficients (b), thus these statistics should refer to the robustness of these models (Prasad et al., 1996). Further, we rigorously tested the relationship between the nutrient foramen and mid-shaft measurements. The descriptive statistics indicated that the mean differences between the circumference measurement from the midshaft level to NF level were slightly higher than anteroposterior measurements at both levels which may be attributed to uneven projections of linea aspera expressed along the diaphysis (Polguj, Bliźniewska, Jędrzejewski, Majos & Topol, 2013) in addition to the increased mediolateral dimension towards the NF level. As expected, the anteroposterior diameter and circumference measurements at the NF level were highly and significantly correlated with their respective measurements at the midshaft (please see the footnote below Table 2). Thus, these measurements should produce nearly identical FL estimates when applied to their respective regression equations, and it would be safe to use the the best bivariate regression model in the case of femoral fragmentation.

Another shortcoming is the use of indirect approach for stature calculation i.e., applying two separate formulae one to estimate FL then inserting it into Raxter's formulae for stature estimation (Raxter et al., 2008), thereby compounding the error (Bidmos, 2009). Direct calculation of equations that relate transverse measurement or segment length to body height showed lower SEE than that would pertain after indirect stature estimation by applying a stature regression formula to estimated bone length (Bidmos, 2008; Bidmos, 2009). Despite the adjusted SEE was high due to consideration of the compound error, the resulting MAD of stature ranged from 3.55 to 4.38 cm. Similarly, Fongkete et al., 2016 presented a comparison between the direct and indirect stature estimation methods based on the same skeletal reference collection in Chiang Mai University that was used by Mahakkanukrauh et al., 2011 to derive stature regression equations using the FL employed in the indirect 2-step method. They found small differences in the values of SEE in both methods signifying comparable performance on the Thai population. Therefore, the presented approach might be considered as a complementary method with reasonable degree of accuracy in absence of regression formulae for femur length estimation and direct estimation of stature from fragments of long bones since past and present populations may vary in body shape and proportions (Shuler et al., 2011).

Conclusions

To combat the recovery of partial remains in archeological settings and to maintain high analytical capabilities in any skeletal assemblages, we report on new standards designed for use on a case-by-case basis to estimate FL from its fragments in Egyptian specimens. In the modern calibration sample, we demonstrated that the dominant NF and the linea aspera, can be used as stable landmarks for collecting the new measurements APNF, and CNF and estimation of FL. The CNF model can be used when only a piece of femoral diaphysis is found. Recommendations are also made for estimating FL from the formula of both THD and APNF variables which presents the highest multiple correlation coefficient and least SEE.

The dominant NF is defined by its larger size than other accessory nutrient foramina that may exist along the diaphysis however, rare anatomical variations could lead to absence of these NF. Our results, combined with previous data on the statistical significance of APD in estimating stature, reveal that the APD model provided the most consistent FL and stature estimates when applied to the ancient test sample. Therefore, good performance of THD +APNF model could be anticipated because small differences are present between the mean of the anteroposterior measurements at both levels. However, caution should be practiced with femoral fragments recovered from the Old Kingdom period because they have, in general, larger femora than the following dynastic periods. These findings have several potential values not only for boosting our ability to analyze body-size variations in ancient Egyptians from the available femoral fragments, but also they provide additional insights on the relative resistance of certain femoral metrics to change through time. Consequently, further research in this area should be pursued.

Data Availability Statement

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Received 28 August 2020; accepted for publication 2 March 2021.

Literature Cited

Abledu, J. K., E. B. Offei, and C. K. Osabutey. 2016. Reconstruction of femoral length from fragmentary femora. *Anat. Cell Biol.* 49:206–209.

Agostini, G. M., and A. H. Ross. 2011. The effect of weight on the femur: A cross sectional analysis. *J. Forensic Sci.* 56:339–343.

Albanese, J., A. Tuck, J. Gomes et al. 2016. An alternative approach for estimating stature from long bones that is not population- or group-specific. *Forensic Sci. Int.* 259:59–68.

Anzellini, A., and J. M. Toyne. 2019. Estimating the stature of ancient high-altitude Andean populations from skeletal remains of the Chachapoya of Peru. *Am. J. Phys. Anthropol.* 171:539–549.

Attia, M. H., F. M. M. Badr El-Dine, M. H. Attia et al. 2020. Sample specific sex estimation method using parameters around dominant nutrient foramina in unknown femur bones collection of contemporary Egyptians. *Int. J. Osteoarchaeol.* 30:583–593.

Auerbach, B. M. 2011. Methods for estimating missing human skeletal element osteometric dimensions employed in the revised fully technique for estimating stature. *Am. J. Phys. Anthropol.* 145:67–80.

Auerbach, B. M., and C. B. Ruff. 2004. Human body mass estimation: A comparison of “morphometric” and “mechanical” methods. *Am. J. Phys. Anthropol.* 125:331–342.

Auerbach, B. M., and C. B. Ruff. 2006. Limb bone bilateral asymmetry: Variability and commonality among modern humans. *J. Hum. Evol.* 50:203–218.

Bass, W. 2005. *Human Osteology: A Laboratory Manual and Field Manual*. 5th ed. Columbia, MO: Missouri Archaeological Society.

Béguelin, M. 2011. Stature estimation in a central Patagonian prehispanic population: Development of new models considering specific body proportions. *Int. J. Osteoarchaeol.* 21:150–158.

Bidmos, M. A. 2008a. Estimation of stature using fragmentary femora in indigenous South Africans. *Int. J. Legal Med.* 122: 293–299.

Bidmos, M. A. 2008b. Stature reconstruction using fragmentary femora in South Africans of European descent. *J. Forensic Sci.* 53:1,044–1,048.

Bidmos, M. A. 2009. Fragmentary femora: Evaluation of the accuracy of the direct and indirect methods in stature reconstruction. *Forensic Sci. Int.* 192:131.e1–131.e5.

Bridgeman, G., and M. Brookes. 1996. Blood supply to the human femoral diaphysis in youth and senescence. *J. Anat.* 188:611–621.

Buck, A. 2010. Recognising anatomical landmarks in fragmented skeletal remains. *Pathology* 4:S26.

Chiba, F., Y. Makino, S. Torimitsu et al. 2018. Stature estimation based on femoral measurements in the modern Japanese population: A cadaveric study using multidetector computed tomography. *Int. J. Legal Med.* 132:1,485–1,491.

Collins, M. J., C. M. Nielsen-Marsh, J. Hiller et al. 2002. The survival of organic matter in bone: A review. *Archaeometry* 44:383–394.

Curate, F., J. Coelho, D. Gonçalves et al. 2016. A method for sex estimation using the proximal femur. *Forensic Sci. Int.* 266:579.e1–579.e7.

Davies, T. G., and J. T. Stock. 2014. The influence of relative body breadth on the diaphyseal morphology of the human lower limb. *Am. J. Hum. Biol.* 26:822–835.

De Groote, I., and L. T. Humphrey. 2011. Body mass and stature estimation based on the first metatarsal in humans. *Am. J. Phys. Anthropol.* 144:625–632.

Elliott, M., H. Kurki, D. A. Weston et al. 2016. Estimating body mass from skeletal material: New predictive equations and methodological insights from analyses of a known-mass sample of humans. *Archaeol. Anthropol. Sci.* 8:731–750.

Feldesman, M. R., and R. L. Fountain. 1996. "Race" specificity and the femur/stature ratio. *Am. J. Phys. Anthropol.* 100:207–224.

Feldesman, M. R., and J. K. Lundy. 1988. Stature estimates for some African Plio-Pleistocene fossil hominids. *J. Hum. Evol.* 17:583–596.

Fongkete, I., P. Singsuwan, S. Prasitwattanaseree et al. 2016. Estimation of stature using fragmentary femur and tibia lengths in a Thai population. *Aust. J. Forensic Sci.* 48:287–296.

Fully, G., and H. Pineau.1960. Détermination de la stature au moyen du squelette. *Ann. Méd. Lég.* 40:145–153.

Giannecchini, M., and J. Moggi-Cecchi. 2008. Stature in archeological samples from central Italy: Methodological issues and diachronic changes. *Am. J. Phys. Anthropol.* 135:284–292.

Gidna, A. O., and M. Domínguez-Rodrigo. 2013. A method for reconstructing human femoral length from fragmented shaft specimens. *Homo.* 64:29–41.

Goldman Osteometric Data Set. https://web.utk.edu/~auerbach/GOLD.htm (accessed July 2020).

Hens, S. M., L. W. Konigsberg, and W. L. Jungers. 2000. Estimating stature in fossil hominids: Which regression model and reference sample to use? *J. Hum. Evol.* 38:767–784.

Holliday, T. W., and C. B. Ruff. 1997. Ecogeographical patterning and stature prediction in fossil hominids: Comment on M.R. Feldesman and R.L. Fountain. *Am. J. Phys. Anthropol*. 103:137–140.

Hoppa, R. D., and K. L. Gruspier. 1996. Estimating diaphyseal length from fragmentary subadult skeletal remains: Implications for palaeodemographic reconstructions of a southern Ontario ossuary. *Am. J. Phys. Anthropol.* 100:341–354.

Jerkovic, I., Ž. Bašic, I. Kruži et al. 2016. Sex determination from femora in late antique sample from Eastern Adriatic coast (Salona necropolis). *Anthropol. Rev.* 79:59–67.

Johnson, V., S. Beckett, and N. Márquez-Grant. 2017. Differentiating human versus non-human bone by exploring the nutrient foramen: Implications for forensic anthropology. *Int. J. Legal Med.* 131:1,757–1,763.

Kawasaki, Y., S. Kinose, K. Kato et al. 2019. Anatomic characterization of the femoral nutrient artery: Application to fracture and surgery of the femur. *Clin. Anat.* 33:479–487.

Kizilkanat, E., N. Boyan, E. T. Ozsahin et al. 2007. Location, number and clinical significance of nutrient foramina in human long bones. *Ann. Anat.* 189:87–95.

Komar, D. A., and W. E. Potter. 2007. Percentage of body recovered and its effect on identification rates and cause and manner of death determination. *J. Forensic Sci.* 52:528–531.

Konigsberg, L. W., S. M. Hens, L. M. Jantz et al. 1998. Stature estimation and calibration: Bayesian and maximum likelihood perspectives in physical anthropology. *Yearb. Phys. Anthropol.* 41:65–92.

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

Mahakkanukrauh, P., P. Khanpetch, S. Prasitwattanaseree et al. 2011. Stature estimation from long bone lengths in a Thai population. *Forensic Sci. Int.* 210:279.e1–279.e7.

Mays, S. 2016. Estimation of stature in archaeological human skeletal remains from Britain. *Am. J. Phys. Anthropol.* 161:646–655.

Mazengenya, P., and M. D. Fasemore. 2015. Morphometric studies of the nutrient foramen in lower limb long bones of adult black and white South Africans. *Eur. J. Anat.* 19:155–163.

McHenry, H. M. 1974. How large were the australopithecines? *Am. J. Phys. Anthropol.* 40:329–340.

Meeusen, R. A., A. M. Christensen, and J. T. Hefner. 2015. The use of femoral neck axis length to estimate sex and ancestry. *J. Forensic Sci.* 60:1,300–1,304.

Meyer, S., N. Frater, R. Seiler et al. 2020. Multidisciplinary studies of heavily fragmented and commingled ancient Egyptian human remains found in KV 40 (Valley of the Kings, Luxor, Egypt): A pragmatic workflow and first results. *J. Archaeol. Sci. Rep.* 29:102069.

Murlimanju, B., K. Prashanth, L. V. Prabhu et al. 2011. Morphological and topographical anatomy of nutrient foramina in the lower limb long bones and its clinical importance. *Australas. Med. J.* 4:530–537.

Nanayakkara, D., A. Vadysinghe, and L. S. Nawarathna. 2018. Prediction of the length of the femur from its fragments in a Sri Lankan population. *Aust. J. Forensic Sci.* 52:303–312.

Nikita, E., and M.-E. Chovalopoulou. 2017. Regression equations for the estimation of stature and body mass using a Greek documented skeletal collection. *Homo.* 68:422–432.

Nissen, N., E. M. Hauge, B. Abrahamsen et al. 2005. Geometry of the proximal femur in relation to age and sex: A cross-sectional study in healthy adult Danes. *Acta Radiol.* 46:514–518.

Pearson, O. M., and D. E. Lieberman. 2004. The aging of Wolff's "law": Ontogeny and responses to mechanical loading in cortical bone. *Am. J. Phys. Anthropol.* 125:63–99.

Polguj, M., K. Bliźniewska, K. Jędrzejewski et al. 2013. Morphological study of linea asperavariations: Proposal of classification and sexual dimorphism. *Folia Morphol. (Warsz.)* 72:72–77.

Pomeroy, E., J. T. Stock, S. Stanojevic et al. 2012. Trade-offs in relative limb length among Peruvian children: Extending the thrifty phenotype hypothesis to limb proportions. *PLoS One* 7:e51795.

Pomeroy, E., and S. R. Zakrzewski. 2009. Sexual dimorphism in diaphyseal cross sectional shape in the medieval Muslim population of Écija, Spain, and Anglo-Saxon Great Chesterford, UK. *Int. J. Osteoarchaeol.* 19:50–65.

Porter, A. M. 1999. The prediction of physique from the skeleton. *Int. J. Osteoarchaeol.* 9:102–115.

Porter, A. M. W. 2002. Estimation of body size and physique from hominin skeletal remains. *Homo.* 53:17–38.

Prasad, R., S. Vettivel, L. Jeyaseelan et al. 1996. Reconstruction of femur length from markers of its proximal end. *Clin. Anat.* 9:28–33.

Raxter, M. H. 2011. Egyptian body size: A regional and worldwide comparison. PhD diss., University of South Florida.

Raxter, M. H., B. M. Auerbach, and C. B. Ruff. 2006. Revision of the Fully technique for estimating statures. *Am. J. Phys. Anthropol.* 130:374–384.

Raxter, M. H., C. B. Ruff, A. Azab et al. 2008. Stature estimation in ancient Egyptians: A new technique based on anatomical reconstruction of stature. *Am. J. Phys. Anthropol.* 136:147–155.

Reynolds, M. S., D. M. MacGregor, C. L. Alston-Knox et al. 2018. Bayesian modeling predicts age and sex are not required for accurate stature estimation from femoral length. *Forensic Sci. Int.* 289:452.e1–452.e14.

Ruff, C. B. 1995. Biomechanics of the hip and birth in early *Homo*. *Am. J. Phys. Anthropol.* 98:527–574.

Ruff, C. B., B. M. Holt, M. Niskanen et al. 2012. Stature and body mass estimation from skeletal remains in the European Holocene. *Am. J. Phys. Anthropol.* 148:601–617.

Ruff, C. B., B. M. Holt, V. Sladek et al. 2006. Body size, body proportions, and mobility in the Tyrolean "Iceman". *J. Hum. Evol.* 51:91–101.

Ruff, C. B., M. Niskanen, H. Maijanen et al. 2019. Effects of age and body proportions on stature estimation. *Am. J. Phys. Anthropol.* 168:370–377.

Ruff, C. B., W. W. Scott, and A. Y. C. Liu. 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., E. Trinkaus, A. Walker et al. 1993. Postcranial robusticity in *Homo*. I: Temporal trends and mechanical interpretation. *Am. J. Phys. Anthropol.* 91:21–53.

Santos, F., and J. A. Lacoste. 2019. Diaphysator: An online application for the exhaustive cartography and user-friendly statistical analysis of long bone diaphyses. *Am. J. Phys. Anthropol.* 169:377–384.

Sciulli, P. W., K. N. Schneider, M. C. Mahaney. 1990. Stature estimation in prehistoric Native Americans of Ohio. *Am. J. Phys. Anthropol*. 83:275–280.

Shuler, K. A., M. E. Danforth, and J. Auerbach. 2011. Challenges in approaches to skeletal stature estimation: An example from prehistoric eastern Mississippi and western Alabama. *Southeast. Archaeol.* 30:377–389.

Simmons, T., R. L. Jantz, and W. M. Bass. 1990. Stature estimation from fragmentary femora: A revision of the Steele method. *J. Forensic Sci.* 35:628–636.

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

Sołtysiak, A. 2015. Vascular grooves on human femora and tibiae as a potential activity-related trait. *Int. J. Osteoarchaeol.* 25:345–351.

Steele, G., and T. McKern. 1969. A method for assessment of maximum long bone length and living stature from fragmentary long bones. *Am. J. Phys. Anthropol.* 31:215–227.

Stock, J. T., and C. N. Shaw. 2007. Which measures of diaphyseal robusticity are robust? A comparison of external methods of quantifying the strength of long bone diaphyses to cross-sectional geometric properties. *Am. J. Phys. Anthropol.* 134:412–423.

Swinscow, T. D., and M. J. Campbell. 1997. *Statistics at Square One.* 9th ed. London: BMJ Publishing Group.

Timonov, P., and A. Fusova. 2018. Reconstruction of femur length from its proximal fragments in a Bulgarian modern population. *Aust. J. Forensic Sci.* 50:403–413.

Tomsová, J., and Z. Shierová. 2016. Skeletal material from Deir El-Medina in the Egyptological collection of the Hrdlička Museum of Man in Prague. *Ann. Náprstek Mus.* 37:41–69.

Torimitsu, S., Y. Makino, H. Saitoh et al. 2015. Stature estimation in Japanese cadavers based on pelvic measurements in three-dimensional multidetector computed tomographic images. *Int. J. Legal Med.* 129:633–639.

Wescott, D. J., and L. Zephro. 2012. Secular change in the femur. *Am. J. Phys. Anthropol.* 147:301.

Wescott, D. J., and L. R. Zephro. 2016. Secular change in the femur diaphyseal biomechanical properties of American whites. *Hum. Biol.* 88:38–46.

Wright, L. E., and M. A. Vásquez. 2003. Estimating the length of incomplete long bones: Forensic standards from Guatemala. *Am. J. Phys. Anthropol.* 120:233–251.

Table 1. The Femoral Measurements

Parameter	Acronym	Reference
Standard femur measurements		
Maximal length of femur	LF	Auerbach & Ruff, (2006)
Vertical head diameter	VHD	Auerbach & Ruff, (2006)
Transverse head diameter	THD	Auerbach & Ruff, (2006)
Femoral neck axis length	FNAL	Curate et al, (2016)
Femoral neck width	FNW	Curate et al, (2016)
Midshaft mediolateral diameter	MLD	Auerbach & Ruff, (2006)
Midshaft antero-posterior shaft diameter	APD	Auerbach & Ruff, (2006)
Mid shaft circumference	CF	Bass 2005
Distal biepicondylar breadth	DEB	Auerbach & Ruff, (2006)
New set of measurements around the dominant nutrient foramen of femur		
Proximal end to nutrient foramen	PENF	Attia et al., 2020
Medio-lateral diameter at nutrient foramen	MLNF	Attia et al., 2020
Antero-posterior diameter at nutrient foramen	APNF	Attia et al., 2020
Circumference at nutrient foramen	CNF	Attia et al., 2020

Table 2. Descriptive Statistics for Femoral Parameters (in mm) Measured in the Whole Training Sample of Contemporary Egyptians (n=64)

	Descriptive statistics					**Intraobserver error**	**Pearson's r Correlation**
	Mean	Median	SD	Min	Max	**ICC**†	**FL**†
FL	432.13	437.50	30.318	360	495	0.995***	1
VHD	42.47	42.00	3.413	35	49	0.993***	0.674***
THD	43.44	43.00	3.903	33	50	0.984***	0.725***
DEB	76.97	76.50	6.299	62	91	0.991***	0.694***
FNW	32.77	33.00	3.706	25	41	0.883*	0.548***
FNAL	90.19	90.00	6.215	77	105	0.994***	0.679***
APD ‡	27.81	28.00	3.585	18	39	0.990***	0.610***
MLD §	26.97	27.00	2.851	15	32	0.994***	0.492***
CF ((	88.38	90.00	9.604	58	120	0.989***	0.596***
APNF ‡	27.61	27.00	3.494	18	39	0.994***	0.610***
MLNF §	27.63	27.50	2.898	18	34	0.994***	0.470***
CNF ((	89.78	90.00	9.543	60	120	0.986***	0.618***
PENF	201.53	190.00	47.806	115	290	0.999***	0.339**

† All are significant at p=0.000 except FNW in ICC; p=0.010, PENF in Pearson's correlation; p= 0.006.

‡ The mean differences between measurements at mid shaft to NF level= -0.2mm, The correlation coefficient r= 0.905, at P=0.000

§ The mean differences between both measurements= 0.66 mm, The correlation coefficient r= 0.886, at P=0.000

((The mean differences between both measurements= 1.8mm, The correlation coefficient r=0.929 at *P*=0.000

Table 3. Descriptive Statistics for Femoral Parameters (in mm) Measured in the Overall Test Sample of Ancient Egyptians (n=73)

	Descriptive statistics					**Pearson's r Correlation**
	Mean	Median	SD	Min	Max	FL
FL	431.61	430	30.69	361	500	1
THD	42.54	42.54	3.32	34.82	49.37	0.723***
DEB†	75.54	75.75	4.82	65	86	0.689***
APD ‡, **	26.05	26.31	3.21	19.25	31.79	0.759***
MLD‡, ***	24.69	24.79	2.26	20.56	30.96	0.581***

† Sample size (n=70)

‡ In comparison to the mean values of the respective measurements in the modern Egyptian sample.

** Significant at p=0.003

***Significant at p<0.00001

Table 4. Linear Regression Models Predicting FL

Model		B	Std. Err	R	R Square	Adjusted R Square	SEE† (SEE%)
1	(Constant)	177.855	35.503	0.674	0.454	0.446	22.576 (5.2%)
	VHD	5.987	0.833				
2	(Constant)	187.341	29.609	0.725	0.526	0.519	21.034 (4.9%)
	THD	5.635	0.679				
4	(Constant)	133.279	41.097	0.679	0.461	0.453	22.428 (5.2%)
	FNAL	3.314	0.455				
5	(Constant)	288.709	23.868	0.610	0.372	0.362	24.223 (5.6%)
	APD	5.157	0.851				
8	(Constant)	286.112	24.308	0.610	0.372	0.361	24.228 (5.6%)
	APNF	5.289	0.874				
10	(Constant)	255.835	28.637	0.618	0.382	0.372	24.026 (5.6%)
	CNF	1.964	0.317				
12	(Constant)	175.220	34.006	0.694	0.481	0.473	22.018 (5.1%)
	DEB	3.338	0.440				
13‡ (All variables)	(Constant)	170.331	28.244	0.769	0.592	0.578	19.688 (4.6%)
	THD	4.346	0.758				
	APNF	2.645	0.846				
14‡ (Mid shaft)	(Constant)	245.397	30.024	0.648	0.420	0.401	23.458 (5.4%)
	APD	4.104	0.947				
	MLD	2.691	1.191				
15‡ (NF)	(Constant)	255.835	28.637	0.618	0.382	0.372	24.026 (5.6%)
	CNF	1.964	0.317				
16‡ (Proximal end only)	(Constant)	133.789	37.352	0.750	0.562	0.548	20.384 (4.7%)
	THD	3.861	1.030				
	FNAL	1.448	0.647				

† SEE, Standard Error of the Estimate; all the coefficients were statistically significantly different from zero at p=0.0000 except FNAL 0.029; % SEE= SEE/mean

‡ Models 13-16 are Stepwise selected (Criteria: Probability-of-F-to-enter <= 0.050, Probability-of-F-to-remove >= 0.100)

Table 5. Assessment of the Performance of the Generic Models on the Reference Sample (in mm)

	N	Accuracy within 1SEE	N	Accuracy within 2 SEE	MD	MAD
VHD	42/64	65.3%	62/64	96.9%	0.009	18.42
THD	43/64	67.19%	62/64	96.9%	0.014	16.51
FNAL	43/64	67.19%	61/64	95.3%	-0.035	18.31
DEB	42/64	65.3%	63/64	98.4%	-0.017	17.35
APD	46/64	71.9%	61/64	95.3%	-0.013	18.86
APNF	44/64	68.8%	62/64	96.9%	-0.26	19.13
CNF	43/64	67.19%	62/64	96.9%	-0.04	19.23
THD+APNF	40/64	62.5%	63/64	98.4%	-0.012	15.92
THD+FNAL	45/64	70.3%	62/64	96.9%	0.032	16.14
APD+MLD	43/64	67.19%	62/64	96.9%	0.013	18.59

Table 6. Validation of the Generic Equations on the Ancient Egyptians Sample (Values Are Reported in mm)†

Model	Sex	N	Accuracy within 1 SEE	N	Accuracy within 2 SEE	MD	MAD
THD	Males	33/45	73.3%	39/45	86.67%	7.96	18.49
	Females	20/28	71.4%	28/28	100%	-0.85	16.04
DEB	Males	30/44	68.2%	39/45	86.67%	8.69	19.29
	Females	20/26	76.9%	26/26	100%	-4.36	14.24
APD	Males	32/45	71.1%	40/45	88.89%	13.82	18.78
	Females	20/28	71.4%	28/28	100%	0.10	15.62
APD+MLD	Males	31/45	68.9%	39/45	86.67%	9.92	17.18
	Females	20/28	71.4%	28/28	100%	0.21	15.24

† Mean values of maximal femur length in males= 445.07 mm and in females=409.96 mm

Table 7. Stature Prediction Using Reconstructed Femur Length Fragments in Males and Females Ancient Egyptians Using a Two-Step Procedure (Values Are Reported in cm)†‡

		Adj. Std Error	N	Accuracy within 1 SEE	N	Accuracy within 2 SEE	MD	MAD
THD	Males	7.96	39/45	86.67%	45/45	100%	1.42	4.08
	Females	7.43	25/28	89.29%	28/28	100%	-0.74	3.73
DEB	Males	8.12	37/44	84.09%	44/44	100%	1.58	4.32
	Females	7.86	22/26	84.61%	26/26	100%	-1.53	3.66
APD	Males	8.64	41/45	91.11%	45/45	100%	2.74	4.15
	Females	8.13	27/28	96.43%	28/28	100%	-0.51	3.65
APD+MLD	Males	8.33	41/45	91.11%	45/45	100%	1.86	3.82
	Females	7.81	27/28	96.43%	28/28	100%	-0.49	3.55

† Mean of reference stature in males= 164 cm and in females=152.39 cm. The stature was corrected for age according to recommendations of Raxter.

‡ Raxter's formula of males: 2.257 (FML)+63.93+/- 3.218; females: 2.340 (FML)+56.99 +/- 2.517

Figures Caption

Figure 1. Scatterplots of data from various measurement (a: THD, b: DEB, c: APD) Vs.FL in the Ancient and modern Egyptian with a fit line and 95% prediction interval for each group: outliers of the ancient Egyptian sample are discussed in the text (Ancient Egyptians, n=73; Modern Egyptians, n=64).

Figure 1.

A)

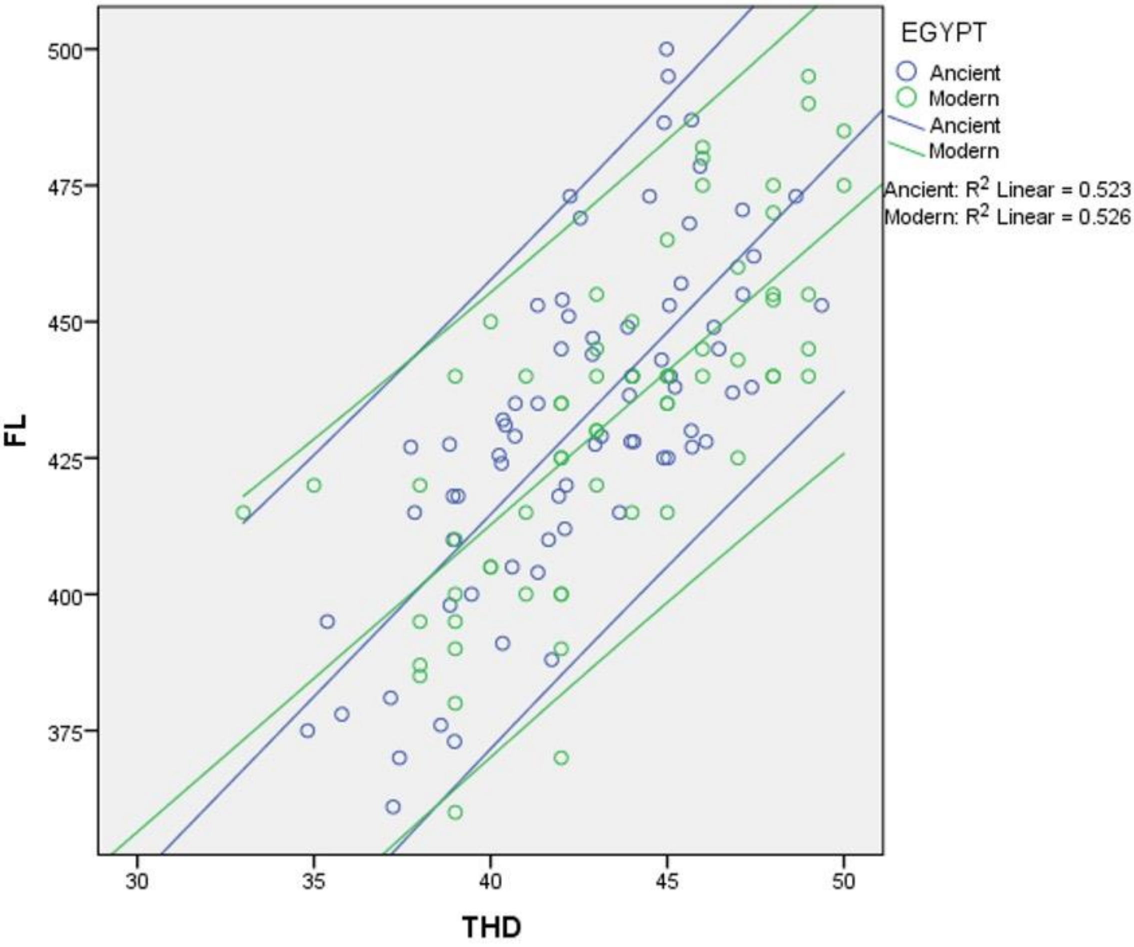

Figure 1. (cont)

B)

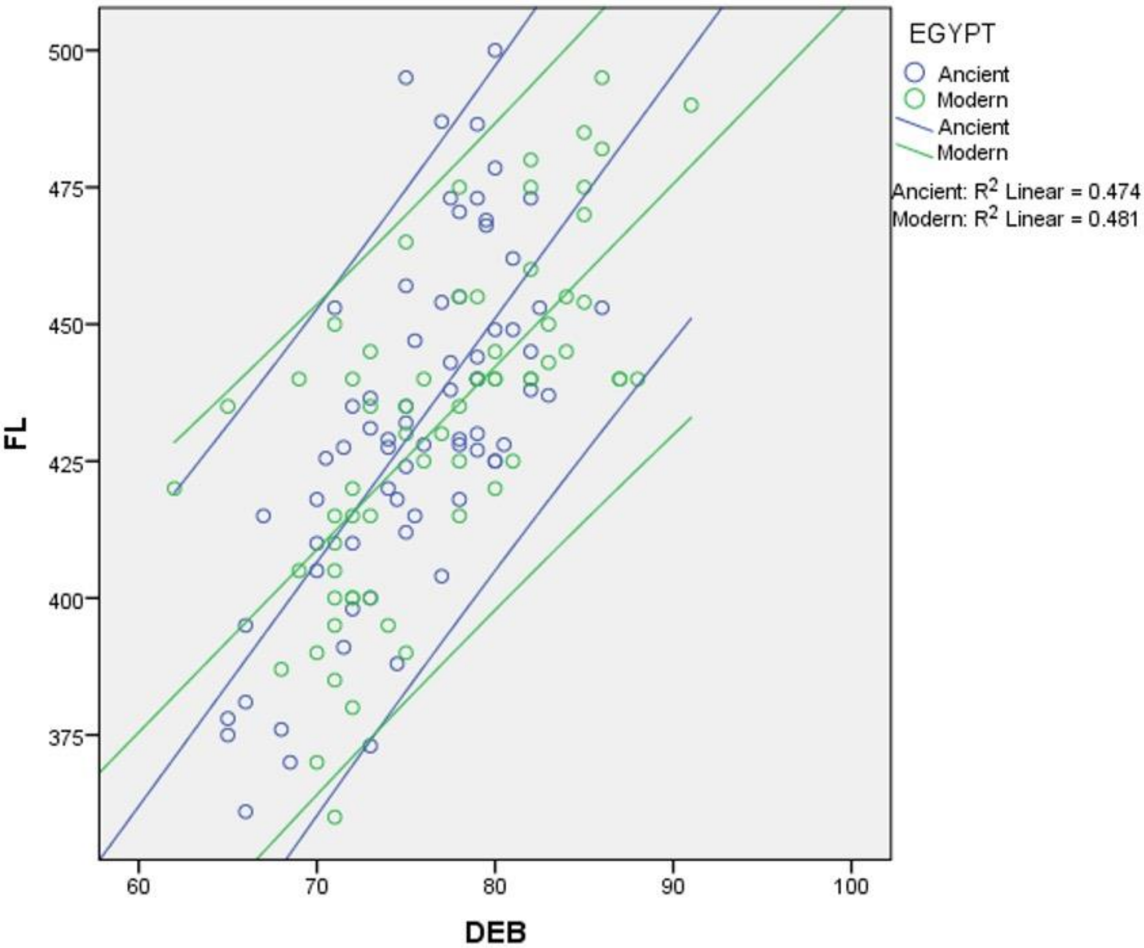

Figure 1. (cont)

C)