Sex estimation using long bones in the largest burial site of the Copper Age: Linear discriminant analysis and random forest

S. Díaz-Navarro [a,*], S. Díez-Hermano [b], M.A. Rojo-Guerra [a], J. Lomba Maurandi [c], C. Valdiosera [d], T. Gunther [e], M. Haber Uriarte [c]

[a] Departament of Prehistory, Archaeology, Social Anthropology and Historiographic Sciences and Techniques. Faculty of Philosophy and Letters. Universidad de Valladolid, Spain
[b] Department of Plant Production and Forest Resources, Sustainable Forest Management Research Institute (iuFOR). Higher Technical School of Agricultural Engineering (ETSIIAA), Universidad de Valladolid, Spain
[c] Departament of Prehistory, Archaeology, Ancient History, Medieval History and Historiographic Sciences and Techniques. Faculty of Letters. Universidad de Murcia, Spain
[d] Laboratory of Human Evolution, Universidad de Burgos, Spain
[e] Human Evolution, Department of Organismal Biology, Uppsala University, Sweden

ABSTRACT

Sex estimation of the individuals in a sample is fundamental for any bioarchaeological study to define a particular demographic assemblage or to classify isolated remains. Long bones are an excellent alternative for sex estimation when the most dimorphic anatomical parts are not preserved or are highly altered. Here we propose a set of discriminant functions and classification models to estimate the sex of prehistoric individuals using linear discriminant analysis and machine learning approaches. Different osteometric variables were taken from the humeri, ulnae, radii, femurs and tibias of a sample of 109 articulated skeletons buried in the collective tomb of Camino del Molino (Region of Murcia, SE-Spain), dated to the 3rd millennium BC. Sex was estimated based on standard anthropological methods and ancient DNA analysis of a control sample. Fifty-two discriminant functions with prediction thresholds higher than 0.8 on the ROC curve were obtained using independent (22) and combined variables (30). The best LDA models for sex prediction were those based on proximal epiphyseal widths or their combination with other variables, reaching values close to 0.98 on the ROC curve. The random forest-based model obtained an accuracy of 0.94 and confirmed the importance of epiphyseal widths in sex classification. This analysis is more comprehensive than univariate LDA, as it allows for ranking the importance of bones in sex discrimination and considers correlations between long bones rather than treating them as independent observations. In contrast, applying LDA to each bone makes it easier to predict the sex of other coeval collections that do not have such a complete sample. This work aims to overcome the scarcity of methods that can be applied to sex estimation of the large volume of isolated remains from Camino del Molino and for other Mediterranean skeletal series from the Late Prehistory with high biological affinity and that share similar environmental conditions.

1. Introduction

The sex assignment of individuals or skeletal fragments is key in bioarchaeological analysis. Its results are essential to completing the population pyramid of a settlement, to associating individuals of one or the other sex with possible burial rituals (their form of deposition, location, treatment of the corpse, offerings, etc.), ways of life (daily activities conditioned by the sex of the subject, access to food resources, sexual biases in food, physical activities, or mobility), diseases or causes of death. If a complete adult skeleton is recovered, its sex can be estimated with up to 95 % reliability, based on indicators of the skull and pelvis, as these are the regions where the differences between sexes are best expressed (Buikstra and Ubelaker 1994; White and Folkens, 2005). These bones are not always preserved, especially if the population under

---

study is archaeological.

Long bones are an excellent alternative for estimating the sex when the most dimorphic anatomical regions are not preserved or are very degraded, as well as when studying a skeletal collection from an altered funerary deposit, a very frequent occurrence in the Late Prehistory of Iberia (Díaz-Navarro et al. 2023a). Recent studies have shown that long bones have greater classification power in sex estimations than the skull (Spradley and Jantz 2011). Pearson and Bell (1917) were pioneers in the development of alternative techniques for sex estimation, although the use of discrimination formulas saw its greatest boom in the second half of the 20th century. Since then, there has been a plethora of work by researchers focused on these studies for populations of different origins, especially using long bones (e.g., Godycki 1957; Hanihara 1958; Steel, 1962; Black 1978; Krogman and Iscan 1986, Işcan and Miller-Shaivitz, 1984; Isçan and Miller-Shaivitz, 1986; Dittrick and Suchey 1986; Carretero et al. 1995; Alemán 1997; González-Reimers et al. 2000; Ríos Frutos 2005; Alemán et al. 2012; Djorojevic et al. 2016; Bidmos and Mazengenya 2021).

Although functions have traditionally been formulated using reference skeletal collections of known sex and age at death, it is increasingly common to find discriminant analyses based on archaeological populations, with sex estimations supported by standard anthropological methods (e.g., Safont et al., 2000; Šlaus and Tomičić 2005; Özer and Katayama 2006; Jiménez-Arenas 2009, 2010; Bašić et al. 2013; Jerković et al. 2016; Novak 2016; Tomczyk et al. 2017; Dabbs, 2020; Chovalopoulou et al., 2018; Wysocka et al. 2023), paleogenomic analyses (Ordóñez et al. 2013) or gender assignments supported by the offerings accompanying the bodily remains (Cavazzuti et al. 2019). However, the latter should be taken with caution since gender is a social construct that does not necessarily obey biological sex.

In the Iberian Peninsula, discriminant analyses of long bones have traditionally been based on populations of known sex and age with minimal margins of error. Of note are those formulated by I. Alemán (1997) on a current Mediterranean population (Granada), the discriminant functions obtained from the reference collection of the Complutense University of Madrid (e.g., Trancho et al. 1997, 2012), or those of Carretero et al. (1995) based on different metric variables of a humerus from a 19th century skeletal series from Coimbra of known sex and age (Portugal). There are also discriminant analyses on peninsular archaeological populations, such as those of the medieval population of Wamba on tibiae (López-Bueis 1995), those of Jiménez-Arenas (2009, 2010) based on femurs and humeri from the Muslim necropolis of La Torrecilla (Granada), or those of Safont et al. (2000) based on the circumference of the diaphysis of the bones of the upper and lower extremities from the late Roman sites of Mas Rimbau and Mas Mallol (Tarragona). This type of analysis has also been frequently applied to pre-Hispanic populations in the Canary Islands (e.g., González-Reimers et al. 2000; Ordóñez et al. 2013). In contrast, no discriminant analyses have thus been carried out on prehistoric populations, perhaps there are not enough individualised and well-preserved skeleton samples to provide acceptable results. An exception could be the discriminant analysis on innominate bones from the necropolis of S'Illot des Porros (Mallorca), dated to 2430 ± 200 years BP (Rissech and Malgosa 1997).

The same picture can be extrapolated to the European continent. Although several methods for sex estimation in long bones have been developed on archaeological collections (Šlaus and Tomičić 2005; Özer and Katayama 2006; Jerković et al. 2016; Tomczyk et al. 2017; Dabbs, 2020; Wysocka et al. 2023), hardly any discriminant functions performed on prehistoric samples have been published. Some exceptions are the formulae developed on cremated bones from Italian Bronze and Iron Age sites (Cavazzuti et al. 2019) or those of Chovalopoulou et al. (2018) on different samples of ancient Greeks —from the 2nd millennium BC to the Middle Ages-. Moreover, some authors have tested the validity of discriminant functions elaborated with modern osteological collections on Palaeolithic (Alonso-Llamazares and Pablos 2019) and Late Prehistoric (Francis et al. 2023) samples.

In light of these circumstances, and after proving the ineffectiveness of applying discriminant functions derived from other Spanish populations to estimate the sex of the Chalcolithic population from Camino del Molino (Caravaca de la Cruz, Murcia, SE-Spain), this paper proposes for the first time a sex estimation method based on individuals chronologically dated from Prehistory, using linear discriminant analysis (LDA) and machine learning classifiers: random forest (RF). Thanks to the large volume of skeletons recovered from this grave and their excellent preservation, we can, on the one hand, conduct this analysis on 109 individualised skeletons and, on the other hand, estimate the sex of the vast number of isolated skeletal remains with great reliability. Furthermore, this work can serve as a reference considering the scarce methods that can be applied to sex estimation in other prehistoric funerary contexts with similar characteristics.

2. The collective burial site of Camino del Molino (Caravaca de la Cruz, SE-Spain)

The collective burial site at Camino del Molino (CMOL) (Fig. 1), was discovered by chance in 2007 during a construction project, and emergency excavation was conducted throughout 2008 (Lomba et al. 2009). The site is in Caravaca de la Cruz, a municipality in the northwest of the Region of Murcia (Fig. 1a), surrounded from west to east by the Argos and Quípar rivers. This site is part of the Subbaetic system and has a very rugged relief and a high altitude, over 800 m in most of its territory (Fig. 1b).

The tomb can be defined as a hypogeum excavated in the travertine with slightly flared walls and a diameter of about 6–7 m (Fig. 1c, d).

The sequence is characterised by a continuous introduction of bodies deposited on the previously prepared floor. As the number of corpses increased, relocations are documented in the form of lateral displacements of bodies to accommodate more individuals. As the perimeter walls collapsed with human remains, numerous complete or partial individuals were relocated towards the centre of the structure and arranged in no apparent order (Haber et al. 2012). Some individuals were recovered in perfect anatomical connection (Fig. 2), others were in the form of skeletal packages -grouped skeletal remains of a probable same individual not articulated-, and others were completely isolated remains as a consequence of the continuous removal and relocation of corpses.

An anthropological study has identified a minimum number of 1348 individuals (30.7 % non-adults and 69.3 % adults), making CMOL the largest prehistoric burial site known to date (Díaz-Navarro et al. 2023b). The funerary record of other large contemporary collective burials points to cemeteries of around 100–400 individuals (Mahieu 1987; Chambon 2003; Etxeberria and Herrasti 2007; Silva 2012) which hints at the magnitude of CMOL.

This burial site comprised 24.8 % (n = 334) infants (0–12 years), 5.9 % (n = 79) juveniles (13–20 years), 30.9 % (n = 417) young adults (21–39 years), 34.7 % (n = 468) middle-aged adults (40–59 years) and 3.7 % (n = 50) elderly adults (over 60 years). In terms of sex estimation, the population included 377 females (27.9 %), 400 males (29.7 %), 62 individuals 'probably' female (4.6 %), 62 'probably' male (4.6 %) and 447 individuals of undetermined sex (33.2) (Díaz-Navarro et al. 2023b).

The 14C dating of 28 articulated skeletons has allowed us to identify two possible contiguous funerary phases covering almost the entire 3rd millennium BC (2971–2711-2451–2251 years cal. BC, 2σ) (Díaz-Navarro et al. 2023b) (S1 Table). The first funerary phase of use started at the beginning of the third millennium (2971–2711 years cal. BC, 2σ) and spanned a maximum of three centuries (0–259/82–377 years cal. BC, 2σ). The second funerary phase was shorter (0–220/20–282 years cal. BC, 2σ), and the end of the sequence is marked between 2451 and 2251 cal. BC 2σ. The transition period between the two phases was 80 years at most.

A final aspect to highlight is the scarce grave goods documented in the burial site, especially considering the large volume of the buried population (Avilés et al. 2012). In fact, there is hardly any direct
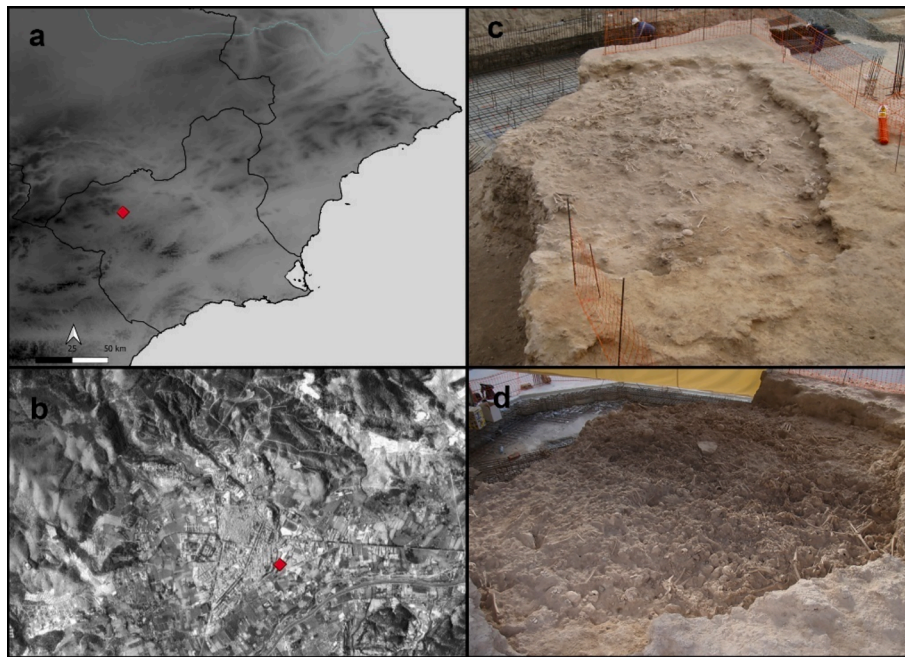
Fig. 1. The collective burial site of CMOL. 1a. Location of the site in the southeast of the Iberian Peninsula. 1b. Aerial photograph of the area around Caravaca de la Cruz, showing the location of the archaeological site. 1c. General photograph of the burial site after surface cleaning. 1d. General view of the funerary deposit during the excavation process.
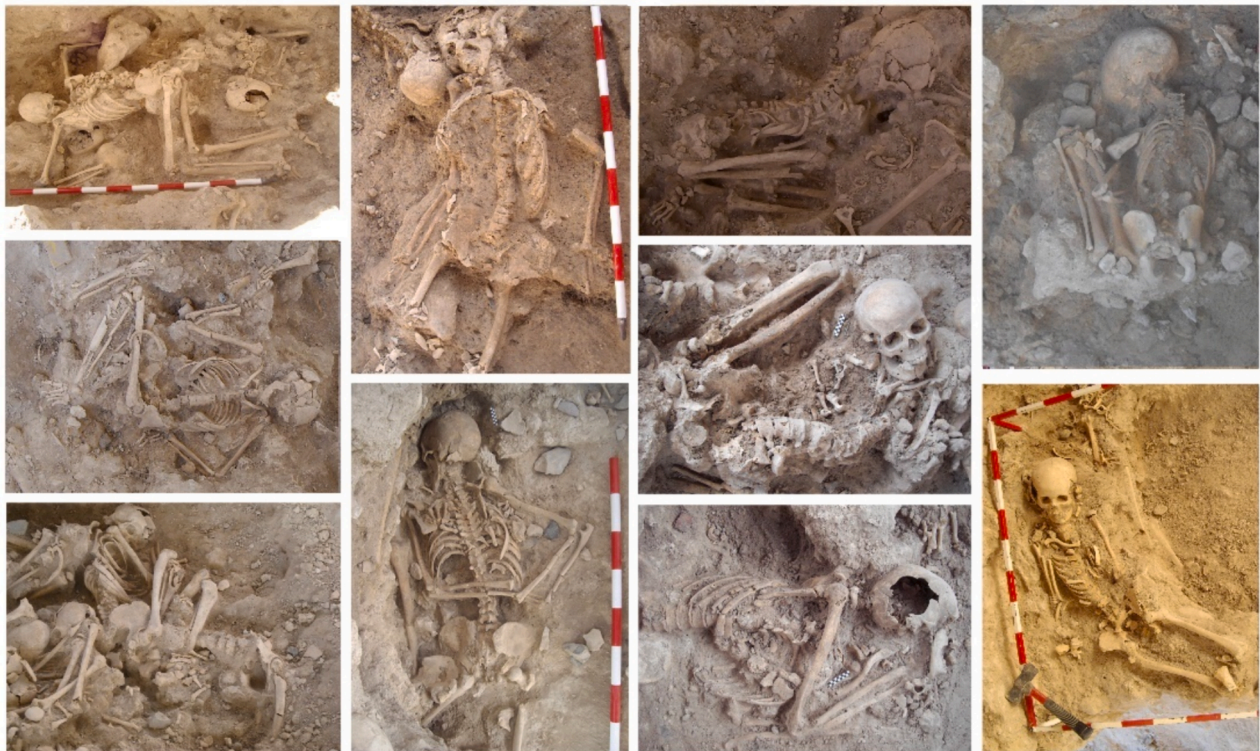
Fig. 2. Photographs of some of the articulated skeletons recovered from the collective burial site at Camino del Molino during fieldwork.

evidence of objects associated with specific individuals. However, the most outstanding aspect is the documentation of 51 canids, mainly of the *Canis lupus familiaris* type, but also a common European fox (*Vulpes vulpes*) and five wolves (*Canis lupus* sp.) (Catagnano 2016).

The osteological study of this population is providing highly relevant data on the ways of life and death of the Chalcolithic Mediterranean societies (Díaz-Navarro et al., 2023b–d).

3. Material and methods

3.1. Sample selection

The humeri, ulnae, radii, femurs and tibiae of 109 articulated skeletons were analysed. Skeletons recovered in perfect anatomical connection, with a preserved pelvis and/or skull for sex estimation, that had completed the skeletal maturation process at the time of death (adults with fused epiphyses) were selected. Skeletons with badly damaged or excessively fragmented cortexes, those exhibiting degenerative or metabolic pathologies, signs of trauma or visible malformations that could alter the results, as well as individuals of undetermined sex were excluded from the analysis.

The sample selected using these criteria represents 65.3 % of the articulated skeletons recovered (N=167) and 11.7 % of the adult individuals from the count and analysis of the isolated CMOL skulls (N=1348). The sample can be considered representative of the total inhumed population, as it includes specimens found in the different areas of the burial site, from all the stratigraphic units of the funerary deposit (SUU 1100, 1104, 1106–1110), and, therefore, from the two contiguous phases of use documented in the chronometric analysis (Díaz-Navarro et al. 2023b) (Fig. 3).

3.2. Sex estimation

Sex estimations of articulated skeletons have been based on an exhaustive analysis of the different cranial, mandibular and pelvic indicators (Buikstra and Ubelaker 1994; Herrmann et al. 1990). Fifty-nine (54.1 %) of the 109 articulated skeletons analysed preserve the skull and pelvis for sexual estimation, 9 only the skull (8.3 %) and 41 (37.6 %) the pelvis (S2 Table). On the skull, the morphology of the external occipital protuberance, mastoid processes, orbital rim, supraciliary arches, nuchal crest and supramastoid crest, and the height of the glabella and frontal bossing has been assessed. On the mandibles, the focus was on the morphology of the chin and gonion, the angle of the mandibular ramus, the mandibular muscle attachments and the size of the teeth. A total of 10 qualitative traits were observed in the pelvis (angle of the sciatic notch; preauricular sulcus; subpubic angle; and morphology of the iliac crest, ischial spine, sacrum, ischium, acetabulum, obturator

foramen and sacroiliac joint) (S2 Table).

In addition, sex estimates were cross-checked by DNA analysis of a control sample of 19 skeletons where biological sex was estimated based on the number of DNA fragments mapping to the sex chromosomes (Mittnik et al. 2016).

3.3. DNA analysis

Ancient DNA analysis was performed on 11 skeletons, and we have data from 8 other individuals previously analysed and published (Olalde et al. 2018; Villalba-Mouco et al. 2021) (S1 Table). Individuals with clear sexual assignments preserving the pelvis and skull, as well as skeletons sexed only from the skull have been selected to validate the effectiveness of estimates based on this anatomical region (S2 Table). We sampled one tooth per individual (S1 Table) for DNA extraction.

Processing of ancient DNA was performed by the SciLifeLab Ancient DNA unit (Uppsala, Sweden). Laboratory procedures are specified in the Supplementary Information. DNA sequences generated for this study are available from the European Nucleotide archive under accession PRJEB77020.

3.4. Linear discriminant analysis

Forty metric variables were taken for each side (Table 1) of the lateralised long bones of the limbs from the 109 skeletons. The fibulae were excluded due to their poor preservation in this sample. For the measurements, the classical procedures proposed by Martin and Saller (1957) also collected in detail by Alemán (1997) were followed, using a King's Foot with an accuracy of 0.02 mm, an osteometric plate with an accuracy of 0.1 mm and millimetric tape (Table 1).

Two independent observers took the measurements in mm using the same methods, so the technical measurement error (TEM) and relative technical measurement error (RTEM) were calculated to assess the intra- and inter-observer errors (Perini et al. 2005).

The data were processed by dividing the sample into subgroups according to the sex and sides. Descriptive analyses were performed for each variable and the Wilcoxon two-sample comparison test (with a significance value of $p \leq 0.05$ and a false discovery rate correction for multiple comparisons) was used to analyse possible statistically
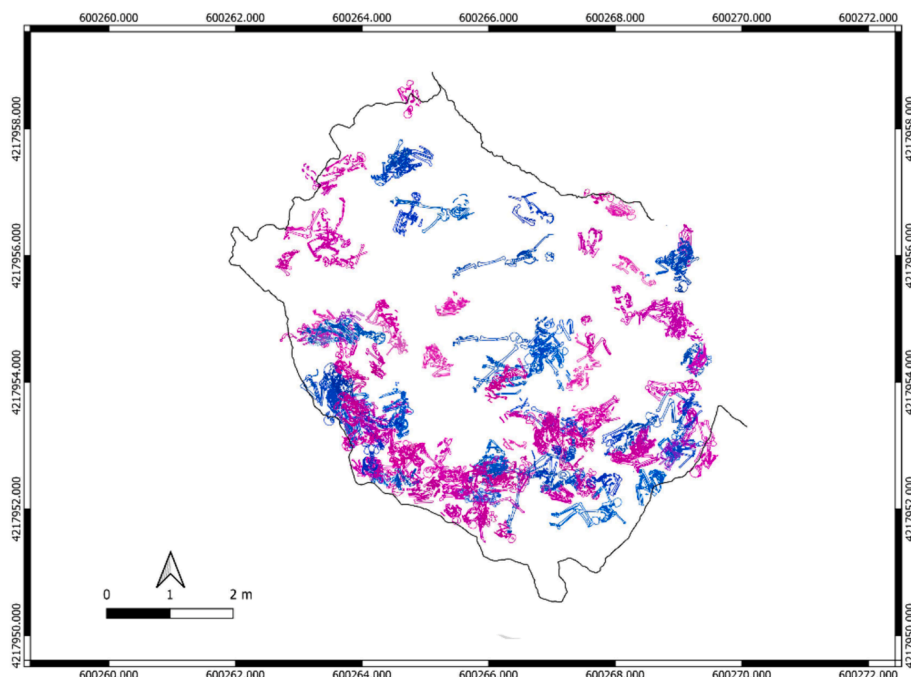

Fig. 3. Digital planimetry (QGIS 3.16 Hannover©) with the articulated skeletons of CMOL burial site that form part of this work (pink: female; blue: male).

Table 1
Information on the osteometric variables considered in this study and their abbreviations, including the technical measurement error (TEM) and relative technical measurement error (RTEM). O1: Observer 1; O2: Observer 2.

Bone	Measure	Abbreviation	O1_TEM	O1_rTEM	O2_TEM	O2_rTEM
Humerus	Maximum Lenght	MLH	1.6051	0.5611	1.3420	0.4684
	Distal epiphyseal breadth	DEBH	0.4198	0.7685	0.4800	0.8782
	Vertical diameter of the head	VDHH	0.5687	1.4538	0.4391	1.1208
	Deltoid V perimeter	DVPH	1.1948	2.1799	1.0941	1.9926
	Midshaft perimeter	MSPH	1.1429	1.9425	1.0656	1.8055
	Maximum midshaft diameter	MAXDH	0.7274	3.7122	0.6981	3.5316
	Minimum midshaft diameter	MINDH	0.5992	3.8101	0.6786	4.3006
	Maximum diameter at deltoid V	MAXDVH	0.7092	3.8462	1.0371	**5.5721**
	Minimum diameter at deltoid V	MINDVH	0.5620	3.7254	0.6455	4.2650
Ulna	Maximum Lenght	MLU	1.2499	0.5171	1.1469	0.4726
	Midshaft perimeter	MPU	0.8810	2.1311	0.8088	1.9582
	Maximum midshaft diameter	MAXDU	0.4216	3.0291	0.5499	3.8956
	Minimum midshaft diameter	MINDU	0.3984	3.4640	0.4415	3.8665
	Distal epiphyseal breadth	DEBU	0.4429	2.5696	0.5064	2.9765
	Minimum perimeter	MINPU	0.7377	2.2810	0.8024	2.4752
Radius	Maximum Lenght	MLR	0.8711	0.3879	1.1465	0.5103
	Maximum diameter of the head	MAXDHR	0.2941	1.4706	0.3296	1.6494
	Minimum diameter of the head	MINDHR	0.2937	1.5439	0.3254	1.7104
	Distal epiphyseal breadth	DEBR	0.6990	**5.2539**	0.7508	**5.5766**
	Maximum midshaft diameter	MAXDR	0.5371	**5.2431**	0.8439	**8.0862**
	Minimum midshaft diameter	MINDR	0.7265	2.4533	0.6559	2.1992
	Midshaft perimeter	MPR	0.8315	2.1844	1.0639	2.7752
	Minimum perimeter	MINPR	0.8803	2.4809	0.8970	2.5066
Femur	Maximum Lenght	MLF	0.8019	0.1974	1.0344	0.2527
	Distal epiphyseal breadth	DEBF	0.6382	0.8855	0.5836	0.8073
	Maximum diameter of the head	MDHF	0.6888	1.7280	0.8131	2.0384
	Medium shaft perimeter	MSPF	0.9065	1.1628	1.0214	1.3096
	Anteroposterior subtrochanteric diameter	APSDF	0.7418	3.1996	1.1004	4.7194
	Transverse subtrochanteric diameter	TSDF	0.7319	2.3826	0.6622	2.1518
	Anteroposterior diameter at midshaft	APDMF	0.6017	2.3674	0.7160	2.8033
	Transverse diameter at midshaft	TDMF	0.5416	2.2542	0.5156	2.1393
Tibia	Maximum Lenght	MLT	1.5479	0.4631	2.0986	0.6267
	Distal epiphyseal breadth	DEBT	0.5903	1.2369	0.5634	1.1764
	Proximal epiphyseal breadth	PEBT	0.6114	0.9038	0.7436	1.0997
	Midshaft perimeter	MSPT	0.9835	1.2736	1.2845	1.6574
	Anteroposterior diameter at midshaft	APDMT	0,.9175	3.1456	0.8997	3.0772
	Transverse diameter at midshaft	TDMT	0.6154	3.1372	0.6681	3.4130
	Nutrient foramen perimeter	NFPT	0,.6595	0.7833	0.7497	0.8890
	Anteroposterior diameter at nutrient foramen	APDNFT	0,.7769	2.4443	0.8294	2.6113
	Transverse diameter at nutrient foramen	TDNFT	0,.6564	3.1374	0.8238	3,.9165

significant differences by side and sex. Side comparisons were considered independent samples instead of paired data, as many individuals didn't have information from both sides available.

The sexual dimorphism index (SDI) was also calculated from a simple formula (Smith 1999) applied to all the variables expressing the difference between the sexes as a percentage [SDI= (Mean of the variable in males/Mean of the variable in females)*100].

LDA was performed independently for upper limb long bone and lower limb long bone variables (models can be replicated from the supplementary discriminant_analysis compressed files). Sex (binary: female or male) was the predicted outcome. One LD function was estimated for each independent variable from one bone, and combinations of two variables were made, either from the same bone or from another bone of the same limb. The LD functions were trained by repeated 10-fold cross-validation (10 times) (Kim 2009) with a 75/25 data split for training/testing. Missing data were removed separately for each body measurement. Model performance was estimated using the area under the ROC curve (Fawcett, 2006) (AUROC, sensitivity versus 1-specificity). True positive rates indicate probability of correctly classifying females. Criteria for model selection were AUROC>0.8 and sample size > 50.

3.5. Random forest

In order to determine the importance of the variables in predicting sex, a random forest model was trained following an out-of-bag approach for error estimation (models can be replicated from the supplementary predict_randomforest compressed files).

The number of trained trees was 1000 and the number of variables tested in each split (mtry) was 15 (selected by grid search optimisation). Unlike the uni- and bivariate discriminant analyses, this model took into account all variables (upper and lower limb) at once.

The significance of the variables was measured by the average reduction of the Gini index. Missing data were replaced by the strategy with highest accuracy in the test set among the following imputation methods: predictive mean matching (PMM), classification and regression trees (CART), lasso linear regression (LASSO) and median value.

All analysis were performed in R v4.2.1 (R Core Team. R, 2022). LDA and Random Forest models were trained and tested using the caret (Kuhn 2008), finalfit (Harrison et al. 2023) and randomForest (Liaw and Wiener 2002) packages. Missing data imputation was performed using the mice (van Buuren and Groothuis-Oudshoorn 2011) package.

4. Results

4.1. Sex estimation

After selecting the sample according to the aforementioned criteria, it was possible to carry out the osteometric analysis of 109 adult individuals, aged between 21 and 59 years at the time of death. Of these skeletons, 66 (60.6 %) were classified as female, 3 (2.7 %) as probably female, 38 (34.9 %) as male, and 2 (1.8 %) as probably male (S2 Table). Priority has been given to the sexual estimates yielded by the study of the pelvis, as they demonstrate greater confidence in the sexual

classification of skeletal collections (Buikstra and Ubelaker 1994; White and Folkens 2005). Therefore, skeletons that do not preserve the pelvis have been classified as 'probable', although the sexual features of the skull are clear.

It is important to note that by chance females are better represented in the total number of articulated skeletons recovered (N=167: 40.7 % females, 25.7 % males, 4.1 % probable females, 3 % probable males, 0.6 % of undetermined sex and 25.7 % non-adults without sexual assignment), while in the total sample of skulls analysed the number of females and males is similar.

The results of the genetic analysis coincide with the anthropological estimates in all cases (S1 and S2 Tables).

4.2. Linear discriminant analysis

The results of the intra- and inter-observer error calculation can be seen in Table 1. All variables have a RTEM (relative index of inter-observer differences) below the acceptance threshold of 0.05, with three exceptions: MAXDR, MINDR and MINDVH. Therefore, these variables have been ruled out from all osteometric analyses. Therefore, 37 variables are used for LDA and RF analysis.

The Wilcoxon two-sample comparison test allows us to conclude that there are no statistically significant differences by laterality in any of the variables (Table 2 and S3 Table). On the other hand, the SDI is greater than 100 in all cases, identifying statistically significant differences in the Wilcoxon two-sample comparison test by sex (Table 3 and S4 Table).

Table 3 shows the large volume of missing cases, which is related to the nature and preservation of the osteological collection. Despite

Table 2
Wilcoxon two-sample comparison test by sides (right and left). n: sample size, Me: median, iqr: interquartile range.

Variables	Left			Right			*p* value
	n	Me	iqr	n	Me	iqr	
MLH	28	279	21.5	31	289	18.5	0.94802365
DEBH	46	54	5	57	54	5	0.94802365
VDHH	34	38	3.75	37	38	7	0.94802365
DVPH	64	54.5	5.25	61	55	6	0.98488572
MSPH	32	57	5.5	35	59	8	0.94802365
MAXDH	31	19	2	36	20	2.25	0.78141412
MINDH	31	15	3	36	15.5	2	0.94802365
MINDVH	63	15	2	59	15	3	0.98488572
MLU	14	233	25.25	28	245	22	0.94802365
MPU	16	42	4	27	41	4	0.94802365
MAXDU	19	14	2	35	14	2	0.98488572
MINDU	19	12	1.5	35	12	1	0.94802365
DEBU	26	16.5	3	39	18	3	0.78141412
MINPU	34	33	5	47	32	4.5	0.98488572
MLR	28	225	27.5	26	222.5	29	0.94802365
MAXDHR	35	20	2	34	19.5	2	0.94802365
MINDHR	35	19	2	34	18.5	2	0.94802365
DEBR	35	30	3.5	39	30	2.5	0.98488572
MPR	31	39	4.5	27	38	4.5	0.94802365
MINPR	54	36	4	56	35	4	0.94802365
MLF	34	401.5	35.75	30	410	33.5	0.94802365
DEBF	24	72.5	10	26	72	7.75	0.94802365
MDHF	61	40	4	57	40	6	0.99136473
MSPF	38	77	10.75	38	77	9.5	0.98488572
APSDF	81	23	3	71	23	3	0.98488572
TSDF	82	30	4	71	31	3	0.94802365
APDMF	37	25	3	38	25	4.75	0.98843445
TDMF	36	24	2.5	38	24	1.75	0.98488572
MLT	28	335.5	26	19	327	34.5	0.94802365
DEBT	45	48	6	41	48	4	0.94802365
PEBT	29	67	8	27	66	4	0.94802365
MSPT	27	78	9	27	76	10	0.98488572
APDMT	28	29.5	5	30	28	4.75	0.98488572
TDMT	28	20	3	30	19.5	3	0.98488572
NFPT	60	84.5	10	53	83	10	0.94802365
APDNFT	56	32	4	51	32	3	0.94802365
TDNFT	56	21	3	51	21	2.5	0.98488572

working with 109 articulated skeletons, osteometric results were obtained for 57 individuals on average, which also allows us to obtain reliable statistical results. In the humerus, all variables have a sample size greater than 50 (mean = 59). Only the variables DEBU and MINPU of the ulna exceed this threshold (mean = 46). The mean of the radius is 54.8, with N>50 for the variables MAXDHR, MINDHR, DEBR and MINPR. The femur variables are the ones with the largest sample (mean = 65.7), except for DEBF which could only be measured in 36 individuals. Finally, the variables DEBT, NFPT, APDNFT and TDNFT of the tibia are the ones that could be measured in more than 50 individuals (Mean = 56.2).

4.2.1. Upper limb long bones variables
Fig. 4 depicts the performance of the LD functions using univariate and bivariate combinations of the long bone measurements from the upper limb of the skeleton as predictors. The performance of the LD functions using univariate and bivariate combinations of long bones measurements of the upper limbs as predictors can be seen in S5 and S6 Tables. Seventeen out of 210 models were selected following predefined criteria (AUROC>0.8 and N>50) listed in Table 4, including formulas and prediction thresholds. Those combined formulae whose prediction is less than that of one of the variables independently have been eliminated from this selection. The VDHH model achieved near perfect classification power (AUROC=0.974). The best performing models included different combinations of MAXDHR and MINDHR and VDHH independently. In total there are 12 independent and 5 combined variables that meet these criteria (Table 4).

4.2.2. Lower limb long bone variables
Fig. 5 exhibits the LD functions using univariate and bivariate combinations of lower limb long bone measurements as predictors. The performance of the LD functions using univariate and bivariate combinations of long bones measurements of the lower limbs as predictors can be seen in S7 and S8 Tables. Thirty-five out of 153 models were selected following predefined criteria (AUROC>0.8 and N>50), which are listed in Table 5 with formulas and prediction thresholds. We have eliminated from the selection those combined formulae whose predictive power is lower than that of the independent variables that form it. It is worth noting that the models that include MDHF, alone or in combination, occupy the top ten positions in the ranking (Fig. 5, Table 5). In total, 25 bivariate and 10 univariate combinations reach accuracy thresholds above 80. Independently, the femur MDHF variable reaches 0.9470. Seven discriminant functions combining two variables exceed 95 points on the ROC curve. The combination of MDHF and MSPF variables of the femur reaches 97 points on the ROC curve.

4.3. Random forest

The random forest-based model obtained an accuracy of 0.94 (95 % CI: [0.79, 0.99], p-value [Accuracy > Non-Information Criteria] = 0.001, Kappa = 0.86) on the test set data. CART imputation method for missing data yielded the best results among all tested techniques (see S9 Table). The confusion matrix can be found in Table 6. DEBH, MDHF and VDHH were the most important body measurements for sex classification (Fig. 6).

5. Discussion

This study provides a comprehensive anthropometric analysis of 37 variables taken from the long bones of a collection of 109 articulated skeletons from the largest prehistoric burial site known to date. The preservation of skeletons in the primary position is unusual in 3rd millennium Iberian collective tombs, where reduced or displaced primary deposits are generally documented in which the remains underwent continuous removal once they became skeletonised (e.g., Etxeberria and Herrasti 2007; Rivera García 2011; Silva 2012; Díaz-Zorita 2017; Díaz-

Table 3

Wilcoxon two-sample comparison test by sex (female and male). n: sample size, Me: median, iqr: interquartile range.

Variables	Female			Male			SDI	*p* value
	n	Me	iqr	n	Me	iqr		
MLH	24	276	12	21	296	23	107.25	0.00001723
DEBH	42	52	3	29	59	4	113.46	0.00000003
VDHH	28	36	2	24	42	4.25	116.67	0.00000004
DVPH	49	53	5	30	58	5	109.43	0.00000078
MSPH	27	55	3	23	62	6.5	112.73	0.00001050
MAXDH	27	19	2	23	21	4	110.53	0.00072736
MINDH	27	15	1	23	17	2.5	113.33	0.00005363
MINDVH	50	14	2	29	16	2	114.29	0.00001050
MLU	18	237.5	20.5	18	257	18.5	108.21	0.00037821
MPU	17	39	4	19	43	3	110.26	0.00316447
MAXDU	26	13	3	19	15	1.5	115.38	0.00395598
MINDU	26	11	2	19	12	0.5	109.09	0.03598819
DEBU	32	16	2	21	19	2	118.75	0.00000601
MINPU	37	30	3	24	34	6	113.33	0.00000240
MLR	28	212	17.5	16	237.5	23.75	112.03	0.00003549
MAXDHR	35	19	1	21	21	1	110.53	0.00000023
MINDHR	35	18	1	21	20	1	111.11	0.00000023
DEBR	35	29	3	18	32	3	110.34	0.00000745
MPR	28	36	4.25	19	41	4.5	113.89	0.00002098
MINPR	48	34	4	25	38	3	111.76	0.00000031
MLF	29	395	24	21	431	25	109.11	0.00001708
DEBF	19	68	4.5	17	77	5	113.24	0.00001873
MDHF	49	38	4	28	43	3.25	113.16	0.00000000
MSPF	35	74	5.5	22	84.5	7.75	114.19	0.00000062
APSDF	61	22	3	36	24.5	3	111.36	0.00000062
TSDF	61	30	3	36	32	3	106.67	0.00000601
APDMF	34	24	3	22	28.5	4	118.75	0.00000601
TDMF	34	23	2	22	26	1.75	113.04	0.00000310
MLT	24	324	24	15	359	31.5	110.80	0.00023186
DEBT	38	46	4	24	51	5.25	110.87	0.00000459
PEBT	25	65	3	19	72	5	110.77	0.00000240
MSPT	23	73	9	19	81	9	110.96	0.00002606
APDMT	26	27	3	19	31	4.5	114.81	0.00019710
TDMT	26	19	3	19	21	2.5	110.53	0.00123906
NFPT	47	80	8.5	30	91.5	10.5	114.38	0.00000009
APDNFT	46	30	2.75	30	34	5	113.33	0.00000104
TDNFT	46	20	2.75	30	22	2.75	110.00	0.00000216

Zorita et al. 2016, 2020; Evangelista 2018). These aspects make CMOL an osteological collection of reference for an in-depth study of the sexual dimorphism of Late Prehistoric populations.

We propose a series of discriminant formulas for long bones to estimate sex in the large volume of disarticulated CMOL remains, as well as in prehistoric collections with characteristics like the study sample. This is one of the few works in which discriminant functions are formulated on prehistoric samples (Safont et al. 2000; Chovalopoulou et al. 2018; Cavazzuti et al. 2019), being the most ancient population analysed for this purpose and the first Iberian example of discriminant analysis on long bones of prehistoric chronology. This must be related to different aspects: (1) the absence of large samples of articulated skeletons of this chronology for the reasons already mentioned; (2) the poor preservation and usual fractionation of prehistoric bone remains which hinders anthropometric analysis; (3) the usual reuse of Late Prehistoric tombs which frequently leads to the identification of human remains from different chrono-cultural periods; and (4) the handicap of making reliable sex estimations in archaeological populations.

CMOL is the perfect site for such an analysis because of its unique characteristics: a recently excavated grave with rigorous recording and exhumation methods (Lomba et al. 2009), that housed an unprecedented osteological collection of well-preserved individuals of both sexes and of all age categories. The extensive series of 14C dating allows us to state that this grave was the regular burial place during 500 years for 10–15 generations of between 100 and 150 people (Díaz-Navarro et al. 2023b). Furthermore, mobility analyses based on ^{87}Sr/^{86}Sr isotopes (Merner 2017) indicate a low proportion of non-local individuals. Furthermore, the preservation of 167 articulated skeletons has made it possible to reliably define the morphological and morphometric

characteristics that best discriminate between sexes and to assign the sex of the bulk of articulated adult skeletons with minimal margins of error. Although nowadays genetics and proteomics are the most solid analytical tools to reach sexual diagnoses, the analysis of skeletal morphology and morphometry is still indispensable when working with adult skeletons, as it allows a fast, economic, direct and reliable analysis of the anatomical characteristics of the bones without destroying, altering or consuming the sample. Recent studies comparing the efficacy of proteomic, genomic and osteological methods in sexing archaeological samples (Buonasera et al. 2020) show high agreement between the three techniques and high confidence in osteological sex estimates when analysing adult individuals whose skeletons preserve the most dimorphic regions. The sex assignments of 19 CMOL individuals from this work and previous publications (Olalde et al. 2018; Villalba-Mouco et al. 2021) using ancient DNA coincide in 100 % of the cases with those made using standard anthropological methods.

This study corroborates that anthropometric variables of long bones are an excellent tool for sexual estimation of ancient human remains and their combination with morphological characters ensures more reliable sexual diagnoses (Işcan and Miller-Shaivitz, 1984; Black 1978; DiBennardo and Taylor, 1979; DiBennardo and Taylor 1979; Dittrick and Suchey 1986; Steyn and İşcan, 1997; Ordóñez et al. 2013).

Morphometric differences between sexes may be due to genetic reasons, geographical and environmental factors affecting growth (nutrition, stress…), or the interaction of any of the above causes (Stini 1969; Trancho et al. 2012). Cultural factors also play a role, such as physical activities or the existence of a radical gender division of labour (Borgognini Tarli and Repetto 1986). It is therefore essential to develop discriminant formulae for different population groups, geographical
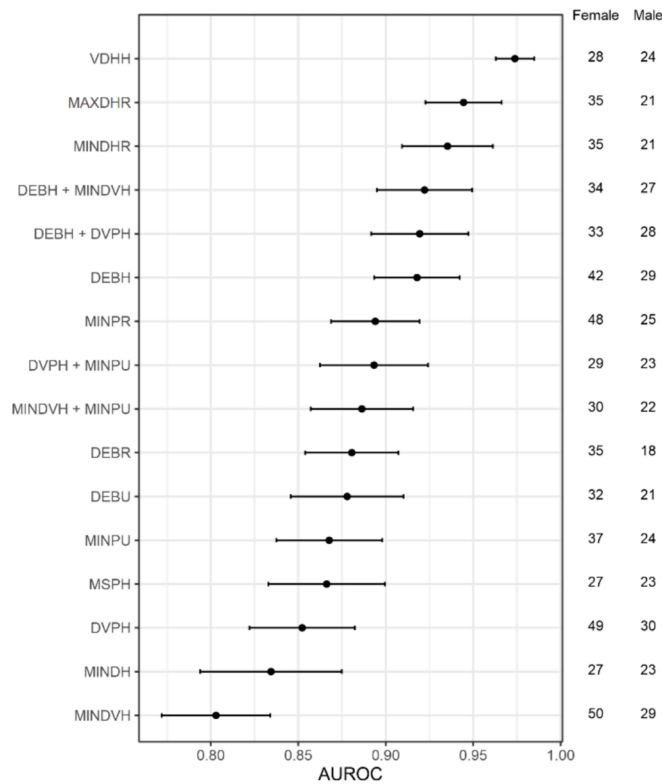
Fig. 4. Ranking of LD functions based on upper body measurements by model performance. Black dots and lines represent the mean AUROCs and minimum and maximum AUROCs obtained across 100 models (10-repetitions, 10-fold CV), respectively. "Female" and "Male" columns reflect sample size.

areas and chronologies.

In this work were obtained 52 linear discriminant functions with prediction thresholds above 0.8 on the ROC curve, with 8 exceeding 0.95 points. Of these, 17 are of the upper extremity and 35 of the lower extremity. Attention has been paid to the classification of isolated remains and variables have been combined to overcome success in sexual diagnosis. Therefore, 22 functions use a single variable and 30 combine two variables from the same bone or from another bone of the same limb. The best LDA models for sex prediction were those based on the

widths of the femur and humerus heads, alone and in combination.

It is traditionally accepted that sex differences are best reflected in the maximum length, as well as in the epiphyseal widths, by adding the measurements of diaphyseal diameters and diaphyseal perimeters (e.g., Black 1978; Işcan and Miller-Shaivitz, 1984; Dittrick and Suchey 1986; Jiménez-Arenas 2010). The above results suggest that the transverse variables are more dimorphic than the longitudinal ones in the CMOL population, which has already been pointed out by other authors (e.g., Alemán et al. 1999; Trancho et al. 1997, 2012), referring to the functional demand produced by physical activity. The epiphyses would be more affected by supporting body weight, while the diaphyses would be more affected by muscular activity itself.

The widths of the proximal epiphyses discriminate better than those of the distal epiphyses in all the long bones of the CMOL sample. This must be related to the fact that the proximal epiphyses contain the sexually dimorphic muscle attachments and transmit weight (France 1988, Kim et al. 2013; Djorojevic et al. 2016), so that the anatomy of the region is undoubtedly affected in terms of size and shape, which has an impact on sexual dimorphism and therefore on the accuracy of the resulting measurement.

The success of the femoral head diameter in discrimination must be related to the fact that this region is the most influenced by the dimorphism of the pelvis by the biomechanical link with the hip, as well as housing sexually dimorphic muscle attachments, which are generally more accentuated in males (Anastopoulou et al. 2014; Djorojevic et al. 2016; Machado et al. 2021). Furthermore, due to its robustness and density, the femur is the anatomical area less susceptible to damage and better preserved than other long bones, especially the proximal end. This variable performed greater accuracy over other features of the femur in a variety of populations studied worldwide (Alemán 1997; Seidemann et al. 1998; Mall et al. 2000; Safont et al. 2000; Purkait and Chandra 2004; Anastopoulou et al. 2014; Djorojevic et al. 2016). In contrast, this variable has been found to be less effective than other femoral variables in other osteological samples (Steyn and Işcan, 1997; Alunni-Perret et al. 2008), suggesting inter-population variability and the need to develop specific standards.

Traditionally, along with the femur, the tibia is one of the bones most used to derive standards for sex determination (Işcan and Miller-Shaivitz, 1984; Alemán 1997; González-Reimers et al. 2000; Ordóñez et al. 2013). In CMOL sample, the tibia variable with the greatest classification power is the circumference at the level of the nutrient foramen, a highly dimorphic area as it is the site of insertion of different muscles. It should be noted that it was not possible to obtain
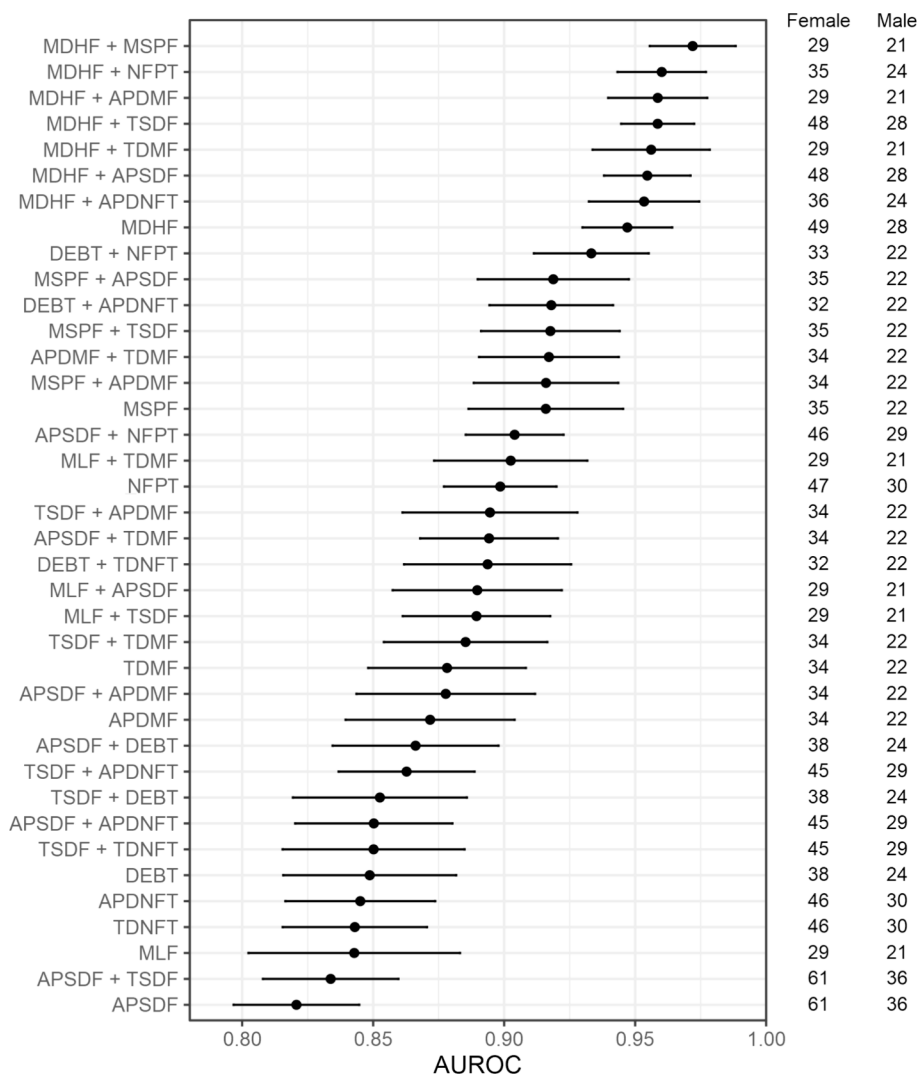
Table 4
Performance of LD functions using univariate and bivariate combinations of measurements of the long bones of the upper limb skeleton as predictors. Female (F) and Male (M) columns depict sample size. ROC, Sens and Spec columns refer to mean values of area under the receiver operator curve, sensibility and specificity. ROCSD, SensSD and SpecSD columns refer to standard deviations relative to the aforementioned average values. Coefficients (Coef) 1 and 2 columns contain the parameters' values for the LD equations (Coef2 is empty when only one variable is considered). Threshold accounts for the decision boundary for sex prediction. Variables marked in bold exceed 90 points on the ROC curve.

Variables	F	M	ROC	Sens	Spec	ROCSD	SensSD	SpecSD	Coef1	Coef2	Threshold
DEBH	42	29	**0.91792**	0.90700	0.73667	0.12290	0.14478	0.24639	0.34277		18.99833
VDHH	28	24	**0.97389**	0.99333	0.75667	0.05539	0.04690	0.24780	0.47451		18.49879
DVPH	49	30	0.85225	0.89700	0.56667	0.15200	0.14334	0.27828	0.26284		14.51136
MSPH	27	23	0.86625	0.88000	0.66500	0.16834	0.21066	0.30475	0.26145		15.34867
MINDH	27	23	0.83444	0.88333	0.62500	0.20386	0.21254	0.30739	0.62423		9.90669
MINDVH	50	29	0.80300	0.82000	0.58000	0.15659	0.15176	0.27266	0.60362		9.16942
DEBU	32	21	0.87799	0.89583	0.57500	0.16245	0.19939	0.33616	0.52725		9.21789
MINPU	37	24	0.86771	0.89250	0.70833	0.15229	0.16679	0.25112	0.37351		12.21465
MAXDHR	35	21	**0.94451**	0.94333	0.81167	0.10991	0.12072	0.27489	0.73199		14.85600
MINDHR	35	21	**0.93535**	0.94250	0.81167	0.13082	0.11707	0.27894	0.74778		14.42143
DEBR	35	18	0.88063	0.92000	0.57000	0.13413	0.14357	0.38284	0.48182		14.58492
MINPR	48	25	0.89404	0.82500	0.67833	0.12728	0.21004	0.30636	0.37324		13.40804
DEBH+DVPH	33	28	**0.91944**	0.90667	0.73500	0.13991	0.15947	0.27123	0.26959	0.06751	18.68951
DEBH+MINDVH	34	27	**0.92222**	0.91667	0.74167	0.13758	0.13914	0.27664	0.30176	0.07818	17.89516
MINDVH+MINPU	30	22	0.88639	0.84333	0.72667	0.14747	0.21945	0.27064	0.00529	0.38961	12.76566
DVPH+MINPR	41	23	0.89313	0.84600	0.67500	0.11623	0.18865	0.29049	0.08585	0.28084	14.80625
DVPH+MINPU	29	23	0.89333	0.85333	0.63500	0.15563	0.21748	0.29842	0.10847	0.25860	14.39776

Fig. 5. Ranking of LD functions based on lower body measurements by model performance. Black dots and lines represent the mean AUROCs and minimum and maximum AUROCs obtained across 100 repetitions (10-repetitions, 10-fold CV), respectively. "Female" and "Male" columns reflect sample size.

discriminant functions for the proximal epiphysis due to its poor preservation in the sample (N=44), although the results are very promising (0.927).

In contrast, this work highlights the importance of upper limb bones variables, such as the proximal epiphyses of the humerus and radius, which express excellent results in the sex classification of CMOL population. The humeral head diameter offers the best results in the sex rankings (0.975) and is the most dimorphic variable in the humerus, a pattern already observed in other osteological collections (France 1988; Carretero et al. 1995; Steyn and İşcan, 1997; Mall et al. 2000; Ríos Frutos 2005; Charisi et al., 2010; Jiménez-Arenas 2010). Some authors point to the cessation of female growth as the main cause of sex differences in the humeral head and is useful for sex diagnosis after the age of 15 years (Rissech et al. 2013). However, radium has been less used in work on sexual dimorphism even though in some osteological collections it has proved to be the most powerful discriminating bone (Mall et al. 2000; Cavazzuti et al. 2019; Nogueira et al. 2023). Again, as with the CMOL sample, it is the proximal radial epiphysis that gives the best overall classification results (Mall et al. 2000; Ríos Frutos 2005; Charisi et al. 2011; Cavazzuti et al. 2019).

A discriminant analysis of CMOL diaphyseal circumferences and diameters yields positive values, but less than 0.9 on the ROC curve, except for the circumference at the femoral midshaft (MSPF). The diaphyses show a strong response to mechanical loading since, unlike the

articular areas in adulthood, they can undergo alterations in the geometry of the cortical bone (Béguelin and González, 2008). This can be explained by the high level of robustness expressed by CMOL females. The comparison of averages of the diaphyseal and robustness indexes of males and females shows that there are no statistically significant differences in any of the long bones, except for the femur (Díaz-Navarro 2022). This suggests that CMOL females attained muscle development on a par with that of males and, therefore, sexual differences are mainly manifested in epiphyseal widths.

Long bone lengths, on the other hand, provide positive results in discriminating CMOL individuals but do not usually exceed 85 points on the ROC curve and, moreover, no long bone reaches the minimum required sample size. In contrast, lengths have been generally very useful variables in the sexing of osteological collections from Iberia (Alemán 1997). Maximum lengths are mainly influenced by variation in the development of bone robustness throughout ontogeny, with a strong genetic component (Cowgill and Hager 2007), although there is always an area of overlap between short male individuals and tall females (Garrido-Varas et al. 2014).

The use of LDA and RF analysis allows to compare different techniques for sexual diagnosis with long bones and to test the validity of machine learning in bioarchaeology. Estimating the sex by combining measurements of different anatomical regions using machine learning classifiers has recently been explored with very promising results

Table 5

Performance of LD functions using univariate and bivariate combinations of measurements of the long bones of the lower limb skeleton as predictors. Female (F) and Male (M) columns depict sample size. ROC, Sens and Spec columns refer to mean values of area under the receiver operator curve, sensibility and specificity. ROCSD, SensSD and SpecSD columns refer to standard deviations relative to the aforementioned average values. Coefficients (Coef) 1 and 2 columns contain the parameters' values for the LD equations (Coef2 is empty when only one variable is considered). Threshold accounts for the decision boundary for sex prediction. Variables marked in bold exceed 90 points on the ROC curve.

Variables	F	M	ROC	Sens	Spec	ROCSD	SensSD	SpecSD	Coef1	Coef2	Threshold
MLF	29	21	0.84278	0.89833	0.72500	0.20352	0.16394	0.31550	0.04199		17.18438
MDHF	49	28	**0.94700**	0.96000	0.75167	0.08634	0.08528	0.26003	0.41203		16.77019
MSPF	35	22	**0.91590**	0.91083	0.70833	0.14880	0.16508	0.28952	0.20636		16.32271
APSDF	61	36	0.82071	0.86691	0.50000	0.12090	0.14465	0.23180	0.52002		12.17927
APDMF	34	22	0.87174	0.90917	0.64000	0.16262	0.14893	0.31757	0.41951		10.92915
TDMF	34	22	0.87819	0.88167	0.73333	0.15172	0.18009	0.30521	0.71091		17.24907
DEBT	38	24	0.84868	0.89417	0.63167	0.16638	0.18647	0.31094	0.31679		15.28308
NFPT	47	30	0.89850	0.86250	0.60000	0.10831	0.18767	0.30704	0.17621		15.12095
APDNFT	46	30	0.84508	0.87050	0.53333	0.14412	0.15733	0.26379	0.38528		12.55115
TDNFT	46	30	0.84300	0.73450	0.73333	0.13878	0.20704	0.23689	0.56977		12.07586
MLF+APSDF	29	21	0.88972	0.86167	0.73167	0.16276	0.21459	0.28809	0.03610	0.15624	18.38018
MLF+TSDF	29	21	0.88944	0.83333	0.62667	0.14197	0.19678	0.31969	0.03388	0.19497	19.87028
MLF+TDMF	29	21	**0.90250**	0.83000	0.73000	0.14744	0.22842	0.32375	0.02676	0.38810	20.37173
MDHF+MSPF	29	21	**0.97194**	0.92500	0.77000	0.08258	0.16476	0.29472	0.31828	0.07691	19.04634
MDHF+APSDF	48	28	**0.95463**	0.92650	0.76667	0.08331	0.10905	0.27010	0.38671	0.06487	17.25504
MDHF+TSDF	48	28	**0.95858**	0.94650	0.74667	0.07030	0.09108	0.25894	0.44141	−0.06430	15.95970
MDHF+APDMF	29	21	**0.95861**	0.93167	0.81500	0.09537	0.14621	0.25170	0.34841	0.13974	17.81621
MDHF+TDMF	29	21	**0.95611**	0.95667	0.90667	0.11279	0.11991	0.24195	0.35217	0.18281	18.77263
MDHF+NFPT	35	24	**0.96014**	0.92000	0.77500	0.08520	0.14599	0.28758	0.30302	0.05946	17.48963
MDHF+APDNFT	36	24	**0.95340**	0.93750	0.76833	0.10622	0.13209	0.25612	0.32803	0.11238	17.07799
APSDF+TSDF	61	36	0.83378	0.81905	0.57667	0.13036	0.14971	0.24495	0.38927	0.15321	13.87831
APSDF+APDMF	34	22	0.87771	0.90417	0.62667	0.17190	0.15415	0.32663	0.11722	0.36094	12.10756
APSDF+TDMF	34	22	0.89424	0.84500	0.69000	0.13254	0.18540	0.29684	0.18609	0.57392	18.21820
APSDF+DEBT	38	24	0.86618	0.87750	0.65500	0.15965	0.17626	0.31360	0.17398	0.26001	16.64345
APSDF+NFPT	46	29	**0.90400**	0.87100	0.60667	0.09414	0.14808	0.28772	0.08253	0.15828	15.55210
APSDF+APDNFT	45	29	0.85025	0.83600	0.58500	0.15127	0.16454	0.29349	0.15279	0.30652	13.61408
TSDF+APDMF	34	22	0.89458	0.87667	0.72667	0.16827	0.18291	0.27981	0.21694	0.32947	15.26597
TSDF+TDMF	34	22	0.88528	0.83500	0.71000	0.15733	0.20343	0.31034	0.16391	0.55510	18.51801
TSDF+DEBT	38	24	0.85257	0.86500	0.62500	0.16745	0.16741	0.31011	0.06494	0.28444	15.75685
TSDF+APDNFT	45	29	0.86275	0.83000	0.56167	0.13081	0.15092	0.30026	0.06989	0.34687	13.50094
TSDF+TDNFT	45	29	0.85017	0.82750	0.65167	0.17508	0.18469	0.27534	0.15859	0.45323	14.58517
APDMF+TDMF	34	22	**0.91708**	0.87917	0.63833	0.13445	0.16343	0.32572	0.26696	0.43106	17.41370
DEBT+NFPT	33	22	**0.93326**	0.90583	0.71833	0.11047	0.15336	0.30309	0.15871	0.12073	18.03847
DEBT+APDNFT	32	22	**0.91799**	0.89917	0.66333	0.11893	0.16036	0.32392	0.20979	0.22443	17.46313
DEBT+ TDNFT	32	22	0.89368	0.87500	0.72000	0.16076	0.18215	0.29194	0.24408	0.22253	16.54057

Table 6

Confusion matrix of random forest model on test data.

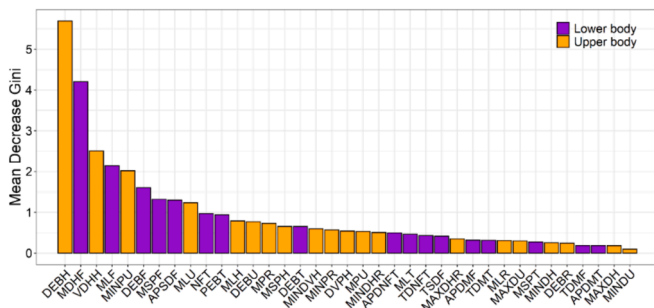| Reference Prediction	Female	Male
Female	20	2
Male	0	10

Fig. 6. Importance of body measurements in random forest model. Bar heights represent the average decrease in Gini index across 1000 trees. Colours indicate whether a measurement comes from lower or upper body.

(Mircea et al. 2015; Navega et al., 2015; Curate et al. 2016, 2017; Nikita and Nikitas, 2020; Bidmos et al. 2023; Knecht et al., 2023; Taskın-Senol et al. 2023) although its usefulness on prehistoric samples has so far not been proven. Unlike univariate LDA, the RF approach uses all variables from all bones simultaneously and allows ranking the importance of long bones in sex discrimination. Therefore, this machine learning approach is more comprehensive than univariate LDA, as it considers correlations between long bones instead of treating them as independent observations. In contrast, performing LDA on each long bone allows researchers to predict sex of individuals in other coeval collections with similar characteristics that do not have such a complete sample. The RF approach confirms the importance of MDHF and VDHH in sex classification although it gives the greatest power of classification to DEBH. The recommended strategy for assigning sex to individuals would follow a consensus majority vote among the best performing models, provided that data are available.

If we compare our RF results with those obtained when analysing long bones from other European prehistoric series (Table 7), we observe

Table 7

Comparison of the classification power of different long bones discriminant analysis from different European prehistoric series.

Archaeological series	Chronology	N	Accuracy	Bone	Reference
CMOL (Spain)	Copper Age	109	94	H, R, U, F, T	This study
Italian collections	Bronze and Iron Age	60	88.3	R	Cavazzuti et al. 2019
Greek collections	II millennium BC to Middle Age	54	74.2	H	Chovalopoulou et al. 2018

that the results obtained are excellent. Previously published compiled functions do not reach 90 points of accuracy in sex classification, whereas in our study we obtain a Random Forest accuracy rate of 0.94 when analysing the data for all variables together.

These results lead to the proposal and discussion of these functions as part of an alternative basis for more reliable sex estimates in the large volume of disarticulated CMOL remains, as well as in other prehistoric skeletal series from southeastern Iberia. Each population is subject to different genetic, cultural and environmental factors, which makes it necessary to use formulas based on populations as homogeneous as possible to the one that is the object of study. Following these criteria, a function can be applied to the very population that was used to produce it, as well as to those populations that show similar sexual dimorphism indexes within the different metric variables to be analysed, with high biological affinity and that share similar environmental conditions (Alemán et al. 1999; Safont et al. 2000). We suggest applying the functions developed here to estimate the sex of Chalcolithic skeletal collections from southeastern Iberia since we have shown that some series with available osteometric data, such as those from Barranco de la Higuera (Murcia) (Font 1980) or Alta Andalucía (Granada) (Jiménez-Brobeil 1988), present very similar sexual dimorphism indices (S10 Table) and, moreover, are spatially and chronologically homogeneous, so we can assume that they lived in very similar environmental conditions. Finally, it is necessary to underscore that the almost total absence of recent osteological studies of prehistoric collections that include osteometric analyses with open access data makes it difficult to compare and test the validity of the data, to gain further insight into inter- and intra-population variability and to elaborate broader discriminant analyses.

6. Conclusions

This study corroborates that anthropometric variables of long bones are an excellent tool for sexual estimation of ancient human remains and their combination with morphological characters guarantees more reliable sexual diagnoses. The analysis of 37 osteometric variables of 109 articulated skeletons from the Chalcolithic site of Camino del Molino has allowed us to obtain 52 univariate and bivariate discriminant functions that reach a classification power of 0.98 (AUROC). These can serve as a sex estimation method for the large volume of isolated CMOL bone remains, as well as a simple alternative method for other homogeneous osteological prehistoric samples. Our study shows that the width of the epiphyses are the variables that best express sexual dimorphism in the study sample, with the femur and humerus being the ones with the best classification results. Finally, the pioneering application of machine learning approaches to samples of this chronology has allowed us to establish a ranking of the usefulness of long bones for sexual estimation, positioning the width of the distal epiphysis of the humerus in first position.

CRediT authorship contribution statement

S. Díaz-Navarro: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **S. Díez-Hermano:** Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **M.A. Rojo-Guerra:** Funding acquisition, Project administration, Resources, Supervision, Writing – review & editing. **J. Lomba Maurandi:** Funding acquisition, Project administration, Resources, Writing – review & editing. **C. Valdiosera:** Funding acquisition, Investigation, Methodology, Software, Writing – original draft, Writing – review & editing. **T. Gunther:** Funding acquisition, Investigation, Methodology, Software, Writing – review & editing. **M. Haber Uriarte:** Data curation, Project administration, Resources, Supervision, Writing – review & editing.

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.

Data availability

I have shared the link to my data/code at the Attach File step

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jasrep.2024.104730.

References

Alemán, I., Botella, M.C., du Souich, P., 1999. Aplicación de las funciones discriminantes en la determinación del sexo. Estud. Antropol. Biol. 9, 221–230.

Alemán, I., 1997. Determinación del sexo en el esqueleto postcraneal. Estudio de una población mediterránea actual (Unpublished PhD Dissertation). Universidad de Granada.

Alemán, I., Botella, M.C., du Souich Henrici, P., 2012. Aplicación de las funciones discriminantes en la determinación del sexo. Estudios De Antropología Biológica 9. https://doi.org/10.22201/iia.14055066p.1999.30818.

Alonso-Llamazares, C., Pablos, A., 2019. Sex estimation from the calcaneus and talus using discriminant function analysis and its possible application in fossil remains. Archaeol. Anthropol. Sci. 11, 4927–4946. https://doi.org/10.1007/s12520-019-00855-y.

Alunni-Perret, V., Staccini, P., Quatrehomme, G., 2008. Sex determination from the distal part of the femur in a French contemporary population. Forensic Sci. Int. 175, 113–117. https://doi.org/10.1016/j.forsciint.2007.05.018.

Anastopoulou, I., Eliopoulos, C., Valakos, E.D., Manolis, S.K., 2014. Application of Purkait's triangle method on a skeletal population from southern Europe. Forensic Sci. Int. 245, 203.e1–203.e4. https://doi.org/10.1016/j.forsciint.2014.10.005.

Avilés, A., Haber, M., Lomba, J., 2012. Enterramiento múltiple del yacimiento calcolítico de Camino del Molino (Caravaca de la Cruz, Murcia). In: Turbon, D., Fañanás, L., Rissech, C., Rosa, A. (Eds.), Biodiversidad Humana y Evolución. Universitat Autònoma de Barcelona, Barcelona, pp. 186–188.

Bašić, Ž., Anterić, I., Vilović, K., Petaros, A., Bosnar, A., Madžar, T., Polašek, O., Anđelinović, Š., 2013. Sex determination in skeletal remains from the medieval Eastern Adriatic coast – discriminant function analysis of humeri. Croat. Med. J. 54, 272–278. https://doi.org/10.3325/cmj.2013.54.272.

Béguelin, M., González, P.N., 2008. Estimación del sexo en poblaciones del sur de Sudamérica mediante funciones discriminantes para el fémur. Rev. Argentina Antropol. Biol. 10, 55–70.

Bidmos, M.A., Mazengenya, P., 2021. Accuracies of discriminant function equations for sex estimation using long bones of upper extremities. Int. J. Leg. Med. 135, 1095–1102. https://doi.org/10.1007/s00414-020-02458-y.

Bidmos, M.A., Olateju, O.I., Latiff, S., Rahman, T., Chowdhury, M.E.H., 2023. Machine learning and discriminant function analysis in the formulation of generic models for sex prediction using patella measurements. Int. J. Leg. Med. 137, 471–485. https://doi.org/10.1007/s00414-022-02899-7.

Black, T.K., 1978. A new method for assessing the sex of fragmentary skeletal remains: Femoral shaft circumference. Am. J. Phys. Anthropol. 48, 227–231. https://doi.org/10.1002/ajpa.1330480217.

Borgognini Tarli, S.M., Repetto, E., 1986. Methodological considerations on the study of sexual dimorphism in past human populations. Hum. Evol. 1, 51–66. https://doi.org/10.1007/BF02437285.

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, Fayetteville.

Buonasera, T., Eerkens, J., de Flamingh, A., Engbring, L., Yip, J., Li, H., et al., 2020. A comparison of proteomic, genomic, and osteological methods of archaeological sex estimation. Sci. Rep. 10, 11897. https://doi.org/10.1038/s41598-020-68550-w.

Carretero, J.M., Lorenzo, C., Arsuaga, J.L., 1995. Análisis multivariante del húmero en la colección de restos identificados de la Universidad de Coimbra (Portugal). Antropol. Port. 13, 139–156.

Catagnano, V., 2016. Aproximación morfométrica y paleogenética al estudio de la variabilidad de canis l. familiaris en la península ibérica desde el neolítico hasta época romana y su contextualización en el ámbito del mediterráneo occidental (Unpublished PhD Dissertation). Universitat Autònoma de Barcelona.

Cavazzuti, C., Bresadola, B., d'Innocenzo, C., Interlando, S., Sperduti, A., 2019. Towards a new osteometric method for sexing ancient cremated human remains. analysis of late bronze age and iron age samples from italy with gendered grave goods. PLoS One 14, e0209423. https://doi.org/10.1371/journal.pone.0209423.

Chambon, P., 2003. Les Morts dans les sépultures collectives néolithiques en France - Du cadavre aux restes ultimes - CNRS Editions, Suppléments à Gallia Préhistoire. CNRS Editions, Paris.

Charisi, D., Eliopoulos, C., Vanna, V., Koilias, C.G., Manolis, S.K., 2011. Sexual dimorphism of the arm bones in a modern greek population. J. Forensic Sci. 56, 10–18. https://doi.org/10.1111/j.1556-4029.2010.01538.x.

Chovalopoulou, M.E., Bertsatos, A., Zoi, I., Manolis, S.K., Valakos, E., 2018. Reliability testing of metric methods for sex determination in ancient skeletal remains in Greece. Mediter. Archaeol. Archaeom. 18, 37.

Cowgill, L.W., Hager, L.D., 2007. Variation in the development of postcranial robusticity: an example from Çatalhöyük, Turkey. Int. J. Osteoarchaeol. 17, 235–252. https://doi.org/10.1002/oa.882.

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 Sci. Int. 266, 579.e1–579.e7. https://doi.org/10.1016/j.forsciint.2016.06.011.

Curate, F., Albuquerque, A., Ferreira, I., Cunha, E., 2017. Sex estimation with the total area of the proximal femur: A densitometric approach. Forensic Sci. Int. 275, 110–116. https://doi.org/10.1016/j.forsciint.2017.02.035.

Dabbs, R.D., 2020. Estimating the sex of Ancient Egyptian skeletal remains: methods from Tell el-Amarna. Bioarchaeol. Near East 14, 11–26.

Díaz-Navarro, S., García-González, R., Cirotto, N., Haber Uriarte, M., 2023a. New insight into prehistoric craft specialisation. Tooth-tool use in the Chalcolithic burial site of Camino del Molino, Murcia, SE Spain. JAS Rep. 50, 104066 https://doi.org/10.1016/j.jasrep.2023.104066.

Díaz-Navarro, S., Haber Uriarte, M., García-González, R., 2023b. Holes in the Head. Double cranial surgery on an individual from the Chalcolithic burial site of Camino del Molino (SE Spain). Int. J. Paleopathol. 43, 22–30. https://doi.org/10.1016/j.ijpp.2023.09.003.

Díaz-Navarro, S., Haber Uriarte, M., Tejedor-Rodríguez, C., Lomba Maurandi, J., 2023c. Emphasising the community: demographic composition of an exceptional tomb—the Chalcolithic burial site of Camino del Molino, Caravaca de la Cruz, Murcia. Archaeol. Anthropol. Sci. 15, 140. https://doi.org/10.1007/s12520-023-01844-y.

Díaz-Navarro, S., Tejedor, C., Santana, J., Díaz-Zorita, M., 2023d. Apuntes para la excavación, análisis e interpretación de los contextos funerarios colectivos de la Prehistoria Reciente. In: Rojo-Guerra, M.A., Díaz-Navarro, S. (Eds.), Las Tumbas y Los Muertos Los Muertos Entre Las Tumbas. Sevilla, Junta de Andalucía, pp. 15–42.

Díaz-Navarro, S. 2022. Caracterización antropológica de una comunidad calcolítica mediterránea. El enterramiento colectivo de Camino del Molino (Caravaca de la Cruz, Murcia). Unpublished PhD Dissertation. Universidad de Valladolid.

Díaz-Zorita, M., 2017. The copper age in south-west Spain: A bioarchaeological approach to prehistoric social organisation. British Archaeological Reports, Oxford.

Díaz-Zorita, M., Jiménez, G., Escudero, J., Robles, S., Lozano, Á., Sánchez Romero, M., et al., 2016. Estudio bioarqueológico de la necrópolis megalítica de El Barranquete (Níjar, almería). Menga 7, 71–98.

Díaz-Zorita, M., Beck, J., Aranda, G., Milesi, L., Sánchez Romero, M., Medina, A.L., et al., 2020. The deposition of human remains inside chalcolithic ditched enclosures: ditch 5 at marroquíes (Jaén, Spain). Eur. J. Archaeol. 23, 330–355. https://doi.org/10.1017/eaa.2020.4.

DiBennardo, R., Taylor, J.V., 1979. Sex assessment of the femur: A test of a new method. Am. J. Phys. Anthropol. 50, 635–637. https://doi.org/10.1002/ajpa.1330500415.

Dittrick, J., Suchey, J.M., 1986. Sex determination of prehistoric central California skeletal remains using discriminant analysis of the femur and humerus. Am. J. Phys. Anthropol. 70, 3–9. https://doi.org/10.1002/ajpa.1330700103.

Djorojevic, M., Roldán, C., Botella, M., Alemán, I., 2016. Estimation of Purkait's triangle method and alternative models for sex assessment from the proximal femur in the Spanish population. Int. J. Leg. Med. 130, 245–251. https://doi.org/10.1007/s00414-015-1201-7.

Etxeberria, F., Herrasti, L., 2007. Los restos humanos del enterramiento de SJAPL: Caracterización de la muestra, tafonomía, paleodemografía y paleopatología. In: San Juan Ante Portam Latinam: Una Inhumación Colectiva Prehistórica En El Valle Medio Del Ebro. Diputación Foral de Álava, Álava, pp. 159–282.

Evangelista L.E.S. 2018. Resting in peace or in pieces? Tomb I and death management in the 3rd millennium BC at the Perdigões Enclosure (Reguengos de Monsaraz, Portugal). Unpublished PhD Dissertation, Universidade de Coimbra.

Fawcett, T., 2006. An introduction to ROC analysis. Pattern Recogn. Lett. 27 (8), 861–874. https://doi.org/10.1016/j.patrec.2005.10.010.

Font, A., 1980. Estudio antropológico de los esqueletos de la Cueva del Barranco de la Higuera. Anales de la Universidad de Murcia 37 (3), 267–290.

France, D.L., 1988. Osteometry at muscle origin and insertion in sex determination. Am. J. Phys. Anthropol. 76, 515–526. https://doi.org/10.1002/ajpa.1330760411.

Francis, S., Makoviychuck, Y., Chavoinik, L., Borgel, S., Pokhojaev, A., Roul, V., Peled, N., May, H., 2023. A new method for sex estimation based on femoral cross-sectional geometry measurements and its validation using recent and ancient populations. Int. J. Leg. Med. 137, 1263–1275. https://doi.org/10.1007/s00414-023-03009-x.

Garrido-Varas, C., Thompson, T., Campbell, A., 2014. Parámetros métricos para la determinación de sexo en restos esqueletales chilenos modernos. Chungará (Arica) 46, 285–294. https://doi.org/10.4067/S0717-73562014000200009.

Godycki, M., 1957. Sur la certitude de détermination du sexe d'après le fémur, le cubitus et l'humérus. Bull. Mémoires Soc. Anthropol. Paris 8, 405–410. https://doi.org/10.3406/bmsap.1957.2688.

González-Reimers, E., Velasco-Vázquez, J., Arnay-de-la-Rosa, M., Santolaria-Fernández, F., 2000. Sex determination by discriminant function analysis of the right tibia in the prehispanic population of the Canary Islands. Forensic Sci. Int. 108, 165–172. https://doi.org/10.1016/S0379-0738(99)00205-4.

Haber, M., Avilés, A., Lomba, J., 2012. Estudio antropológico preliminar de los restos humanos calcolíticos del enterramiento múltiple de Camino del Molino (Caravaca de la Cruz, Murcia). In: Turbon, D., Fañanás, L., Rissech, C., Rosa, A. (Eds.), Biodiversidad Humana y Evolución. Universitat Autònoma de Barcelona, Barcelona, pp. 236–242.

Hanihara, K., 1958. Sexual diagnosis of japanese long bones by means of discriminant function. J. Anthropol. Soc. Nippon 66, 187–196. https://doi.org/10.1537/ase1911.66.187.

Harrison, E., Drake, T., Ots, R. 2023. Finalfit: Quickly Create Elegant Regression Results Tables and Plots when Modelling. https://CRAN.R-project.org/package=finalfit.

Herrmann, B., Grupe, G., Hummel, S., 1990. Prähistorische anthropologie: leitfaden der feld- und labormethoden. Springer, Berlin.

Işcan, M,Y, Miller-Shaivitz, P.1986 Sexual Dimorphism in the Femur and Tibia. In: Reichs KJ, editor. Forensic Osteology: Advances in Identification of Human Remains. Springfield, Ill., U.S.A: Charles C Thomas Pub.

Işcan, M.Y., Miller-Shaivitz, P., 1984. Determination of sex from femur in blacks and whites. Coll. Antropol. 8, 169–177.

Jerković, I., Bašić, Ž., Kružić, I., Anđelinović, Š., 2016. Sex determination from femora in late antique sample from Eastern Adriatic coast (Salona necropolis). Anthropol. Rev. 79, 59–67. https://doi.org/10.1515/anre-2016-0005.

Jiménez-Arenas, J.M., 2009. Discriminación de sexo en una población medieval del sur de la Península Ibérica. Cuad. Prehist. Arqueol. Univ. Granada 19, 463–477. https://doi.org/10.30827/cpag.v19i0.201.

Jiménez-Arenas, J.M., 2010. Discriminación de sexo en una población medieval del sur de la península ibérica mediante el uso de variables simples. Int. J. Morphol. 28, 667–672.

Jiménez-Brobeil SA. 1988. Estudio antropológico de las poblaciones neolíticas y de la edad del cobre en la alta Andalucia. Unpublished PhD Dissertation. Universidad de Granada.

Kim, J.H., 2009. Estimating classification error rate: Repeated cross-validation, repeated hold-out and bootstrap. Comput. Stat. Data Anal. 53 (11), 3735–3745. https://doi.org/10.1016/j.csda.2009.04.009.

Kim, D.I., Kwak, D.S., Han, S.H., 2013. Sex determination using discriminant analysis of the medial and lateral condyles of the femur in Koreans. Forensic Sci. Int. 233, 121–125. https://doi.org/10.1016/j.forsciint.2013.08.028.

Knecht, S., Santos, F., Ardagna, Y., Alunni, V., Adalian, P., Nogueira, L., 2023. Sex estimation from long bones: A machine learning approach. Int. J. Leg. Med. 137, 1887–1895. https://doi.org/10.1007/s00414-023-03072-4.

Krogman, W., Iscan, M.Y. 1986. Human Skeleton in Forensic Medicine. Springfield: Charles C Thomas Pub.

Kuhn, M., 2008. Building predictive models in R using the caret package. J. Stat. Softw. 28 (5), 1–26. https://doi.org/10.18637/jss.v028.i05.

Liaw, A., Wiener, M., 2002. Classification and regression by randomforest. R News 2 (3), 18–22.

Lomba, J., Martínez, M.L., Martínez, F.R., Fernández, A.A., 2009. El enterramiento múltiple, calcolítico, de Camino del Molino (Caravaca, Murcia): Metodología y primeros resultados de un yacimiento excepcional. Trab. Prehist. 66, 143–160.

López-Bueis, I., 1995. Dimorfismo sexual de la tibia: estudio biométrico y paleopatológico (Unpublished PhD Dissertation). Universidad Complutense de Madrid.

Machado, D., Garcel, F., Pérez Pérez, V., 2021. Determinación del sexo a partir del fémur mediante funciones discriminantes, en cubanos de ascendencia hispánica. Rev. Int. Antropol. Odontol. Forense 4, 6–14.

Mahieu, É., 1987. L'hypogée des Boileau (Vaucluse). Vers une meilleure connaissance des rites funéraires du Néolithique provençal. Bull. Soc. Prehist. Fr. 84, 5–7. https://doi.org/10.3406/bspf.1987.9804.

Mall, G., Graw, M., Gehring, K.-D., Hubig, M., 2000. Determination of sex from femora. Forensic Sci Int, Selected papers from the 4th International Symposium Advances in 113, 315–321. doi: 10.1016/S0379-0738(00)00240-1.

Martin, R., Saller, K., 1957. Lehrbuch der Anthropologie, in systematischer Darstellung. Fischer, Stuttgart.

Merner C. 2017. Reconstructing Southeast Iberian copper age mobility: a strontium isotope analysis of the Camino del Molino mass burial. Unpublished PhD Master Dissetation, Memorial University of Newfoundland.

Mircea, I.-G., Czibula, G., Petrusel, M.-R., 2015. Sex identification in archaeological remains using decision tree learning. Studia univ. Babeș–BOLYAI, INFORMATICA LX(2), 92-103.

Mittnik, A., Wang, C.-C., Svoboda, J., Krause, J., 2016. A molecular approach to the sexing of the triple burial at the upper paleolithic site of dolní věstonice. PLoS One 11, e0163019. https://doi.org/10.1371/journal.pone.0163019.

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. Int. J. Leg. Med. 129, 651–659. https://doi.org/10.1007/s00414-014-1070-5.

Nikita, E., Nikitas, P., 2020. On the use of machine learning algorithms in forensic anthropology. Leg. Med. 47, 101771 https://doi.org/10.1016/j.legalmed.2020.101771.

Nogueira, L., Santos, F., Castier, F., Knecht, S., Bernardi, C., Alunni, V., 2023. Sex assessment using the radius bone in a French sample when applying various statistical models. Int. J. Leg. Med. 137, 925–934. https://doi.org/10.1007/s00414-023-02981-8.

Novak, M., 2016. Sex assessment using the femur and tibia in medieval skeletal remains from ireland: Discriminant function analysis. Coll. Antropol. 40, 17–22.

Olalde, I., Brace, S., Allentoft, M.E., Armit, I., Kristiansen, K., Booth, T., et al., 2018. The Beaker phenomenon and the genomic transformation of northwest Europe. Nature 555, 190–196. https://doi.org/10.1038/nature25738.

Ordóñez, A.C., Arnay-de-la-Rosa, M., Fregel, R., Trujillo-Mederos, A., Pestano, J., González-Reimers, E., 2013. Genetic sexing to determine the optimal discriminant functions for the analysis of archaeological remains from El Hierro (Canary Islands). JAS 40, 4411–4419. https://doi.org/10.1016/j.jas.2013.06.025.

Özer, İ, Katayama, K., 2006. Sex determination using the femur in an ancient Anatolian population. Anthropol. Anz. 64, 389–398.

Pearson, K., Bell, J., 1917. A Study of the Long Bones of the English Skeleton. Cambridge University Press, Cambridge.

Perini, T.A., de Oliveira, G.L., dos S Ornellas, J., de Oliveira, F.P., 2005. Technical error of measurement in anthropometry. Rev. Bras. Med. Esporte 11, 81–85. https://doi.org/10.1590/S1517-86922005000100009.

Purkait, R., Chandra, H., 2004. A study of sexual variation in Indian femur. Forensic Sci. Int. 146, 25–33. https://doi.org/10.1016/j.forsciint.2004.04.002.

R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2022. https://www.R-project.org/.

Ríos Frutos, L., 2005. Metric determination of sex from the humerus in a Guatemalan forensic sample. Forensic Sci. Int. Sex Body Size 147, 153–157. https://doi.org/10.1016/j.forsciint.2004.09.077.

Rissech, C., López-Costas, O., Turbón, D., 2013. Humeral development from neonatal period to skeletal maturity—Application in age and sex assessment. Int. J. Leg. Med. 127, 201–212. https://doi.org/10.1007/s00414-012-0713-7.

Rissech, C., Malgosa, A., 1997. Sex prediction by discriminant function with central portion measures of innominate bones. HOMO- J. Compar. Hum. Biol. 48, 22–32.

Rivera García NA. 2011. Impacto biológico y cultural del Neolítico en poblaciones del Norte de la Península Ibérica. Unpublished PhD Dissertation. Universidad del País Vasco.

Safont, S., Malgosa, A., Subirà, M.E., 2000. Sex assessment on the basis of long bone circumference. Am. J. Phys. Anthropol. 113, 317–328. https://doi.org/10.1002/1096-8644(200011)113:3<317::AID-AJPA4>3.0.CO;2-J.

Seidemann, R.M., Stojanowski, C.M., Doran, G.H., 1998. The use of the supero-inferior femoral neck diameter as a sex assessor. Am. J. Phys. Anthropol. 107, 305–313. https://doi.org/10.1002/(SICI)1096-8644(199811)107:3<305::AID-AJPA7>3.0.CO;2-A.

Silva, A.M., 2012. Antropologia funerária e paleobiologia das populações portuguesas (litorais) do neolítico final/calcolitico. Fundação Calouste Gulbenkian-Fundação para a Ciência e a Tecnologia, Lisboa.

Šlaus, M., Tomičić, Ž., 2005. Discriminant function sexing of fragmentary and complete tibiae from medieval Croatian sites. Forensic Sci. Int. Sex Body Size 147, 147–152. https://doi.org/10.1016/j.forsciint.2004.09.073.

Smith, R.J., 1999. Statistics of sexual size dimorphism. J. Hum. Evol. 36, 423–458. https://doi.org/10.1006/jhev.1998.0281.

Spradley, M.K., Jantz, R.L., 2011. Sex estimation in forensic anthropology: Skull versus postcranial elements. J. Forensic Sci. 56, 289–296. https://doi.org/10.1111/j.1556-4029.2010.01635.x.

Steel, F.L.D., 1962. The sexing of long bones, with reference to the st bride's series of identified skeletons. JRAI 92, 212–222. https://doi.org/10.2307/2844259.

Steyn, M., İşcan, M.Y., 1997. Sex determination from the femur and tibia in South African whites. Forensic Sci. Int. 90, 111–119. https://doi.org/10.1016/S0379-0738(97)00156-4.

Stini, W.A., 1969. Nutritional stress and growth: Sex difference in adaptive response. Am. J. Phys. Anthropol. 31, 417–426. https://doi.org/10.1002/ajpa.1330310316.

Taskın-Senol, G., Kürtül, I., Ray, A., Ahmetoglu, G., 2023. Sex determination by the machine learning algorithms through using morphometric measurements of the carpal, metacarpal, and phalangeal bones. Int. J. Morphol. 41, 1267–1272. https://doi.org/10.4067/S0717-95022023000401267.

Tomczyk, J., Nieczuja-Dwojacka, J., Zalewska, M., Niemiro, W., Olczyk, W., 2017. Sex estimation of upper long bones by selected measurements in a Radom (Poland) population from the 18th and 19th centuries AD. Anthropol. Rev. 80, 287–300. https://doi.org/10.1515/anre-2017-0019.

Trancho, G.J., Robledo, B., López-Bueis, I., Sánchez, J.A., 1997. Sexual determination of the femur using discriminant functions. Analysis of a spanish population of known sex and age. J. Forensic Sci. 42, 181–185. https://doi.org/10.1520/JFS14096J.

Trancho, G.J., Robledo, B., Sánchez, J.A., 2012. Dimorfismo sexual del húmero en una población española de sexo y edad conocidos. In: Turbon, D., Fañanás, L., Rissech, C., Rosa, A. (Eds.), Biodiversidad Humana y Evolución. Universidad de Barcelona, Barcelona, pp. 364–369.

Buuren, S. van, Groothuis-Oudshoorn, K., 2011. mice: Multivariate Imputation by Chained Equations in R. Journal of Statistical Software 45, 1–67. doi: 10.18637/jss.v045.i03.

Villalba-Mouco, V., Oliart, C., Rihuete-Herrada, C., Childebayeva, A., Rohrlach, A.B., Fregeiro, M.I., et al., 2021. Genomic transformation and social organization during the copper age-bronze age transition in southern Iberia. Sci. Adv. 7 https://doi.org/10.1126/sciadv.abi7038.

White, T.D., Folkens, P.A., 2005. The Human Bone Manual. Academic Press, Boston.

Wysocka, J., Cieślik, A.I., Danel, D.P., 2023. Sex estimation using measurements of the proximal femur in a historical population from Poland. Anthropol. Rev. 86, 37–49. https://doi.org/10.18778/1898-6773.86.1.04.