Science & Justice 61 (2021) 555–563

Contents lists available at ScienceDirect

Science & Justice

journal homepage: www.elsevier.com/locate/scijus

Technical Note

Testing the validity of population-specific sex estimation equations: An evaluation based on talus and patella measurements

Lara Indra [a,*], Werner Vach [b], Jocelyne Desideri [c], Marie Besse [c], Sandra L. Pichler [b]

[a] Institute of Forensic Medicine, University of Bern, Sulgenauweg 40, 3007 Bern, Switzerland
[b] Integrative Prehistory and Archaeological Science, University of Basel, Spalenring 145, 4055 Basel, Switzerland
[c] Laboratoire d'archéologie préhistorique et anthropologie, University of Geneva, 66, Boulevard Carl-Vogt, 1211 Genève 4, Switzerland

ARTICLE INFO

Keywords:
Metric sex estimation
Discriminant function analysis
Talus
Patella
Forensic anthropology
Forensic identification
Swiss SIMON Identified Skeletal Collection

ABSTRACT

Sex estimation is essential for forensic scientists to identify human skeletal remains. However, the most sexually dimorphic elements like pelvis or skull are not always assessable. Osteometric analyses have proven useful in sex estimation, but also to be population specific. The main purpose of this study was to test the validity of contemporary Greek and Spanish discriminant functions for the talus and the patella, respectively, on a Swiss skeletal sample and to quantify the utility of the measurements as a novel approach in osteometric sex assessment.

Four talus and three patella measurements on dry bone were obtained from 234 individuals of the modern cemetery SIMON Identified Skeletal Collection. The previously derived discriminant functions were applied, accuracies determined, the utility of the different measurements was assessed and new multivariable equations constructed.

Accuracies varied between 67% and 86% for talus and 63% and 84% for patella, similar to those reported by the original studies. Multivariable equations should be preferred over equations based on single measurements and combining the most significant measurements rather than using several variables obtained the best possible accuracy. The new discriminant functions did not provide a substantial improvement to the original ones. The overall utility of talus and patella is limited, allowing sex estimation with sufficient certainty only in a small proportion of individuals.

Discriminant functions developed in contemporary Greek or Spanish populations are in principle applicable also to Swiss contemporary populations. We recommend that at present existent studies of this type should be validated and tested rather than developing new formulas.

1. Introduction

Sex estimation is of central importance in analysing human skeletal remains and one of the first steps in establishing a biological profile and in the process of identification, both in individual and demographic studies [1]. In forensic cases for example, the estimation of sex eliminates around 50% of individuals from the list containing missing persons [2]. There are several approaches to estimate sex in skeletal remains, including morphologic, metric, geometric morphometric and molecular methods on either dry bone or via imaging techniques like 2D radiographs, computed tomography (CT), and magnetic resonance imaging (MRI) [3]. Metric sex estimation of skeletal remains is highly valuable in court testimonies because it is a standardised, statistically approved method quantified by error rates [4]. Collection of measurements requires little previous experience while still achieving high inter-rater agreement [5]. In the metric approach, sex can then be evaluated statistically, such as by discriminant function analysis (DFA). DFA is helpful in cases of highly decomposed, fragmented, badly preserved or commingled remains, or in cases where the most useful bones in sex assessment, such as the pelvis or cranium, are absent [6]. Use of DFA for sexing postcranial elements increased in the last decade with applications on dry bone [e.g. 7-11], in geometric morphometrics [12], magnetic resonance imaging (MRI) [13], digital 2D-radiographs [14] and computed tomography (CT) [e.g. 15–17]. Although some of these

* Corresponding author.
E-mail addresses: lara.indra@irm.unibe.ch (L. Indra), werner.vach@unibas.ch (W. Vach), jocelyne.desideri@unige.ch (J. Desideri), marie.besse@unige.ch (M. Besse), sandra.pichler@unibas.ch (S.L. Pichler).

https://doi.org/10.1016/j.scijus.2021.06.011
Received 30 October 2020; Received in revised form 27 May 2021; Accepted 30 June 2021
Available online 2 July 2021

studies obtain accuracies above 90%, Dorado-Fernández et al. [18] found that results are method-dependent, thus making comparison difficult.

In addition, since sexual dimorphism is population specific, DFAs established for one population may be inaccurate for another [e.g. 10,19–21]. However, Alonso-Llamazares and Pablos [10] state that if the populations from which the formulas were derived exhibit similar proportions to the group the formulas are applied to, these formulas can still obtain high accuracy rates. Rather than developing new population-specific formulas, existent studies should therefore be validated and checked for their applicability to other groups. Yet, only a few inter-population validation studies of metric sex estimation approaches have been published, covering just a small range of postcranial skeletal elements [e.g. 10,22–24]. With this study we aim precisely at validating two specific studies of Peckmann et al. [25,26] presenting sex estimation on the talus in a Greek and on the patella in a Spanish sample and testing whether these formulas are also valid when applied to an outgroup, i.e., individuals from the SIMON Identified Skeletal Collection (Switzerland). The Greek and Spanish sample formulas will be applied to the Swiss sample and validated by considering accuracy rates. In addition, we will quantify the utility of the measurements considered and create new equations for the Swiss (central European) population.

Foot bones are often well preserved, especially in forensic cases where feet are sometimes covered and protected by socks and shoes [27]. The calcaneus for example is a useful bone in sex estimation due to its weight-bearing nature [e.g. 10,27,28]. However, we selected the talus because of its size and compact morphology, good preservation and its higher accuracy rates in sex estimation compared to the calcaneus [e.g. 10,29]. The talus has further shown to be useful for sex-specific stature estimation by linear regression equations [30]. Previous metric studies on sex estimation of the talus reported accuracy rates of 80% and above [e.g. 5 for South African Whites; 8 for Portuguese; 10,29 for American Blacks and Whites; 31 for European Americans; 32 for Koreans; 33 for northern and southern Italians].

The patella also shows good resistance to postmortem changes, it is often recovered in forensic cases and its position between the sexually dimorphic femur and tibia suggests dimorphism as well [34,35]. Metric sex estimation on the patella returned varying accuracies, but all studies produced at least one equation that achieved 80% accuracy or more on dry bone [e.g. 34 for southern Italians; 36 for South African whites; 37 for Iranians; 38 for African Americans; 22 for prehistoric Germans] or using imaging techniques [e.g. 39 for Japanese; 40 for Turks; 13 for Iranians; 17 for Chinese]. In a meta-analytic review, Dorado-Fernández et al. [18] show that maximum width, height and thickness of the patella serve well for sex estimation and that the reviewed studies achieve a high heterogeneity in their outcomes.

2. Material and methods

2.1. Material

The study sample comprised individuals from the SIMON Identified Skeletal Collection housed at the Laboratory of prehistoric archaeology and anthropology at the University of Geneva in Switzerland. The collection holds a total of 495 named skeletons from modern decommissioned cemeteries from Canton Vaud, Switzerland, for whom the date of birth and death and occupation are known [40-42]. The SIMON Identified Skeletal Collection is an established repository of human skeletal remains in accordance with Swiss legislation. It was set up as a reference collection with the specific purpose of supporting research and for teaching. The curatorial committee granted us access to the collection for our non-invasive study, taking ethical issues into account (for further information about the collection see Supplementary Text S1, Figure S1 and Table S1).

The study sub-sample consists of 117 males and females (n = 234), who died between 1930 and 1960 with the earliest birth in 1856 and the latest in 1927. The females represent all females from the SIMON collection for whom both talus and patella could be analysed, while the males were matched to the female sample by number and age. The 234 individuals were distributed among three age groups from 20 to 60 + years. Distribution of males and females among the age groups was approximately equal with 33 females and males in age group 20–39, 45 males and 41 females in age group 40–59 and 39 males and 43 females in age group 60 + years (Fig. 1). The mean age for both, males and females, is 52 years and the median for both is 53 years (males between 20 and 79 years, females between 20 and 85).

2.2. Measurements

We used primarily left-side tali and patellae since the studies of Peckmann et al. [25,26] found no statistically significant differences between measurements of left and right-side bones. In 16 cases where the left patella was either absent or had to be excluded due to severe damage or pathologies, we used the right-side patella instead. All measurements were taken blindly by one person (L.I.) with a digital sliding caliper graduated to 0.01 mm and without prior knowledge of the demographic data. The measurements were taken following the instructions of Martin (1914) [in 44] and Martin's established numeration was also used. Of the six measurements used by Peckmann et al. [26] on the patella and the nine measurements used by Peckmann et al. [25] on the talus, we selected those returning high accuracy rates in sex estimation in the original publications [25,26] and because they are frequently used in sexing studies. For the talus, we measured talar length (Tal1), talar width (Tal2), trochlear length (Tal4) and trochlear breadth (Tal5). For the patella, maximum height (Pat1), maximum breadth (Pat2) and thickness (Pat3) measurements were taken. Fig. 2 depicts the measurement locations on talus and patella.

2.3. Statistical methods

Sex estimation using the four single measurements of the talus and the three single measurements of the patella were based on the sectioning points from the original publications, shown in Table 1. Individuals with a measurement above the corresponding sectioning point were rated as male, those below were rated as female.

For the sex estimation based on groups of measurements, we used two discriminant functions, one for the talus and one for the patella. The following linear combinations are described in the original publications [25,26] and are to be compared with 0 (talus equation) and 0.376 (patella equation):

Talus equation y = (0.239*Tal2) + (0.155*Tal1) – 18.442

Patella equation y = (0.245*Pat1) + (0.102*Pat2) + (0.024*Pat3) – 14.741

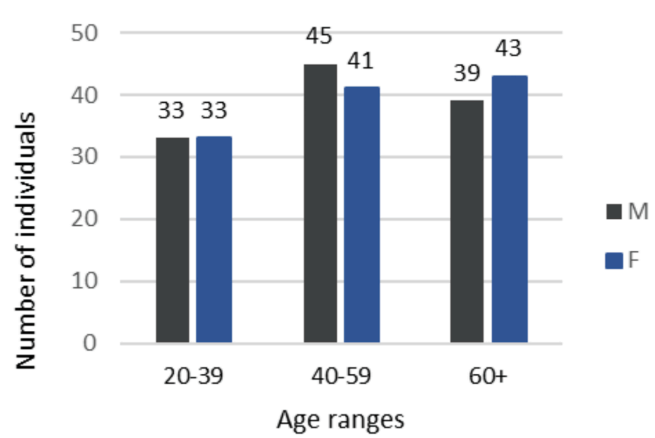

Fig. 1. Distribution of male and female individuals among the three age ranges of 20–39, 40–59 and 60 + years.

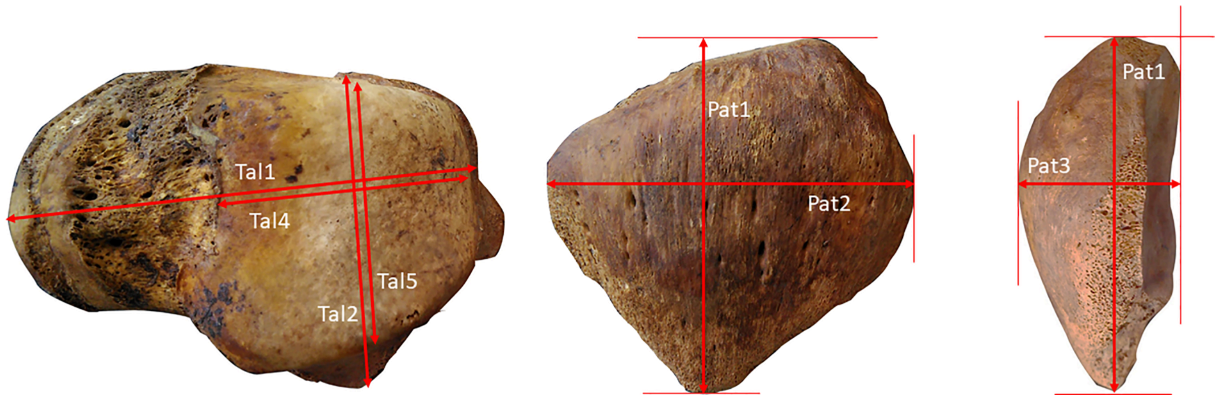

Fig. 2. Superior view of the talus (left), anterior and lateral (middle and right) view of the patella with measurements marked after Martin 1914 in Bräuer [43].

Table 1
Number of individuals measured (N), mean values in mm and standard deviations stratified by sex for the seven traits in the original [25,26] and our study and sectioning points obtained in this study.

	Male						Female						Both
	Original study			Our study			Original study			Our study			
Variable	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD	Sectioning point (mm)
Tal1	81	61	3.93	117	56	2.84	71	53.9	2.9	117	50.2	2.59	57.43
Tal2	81	42.3	2.58	117	43.7	2.45	71	37.6	2.07	117	38.8	2.39	39.94
Tal4	81	34.6	2.73	117	37	2.6	66	31	2.1	117	32.8	2.13	32.78
Tal5	49	32.2	1.96	117	33.4	1.93	71	28.6	1.52	117	29.8	1.75	30.42
Pat1	55	42.9	3.03	117	44.2	3.09	51	37.9	3	117	38.7	2.54	40.4
Pat2	55	44.6	3.28	117	45.3	3.28	51	40.3	2.94	117	40.3	2.85	42.46
Pat3	55	20.3	1.95	117	22.7	2.07	50	18.1	1.78	117	20.3	1.7	19.23

The statistical analyses of the data were performed using IBM SPSS software program version 21.0.

To assess intra-observer variation, the first ten individuals were re-measured right after they had been measured first (short-term reproducibility) and the measurements of the first twenty individuals were retaken after the scoring of all individuals was complete and after a minimum break period of one week (long-term reproducibility). Both kinds of reproducibility were assessed using Bland-Altman-plots for visualisation and by computing limits of agreement and the absolute technical error of measurement (TEM). Bland-Altman plots show the distribution of the observed differences in relation to the observed mean from each pair of measurements, and the limits of agreement describe the range in which we can find 95% of the observed differences [44]. The absolute TEM reflects the standard error of a single measurement [45].

The distributions of variables are described by means and standard deviations. In addition, the distributions in our study population are visualised by histograms. In order to facilitate comparison, the histograms are further overlaid with the normal distribution curve of the original studies [25,26], based on the published means and standard deviations.

We calculated accuracy rates, defined as the frequency of correct sex identifications, from application of the sectioning points and the two multivariable equations provided in the original publications [25,26]. Values are given for males and females separately as well as for the combination of both. In comparing the accuracy rates between the original studies [25,26] and our study, we consider only apparent accuracy rates and not cross validated accuracy rates, since Peckmann et al. [25,26] reported cross-validated accuracy rates larger than apparent rates without explaining this procedure in the respective papers.

The utility of each measurement was determined as the fraction of subjects for whom we can determine the correct sex with a posterior probability of at least 95%. This posterior probability was computed for each single individual using the formula phi(y, mean_male, sd_male) / phi(y, mean_male, sd_male) + phi(y, mean_female, sd_female) with y denoting the measurement and phi the density of a normal distribution. Results above 0.95 indicate a male, values below 0.05 a female individual. Means and standard deviations were calculated from our data since the observed discrepancies between our findings and those of the original studies invalidates the posterior probabilities based on the original sample.

Partial correlation was used to evaluate the relationship between the talus measurements, patella measurements and between the measurements of talus and patella adjusted for sex.

We further developed new equations for each bone, depending on the four talus and on the three patella measurements, and an equation for both bones combined using full and stepwise linear discriminant analysis (procedure DISCRIMANT in SPSS 21.0). The cross-validated accuracy of these equations was determined by a leave-one-out-cross-validation [46].

3. Results and discussion

3.1. Intra-observer reproducibility

Figs. 3 and 4 show the Bland-Altman plots for the short-term and long-term reproducibility. The limits of agreement are −1.04 to 1.14 for the short-term measurements in the magnitude of 1 mm for raw measurements in the range between 20 and 60 mm, and similar for the long-term reproducibility with −1.04 to 0.99. The results of the Bland-Altman plots for short-term and long-term reproducibility indicate a high reproducibility and no short-term memory effects in handling the caliper. The absolute TEM values were 0.37 (talus) and 0.42 (patella) for the short-term, and 0.35 (talus) and 0.39 (patella) for the long-term comparison. These values lie within the acceptable range.

3.2. Distribution of variables

Table 1 presents the mean values and standard deviations of our data

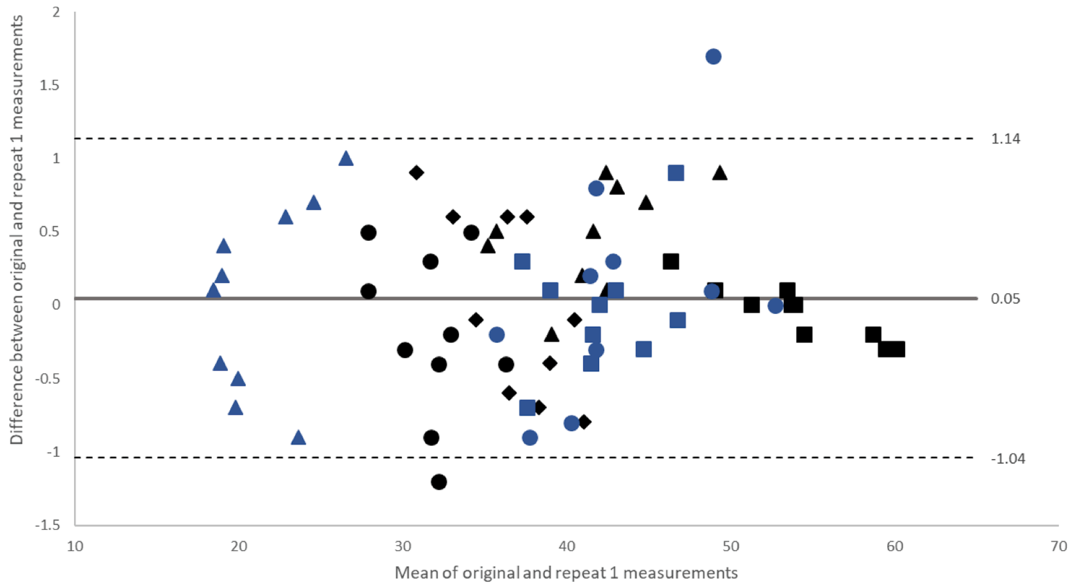

Fig. 3. Bland-Altman plot with data of the original and the repeat 1 measurements showing short-term reproducibility. Talus is black (Tal1 square, Tal2 triangle, Tal4 rhomboid, Tal5 circle), patella blue (Pat1 square, Pat2 circle, Pat3 triangle).

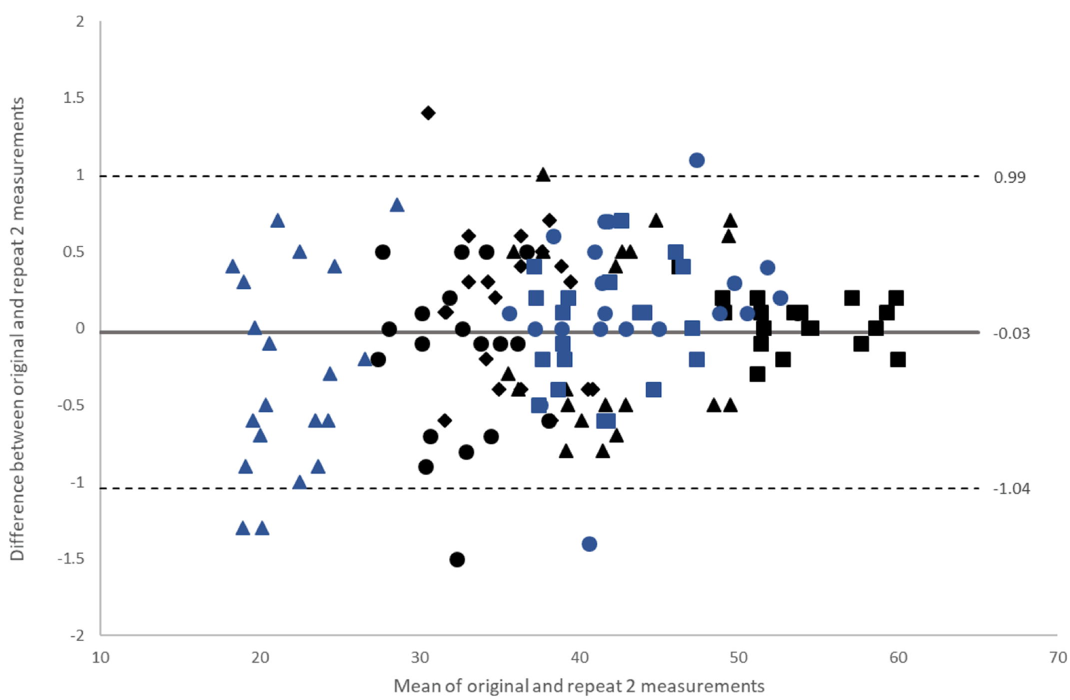

Fig. 4. Bland-Altman plot with data of the original and repeat 2 measurements showing long-term reproducibility. Talus is black (Tal1 square, Tal2 triangle, Tal4 rhomboid, Tal5 circle), patella blue (Pat1 square, Pat2 circle, Pat3 triangle).

and the original studies [25,26], separated by sex and for each measurement. Corresponding to Table 1, Figs. 5 and 6 depict the full distribution of the four talus and the three patella measurements, respectively, from our study compared to the normal curves depicting the distributions in Peckmann et al. [25,26]. When using existing discriminant functions in a new sample, we make the implicit assumption that the sex specific distributions of variables are similar in both the new and original samples / populations [47]. In our study, this assumption was not perfectly met as we observed a tendency towards larger mean values compared to the original studies [25,26]. The only exception is Tal1 for both males and females, where the original study (61 mm for males, 53.9 mm for females) [25] presents distinctly higher values than our study (56 mm males, 50.2 mm females). This discrepancy is apparently caused by the definition of Tal1 used by Peckmann et al. [25], which deviates from Martin's Tal1 [43] and returns larger values. We strictly followed the description given by Martin, thus measuring smaller distances for Tal1 in the Swiss sample. No systematic differences in standard deviations between the original study and our study were observed.

Comparing our sex-specific mean values with further studies, we find that especially trochlear distances of the talus are generally larger in our sample for males and females than other samples. For example, the mean trochlear length in our study is 37/32.8 mm for males and females, respectively, while they are 35.5/32.3 mm in South African Whites [5],

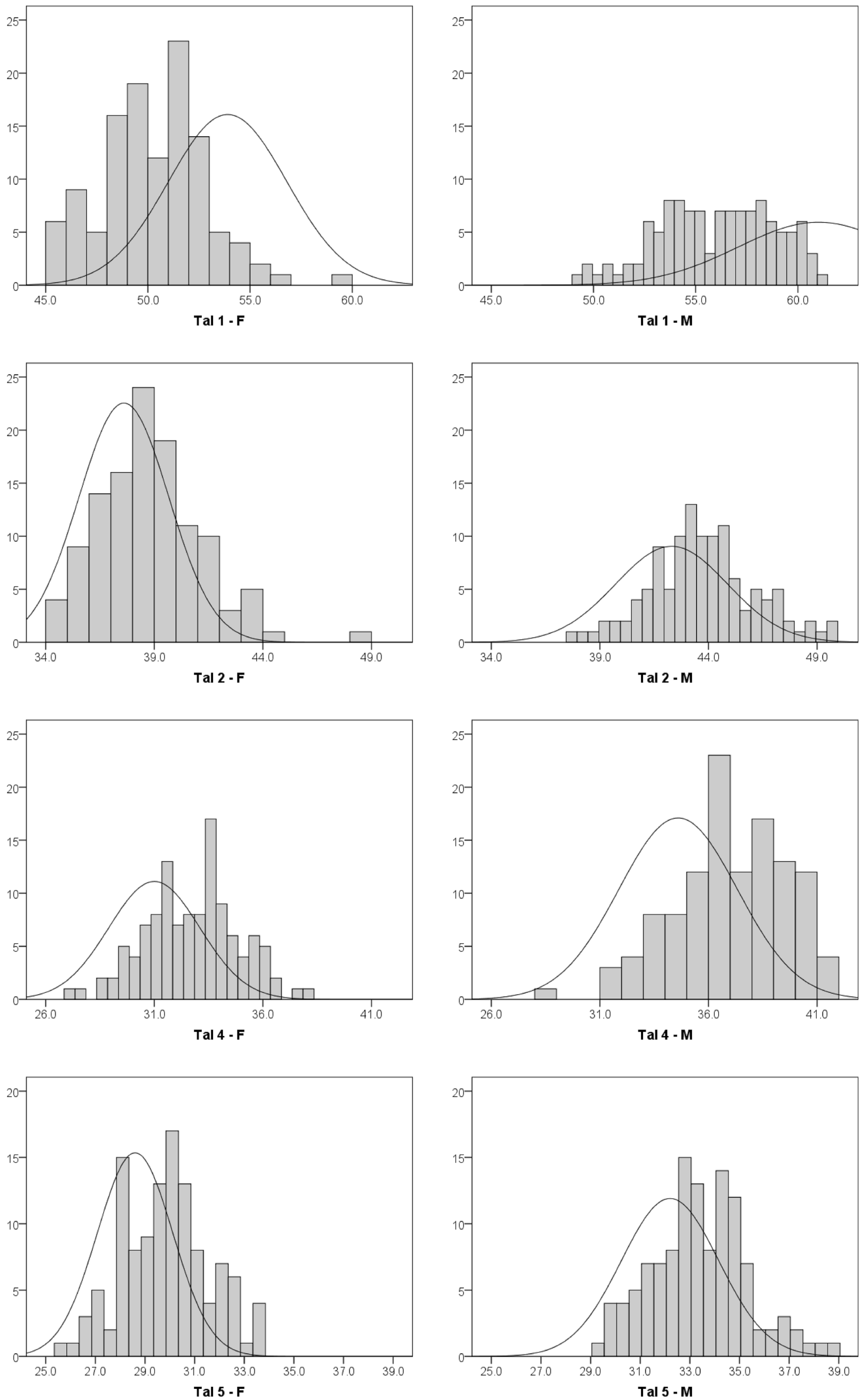

Fig. 5. Histograms with measurements of the four talus measurements from our study overlaid with the distributions observed in Peckmann et al. [25] assuming a normal distribution. Left side females, right side males, x-axis measurements, y-axis frequency.

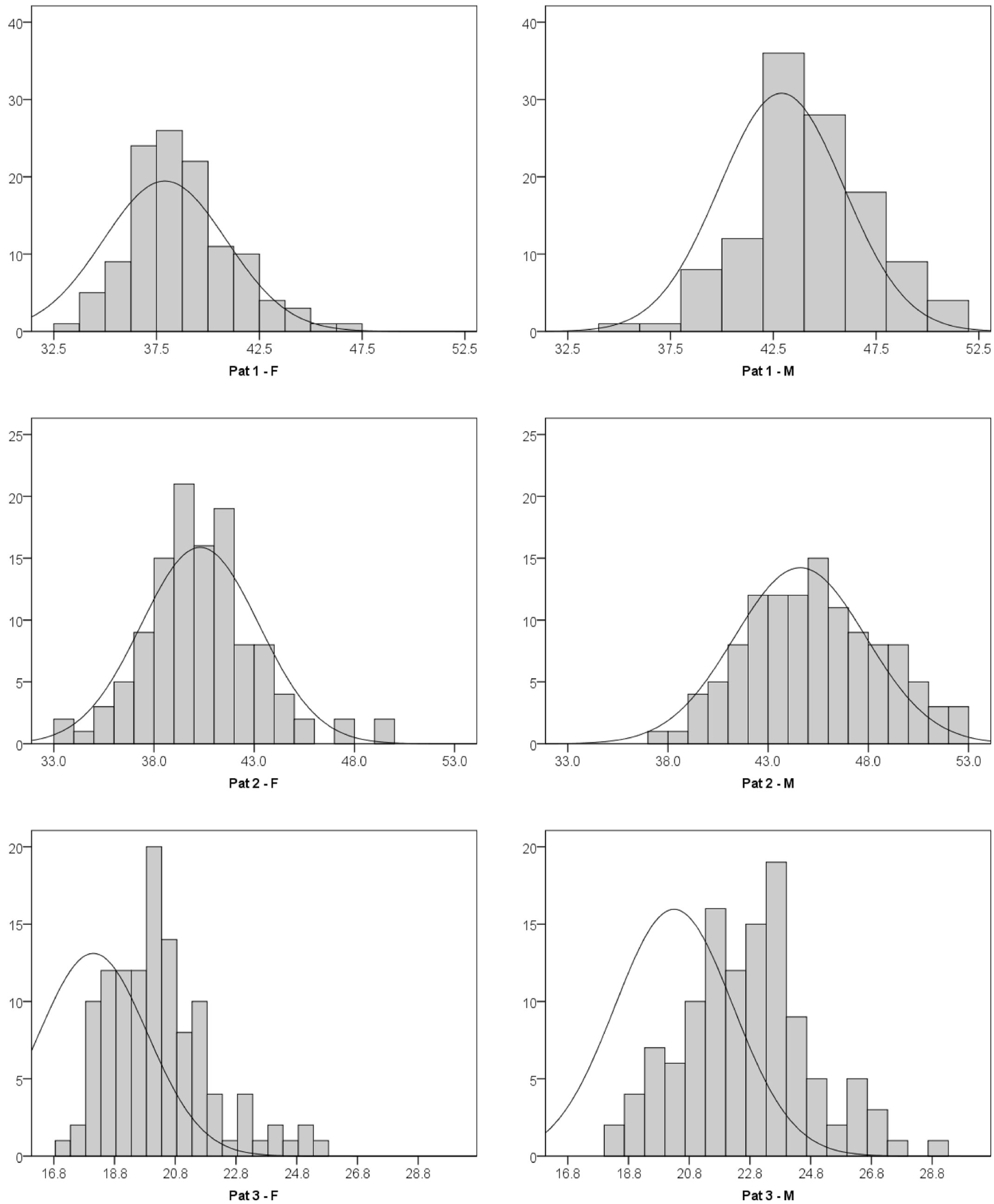

Fig. 6. Histograms with measurements of the three patella measurements from our study overlaid with the distributions observed in Peckmann et al. [26] assuming a normal distribution. Left side females, right side males, x-axis measurements, y-axis frequency.

33.3/30.8 mm in Koreans [32], and 35.0/31.1 mm in North American Whites [10]. We also note that the patella thickness values are usually larger in our sample (22.7/20.3 mm for males and females, respectively) than previously reported on dry bone, e.g. 20.4/18.4 mm for South African Whites [19], 20.4/18.3 mm for Southern Italians [34], and 21.9/20.3 mm for Iranians [36]. Because the measurement technique used can affect the results [18], we did not compare our dry bone sample to samples measured by imaging techniques, such as MRI or CT [e.g. 13,17,39].

3.3. Accuracy rates

Table 2 depicts the apparent accuracy rates for the different measurements and two functions in Peckmann et al. [25,26] and our study. The combined accuracies of the talus measurements tend to be similar or

Table 2
The accuracy rates for the seven single traits and the two multivariable rules in our study and as reported in Peckmann et al. [25, 26], shown sex-specific as well as combined. Only original data results were used.

Variable	Original study (%)			Our study		
	M (%)	F (%)	Combined (%)	M (%)	F (%)	Combined (%)
Tal1	82.7	87.3	85.0	35.04	99.15	67.09
Tal2	76.5	88.7	82.6	94.02	73.50	83.76
Tal4	76.5	81.7	79.1	93.16	49.57	71.37
Tal5	81.5	85.9	83.7	93.16	65.81	79.49
Multivariable	81.1	88.9	84.7	77.78	94.87	86.32
Pat1	–	–	–	89.74	75.21	82.48
Pat2	–	–	–	80.34	82.05	81.20
Pat3	–	–	–	96.58	29.06	62.82
Multivariable	81.8	78.0	80.0	78.63	88.89	83.76

slightly higher in the original study (79.1–85.0%) [25] compared to ours (67.1–86.3%). The only exception was Tal1 with a distinctly lower accuracy of 67.1%, reflecting the substantial difference in mean values. Overall, in our study, single measurement accuracies in males were higher than in females (except for Tal1), reflecting the general tendency to larger values and thus favouring the correct classification of males. However, the multivariable equations are more accurate in females than in males as shown in Table 2.

It is interesting to note that despite the observed differences in sex specific mean values between our study and Peckmann et al. [25,26], we obtained accuracies comparable to the original studies for several measurements (Tal2, Pat1 and Pat2 with accuracy rates above 80%) and in particular for the multivariable equations (86.32% for the talus and 83.76% for the patella).

Our single measurement accuracies obtained lower accuracies (67.1–83.8%) than most of the published literature for talus measurements [10,25,31]. We find the main difference for talar values in Tal1, which obtains the highest accuracy in most studies with up to 90% [10,31]. For the patellar height and breadth (Pat1 and Pat2) we obtained accuracies around 80%, similar to most of the previously published studies [e.g. 13,17,19,22,38]. Our Pat3 accuracy however is generally smaller than the reported values, reflecting the larger measurements taken in our sample compared to the original study [26]. Nevertheless, our multivariable equations resulted in accuracies of 86.3% for talus and 83.8% for patella, similar as reported for both, talus and patella [e.g. 5,13,19,26,31].

3.4. Utility

Table 3 presents the utilities of the seven measurements explained by the number of individuals holding posterior probabilities above 95%. Despite accuracies of 80% and above obtained in this study, with a single measurement sex estimation with a posterior probability of 95% and above can only be made on average in 31.4% of the subjects measuring tali, and on average in 15.4% of the subjects measuring patellae. The multivariable equations show utilities of 46.15% for the talus, 35.04% for the patella and 53.42% for the multivariable equation including all seven traits combined. Actually, these numbers are yet too optimistic, as they refer to the subpopulation of subjects with both patella and talus to be measurable. In our study, several individuals were excluded because either talus or patella were not preserved. This underlines that patella and talus can contribute to sex estimation in combination with other traits, but that they are insufficient when used alone.

Table 3
Utility rates for the seven single traits and the two multivariable rules.

Method	Variable	Number of individuals	Above 95% N	Mean	Utility (%)
Single traits	Tal1	234	90	98.19	38.46
	Tal2	234	74	98.15	31.62
	Tal4	234	55	97.25	23.50
	Tal5	234	75	98.61	32.05
	Average				31.41
	Pat1	234	57	97.71	24.36
	Pat2	234	38	97.30	16.24
	Pat3	234	13	98.48	5.56
	Average				15.39
Multivariable rules	Talus	234	108	98.57	46.15
	Patella	234	82	97.73	35.04
	Talus and patella	234	125	98.69	53.42

3.5. Correlations

Table 4 presents the partial correlation coefficients for all six pairs of talus combinations and all three pairs of patella combinations, while controlling for sex. Moderate (Tal4/Tal5) to strong (Tal1/Pat1) positive correlations can be seen for all combinations of variables. The strongest correlation for either talus or patella is present between length and breadth of the bone and the strongest correlation among the two bones combined is the total length of both (see also Supplementary Figures S2 and S3 for scatter diagrams of the partial correlations). These high correlations explain why the multivariable equations tend to improve the accuracy only to a moderate degree when compared to single measurements.

3.6. Development of new multivariable equations

New full discrimination functions based on all four talus (function 1), on all three patellar (function 2) and on all seven measurements combined (function 3) resulted in linear combinations as shown in Table 5, equations 1 to 3.

The average cross-validated accuracy rates were all in the magnitude of 85% to 88% and hence only slightly larger than those obtained by the previously published equations. All accuracy rates were slightly higher in females than in males, similar to the original studies of Peckmann et al. [25,26].

Using stepwise discrimination analysis for talus and patella separately and combined, the stepwise function 4 for only talus includes variables Tal1, Tal2 and Tal5, the stepwise function 5 for only patella includes Pat1 and Pat2 and the stepwise function 6 for both, talus and patella, includes Pat1, Tal1, Tal4 and Tal5, as shown in Table 5, equations 4 to 6.

The combined cross-validated accuracy rates were similar to those based on the full approach without variable selection, indicating that it is not necessary to measure all traits to obtain an optimal accuracy. This can be explained by the substantial correlation among the variables measured.

Since our accuracy rates from the newly developed multivariate equations are only slightly higher than those we obtained when we applied the equations from the original publications [25,26] to our sample, they still correlate with the majority of published studies as summarised above.

4. Conclusion

This study corroborates the potential role of the talus and patella in sex estimation based on skeletal remains and demonstrates that equations for sex estimation developed in a contemporary Greek and Spanish population are in principle applicable to a Swiss population from the 19th/20th century. We obtained accuracies of 67.1–86.3% for the talus and 62.8–83.8% for the patella. Our study reminds us that differences in population means between different samples are not unlikely, and that this may imply some sex-based asymmetry in the accuracy, even if the average accuracy is acceptable. Multivariable equations can overcome

Table 4
Partial correlation coefficients (r) adjusting for sex for all combinations between talus, patella and between both traits. All values are significant at a level $\alpha = 0.05$.

	Tal1	Tal2	Tal4	Tal5	Pat1	Pat2	Pat3
Tal1	–	0.705	0.643	0.642	0.743	0.679	0.631
Tal2	0.705	–	0.571	0.667	0.739	0.708	0.599
Tal4	0.643	0.571	–	0.491	0.639	0.574	0.544
Tal5	0.642	0.667	0.491	–	0.687	0.650	0.575
				Pat1	–	0.639	0.514
				Pat2	0.639	–	0.582
				Pat3	0.514	0.582	–

Table 5
Direct (1–3) and stepwise (4–6) discriminant function equations for talus only, patella only and both combined.

No.	Variables	Unstandardized coefficients	Wilk's Lambda	Apparent Accuracy(%)			Cross-validated Accuracy (%)		
				M	F	Combined	M	F	Combined
Direct									
1	Tal1	0.147	0.419[a]	86.3	88.9	87.6	86.3	88.9	87.6
	Tal2	0.121							
	Tal4	0.082							
	Tal5	0.152							
	Constant	−20.482							
2	Pat1	0.257	0.495[a]	82.9	88.0	85.5	82.1	88.0	85.0
	Pat2	0.092							
	Pat3	0.057							
	Constant	−15.811							
3	Tal1	0.097	0.385[a]	88.9	91.5	90.2	85.5	89.7	87.6
	Tal2	0.057							
	Tal4	0.084							
	Tal5	0.129							
	Pat1	0.126							
	Pat2	0.039							
	Pat3	−0.015							
	Constant	−21.110							
Stepwise									
4	Tal1	0.182	0.424[a]	86.3	89.7	88.0	86.3	88.0	87.2
	Tal2	0.139							
	Tal5	0.160							
	Constant	−20.467							
5	Pat1	0.267	0.497[a]	83.8	87.2	85.5	82.9	87.2	85.0
	Pat2	0.107							
	Constant	−15.671							
6	Pat1	0.155	0.389[a]	88.0	90.6	89.3	87.2	89.7	88.5
	Tal1	0.117							
	Tal4	0.092							
	Tal5	0.161							
	Constant	−20.924							

[a] Significant ($p < 0.001$).

this issue to some degree, yet we find that fewer variables than used both in this and the original studies seem to suffice to reach the accuracies observed. Overall, talus and patella measurements alone are not sufficient to estimate sex in all individuals, and their role must be perceived as mainly supplementary to other traits. Attention is necessary with regard to the consistency of metric trait definitions across studies. The results we attained in this study clearly underscore the necessity to evaluate and validate existing studies rather than presenting new formulas. As demonstrated here, varying definitions or slightly aberrant interpretations of measurements are only detected in extensive validation studies. Conducting formula validations will decrease the number of independent, single studies without validation of their correctness and increase the quality and credibility of the existent data. Instead of discarding the osteometric method entirely in an identification case of unknown geographic origin, forensic anthropologists could then use formulas from other areas or samples and be justified by the probabilities obtained in such validation studies.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CRediT authorship contribution statement

Lara Indra: Formal analysis, Investigation, Writing - original draft, Visualization. **Werner Vach:** Conceptualization, Methodology. **Jocelyne Desideri:** Resources. **Marie Besse:** Resources. **Sandra L. Pichler:** Conceptualization, Methodology, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would like to thank the staff of the Laboratoire d'archéologie préhistorique et anthropologie at the University of Geneva for their help and organisation during the process of data collection.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scijus.2021.06.011.

References

[1] K.J. Reichs, Forensic Osteology. Advances in the Identification of Human Remains, CC Thomas, Springfield, IL, 1997.
[2] S.K. Rowbotham, Anthropological estimation of sex, in: S. Blau, D. Ubelaker (Eds.), Handbook of Forensic Anthropology and Archaeology, Routledge, New York, 2016, pp. 261–272.
[3] K. Krishan, P.M. Chatterjee, T. Kanchan, S. Kaur, N. Baryah, R.K.A. Singh, A review of sex estimation techniques during examination of skeletal remains in forensic anthropology casework, Forensic Science International 261 (165) (2016) e161–165.e168, https://doi.org/10.1016/j.forsciint.2016.02.007.
[4] M.K. Moore, Sex estimation and assessment, in: E.A. DiGangi, M.K. Moore (Eds.), Research Methods in Human Skeletal Biology, Academic Press, Waltham, 2013, pp. 91–116.

[5] M.A. Bidmos, M.R. Dayal, Sex determination from the talus of south african whites by discriminant function analysis, The American Journal of Forensic Medicine and Pathology 24 (2003) 322-8. https://doi.org/10.1097/01.paf.0000098507.78553.4a.

[6] F. Introna, G. Di Vella, C.P. Campobasso, Sex determination by discriminant analysis of patella measurements, Forensic Science International 95 (1) (1998) 39–45, https://doi.org/10.1016/S0379-0738(98)00080-2.

[7] M.K. Spradley, R.L. Jantz, Sex estimation in forensic anthropology: skull versus postcranial elements, Journal of Forensic Sciences 56 (2011) 289.

[8] D. Navega, R. Vicente, D.N. Vieira, A.H. Ross, E. Cunha, Sex estimation from the tarsal bones in a Portuguese sample: a machine learning approach, International Journal of Legal Medicine 129 (3) (2015) 651–659, https://doi.org/10.1007/s00414-014-1070-5.

[9] M.K. Moore, E.A. DiGangi, F.P. Nino Ruiz, O.J. Hidalgo Davila, C. Sanabria Medina, Metric sex estimation from the postcranial skeleton for the Colombian population, Forensic Science International 262 (2016) 286 e281–288. https://doi.org/10.1016/j.forsciint.2016.02.018.

[10] C. Alonso-Llamazares, A. Pablos, Sex estimation from the calcaneus and talus using discriminant function analysis and its possible application in fossil remains, Archaeological and Anthropological Sciences 11 (9) (2019) 4927–4946, https://doi.org/10.1007/s12520-019-00855-y.

[11] M.A. Bidmos, A.A. Adebesin, P. Mazengenya, O.I. Olateju, O. Adegboye, Estimation of sex from metatarsals using discriminant function and logistic regression analyses, Australian Journal of Forensic Science (2020) 1–14, https://doi.org/10.1080/00450618.2019.1711180.

[12] R. Sorrentino, M.G. Belcastro, C. Figus, N.B. Stephens, K. Turley, W. Harcourt-Smith, T.M. Ryan, S. Benazzi, E. Nikita, Exploring sexual dimorphism of the modern human talus through geometric morphometric methods, PLoS One 15 (2) (2020) e0229255, https://doi.org/10.1371/journal.pone.0229255.

[13] E. Rahmani, S. Mohammadi, A. Babahajian, K. Rahmani, V. Yousefinejad, Anthropometric characteristics of patella for sex estimation using magnetic resonance images, Forensic Imaging 23 (2020) 200412, https://doi.org/10.1016/j.fri.2020.200412.

[14] R. DeSilva, A. Flavel, D. Franklin, Estimation of sex from the metric assessment of digital hand radiographs in a Western Australian population, Forensic Science International 244 (314) (2014) e311–317, https://doi.org/10.1016/j.forsciint.2014.08.019.

[15] M. Djorojevic, C. Roldán, P. García-Parra, I. Alemán, M. Botella, Morphometric sex estimation from 3D computed tomography os coxae model and its validation in skeletal remains, International Journal of Legal Medicine 128 (5) (2014) 879–888, https://doi.org/10.1007/s00414-014-1033-x.

[16] S. Torimitsu, Y. Makino, H. Saitoh, A. Sakuma, N. Ishii, D. Yajima, G.o. Inokuchi, A. Motomura, F. Chiba, R. Yamaguchi, M. Hashimoto, Y. Hoshioka, H. Iwase, Sex estimation based on scapula analysis in a Japanese population using multidetector computed tomography, Forensic Science International 262 (2016) 285.e1–285.e5, https://doi.org/10.1016/j.forsciint.2016.02.023.

[17] M.-J. Zhan, C.-L. Li, F. Fan, K. Zhang, Y.-J. Chen, Z.-H. Deng, Estimation of sex based on patella measurements in a contemporary Chinese population using multidetector computed tomography: An automatic measurement method, Legal Medicine 47 (2020) 101778, https://doi.org/10.1016/j.legalmed.2020.101778.

[18] E. Dorado-Fernández, D.A. Cáceres-Monllor, M.F. Carrillo-Rodríguez, B. Perea-Pérez, M. Botella-López, A meta-analytic review for the patella sexual dimorphism assessment, International Journal of Morphology 38 (4) (2020) 933–939.

[19] M.A. Bidmos, N. Steinberg, K.L. Kuykendall, Patella measurements of South African whites as sex assessors, Homo 56 (1) (2005) 69–74, https://doi.org/10.1016/j.jchb.2004.10.002.

[20] T.D. White, M.T. Black, P.A. Folkens, Human Osteology, Academic Press, Burlington, MA, 2012.

[21] T. Simmons, W.D. Haglund, Anthropology in a forensic context, in: J. Hunter, M. Cox (Eds.), Forensic Archaeology: Advances in Theory and Practice, Routledge, Oxon, UK, 2005, pp. 159–176.

[22] A. Kemkes-Grottenthaler, Sex determination by discriminant analysis: an evaluation of the reliability of patella measurements, Forensic Science International 147 (2-3) (2005) 129–133, https://doi.org/10.1016/j.forsciint.2004.09.075.

[23] M.C. Cuzzulin, F. Curate, A. Rodrigues Freire, S.T. Costa, F. Bevilacqua Prado, E. Daruge Junior E. et al., Validation of anthropological measures of the human femur for sex estimation in Brazilians, Australian Journal of Forensic Science (2020). https://dx.doi.org/10.1080/00450618.2020.1729411.

[24] P. Selliah, F. Martino, M. Cummaudo, L. Indra, L. Biehler-Gomez, C.P. Campobasso, C. Cattaneo, Sex estimation of skeletons in middle and late adulthood: reliability of pelvic morphological traits and long bone metrics on an Italian skeletal collection, International Journal of Legal Medicine 134 (5) (2020) 1683–1690, https://doi.org/10.1007/s00414-020-02292-2.

[25] T.R. Peckmann, K. Orr, S. Meek, S.K. Manolis, Sex determination from the talus in a contemporary Greek population using discriminant function analysis, Journal of Forensic and Legal Medicine 33 (2015) 14–19, https://doi.org/10.1016/j.jflm.2015.03.011.

[26] T.R. Peckmann, S. Meek, N. Dilkie, A. Rozendaal, Determination of sex from the patella in a contemporary Spanish population, Journal of Forensic and Legal Medicine 44 (2016) 84–91, https://doi.org/10.1016/j.jflm.2016.09.007.

[27] O. Ekizoglu, E. Inci, F.B. Palabiyik, I.O. Can, A. Er, M. Bozdag, I.E. Kacmaz, E. F. Kranioti, Sex estimation in a contemporary Turkish population based on CT scans of the calcaneus, Forensic Science International 279 (2017) 310.e1–310.e6, https://doi.org/10.1016/j.forsciint.2017.07.038.

[28] D. Nathena, E. Michopoulou, E.F. Kranioti, Sexual dimorphism of the calcaneus in contemporary Cretans, Forensic Science International 277 (2017) 260.e1–260.e8, https://doi.org/10.1016/j.forsciint.2017.04.005.

[29] D.G. Steele, The estimation of sex on the basis of talus and calcaneus, American Journal of Physical Anthropology 45 (3) (1976) 581–588, https://doi.org/10.1002/ajpa.1330450323.

[30] T.D. Holland, Estimation of adult stature from the calcaneus and talus, American Journal of Physical Anthropology 96 (3) (1995) 315–320, https://doi.org/10.1002/(ISSN)1096-8644.

[31] S.M. Harris, D.T. Case, Sexual dimorphism in the tarsal bones: implications for sex determination, Journal of Forensic Science 57 (2) (2012) 295–305, https://doi.org/10.1111/j.1556-4029.2011.02004.x.

[32] U.Y. Lee, S.H. Han, D.K. Park, Y.S. Kim, D.I. Kim, I.H. Chung et al., Sex Determination from the Talus of Koreans by Discriminant Function Analysis, Journal of Forensic Science 57(1) (2012) 166–171. https://doi.org/10.1111/j.1556-4029.2011.01914.x.

[33] E. Gualdi-Russo, Sex determination from the talus and calcaneus measurements, Forensic Science International 171 (2-3) (2007) 151–156, https://doi.org/10.1016/j.forsciint.2006.10.014.

[34] F. Introna, G. Di Vella, C.P. Campobasso, Sex determination by discriminant analysis of patella measurements, Forensic Science International 95 (1998) 39–45, https://doi.org/10.1016/S0379-0738(98(00080-2.

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

[36] M. Akhlaghi, A. Sheikhazadi, A. Naghsh, G. Dorvashi, Identification of sex in Iranian population using patella dimensions, Journal of Forensic and Legal Medicine 17 (3) (2010) 150–155, https://doi.org/10.1016/j.jflm.2009.11.005.

[37] T.R. Peckmann, B. Fisher, Sex estimation from the patella in an African American population, Journal of Forensic and Legal Medicine 54 (2018) 1–7, https://doi.org/10.1016/j.jflm.2017.12.002.

[38] T. Michiue, A.M. Hishmat, S. Oritani, K. Miyamoto, M.F. Amin, T. Ishikawa, H. Maeda, Virtual computed tomography morphometry of the patella for estimation of sex using postmortem Japanese adult data in forensic identification, Forensic Science International 285 (2018) 206.e1–206.e6, https://doi.org/10.1016/j.forsciint.2017.11.029.

[39] H. Yasar Teke, Ö. Ünlütürk, E. Günaydin, S. Duran, S. Özsoy, S., Determining gender by taking measurements from magnetic resonance images of the patella, Journal of Forensic and Legal Medicine 58 (2018) 87–92, https://doi.org/10.1016/j.jflm.2018.05.002.

[40] I. Gemmerich, Création d'une collection de squelettes récents provenant de cimetières désaffectés du canton de Vaud (Suisse): âge, sexe et liens familiaux connus: présentation de l'organisation du projet [Creation of a collection of modern skeletons from disused cemeteries in the canton of Vaud (Switzerland): known age, sex and kinship: presentation of the organization of the project], Bull Centre Genevois Anthropol 4 (1994) 103–104.

[41] G. Perréard Lopreno, S. Eades, Une démarche actualiste en paléoanthropologie: la collection de squelettes de référence [An up-to-date approach in palaeoanthropology: the skeletal reference collection], in: M. Besse, L.I. Stahl Gretsch, P. Curdy (Eds.) ConstellaSion: hommage à Alain Gallay. Cahiers d'archéologie romande 95 (2003) 463–472.

[42] G. Perréard Lopreno, Adaptation structurelle des os du membre supérieur et de la clavicule à l'activité [Structural adaptation of the bones of the upper limb and clavicle to activity], PhD thesis, Université de Genève, Genève, 2007.

[43] G. Bräuer, . Osteometrie [Osteometrics], in: R. Knussmann (Eds.), Anthropologie. Handbuch der Vergleichenden Biologie des Menschen I/1 [Anthropology. Handbook of comparative human biology I/1], Gustav Fischer Verlag, Stuttgart, 1988, pp. 160–232.

[44] J. Martin Bland, DouglasG. Altman, Statistical methods for assessing agreement between two methods of clinical measurement, The Lancet 327 (8476) (1986) 307–310, https://doi.org/10.1016/S0140-6736(86)90837-8.

[45] T. Adão Perini, G. Lameira de Oliveira, J. dos Santos Ornellas, F. Palha de Oliveira, Technical error of measurement in anthropometry, Revista Brasileira de Medicina do Esporte 11 (2005) 81–85.

[46] IBM SPSS Statistics for Windows, Version 21.0. Armonk, NY, IBM Corp, 2012.

[47] M. Pietrusewsky, Metric analysis of skeletal remains: Methods and applications, in: M.A. Katzenberg, S.R. Saunders (Eds.), Biological Anthropology of the Human Skeleton, John Wiley & Sons, New Jersey, 2008, pp. 487–532.