Anthropol. Anz. 81/2 (2024), 139–151
J. Biol. Clin. Anthropol.
Published online 14 August 2023, published in print March 2024

Article

Comparing morphological (*os coxae*) and metric (long bone length) sex estimation methods in archaeological collections

Ana Curto*[,1,2], Sónia Ferro, Cláudia Santos, Marina Lourenço[3,4], Geraldine E. Fahy, Daniela Anselmo, Teresa Fernandes[2,3]

[1] Laboratório HERCULES, Palácio do Vimioso, Largo Marquês de Marialva, University of Évora, 7000-809 Évora, Portugal
[2] Department of Biology, Pólo da Mitra, University of Évora, 7002-554 Évora, Portugal
[3] Research Centre for Anthropology and Health, Department of Life Sciences, University of Coimbra, 3000-456 Coimbra, Portugal
[4] University of Coimbra, Centre for Functional Ecology, Laboratory of Forensic Anthropology, Department of Life Sciences, Calçada Martim de Freitas, 3000-456 Coimbra, Portugal
* Corresponding author: ana.curto@uevora.pt

With 8 tables

Abstract: *Objectives:* The aim of this study is to evaluate the internal consistency of sex estimation using metric (long bone length) and morphological (*os coxae*) methodologies from different bones and across different archaeological populations from different regions. *Materials and Methods:* Sex was estimated using characteristics of the pelvis and compared with sex estimations using long bone length. Portuguese (659 females; 906 males) and English (141 females; 277 males) archaeological collections were analysed in this study. A set of long bone length functions were developed using one of the archaeological collections (531 females; 600 males) and its coincidence with sex estimated from the pelvis was compared to the coincidence between the pelvis and long bone length sex estimations using functions developed from contemporary collections. Intra- and inter-observer errors were calculated, as well as the sexual dimorphism index for each bone and osteological collection. *Results:* The accuracy of the developed functions and the other methods tested is highly variable, ranging between 25 and 100%. The accuracy of the standard forensic methods varied between collections and analysed bones. *Discussion:* This study reinforces that long bone length is highly population-specific, even between samples of close chronology and geography. Metric methods are good options to strengthen the sex estimations, but they need to be carefully chosen and always report the estimated probability of being male or female in either forensic or archaeological analysis.

Keywords: demography; sexual dimorphism; osteometric; logistic regression

1 Introduction

Estimating human sex is one of the first steps in reconstructing a biological profile of skeletonised remains, both in forensic and archaeological contexts. Sexual dimorphism in bone dimensions develops just before sexual maturity and is related to faster growth in females and an extension of the pre-pubertal growth phase in males (e.g. Willner & Martin 1985; Tanner 1990; Roche 1992; Gasser et al. 2000). Routinely, anthropologists incorporate morphological traits of the pelvis and the cranium into their sex estimations (e.g. Sauter & Privat 1955; Ferembach et al. 1980; Buikstra & Mielke 1985; Uytterschaut 1986; Buikstra & Ubelaker 1994; Bruzek 2002; Bruzek & Murail 2006). When the pelvis and cranium are not available for sex estimations, alternative methods rely on metric analyses of other bones (e.g. Pons 1955; Işcan & Shihai 1995; Silva 1995; Steyn & Işcan 1997; Mall et al. 2000; Wasterlain & Cunha 2000; Mall et al. 2001; Bašić et al. 2013).

Methods relying on bone measurements are highly population-specific and their accuracy can vary regionally and chronologically (e.g. Bidmos & Dayal 2004). Bone length can reveal developmental trends, environmental stress (such as nutritional deficits) and evolutionary relationships (e.g. Bogin 1999; Johnston & Padez 1999; Padez 2003; Padez 2007; Moore & Ross 2013), and are also largely affected by secular changes in stature (e.g. Cardoso & Gomes 2009). Cardoso & Gomes (2009) analysed approximately 7000 years of mean stature in Portugal. They observed a slow increase in stature from prehistory to the Middle Ages, followed by a negative trend to the late 19th century and a rapid increase during the second half of the 20th century. This has

© 2023 E. Schweizerbart'sche Verlagsbuchhandlung, 70176 Stuttgart, Germany
DOI: 10.1127/anthranz/2023/1730

0003-5548/2023/1730 $ 3.25

also been recorded in other studies across Europe (e.g. Arcini et al. 2014).

A few studies have developed metric sex estimation methods for skeletal samples from Portugal. These methods were developed mainly from the Coimbra Identified Skeletal Collection, which dates from the late 19th to the early 20th centuries (e.g. Cunha 1993) and has not been frequently tested in archaeological collections. From Coimbra Identified Skeletal Collection, Pons (1955) examined the sexual dimorphism in the femur using several measurements. Silva (1995) assessed sex using the maximum length of the calcaneus and talus. Wasterlain & Cunha (2000) developed methods to diagnose sex based on the vertical diameter of the femoral head and the vertical diameter of the humeral head. Wasterlain (2000) also analysed the sexual dimorphism in the maximum length of the radius. Using the Identified Skeletal Collection of Museu Bocage of Lisbon University (Luís Lopes Collection), Cardoso & Cunha (2000) evaluated the sexual dimorphism in the upper limb proportions. The epicondylar breadth of the humerus was also explored as a sex discriminant in the Coimbra Identified Skeletal Collection and tested in the Luís Lopes Collection (Albanese et al. 2005). More recently, Curate et al. (2016) also used the Luis Lopes Collection to estimate sex from the proximal femur. Garcia (2012) tested the validity of the circumference at the nutrient foramen of the tibia for sex in the Lisbon Collection of Identified Skeletons and in a medieval collection from Leiria (Portugal). Vertebrae have also been used to estimate sex (Gama et al. 2015; Rozendaal et al. 2020) as well as machine-learning classifiers (Navega et al. 2015; Curate et al. 2017).

The aim of this study is to test whether sex estimation using long bone length methods corresponds to sex estimation using *os coxae* morphology in archaeological collections. We intend to test if metric methodologies (using long bone length) developed from contemporary skeletal collections correspond to morphology methodologies (*os coxae*) in the same way and across different populations from different geographies. Another objective of this study is to compare equations developed from contemporary and archaeological skeletal collections.

2 Material and methods

This study compares sex estimation methods, developed from the long bone length of contemporary and archaeological collections, in Portuguese and English archaeological samples (Table 1). The data from the Portuguese samples (Tomar, Lisbon, Estremoz, Sintra and Santarém) was collected by the authors during archaeological excavations, and the data from the English samples (St Mary Graces, Spital Square, East Smithfield Black Death cemetery, St Benet Sherehog) were accessed online at the Welcome Osteological Research Database (WORD Database 2009). The English archaeologi-

Table 1. Collections used for this study.

Country	City	Site	ID	Chronology	Number of females	Number of males	Total	
Portugal	Tomar	Santa Maria do Olival	SMOL.A	11th–18th centuries	32	155	187	1,268
			SMOL.B		531	600	1,131	
	Estremoz	RossioMarquês de Pombal	RMPE	13th–15th centuries	18	36	54	
	Santarém	Igreja de Santa Cruz	ISCR	14th–18th centuries	20	40	60	
	Lisbon	Rua dos Lagares	RL	? – 15th centuries	47	50	97	
	Sintra	São Miguel de Odrinhas	SMO	11th–15th centuries	11	25	36	
United Kingdom	London	St Mary Graces	MSMG	14th–16th centuries	33	69	102	418
		Spital Square	MSS	12th–14th centuries	17	30	47	
		East Smithfield Black Death cemetery	MESBD	14th century	61	119	180	
		St Benet Sherehog	PMSBS	13th–17th centuries	30	59	89	
Total					568	1,118	1,686	

It was not possible to measure all bones for all the individuals.

cal samples were added to this study to further explore sex estimation between populations from different regions.

Sex was estimated for the Portuguese samples using the methods developed by Phenice (1969) and Buikstra & Ubelaker (1994) using the pelvis. Cranial features were not included as the crania were poorly preserved, especially for some of the collections. If the pelvis features were non-conclusive the individuals were classified as undetermined and not used for this study. Only individuals with clear sex estimations were analysed.

The long bone length was chosen for this study as it is the most registered measurement during archaeological excavations. The ulna was not considered as fewer measurements from this bone were available. The maximum length of each long bone (humerus, femur, radius, tibia) was measured using an osteometric board (Martin 1928; Buikstra & Ubelaker 1994). Not all long bones were present for each skeleton, so the sample size varies depending on the bone analysed. Measurements were preferentially taken from the left side but when it was not possible measurements from the right side were used. Bones with healed fractures, other lesions or taphonomic damage that could affect the maximum length were excluded from this study.

Intra- and inter-observer technical error of measurement (%TEM) and coefficient of reliability (R) were calculated for 80 randomly selected bones (20 of each bone type), following the suggestions of Ulijaszek & Kerr (1999). Descriptive statistical analysis was carried out for each long bone, generating mean values and standard deviations for all the samples analysed. The length of the long bones of males was compared to the length of the long bones from females, using independent samples t-test, as a first assessment as to whether long bone length in these samples was sexually dimorphic. The sexual dimorphism index (SDI) was also calculated (Borgognini Tarli & Repetto 1986) to better understand the difference in the bone length mean between females and males for the different collections.

Sex was estimated for the different samples using functions developed from the maximum length of the humerus (Kranioti & Michalodimitrakis 2009; Charisi et al. 2011), radius (Wasterlain 2000; Charisi et al. 2011) and tibia (Bruzek 1995; Kranioti et al. 2017). Sex estimation using bone length was then compared to sex estimation using *os coxae* (Phenice 1969; Buikstra & Ubelaker 1994) and a percentage of coincident estimations was calculated.

A set of long bone length functions were developed for Tomar's sample to estimate sex. The sex of these individuals is unknown; therefore, they were classified as male or female using sex estimations based on morphological features of *os coxae* (Phenice 1969; Buikstra & Ubelaker 1994). The *os coxae* are considered less population-specific than other bones (Steyn & Patriquin 2009), still, some studies have reported accuracies below 88% (Lovell 1989; Ubelaker & Volk 2002; Steyn & Patriquin 2009) which bias the functions developed from Tomar's collection as the sex was estimated using feature from the *os coxae*.

Tomar's archaeological excavation occurred in two phases (2007/2008 and 2008/2009). Skeletons from the second phase (SMOL.B), corresponding to areas 13 to 20 at the archaeological site, were used for the discriminant function analysis and skeletons from the first phase (SMOL.A) were used to test the functions. Individuals from SMOL.B group were buried closer to the church, potentially having higher socioeconomic status them those buried further from the church (SMOL.A). Data were collected from 1131 adult skeletons (531 females; 600 males) from SMOL.B group, but it was not possible to measure all long bones for all the individuals due to taphonomic degradation. Age was not considered given the difficulty of accurately estimating age in adult individuals (see Merrit 2017).

Following this, logistic regression was used to develop a function for sex estimation:

$$\log\left(\frac{\mu}{1-\mu}\right) = \beta_0 + \sum_i \beta_i x_i \qquad (1)$$

where,
μ – estimated mean for the population
β_0 – intercept term
β_i – slope (expected increment in the response per unit change in x)
x_i – full length of the bone

By applying the expression:

$$p = \frac{1}{1+e^{-(\beta_0+\sum_i \beta_i x_i)}} \qquad (2)$$

where,
p – the probability of being male
β_0 – intercept term
β_i – slope (expected increment in the response per unit change in x)
x_i – full length of the bone

It is also possible to calculate the probability of being male for a specific measurement (x).

The cut-off values were calculated using the unstandardized canonical discriminant functions at group centroids evaluated at group means. To facilitate the use in single bones the cut-off values were also calculated in mm for each analysed bone.

Data were analysed in IBM SPSS® 22.

3 Results

3.1 Intra- and inter-observer errors

The %TEM is below 5% for all the bones measured by the same researcher (intra-observer technical error of measurement) and the R is higher than 0.95 for all the measurements (Table 2), except the tibia's maximum length (R = 0.94). These values are within the acceptable observer error (%TEM < 5%; R > 0.95) suggested by Ulijaszek & Kerr

Table 2. Intra- and inter-observer technical error of measurement (TEM) and coefficient of reliability (R) for each variable.

		TEM	%TEM	R
Intra-observer (N = 20)	Humerus	0.68	1.23	0.96
	Radius	0.45	0.46	0.99
	Femur	0.59	0.89	0.97
	Tibia	0.73	1.49	0.94
Inter-observer (N = 20)	Humerus	0.73	1.78	0.85
	Radius	0.56	0.89	0.92
	Femur	0.63	1.26	0.90
	Tibia	1.22	2.24	0.68

Acceptable human observer error: %TEM < 5% and R > 0.95.

(1999). For the measurements taken by two independent researchers (inter-observer) the %TEM is lower than 5% and the R varied between 0.68 and 0.92 (Table 2).

3.2 Descriptive statistics

The descriptive statistics and the univariate F ratio to measure the differences between both sexes in the different samples are shown in Table 3. The differences between sexes were significant ($p < 0.05$) for most measurements, except for the tibia in the RMPE ($p = 0.06$) and the SMO ($p = 0.49$) collections. The sexual dimorphism index is variable between the samples, but in general, the radius has the highest index (Table 3). The RMPE collection has the lowest SDI for all the analysed bones, with the SMO and ISCR collections also showing low SDI.

3.3 Discriminant functions

Table 4 presents the coefficients for sex estimation, calculated from SMOL.B sample, and the cut-off points. The four long bones measurements were combined in eleven different functions (A to K; Table 4). These coefficients (β_0 – intercept term; β_i – slope; Table 4) can also be used to calculate the probability of being male. Applying expression b) to function A, for example, for an individual whose humerus length is 351 mm, radius length is 254 mm, femur length is 486 mm and tibia is 399 mm the procedure follows:

$$p = \frac{1}{1 + e^{-(-23.692+(0.022\text{x}351)+(0.033\text{x}254)+(0.009\text{x}486)+(0.016\text{x}399))}} = 0.96$$

The value calculated above suggests that there is a probability of 96% of the individual being male.

The Akaike Information Criterion (AIC) shows that the model of function A is the one that fits the data the best, followed by the models of functions F and G (Table 5). Classification accuracy for the discriminant functions analysed ranged from 82.8% (function J) to 92% (function G; Table 5) and give better sex coincidence for the females than males (Table 5). Of the single bones, the humerus (85.8%) and the tibia (85.7%) are the most effective despite the radius having the highest SDI (Table 4) and a lower AIC value (Table 5).

3.4 Coincident sex estimations

When applying these functions to the other collections (Tables 5, 6 and 7) the percentages of coincident sex estimations (%CE) are highly variable among the different samples and analysed bones. Using more than one bone, the percentage of coincident sex estimations ranged between 25 and 100% in females and between 30 and 100% in males (Table 6). Among the Portuguese collections, SMOL.A showed a %CE larger than 80%, for both males and females for functions A, C and G. In the RL collection only function F coincidentally estimated sex for more than 80% in both males and females. SMO presented a %CE larger than 80%, for both males and females for functions C, D and F. None of the functions presented a %CE larger than 80% in both sexes for RMPE and ISCR. Of the English collections, functions A to G showed higher %CE for PMSBS, with only function F having a %CE lower than 80% for both sexes. A %CE higher than 80% for both sexes was only observed when using function D for MSMG and function E for MESBD. For MSS none of the functions presented a %CE larger than 80% in both sexes. It is important to note the small sample sizes used for these functions.

In the upper limb (Table 7), the coincident sex estimations in pooled sex varied between 74 and 88% in the Portuguese collections and between 82 and 91% in the English collections using the humerus length. Using the radius, the %CE in pooled sex varied between 66 and 89% in the Portuguese collections and between 83 and 93% in the English collections. Generally, there are higher %CE in the English than in the Portuguese collections and these percentages varied between males and females.

While the humerus length shows coincident percentages over 80% for females from SMOL.A, SMOL.B, SMO, ISCR, MESBD and PMSBS, these percentages were lower than 70% for RL, RMPE, MSMG and MSS (Table 7). Using the humerus, Kranioti and Michalodimitrakis's (2009) function showed a %CE higher than 80% for both sexes in SMO, ISCR, MESBD and PMSBS. The humerus length function presented by Charisi et al. (2011) showed a %CE higher than 80% for both sexes in SMO, ISCR, MESBD and PMSBS. Tomar's function presented %CE higher than 80% for both sexes only when used for SMOL.B.

The %CE also varied in the radius, with sex estimation coincident in females below 70% for RL, RMPE, ISCR (except for Tomar's function), MSMG (with the exception of Tomar's function), MSS and PMSBS (with the exception of Tomar's function). Both Charisi's et al. (2011) and Wasterlain's (2000) functions using the radius had higher than 80% coincident sex estimation for both sexes in SMOL.A. Tomar's function using the radius showed a %CE higher than 80% for both sexes in SMOL.A, SMOL.B, ISCR, MSMG and PMSBS.

Table 3. Sample size (n), mean, standard deviation (SD), *F* ratios, *p*-values and sexual dimorphism index (SDI) for the long bones of the collections tested (full names of the archaeological sites in Table 1).

		Maximum length	**Females**			**Males**			***F* ratio**	***p*-value**	**SDI**
			n	**Mean**	**SD**	**n**	**Mean**	**SD**			
Portuguese samples	SMOL.A	Humerus	12	290.17	8.03	82	319.44	18.21	29.94	<0.001	1.10
		Radius	5	210.20	11.30	15	233.93	10.51	18.50	<0.001	**1.11**
		Femur	29	404.03	15.73	117	445.91	25.10	73.37	<0.001	1.10
		Tibia	29	328.21	13.63	116	362.54	32.97	30.04	<0.001	1.10
	SMOL.B	Humerus	339	286.01	14.32	388	317.77	16.30	721.31	<0.001	1.11
		Radius	115	213.63	13.53	156	241.24	15.28	238.01	<0.001	**1.13**
		Femur	385	402.37	19.52	457	440.13	21.92	685.08	<0.001	1.09
		Tibia	349	330.68	16.89	436	366.02	19.67	708.29	<0.001	1.11
	RL	Humerus	35	292.80	29.60	32	320.81	21.14	19.54	<0.001	1.10
		Radius	33	218.79	24.96	35	246.26	11.99	34.08	<0.001	**1.13**
		Femur	36	405.50	31.20	43	447.74	21.64	50.10	<0.001	1.10
		Tibia	34	330.38	24.38	41	371.66	18.98	67.92	<0.001	1.12
	RMPE	Humerus	13	298.15	20.29	28	317.89	16.16	11.25	0.002	1.07
		Radius	17	221.35	11.95	31	238.45	11.87	22.68	<0.001	**1.08**
		Femur	14	419.64	21.75	28	441.36	23.97	8.13	0.007	1.05
		Tibia	13	346.46	17.16	31	359.65	21.36	3.88	0.055	1.04
	SMO	Humerus	8	294.53	14.84	9	337.61	15.52	33.99	<0.001	**1.15**
		Radius	4	219.38	8.08	14	239.93	12.70	9.17	0.008	1.09
		Femur	8	424.56	15.27	18	454.86	28.35	7.98	0.009	1.07
		Tibia	5	345.70	27.68	15	375.33	27.02	4.46	0.490	1.09
	ISCR	Humerus	15	297.33	11.47	28	324.11	21.94	19.34	<0.001	**1.09**
		Radius	12	220.17	5.51	25	240.68	14.38	22.54	<0.001	**1.09**
		Femur	16	419.06	9.17	31	444.68	23.98	16.84	<0.001	1.06
		Tibia	17	339.06	12.15	28	363.39	22.69	16.56	<0.001	1.07
English samples	MSMG	Humerus	8	300.00	14.72	36	332.39	12.01	43.92	<0.001	1.11
		Radius	16	218.75	12.69	26	245.31	9.89	57.49	<0.001	**1.12**
		Femur	18	425.11	14.13	33	459.39	20.66	38.19	<0.001	1.08
		Tibia	19	336.74	14.11	42	372.36	19.87	49.53	<0.001	1.11
	MSS	Humerus	17	297.82	15.13	27	328.15	13.59	47.58	<0.001	**1.10**
		Radius	12	223.08	10.66	23	246.22	12.48	29.78	<0.001	**1.10**
		Femur	12	414.42	25.45	23	452.91	19.58	24.79	<0.001	1.09
		Tibia	9	340.89	25.53	25	363.32	16.06	9.34	0.004	1.07
	MESBD	Humerus	34	295.97	13.51	64	323.95	17.93	63.51	<0.001	1.09
		Radius	27	217.00	10.62	51	240.84	11.22	82.65	<0.001	**1.11**
		Femur	24	408.71	18.53	63	452.33	22.02	74.06	<0.001	**1.11**
		Tibia	37	329.51	15.34	85	358.60	17.86	74.22	<0.001	1.09
	PMSBS	Humerus	18	296.00	15.36	31	329.61	15.16	56.06	<0.001	1.11
		Radius	19	215.74	11.26	35	241.51	10.85	67.71	<0.001	**1.12**
		Femur	15	415.87	23.74	18	458.11	22.32	27.67	<0.001	1.10
		Tibia	22	328.41	11.07	39	368.74	23.14	58.90	<0.001	**1.12**

Table 4. Estimated coefficients for sex estimation (β_0 – intercept term; β_i – slope).

		A	B	C	D	E	F	G	H	I	J	K
N		175	529	242	711	547	203	228	727	271	842	785
β_i	Humerus	0.022	0.039	0.043		0.044			0.065			
	Radius	0.033		0.037			0.037	0.042		0.069		
	Femur	0.009	0.010		0.027			0.032			0.048	
	Tibia	0.016	0.018		0.030	0.023	0.036					0.054
β_0		–23.692	–22.577	–21.762	–21.752	–21.585	–21.397	–23.194	–19.693	–15.764	–20.274	–18.949

A – humerus+radius+femur+tibia; B – humerus+femur+tibia; C – humerus+radius; D – femur+tibia; E – humerus+tibia; F – radius+tibia; G – radius+femur; H – humerus; I – radius; J – femur; K – tibia.

Table 5. Classification accuracy for original and cross-validation for females, males, and pooled sex.

	Cut-off value	Cut-off value (mm)	N			Predicted Group Membership						Akaike Information Criterion (AIC)
						Original group (%)			Cross-validated (%)			
			Females	Males	Total	Females	Males	Total	Females	Males	Total	
A	–0.269		69	106	175	94.2	91.5	92.6	91.3	90.6	90.9	93.10
B	–0.127		234	295	529	91.5	88.1	89.6	91.5	88.1	89.6	354.77
C	–0.169		104	138	242	91.3	90.6	90.9	91.3	90.6	90.9	165.32
D	–0.103		320	391	711	88.1	84.9	86.4	88.1	84.9	86.4	528.65
E	–0.121		243	304	547	89.3	88.2	88.5	89.3	88.2	88.5	375.80
F	–0.240		81	122	203	93.8	88.5	90.6	93.8	88.5	90.6	118.52
G	–0.257		90	138	228	94.4	90.6	92.1	94.4	90.6	92.1	145.91
H	–0.068	301.9	339	388	727	88.8	83.2	85.8	88.8	83.2	85.8	569.39
I	–0.144	226.4	115	156	271	86.1	82.1	83.8	86.1	82.1	83.8	354.18
J	–0.078	420.8	385	457	842	85.7	80.3	82.8	85.7	80.3	82.8	753.59
K	–0.106	349.0	349	436	785	88.5	83.5	85.7	88.5	83.5	85.7	631.12

A – humerus+radius+femur+tibia; B – humerus+femur+tibia; C – humerus+radius; D – femur+tibia; E – humerus+tibia; F – radius+tibia; G – radius+femur; H – humerus; I – radius; J – femur; K – tibia.

For the lower limb (Table 8), the %CE using the femur's maximum length was only calculated using Tomar's function, which varied between 74 and 89% for pooled sexes. However, this function showed very low coincident sex estimations for females, particularly in the collections SMO (25%), ISCR (33%), MSMG (33%) and MSS (50%). Using the tibia, the coincident sex estimations in pooled sex varied between 73 and 87% in the Portuguese collections and between 71 and 87% in the English collections.

Tomar's femur function showed a %CE higher than 80% for both sexes only in SMOL.B. As observed for the femur, this %CE using the tibia was not balanced for males and females, but overall showed higher percentages than those observed for the femur. Bruzek's (1995) function showed %CE higher than 80% for both sexes in MESBD and PMSBS. The function presented by Kranioti et al. (2017) displayed %CE higher than 80% for both sexes in MSMG and Tomar's function in SMOL.B MESBD and PMSBS.

4 Discussion

Previous research (e.g. Mall et al. 2001) suggests that when only one of the long bones is present, the most accurate results derive from the radius. The same is observed in the collections analysed in this study. The SDI is consistently higher for the radius than the other bones (Table 3), except SMO, for which the humerus has the highest SDI. The SDI calculated for each bone of SMOL.B sample (Table 3) is comparable with the predicted cross-validated percentage (Table 5), except for the radius. The radius has the highest SDI and the femur the lowest, but their predicted cross-validated percentage is similar (radius = 83.8%; femur = 82.8%; Table 5). This accuracy similarity, despite the different SDI, can result from a much smaller sample of radii than femora, for developing the functions and testing them. This is also supported by the lower AIC value for the model using the radius (function I) out of the models that only use one measurement, suggesting that this model fits the data better (Table 5). Still, it is inter-

Table 6. Percentage of coincident sex estimations (%CE) using the functions developed from SMOL.B sample.

				A	B	C	D	E	F	G
Female	Portuguese	SMOL.A	N	1	12	1	27	12	4	4
			%CE	**100**	100	**100**	100	100	100	**100**
		RL	N	21	25	29	31	25	24	27
			%CE	62	76	59	93	76	**92**	74
		RMPE	N	9	9	13	11	9	12	12
			%CE	56	56	62	83	67	58	64
		SMO	N	2	4	3	5	4	2	3
			%CE	50	75	**100**	**80**	75	**100**	67
		ISCR	N	10	11	12	15	12	10	11
			%CE	60	64	83	100	67	100	73
	English	MSMG	N	4	4	7	14	5	14	10
			%CE	25	25	57	**100**	40	79	60
		MSS	N	5	6	12	6	9	6	10
			%CE	60	67	67	92	56	50	60
		MESBD	N	11	13	18	20	21	22	16
			%CE	73	77	72	100	**81**	95	69
		PMSBS	N	10	10	16	14	16	16	12
			%CE	**90**	**90**	**81**	**100**	**88**	100	**83**
Male	Portuguese	SMOL.A	N	9	49	13	97	56	10	11
			%CE	**100**	73	**85**	55	71	70	**100**
		RL	N	23	26	26	36	26	30	33
			%CE	100	100	100	55	96	**87**	91
		RMPE	N	23	23	28	26	25	27	25
			%CE	78	74	82	30	76	74	84
		SMO	N	5	6	7	12	7	8	9
			%CE	100	100	**86**	**86**	100	**88**	100
		ISCR	N	14	18	21	25	19	18	20
			%CE	79	78	76	52	79	65	79
	English	MSMG	N	11	16	22	25	22	18	19
			%CE	100	94	100	**81**	91	89	95
		MSS	N	17	20	23	20	23	20	19
			%CE	94	90	91	64	91	90	95
		MESBD	N	15	26	37	51	43	41	22
			%CE	93	85	95	64	**84**	73	86
		PMSBS	N	9	10	23	16	22	23	15
			%CE	**100**	**100**	**100**	**100**	**91**	78	**93**

A – humerus+radius+femur+tibia; B – humerus+femur+tibia; C – humerus+radius; D – femur+tibia; E – humerus+tibia; F – radius+tibia; G – radius+femur; H – humerus; I – radius; J – femur; K – tibia.
Full names of the archaeological sites in Table 1.

esting to note that function G, which uses the radius and the femur, gave the best sex estimation accuracy and function F, using the radius and the tibia, is the second-best fit for the data (Table 5). This can be related to the sex bias observed in Tomar's functions. Function G, for example, correctly estimated 94.4% of the females and 90.6% of the males, which may be related to the sample size of females (n = 90) being smaller than that of males (n = 138).

Using a larger number of bones to estimate sex does not necessarily mean better sex estimations, as suggested by the

Table 7. Upper limb percentages of coincident estimations (%CE) for the different methods by sex and pooled sexes.

			Humerus				Radius			
			N	Kranioti & Michalodimitrakis 2009	Charisi et al. 2011	Tomar	N	Charisi et al. 2011	Wasterlain 2000	Tomar
				%CE	%CE	%CE		%CE	%CE	%CE
Female	Portuguese	SMOL.A	12	100	100	100	5	80	80	100
		SMOL.B	339	93	93	88	115	77	79	88
		RL	35	66	66	60	33	52	52	73
		RMPE	13	62	62	62	17	53	53	59
		SMO	8	88	88	75	4	50	75	75
		ISCR	15	87	87	67	12	67	67	82
	English	MSMG	8	50	50	50	16	56	63	81
		MSS	17	65	65	59	12	58	58	67
		MESBD	34	82	82	74	27	70	70	74
		PMSBS	18	83	83	61	19	68	68	89
Male	Portuguese	SMOL.A	82	71	71	77	15	80	80	80
		SMOL.B	388	77	77	81	156	91	90	83
		RL	32	88	88	94	35	100	97	94
		RMPE	28	79	79	82	31	90	90	84
		SMO	9	89	89	100	14	93	93	93
		ISCR	28	82	82	79	25	92	92	82
	English	MSMG	36	97	97	100	26	100	100	96
		MSS	27	93	93	100	23	100	96	91
		MESBD	63	89	89	95	51	96	96	92
		PMSBS	31	94	94	97	35	100	100	94
Pooled sex	Portuguese	SMOL.A	94	74	74	**80**	20	80	80	**85**
		SMOL.B	727	85	85	**86**		**84**	**84**	**84**
		RL	67	**76**	**76**	**76**	68	76	75	**84**
		RMPE	41	73	73	**76**	48	**77**	**77**	75
		SMO	17	**88**	**88**	**88**	18	83	88	**89**
		ISCR	43	**84**	**84**	77	47	66	66	**86**
	English	MSMG	44	89	89	**91**	42	83	86	**90**
		MSS	44	82	**84**	**84**	35	**86**	83	83
		MESBD	97	87	87	**88**	78	**87**	**87**	86
		PMSBS	49	**90**	**90**	84	54	89	89	**93**

Full names of the archaeological sites in Table 1.

AIC values for the different models (Table 5). The model with the lowest AIC and therefore the one that best fits the data is function A (Bozdogan 1987), which uses the measurements of all the bones analysed. However, the model of function F is the second-best fit for the data and only uses two measurements, the maximum lengths of the radius and the tibia. In this study, the cross-validated percentage of predicted group membership (Table 5) was higher for function G (92.1%), which combined the length of the radius and femur, than for function A (90.9%), which took into consideration the length of the four bones. Again, this is probably due to the smaller sample size using function A than function G, as the model of function A fits the data the best. Despite the radius having the higher SDI in SMOL.B sample (Table 3), out of the four analysed bones, the radius did not display the highest cross-validated percentage of predicted group membership (Table 5), which can be related to the smaller sample size for this bone as the AIC value is lower for this model than the others using single bones. The femur, which has the lowest SDI (Table 3) also shows the lowest highest cross-validated

Table 8. Lower limb percentages of coincident sex estimations (%CE) for the different methods by sex and pooled sexes.

			Femur		**Tibia**			
			N	**Tomar**	**N**	**Bruzek 1995**	**Kranioti et al. 2017**	**Tomar**
				%CE		**%CE**	**%CE**	**%CE**
Female	Portuguese	SMOL.A	119	73	118	70	90	70
		SMOL.B	385	86	349	70	92	89
		RL	36	67	34	71	88	76
		RMPE	14	57	13	54	62	54
		SMO	8	25	5	40	60	60
		ISCR	15	33	16	53	88	63
	English	MSMG	18	33	20	70	85	70
		MSS	12	50	9	44	67	67
		MESBD	25	64	37	86	95	92
		PMSBS	15	60	22	91	95	95
Male	Portuguese	SMOL.A	117	89	116	84	61	82
		SMOL.B	457	80	436	89	76	84
		RL	43	86	41	95	75	93
		RMPE	28	86	31	81	68	77
		SMO	18	89	15	93	80	93
		ISCR	31	90	28	82	68	82
	English	MSMG	33	94	42	90	86	90
		MSS	23	100	25	88	80	80
		MESBD	62	95	85	87	61	86
		PMSBS	18	94	39	85	79	82
Pooled sex	Portuguese	SMOL.A	236	85	234	**77**	76	**77**
		SMOL.B	842	83	785	83	81	**86**
		RL	79	80	75	84	**87**	84
		RMPE	42	74	44	73	**86**	70
		SMO	26	77	20	**80**	75	**80**
		ISCR	46	79	44	73	**75**	73
	English	MSMG	51	76	62	84	**85**	84
		MSS	35	83	34	**76**	**76**	**76**
		MESBD	99	89	122	**87**	71	75
		PMSBS	33	79	61	**87**	85	**87**

Full names of the archaeological sites in Table 1.

percentage of a predicted group membership and the highest AIC value (Table 5). RMPE and ISCR displayed the lowest SDI (Table 3) out of the Portuguese collections, which can explain why none of the %CE using multiple bones (Table 6) presented values higher than 80% for both sexes. RMPE also showed a particularly low %CE using single upper (Table 7) and lower (Table 8) limb bones.

Overall, the %CE were higher for single (Table 7 and 8) than for multiple bone functions (Table 6), which is probably related to the smaller sample sizes allowing to measure more than one bone.

Both Kranioti & Michalodimitrakis's (2009) and Charisi et al. (2011) functions to estimate sex from the humerus maximum length showed similar results with higher %CE for SMO, ISCR, MESBD and PMSBS (Table 7). These results can be partially explained by the high SDI calculated for the humerus (Table 3), but not entirely, as other collections have even higher SDI values for this bone than ISCR and MESBD. The %CE from Tomar's function is only higher than 80% for both sexes when used for SMOL.B, the sample from which the function was derived, and not when used for SMOL.A despite both samples being from the same archaeological site. This difference between SMOL.A and SMOL.B can be related to the different sample sizes or social status, as SMOL.A represents an area further away from the church than SMOL.B (Curto 2019). People of higher socioeconomic status were buried inside or closer to the church than those of lower status (e.g. Binski 1996; Daniell 1997).

In general, there are higher %CE in the English than in the Portuguese collections (Tables 6 and 7), despite the functions tested being developed from South European populations (Wasterlain 2000; Kranioti & Michalodimitrakis 2009; Charisi et al. 2011). For metric methods, the sample's mean compared to that of the methods might be more appropriate than its temporal or geographic location. It is desirable to calculate the probability of being male or female and classify the individual as undetermined sex if the results are not conclusive as well as choose methods with similar mean measurements to those observed in the sample under study. It is also important to note the differences in chronology between collections usually used to develop methods and archaeological collections. In Portugal, from pre-history to the Middle Ages there was a slow increase in stature, followed by a negative trend until the late 19th century and a rapid increase during the 2nd half of the 20th century (Cardoso & Gomes 2009). This means that individuals from mediaeval and post-mediaeval were taller than the ones from the Portuguese Collections of Identified Skeletons used to develop metric methodologies.

Charisi's et al. (2011) functions gave a higher %CE for different collections when using the humerus or the radius (Table 7). This suggests that even within the same sample it may be important to choose a method for each bone being studied, one by one. The different chronology could be related to the different accuracy of Tomar's functions in other collections. However, other methods tested in this study gave better coincident sex estimations than Tomar's functions (e.g. humerus) for some collections despite all of them having used more recent osteological assemblages. Therefore, the chronological period may not be the most important aspect when choosing a metric method to estimate sex. Body size sex dimorphism differs between populations in a complex manner. Height reflects genetics and environmental factors (e.g. Silventoinen 2003). Manifestations of sexual dimorphism in the growth process such as different timings of growth cessation (bimaturatism) and different rates of growth between the sexes (Badyaev 2002), affect the human body measurements. Other factors such as activity patterns (Krishan et al. 2016) and general secular trends (e.g. Cardoso & Gomes 2009; Godde 2015) can also have an impact on sexual dimorphism.

The variability in SDI between the collections under study might be related with sexual differences in environmental factors such as morbidity, nutrition, and gender specific workload, as well as regional secular trends and genetic admixture. Low stature sexual dimorphism has also been related with food security and higher female status (Gleeson & Kushnick 2018). In the future, I better knowledge about the history and lifestyle of the different populations under study would help identifying specific factors that might have played a role on the differences on SDI and the different equations' accuracy.

Study limitations

The most important limitation of this study was the fact that these collections do not consist of identified skeletons, from which we know their sex. To develop the functions for sex estimation using the SMOL.B sample, the sex of these individuals was estimated through morphological methods based on the pelvis. Therefore, the sex of some individuals, both those used to build the functions and the other ones being tested, may have not been correctly estimated to begin with. Even though sex estimations using the *os coxae* have high accuracies, even in archaeological contexts. An example is the genetically sexed skeletons from the 13th–16th century to which sex was estimated with a 95.7% accuracy based on morphological features of the *os coxae* (Inskip et al. 2019). Still, populational discrepancies in sex estimations based on pelvic features have been recorded by several researchers (e.g., Walker 2005; Spradley & Jantz 2011). Non-binary sex (Roca-Rada et al. 2022), even if rare, also plays a role in sex estimation limitations.

It is also important to note the small sample sizes for some samples when estimating sex (Tables 6 and 7). The bone length depends on genetic factors, physiological stress, nutrition and secular trends (e.g. Cowgill & Hager 2007; Gustafsson et al. 2007; Cardoso & Gomes 2009). Epiphyseal dimensions, on the other hand, are more prone to change as a response to intense physical activity (Forriol & Shapiro 2005; Carlson et al. 2007) but have been suggested to discriminate between sexes (e.g., İşcan et al. 1998; Frutos 2005; Charisi et al. 2011). It would also be interesting to compare sex estimated using *os coxae* to that estimated using epiphyses width and head diameter.

5 Conclusions

The results in this study clearly show population-specific sexual dimorphism and especially for the lower limb, how the higher number of males, compared to females, biased the general (pooled sex) percentage of coincident sex estimation.

The percentage of coincident sex estimations varied greatly among the tested functions depending on the collection, however, it does not seem to be related to broad geography (Portugal vs England) but rather to each individual collection. Our results support previously published findings that have shown sexual dimorphism of long bones to be population specific (e.g. Bidmos & Dayal 2004) even within similar geographic regions (e.g. Portuguese collections).

Forensic metric methods for sex estimation based on long bone length can be used in archaeological samples, particularly when used to confirm sex. However, they should be used with care and results should always be accompanied by probability estimates of being male or female. We suggest taking into consideration the mean bone length of the sample, by sex, and how it relates to the mean of the samples used to develop the methods. It is also important to consider the SDI of the measurements within the samples.

Tomar's functions can be used to estimate sex in other samples, in the absence of bones more sexually dimorphic, particularly if accompanied by a very high or low probability of being male. However, these functions will be best used as a metric confirmation of sex when the features from *os coxae* and cranium provide ambiguous results.

Acknowledgements: This work was supported by FCT – Fundação para a Ciência e a Tecnologia (FCT) [2020.02110. CEECIND]. The authors also acknowledge the financial support of the UIDB/04449/2020 and UIDP/04449/2020 projects, funded by Fundação para a Ciência e a Tecnologia (FCT), Portugal, and by the European Regional Development Fund. We thank ERA Arqueologia, S.A. for kindly allowing the access to the unpublished data from Rua dos Lagares 74, Lisbon.

References

Arcini, C., Ahlström, T., & Tagesson, G. (2014). Variations in diet and stature: Are they linked? Bioarchaeology and paleodietary Bayesian mixing models from Linköping, Sweden. *International Journal of Osteoarchaeology, 24*(4), 543–556. https://doi.org/10.1002/oa.2247

Badyaev, A. V. (2002). Growing apart: An ontogenetic perspective on the evolution of sexual size dimorphism. *Trends in Ecology & Evolution, 17*(8), 369–378. https://doi.org/10.1016/S0169-5347(02)02569-7

Bašić, Ž., Anterić, I., Vilović, K., Petaros, A., Bosnar, A., Madžar, T., … Anđelinović, Š. (2013). Sex determination in skeletal remains from the medieval Eastern Adriatic coast – discriminant function analysis of humeri. *Croatian Medical Journal, 54*(3), 272–278. https://doi.org/10.3325/cmj.2013.54.272 PMID: 23771758

Bidmos, M. A., & Dayal, M. R. (2004). Further evidence to show population specificity of discriminant function equations for sex determination using the talus of South African blacks. *Journal of Forensic Sciences, 49*(6), 1165–1170. https://doi.org/10.1520/JFS2003431 PMID:15568686

Binski, P. (1996). *Medieval death: Ritual and representation*. London: British Museum Press.

Bogin, B. (1999). Evolutionary perspective on human growth. *Annual Review of Anthropology, 28*(1), 109–153. https://doi.org/10.1146/annurev.anthro.28.1.109 PMID:12295621

Bozdogan, H. (1987). Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions. *Psychometrika, 52*(3), 345–370. https://doi.org/10.1007/BF02294361

Bruzek, J. (2002). A method for visual determination of sex, using the human hip bone. *American Journal of Physical Anthropology, 117*(2), 157–168. https://doi.org/10.1002/ajpa.10012 PMID:11815949

Bruzek, J., & Murail, P. (2006). *Methodology and reliability of sex determination from theskeleton. Forensic anthropology and medicine.* Springer.

Buikstra, J. E., & Mielke, J. L. (1985). Demography, diet, and health. In R. I. Gilbert & J. H. Mielke (Eds.), *Analysis of prehistoric diets* (pp. 360–422). Academic Press.

Buikstra, J. E., & Ubelaker, D. H. (1994). *Standards for data collection from human skeletal remains: proceedings of a seminar at the Field Museum of Natural History*. Arkansas Archeological Survey Research Series 1. Fayetteville, AL: Arkansas.

Cardoso, H. V., & Cunha, E. (2000). Sexual dimorphism in upper limb skeletal proportions. *Biométrie Humaine et Anthropologie, 18*(1-2), 55–61.

Cardoso, H. F. V., & Gomes, J. E. A. (2009). Trends in adult stature of peoples who inhabited the modern Portuguese territory from the Mesolithic to the late 20th century. *International Journal of Osteoarchaeology, 19*(6), 711–725. https://doi.org/10.1002/oa.991

Carlson, K. J., Grine, F. E., & Pearson, O. M. (2007). Robusticity and sexual dimorphism in the postcranium of modern hunter-gatherers from Australia. *American Journal of Physical Anthropology, 134*(1), 9–23. https://doi.org/10.1002/ajpa.20617 PMID:17503447

Charisi, D., Eliopoulos, C., Vanna, V., Koilias, C. G., & Manolis, S. K. (2011). Sexual dimorphism of the arm bones in a modern greek population. *Journal of Forensic Sciences, 56*(1), 10–18. https://doi.org/10.1111/j.1556-4029.2010.01538.x PMID:20840296

Cowgill, L. W., & Hager, L. D. (2007). Variation in the development of postcranial robusticity: An example from Çatalhöyük, Turkey. *International Journal of Osteoarchaeology, 17*(3), 235–252. https://doi.org/10.1002/oa.882

Cunha, E. (1993). Testing identification records: Evidence from the Coimbra identified skeletal collections (19th–20th centuries). *American Journal of Physical Anthropology,* (Suppl., 16), 77–77.

Curate, F., Umbelino, C., Perinha, A., Nogueira, C., Silva, A. M., & Cunha, E. (2017). Sex determination from the femur in Portuguese populations with classical and machine-learning classifiers. *Journal of Forensic and Legal Medicine, 52*, 75–81. https://doi.org/10.1016/j.jflm.2017.08.011 PMID:28866285

Curate, F., Coelho, J., Gonçalves, D., Coelho, C., Ferreira, M. T., Navega, D., & Cunha, E. (2016). A method for sex estimation using the proximal femur. *Forensic Science International, 266*, 579.e1–579.e7. https://doi.org/10.1016/j.forsciint.2016.06.011 PMID:27373600

Curto, A. (2019). *The Impact of Diet and Health on Bone Stable Isotope Ratios: A Comparative Study*. United Kingdom: University of Kent.

Daniell, C. (1997). *Death and burial in medieval England*, 1066–1550. London: Routledge. https://doi.org/10.4324/9780203979174

Ferembach, D., Schwindezky, I., & Stoukal, M. (1980). Recommendation for age and sex diagnoses of skeletons. *Journal of Human Evolution, 9*(7), 517–549. https://doi.org/10.1016/0047-2484(80)90061-5

Forriol, F., & Shapiro, F. (2005). Bone development: Interaction of molecular components and biophysical forces. *Clinical Orthopaedics and Related Research, 432*, 14–33. https://doi.org/10.1097/01.blo.0000156001.78631.e9 PMID:15738800

Gama, I., Navega, D., & Cunha, E. (2015). Sex estimation using the second cervical vertebra: A morphometric analysis in a documented Portuguese skeletal sample. *International Journal of Legal Medicine, 129*(2), 365–372. https://doi.org/10.1007/s00414-014-1083-0 PMID:25212127

Garcia, S. (2012). Is the circumference at the nutrient foramen of the tibia of value to sex determination on human osteological collections? Testing a new method. *International Journal of Osteoarchaeology, 22*(3), 361–365. https://doi.org/10.1002/oa.1202

Gasser, T., Sheehy, A., Molinari, L., & Largo, R. H., & the T. Gasser, A. Sheehy, L. Molinari. (2000). Sex dimorphism in growth. *Annals of Human Biology, 27*(2), 187–197. https://doi.org/10.1080/030144600282299 PMID:10768423

Gleeson, B. T., & Kushnick, G. (2018). Female status, food security, and stature sexual dimorphism: Testing mate choice as a mechanism in human self-domestication. *American Journal of Physical Anthropology, 167*(3), 458–469. https://doi.org/10.1002/ajpa.23642 PMID:30159867

Godde, K. (2015). Secular trends in cranial morphological traits: A socioeconomic perspective of change and sexual dimorphism in North Americans 1849-1960. *Annals of Human Biology, 42*(3), 253–259. https://doi.org/10.3109/03014460.2014.941399 PMID:25156659

Gustafsson, A., Werdelin, L., Tullberg, B. S., & Lindenfors, P. (2007). Stature and sexual stature dimorphism in Sweden, from the 10th to the end of the 20th century. *American Journal of Human Biology, 19*(6), 861–870. https://doi.org/10.1002/ajhb.20657 PMID:17712787

Işcan, M. Y., & Shihai, D. (1995). Sexual dimorphism in the Chinese femur. *Forensic Science International, 74*(1-2), 79–87. https://doi.org/10.1016/0379-0738(95)01691-B PMID:7665135

Inskip, S., Scheib, C. L., Wohns, A. W., Ge, X., Kivisild, T., & Robb, J. (2019). Evaluating macroscopic sex estimation methods using genetically sexed archaeological material: The medieval skeletal collection from St John's Divinity School, Cambridge. *American Journal of Physical Anthropology, 168*(2), 340–351. https://doi.org/10.1002/ajpa.23753 PMID:30575013

Johnston, F., & Padez, C. (1999). Secular trends in male adult height 1904-1996 in relation to place of residence and parent's educational level in Portugal. *Annals of Human Biology, 26*(3), 287–298. https://doi.org/10.1080/030144699282787 PMID:10355499

Kranioti, E. F., & Michalodimitrakis, M. (2009). Sexual dimorphism of the humerus in contemporary Cretans – A population-specific study and a review of the literature. *Journal of Forensic Sciences, 54*(5), 996–1000. https://doi.org/10.1111/j.1556-4029.2009.01103.x PMID:19627412

Kranioti, E. K., García-Donas, J. G., Almeida Prado, P. S., Kyriakou, X. P., & Langstaff, H. C. (2017). Sexual dimorphism of the tibia in contemporary Greek-Cypriots and Cretans: Forensic applications. *Forensic Science International, 271*, 129.e1–129.e7. https://doi.org/10.1016/j.forsciint.2016.11.018 PMID:27919515

Krishan, K., Chatterjee, P. M., Kanchan, T., Kaur, S., Baryah, N., & Singh, R. K. (2016). A review of sex estimation techniques during examination of skeletal remains in forensic anthropology casework. *Forensic Science International, 261*, 165.e1–165.e8. https://doi.org/10.1016/j.forsciint.2016.02.007 PMID:26926105

Legates, D. R., & Davis, R. E. (1997). The continuing search for an anthropogenic climate change signal: Limitations of correlation-based approaches. *Geophysical Research Letters, 24*(18), 2319–2322. https://doi.org/10.1029/97GL02207

Lovell, N. C. (1989). Test of Phenice's technique for determining sex from the os pubis. *American Journal of Physical Anthropology, 79*(1), 117–120. https://doi.org/10.1002/ajpa.1330790112 PMID:2750876

Mall, G., Graw, M., Gehring, K., & Hubig, M. (2000). Determination of sex from femora. *Forensic Science International, 113*(1-3), 315–321. https://doi.org/10.1016/S0379-0738(00)00240-1 PMID:10978643

Mall, G., Hubig, M., Büttner, A., Kuznik, J., Penning, R., & Graw, M. (2001). Sex determination and estimation of stature from the long bones of the arm. *Forensic Science International, 117*(1-2), 23–30. https://doi.org/10.1016/S0379-0738(00)00445-X PMID:11230943

Martin, R. (1928). *Lehrbuch der Anthropologie in systematischer Darstellung: mit besonderer Berücksichtigung der anthropologischen Methoden für Studierende Ärzte und Forschungsreisende (Vol. 3).* Stuttgart: G. Fischer.

Merritt, C. E. (2017). Inaccuracy and bias in adult skeletal age estimation: Assessing the reliability of eight methods on individuals of varying body sizes. *Forensic Science International, 275*, 315.e1–315.e11. https://doi.org/10.1016/j.forsciint.2017.03.003 PMID:28359575

Moore, M. K., & Ross, A. H. (2013). Stature estimation. In E. A. Di Gangi & M. K. Moore (Eds.), *Research Methods in Human Skeletal Biology* (pp. 151–179). Academic Press. https://doi.org/10.1016/B978-0-12-385189-5.00006-6

Navega, D., Vicente, R., Vieira, D. N., Ross, A. H., & Cunha, E. (2015). Sex estimation from the tarsal bones in a Portuguese sample: A machine learning approach. *International Journal of Legal Medicine, 129*(3), 651–659. https://doi.org/10.1007/s00414-014-1070-5 PMID:25186617

Padez, C. (2003). Secular trend in stature in the Portuguese population (1904-2000). *Annals of Human Biology, 30*(3), 262–278. https://doi.org/10.1080/0301446031000064530 PMID:12850960

Padez, C. (2007). Secular trend in Portugal. *Journal of Human Ecology (Delhi, India), 22*(1), 15–22. https://doi.org/10.1080/09709274.2007.11905993

Phenice, T. W. (1969). A newly developed visual method of sexing the os pubis. *American Journal of Physical Anthropology, 30*(2), 297–301. https://doi.org/10.1002/ajpa.1330300214 PMID:5772048

Pons, J. (1955). The sexual diagnosis of isolated bones of the skeleton. *Human Biology, 27*(1), 12–21. PMID:14353506

Ríos Frutos, L. (2005). Metric determination of sex from the humerus in a Guatemalan forensic sample. *Forensic Science International, 147*(2-3), 153–157. https://doi.org/10.1016/j.forsciint.2004.09.077 PMID:15567620

Roca-Rada, X., Tereso, S., Rohrlach, A. B., Brito, A., Williams, M. P., Umbelino, C., … Teixeira, J. C. (2022). A 1000-year-old case of Klinefelter's syndrome diagnosed by integrating morphology, osteology, and genetics. *Lancet, 400*(10353), 691–692. https://doi.org/10.1016/S0140-6736(22)01476-3 PMID:36030812

Roche, A. F. (1992). *Growth, maturation, and body composition: the Fels Longitudinal Study 1929-1991 (No. 9)*. Cambridge University Press. https://doi.org/10.1017/CBO9780511661655

Rozendaal, A. S., Scott, S., Peckmann, T. R., & Meek, S. (2020). Estimating sex from the seven cervical vertebrae: An analysis of two European skeletal populations. *Forensic Science International, 306*, 110072. https://doi.org/10.1016/j.forsciint.2019.110072 PMID:31812819

Sauter, M. R., & Privat, F. (1955). Sur un nouveau procédé métrique de détermination sexuelle du bassin osseux. *Bulletin de la Société Suisse d'Anthropologie et d'Ethnologie, 31*, 60–81.

Silva, A. M. (1995). Sex assessment using the calcaneus and talus. *Antropologia Portuguesa, 13*, 107–119.

Silventoinen, K. (2003). Determinants of variation in adult body height. *Journal of Biosocial Science, 35*(2), 263–285. https://doi.org/10.1017/S0021932003002633 PMID:12664962

Spradley, M. K., & Jantz, R. L. (2011). Sex estimation in forensic anthropology: Skull versus postcranial elements. *Journal of Forensic Sciences, 56*(2), 289–296. https://doi.org/10.1111/j.1556-4029.2010.01635.x PMID:21210801

Steyn, M., & Işcan, M. Y. (1997). Sex determination from the femur and tibia in South African whites. *Forensic Science International, 90*(1-2), 111–119. https://doi.org/10.1016/S0379-0738(97)00156-4 PMID:9438371

Steyn, M., & Patriquin, M. L. (2009). Osteometric sex determination from the pelvis – Does population specificity matter? *Forensic Science International, 191*(1-3), 113.e1–113.e5. https://doi.org/10.1016/j.forsciint.2009.07.009 PMID:19665855

Tanner, J. M., & Tanner, J. M. (1990). *Foetus into man: Physical growth from conception to maturity*. Harvard University Press.

Borgognini Tarli, S. B., & Repetto, E. (1986). Methodological considerations on the study of sexual dimorphism in past human populations. *Human Evolution, 1*(1), 51–66. https://doi.org/10.1007/BF02437285

Ubelaker, D. H., & Volk, C. G. (2002). A test of the phenice method for the estimation of sex. *Journal of Forensic Sciences, 47*(1), 19–24. https://doi.org/10.1520/JFS15200J PMID:12064650

Ulijaszek, S. J., & Kerr, D. A. (1999). Anthropometric measurement error and the assessment of nutritional status. *British Journal of Nutrition, 82*(3), 165–177. https://doi.org/10.1017/S0007114599001348 PMID:10655963

Uytterschaut, H. T. (1986). Sexual dimorphism in human skulls. A comparison of sexual dimorphism in different populations. *Human Evolution, 1*(3), 243–250. https://doi.org/10.1007/BF02436582

Walker, P. L. (2005). Greater sciatic notch morphology: Sex, age, and population differences. *American Journal of Physical Anthropology, 127*(4), 385–391. https://doi.org/10.1002/ajpa.10422 PMID:15693026

Wasterlain, R. S. N. (2000). *Morphé: Análise das proporções entre os membros, dimorfismo sexual e estatura de uma amostra da colecção de esqueletos identificados do museu antropológico da Universidade de Coimbra.* (Unpublished master's thesis). University of Coimbra, Coimbra.

Wasterlain, R. S. N., & Cunha, E. (2000). Comparative performance of femur and humerus epiphysis for sex diagnosis. *Biométrie Humaine et Anthropologie, 18*(1-2), 9–13.

White, T. D., Black, M. T., & Folkens, P. A. (2011). *Human Osteology*. Academic Press.

Willey, P., Galloway, A., & Snyder, L. (1997). Bone mineral density and survival of elements and element portions in the bones of the Crow Creek massacre victims. *American Journal of Physical Anthropology, 104*(4), 513–528. https://doi.org/10.1002/(SICI)1096-8644(199712)104:43.0.CO;2-S PMID:9453699

Willner, L., & Martin, R. (1985). Some basic principles of mammalian sexual dimorphism. In J. Ghesquiere, R. Martin, & F. Newcombe (Eds.), *Human Sexual Dimorphism* (pp. 1–42). London: Taylor and Francis.

WORD Database (2009). *Welcome Osteological Record Database, Museum of London Centre for Human Bioarchaeology* [Online]. Available from: http://www.museumoflondon.org.uk/English/Collections/OnlineResources/CHB/Resources/ [Accessed 10 October 2020]

Manuscript received: 11 April 2023
Revisions requested: 15 June 2023
Revised version received: 29 June 2023
Manuscript accepted: 30 June 2023