Archivio per l'Antropologia e la Etnologia - Vol. CLV (2025)
www.antropologiaetnologia.it
www.fupress.com/aae
ISSN 0373-3009

An assessment of different metrical methods of sex estimation using long bones

Jade Armann[1*], Jacopo Moggi Cecchi[2]

[1]*Université de Toulouse*
[2]*Dipartimento di Biologia, Laboratori di Antropologia - Università degli Studi di Firenze*
*E-mail: jadearmann@gmail.com

OPEN ACCESS

Citation: Armann, J., Moggi Cecchi, J. (2025). An assessment of different metrical methods of sex estimation using long bones, *Archivio per l'Antropologia e la Etnologia*, 155, 13-28. doi: https://doi.org/10.36253/aae-3837

Published: December 1, 2025

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Competing Interests: The Author(s) declare(s) no conflict of interest.

Abstract. Sex estimation from long bones metrics represents an alternative in an archaeological context, especially when the pelvic bone or cranial bones are too fragmented. This study presents an evaluation of univariate and bivariate methods, applied to an Italian reference collection housed at the Museum of Anthropology of the University of Florence, with measurements recorded in the Goldman Data Set. This collection consists of 76 individuals who died between 1900 and 1901. A total of 19 measurements were taken from long bones, including the humerus, radius, femur, and tibia. By studying sexual dimorphism from long bones using T-tests and MANOVA, the differences between male and female means were significant for many variables. The anteroposterior diameter of the radius (RAPD) was found to be the most sexually dimorphic variable ($p<0.001$), while the combination of this same variable and the maximum length of the radius (RML) was found to have the highest discriminatory power ($p<0.001$). Following a review of existing methods, those based on the most sexually dimorphic variables identified were selected. Regarding the univariate and bivariate methods, between 79% and 86% of individuals were correctly classified, suggesting that the methods evaluated, initially developed on non-Italian populations, could be potentially applied to an Italian sample.

Keywords: sexual dimorphism, italian population, Goldman Data, set, archaeology.

Introduction

Biological sex estimation is one of the first and most important steps in establishing an individual's biological profile, whether in an archaeological or forensic context. This estimation is traditionally based on the morphological analysis of bones that possess a high degree of sexual dimorphism. Among these, the coxal bone is considered the most reliable, offering an accuracy of around 95% (Phenice, 1969; Bruzek, 2002). For this purpose, the skull is also commonly used, with methods generally providing over 80% accuracy (Spradley and Jantz, 2011).

However, in many archaeological situations, sex estimation is often compromised due to the absence or fragmentation of these bones. In such cases, estimation becomes unreliable, representing a limitation for the further determination of the biological profile. To overcome this, numerous morphometric studies turn to other skeletal elements to estimate sex, particularly the long bones of the upper and lower limbs (for example: Mall *et al.*, 2001; Frutos, 2005; Curate *et al.*, 2016). In certain situations, these bone elements represent the only alternatives when the coxal bone and skull cannot be used. Such studies, developed on specific samples or populations, generally use statistical methods, like discriminant analysis (DFA) or logistic regression models, and often propose population-specific sex estimation equations for a variety of bones.

However, there is a limitation to the use of these methods, as often they cannot be applied to populations for which they have not been developed. In fact, these sex estimation equations are developed by taking into account the specific characteristics of a population, in particular sexual dimorphism, which can be influenced by genetic, environmental and socio-cultural factors (Ruff, 1994; Cox *et al.*, 2023). As a result, these methods are not necessarily reliable when applied to other populations, where dimorphism may be expressed differently, as they can lead to estimation errors if they are not validated or adapted to the new sample studied.

The present study highlights the importance of evaluating sex estimation methods to determine their reliability when applied to different populations.

In contrast to many recent studies that have focused on the analysis of dimorphic traits from a single bone in Italian archaeological samples (for example: Kranioti and Apostol, 2014; Vassallo *et al.*, 2021), this study adopts a more comprehensive approach by incorporating osteological measurements from several long bones, such as the humerus, radius, femur, and tibia.

The study is divided in different parts: (1) identifying the most sexually dimorphic variable of the Italian reference collection; (2) comparing and evaluating different univariate sex estimation methods, developed for specific populations, based on the most dimorphic variable identified; and (3) comparing and evaluating different bivariate sex estimation methods based on the most dimorphic combinations of two variables.

Specifically, it assesses the applicability of both univariate and bivariate methods, originally developed for non-Italian reference groups, on an identified Italian skeletal collection. By doing so, it aims to provide insight into how well these methods perform on Italian remains.

Material

The study is based on the variables and measurements available in the Goldman Data Set (Auerbach, 2014), created and completed by Dr. Benjamin Auerbach and freely available at: https://web.utk.edu/~auerbach/GOLD.htm.

The Data Set comprises osteometric measurements from 1,538 human skeletons for which sex and age have been estimated. For the purposes of this study, only data from an identified Italian collection (Syracusan collection) was used. This collection, housed at the Museum of Anthropology of the University of Florence, consists of 76 individuals who died between 1900 and 1901. Recovered in 1909 from the municipal cemetery of Syracuse by Dr. Francesco Aurelio Favara, it represents a valuable reference collection, as many identifying elements are known and available, such as age, sex, and for some individuals, cause of death.

As such, the Syracuse collection offers opportunities to explore a range of hypotheses and characteristics, and is particularly interesting for the development of new osteological studies on an Italian population.

The sample from the Goldman Data Set includes data from 12 females and 19 males. For each individual, a total of 5 measurements were taken from the humerus, 3 from the radius, 7 from the femur, 4 from the tibia and 3 from the pelvis (Tab. 1). However, pelvic measurements were not included in the present study.

VARIABLE	OSTEOMETRIC DESCRIPTION	REFERENCE
HML	Humerus Maximum Length	Martin (1928) #1
HEB	Humerus Epicondylar Breadth	Martin (1928) #4a
HHD	Humerus Head Diameter	Martin (1928) #10
HMLD	Humerus 50% Diaphyseal Mediolateral Diameter	Martin (1928) #6b
HAPD	Humerus 50% Diaphyseal Anteroposterior Diameter	Martin (1928) #6c
RML	Radius Maximum Length	Martin (1928) #1
RMLD	Radius 50% Diaphyseal Mediolateral Diameter	Martin (1928) #4a
RAPD	Radius 50% Diaphyseal Anteroposterior Diameter	Martin (1928) #5

FML	Femur Maximum Length	Martin (1928) #1
FBL	Femur Bicondylar Length	Martin (1928) #2
FEB	Femur Epicondylar Mediolateral Breadth	Martin (1928) #21
FAB	Femur Distal Articular (Bicondylar) Mediolateral Breadth	Auerbach (2014)
FHD	Femur Head Anteroposterior Diameter	Martin (1928) #19
FMLD	Femur 50% Diaphyseal Mediolateral Diameter	Martin (1928) #7
FAPD	Femur 50% Diaphyseal Anteroposterior Diameter	Martin (1928) #6
TML	Tibia Maximum Length	Martin (1928) #2
TPB	Tibia Plateau Mediolateral (Bicondylar) Breadth	Martin (1928) #3
TMLD	Tibia 50% Diaphyseal Mediolateral Diameter	Auerbach (2014)
TAPD	Tibia 50% Diaphyseal Anteroposterior Diameter	Auerbach (2014)

Tab. 1. *Description of the different measurements from the Goldman Osteometric Data Set (Auerbach 2014) used in this study and their corresponding references.*

Methods

Statistics

The statistical analysis was carried out using R studio, version 4.4.2 (2024-10-31). With the data derived from the Goldman Data Set, descriptive statistics, including the mean, standard deviation and 95% confidence interval were calculated for each variable and both sexes.

Preliminary analysis

An initial approach was conducted to determine whether there were any significant differences between the measurements from right and left sides, in order to determine which laterality would be used in subsequent analyses. Normality of the distribution for each side was tested using the Shapiro-Wilk test ($P > 0.05$). Paired t-tests were used when the data met the normality, but when the normality was not met, the non-parametric Wilcoxon tests were then applied. When no significant difference was observed for a given variable ($P > 0.05$), only the left side was retained. However, if both sides showed a significant difference ($P < 0.05$), both were included in further analysis.

As the next step, the normality of each variable was checked for both males and females. T-tests were then performed to assess sexual dimorphism for each variable, with significance considered at $P < 0.05$. Effect sizes were estimated using Cohen's d tests. Indeed, this test was performed to estimate the magnitude of differences between two groups, according to reference thresholds (small: 0.2, medium: 0.5, large: 0.8) (Cohen, 1988), with females defined as group 1 and males as group 2.

A Bonferroni correction was also applied to the p-values obtained in the different t-tests to limit the risk of type I errors associated with multiple comparisons, since the t-tests were applied individually to each variable. As a result, only the variables for which the corrected p-value remained significant were considered sexually dimorphic in this study.

Univariate analysis

Concerning the evaluation of univariate sex estimation methods employed in this study, the selection of methods was based on several criteria. First, the most sexually dimorphic variable of the Syracusan collection, using the data of the Goldman Data Set was identified. Next, a literature review was carried out to identify papers describing univariate sex estimation methods that used and described the exact same variable. Furthermore, only publications that presented equations or coefficients derived from statistical models were retained. Then, the selected sex estimation methods were applied to the Syracusan collection sample in order to determine which one would offer the best accuracy. These methods are taken from the following publications: Mokoena *et al.* (2019) developed from a mixed ancestry South African population, Duangto *et al.* (2020) developed from a Thai population and Waghmare *et al.* (2011) based on an Indian population.

Different confusion matrices were used in this study to evaluate the performance of the methods by comparing predicted values to actual values (Sathyanarayanan and Tantri, 2024). This makes it possible to analyze the models' ability to correctly distinguish between the different classes, in this case, male and female. They were performed to evaluate sensitivity (ability to correctly identify positive cases), precision (proportion of correct positive predictions), F1-score (balance between precision and sensitivity) and accuracy (overall rate of correct predictions) for each class, to ensure the reliability and validity of the methods selected (Tab. 2).

METRIC	INTERPRETATION
Sensitivity (Female)	Proportion of actual females correctly identified among all real females
Sensitivity (Male)	Proportion of actual males correctly identified among all real males
Precision (Female)	Proportion of predicted females who are actually females
Precision (Male)	Proportion of predicted males who are actually males
F1-score (Female)	Harmonic mean of sensitivity and precision for the female class
F1-score (Male)	Harmonic mean of sensitivity and precision for the male class
Accuracy	Proportion of correct predictions among all predictions

Tab. 2. *Description of the different confusion matrices used to evaluate the sex estimation methods selected.*

Bivariate analysis
In order to estimate the best combinations of two variables for sex estimation, a multivariate analysis of variance (MANOVA) was performed to assess potential correlations between the variables and the sex. All variables related to bones from the left side were considered for the combinations. Only two-variable combinations were tested, since the inclusion of more than two variables would not be robust due to the sample size.

Before applying the MANOVA, the following assumptions were checked: univariate normality using the Shapiro-Wilk test, multivariate normality via Mardia's test, homogeneity of variance-covariance matrices between groups using Box's M test and Pearson's correlation coefficients to avoid multicollinearity. Also, only combinations with a p-value less than 0.001 were retained to ensure high statistical confidence.

After identifying the best combinations of two variables, and following the same procedure as for the univariate analysis, a literature review was conducted to identify bivariate sex estimation methods. Publications using the same measurements and proposing equations for the two-variable combinations were selected. Moreover, to facilitate the literature review, preference was given to combinations using variables from the same bone. The following publications were selected: Mokoena *et al.* (2019), developed from a mixed ancestry South African population and also used in the univariate analysis, and Tise *et al.* (2012) based on a Hispanic sample, defined by the 2010 U.S. Census as «*a person of Cuban, Mexican, Puerto Rican, South or Central American, or other Spanish culture or origin regardless of race*».

As for the univariate analysis, confusion matrices including sensitivity, specificity, precision and F1-score were estimated to assess the performance and reliability of the methods selected (Tab. 2).

Results

Univariate

Determination of the most dimorphic variable
When looking at the variables that are significantly different between the male and female groups, the results of Cohen's d test are all negative, meaning that the variable is generally greater in males than females. Furthermore, the absolute values of the results range from -1.25 for LRMLD to -2.17 for LRAPD, which corresponds to an important effect (> 0.8), indicating a marked difference between the two groups (Tab. 3).

VARIABLE	T-VALUE	P-VALUE (BONFERRONI CORRECTED)	COHEN'S D
LRAPD	-6.26	p < 0.001	-2.17
RTAPD	-5.55	p < 0.001	-1.92
RHMLD	-5.33	p < 0.001	-1.84
RHEB	-5.59	p < 0.001	-1.83
LHAPD	-5.62	p < 0.001	-1.80
RRMLD	-4.90	p < 0.05	-1.70
LHEB	-4.90	p < 0.05	-1.68
LHMLD	-4.99	p < 0.001	-1.67
LTAPD	-4.09	p < 0.05	-1.45
LRML	-4.08	p < 0.05	-1.43
RRML	-4.05	p < 0.05	-1.43
LTPB	-3.65	p < 0.05	-1.32
RFAB	-3.83	p < 0.05	-1.26
LRMLD	-3.72	p < 0.05	-1.25

Tab. 3. *Ranking of significant variables based on the results of the Cohen's d test. The letter preceding each variable indicates laterality: L= left, R = right.*

Among the variables contained in the Goldman Data Set, the one that appears to be the most dimorphic, taking into account the Bonferroni-corrected p-value ($p < 0.001$) as well as the Cohen's d test value (-2.17), is the anteroposterior diameter at the middle of the diaphysis of the left radius (LRAPD). This variable was therefore considered the most sexually dimorphic variable for the evaluation of univariate sex estimation methods on the Syracusan collection.

Evaluation of the univariate sex estimation methods

Several calculations were used to evaluate the three univariate sex estimation methods using the LRAPD variable. In terms of accuracy, the methods by Duangto *et al.* and that by Waghmare *et al.* (based on Thai and Indian populations) have the highest value of 86%. Both methods therefore correctly assign 86% of the sexes. The method by Mokoena *et al.* (based on Mixed ancestry South African population) is only able to assign the correct sex to 79% of the sample (Fig. 1).

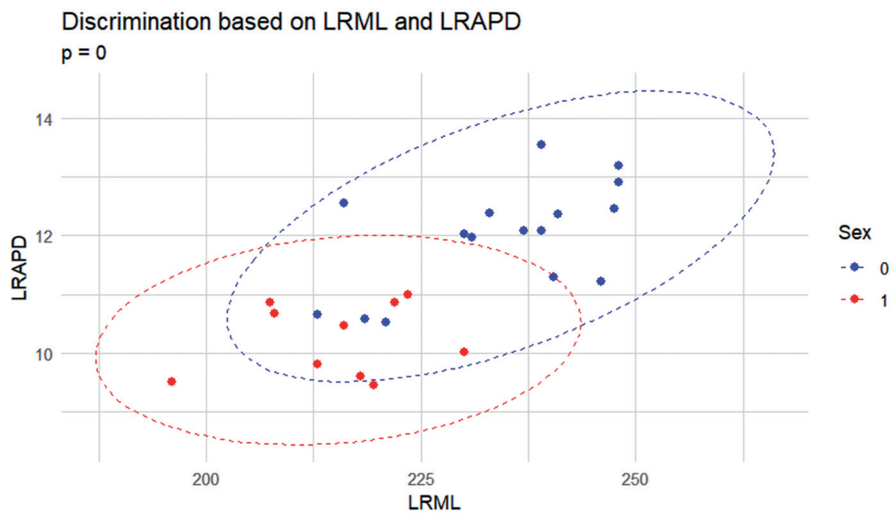

Fig. 1. *Scatter plot of LRML vs. LRAPD with 95% confidence ellipses around group centroids. Blue dots represent males (=0) and red dots represent females (=1).*

The confusion matrices, with values ranging from 0 to 1, show a sensitivity for the female class of 1, 0.92 and 0.67 associated with the methods from Duangto *et al.*, Waghmare *et al.*, and Mokoena *et al.* respectively (Tab. 4). For the male class, the sensitivity is 0.75, 0.82, and 0.88 in the same order. Precision values for the female class are 0.75 for Duangto *et al.*, 0.79 for Waghmare *et al.*, and 0.80 for Mokoena *et al.*, while for the male class they are 1, 0.93, and 0.79 respectively. The F1 scores for the female class are 0.86, 0.85 and 0.73 for the Duangto *et al.*, Waghmare *et al.* and Mokoena *et al.* methods, and 0.87, 0.87, 0.83 for the male class.

METHOD	REFERENCE POPULATION	CLASS	SENSITIVITY	PRECISION	F1-SCORE	ACCURACY
Waghmare *et al.*, 2011	Indian	Female	0.92	0.79	0.85	0.86
		Male	0.82	0.93	0.87	0.86
Duangto *et al.*, 2020	Thai	Female	1	0.75	0.86	0.86
		Male	0.75	1	0.87	0.86
Mokoena *et al.*, 2019	Mixed ancestry South African	Female	0.67	0.80	0.73	0.79
		Male	0.88	0.79	0.83	0.79

Tab. 4. *Confusion matrices results for each univariate method evaluated.*

Bivariate

Determination of the most dimorphic combination of two variables

The MANOVA analysis shows significant sex differences for 46 combinations of two variables ($P < 0.001$). These pairs were ranked in ascending order of p-value, with the first representing the most significant. Among the five most sexually dimorphic, the most discriminant is the one combining the anteroposterior diameter at the middle of the diaphysis of the left radius (LRAPD) and the mediolateral (bicondylar) breadth of the tibial plateau (LTPB) (Wilk's Lambda = 0.417, $p < 0.001$) (Fig. 1). This is closely followed by the combination of the maximum length of the radius (LRML) and the anteroposterior diameter at the middle of the diaphysis of the left radius (LRAPD) (Wilk's Lambda = 0.422, $p < 0.001$). As explained in the Methods section, the LRML + LRAPD pair is considered the most relevant for evaluating bivariate sex estimation methods.

VARIABLES	WILKS' LAMBDA	P-VALUE
LRAPD + LTPB	0.417	p < 0.001
LRML + LRAPD	0.422	p < 0.001
LRAPD + LFAB	0.430	p < 0.001
LRAPD + LFML	0.432	p < 0.001
LRAPD + LFBL	0.435	p < 0.001

Tab. 5. *Ranking of the five most significant combinations of two variables. The letter preceding each variable indicates laterality: L= left, R = right.*

Evaluation of the bivariate sex estimation methods

Regarding the results for the two bivariate sex estimation methods based on the LRML and LRAPD combination, the accuracy is approximately 83% for both the method from Mokoena *et al.* (Mixed ancestry South African), and Tise *et al.* (Hispanic). These methods correctly classify 83% of the sexes. According to the confusion matrices, Mokoena *et al.*'s method shows a sensitivity of 0.92 for the female class and 0.76 for the male class, while Tise *et al.*'s method reaches a sensitivity of 1 for the female class and 0.71 for the male class. Precision values for the female class are 0.73 for Mokoena *et al.*, and 0.71 for Tise *et al.*, while for the male class they are 0.93 and 1 respectively. Finally, the method proposed by Tise *et al.* shows an F1-score of 0.83 for both classes, while the method from Mokoena *et al.* shows 0.81 for the female class and 0.84 for the male class.

METHOD	REFERENCE POPULATION	CLASS	SENSITIVITY	PRECISION	F1-SCORE	ACCURACY
Tise *et al.*, 2012	Hispanic	Female	1	0.71	0.83	0.83
		Male	0.71	1	0.83	0.83
Mokoena *et al.*, 2019	Mixed ancestry South African	Female	0.92	0.73	0.81	0.83
		Male	0.76	0.93	0.84	0.83

Tab. 6. *Confusion matrices results for each bivariate evaluated method.*

Discussion

Univariate

Most of the variables from the upper limbs, particularly radius and humerus, show strong differences between males and females. The LRAPD appears to be the most reliable indicator of sexual dimorphism in the Syracusan collection, which is consistent with previous studies that considered the radius as highly dimorphic (Mall *et al.*, 2001). Moreover, other studies mention the RAPD as one of the best sex indicators, with an average accuracy of 93.65% (Duangto *et al.*, 2020).

The different univariate methods applied show high accuracy on the RAPD variable. Of the three methods, the models appear to be capable of correctly identifying both sexes overall, since their accuracy rates range from 79% for the method developed by Mokoena *et al.* (Mixed ancestry South African) to 86% for the methods from Duangto *et al.* (Thai) and Waghmare *et al.* (Indian). However, when we study the sensitivity (true positive rate) of each model, it would appear that there is heterogeneity in the classification of males and females. Indeed, the model developed by Duangto *et al.* correctly identifies female individuals with a sensitivity of 1, but its sensitivity for males is lower (0.75), indicating less accurate detection of male individuals. As for the method from Waghmare *et al.*, it seems to perform better in classifying females (sensitivity = 0.92) than males (sensitivity = 0.82). For the Mokoena *et al.* method, it tends to detect males more easily (sensitivity = 0.88) than females (sensitivity = 0.67). Regarding precision, the Mokoena *et al.* method shows the highest value for the female class (0.80) meaning that 80% of individuals predicted as female are truly female. For the male, the best precision is for the method from Duangto *et al.* where 100% of male predicted are truly male. The F1 score assesses the balance between precision and sensitivity, and in this case, the method proposed by Duangto *et al.* offers the best compromise.

Bivariate

When studying the combinations of two dimorphic variables, the one based on the LRML and LRAPD is the most interesting, as it involves two measurements from the same bone. As for the univariate analysis, this combination doesn't appear to be trivial, as several sources mention these specific two variables as dimorphic and the radius as being interesting for bivariate analysis in several populations (Spradley and Jantz, 2011; Tomczyk *et al.*, 2017; Curate *et al.*, 2020).

The two bivariate sex estimations methods applied showed high accuracy on the Syracusan sample. Both Mokoena *et al.* (Mixed ancestry South African) and Tise *et al.* (Hispanic) methods achieve an accuracy of 83%. The confusion matrices also allowed us to evaluate the performance by studying the sensitivity (true positive rate) of each model. The method from Tise *et al.* demonstrates a very good ability to identify female individuals (sensitivity = 1) and indicates a lower detection of male individuals (sensitivity = 0.71). In contrast, compared to the Tise *et al.* method, the one from Mokoena *et al.* has a lower sensitivity for the female class (= 0.92), and a higher sensitivity for the male class (= 0.76). Regarding precision, the Mokoena *et al.* method again shows the highest value for the female class (= 0.73), while the Tise *et al.* method achieves the highest value for the male class (= 1) meaning that 100% of male predictions are truly male for this method. Moreover, the F1-scores for both methods are comparable, suggesting a similarly balanced performance in sex estimation. It's important to note that the sample size doesn't allow us to fully confirm the robustness of the MANOVA test. However, the results obtained from the bivariate analyses are generally consistent with those reported in the literature (Smal, 2020).

Conclusion

This study aims to evaluate the performance of univariate and bivariate sex estimation methods based on non Italian samples when applied to an Italian reference osteological collection. By identifying the most sexually dimorphic variables on several long bones from the Italian collection, it's possible to determine whether methods developed for non-Italian populations can be applied to Italian samples.

Regarding the univariate analyses, the methods from Duangto *et al.* (Thai) and Waghmare *et al.* (Indian) based on the anteroposterior diameter of the radius (RAPD) show overall accuracy rates above 86%, which can be considered, in our opinion, as a good threshold to determine whether a method may be applicable (İşcan and Steyn, 2013; Capitaneanu *et al.*, 2017; Ujaddughe *et al.*, 2025). The high accuracy may be due to similarities in sexual dimorphism between the Italian sample and the Indian or Thai populations regarding the radius. On the other hand, regarding the bivariate analyses, both Mokoena *et al.* (Mixed ancestry South African) and Tise *et al.* (Hispanic) methods based on the combination of the maximum length of the radius (RML) and the anteroposterior diameter of the radius (RAPD) show comparable but slightly lower accuracy rates, falling just below the 85% threshold. So in this case, the classification errors are somewhat higher. Despite the fact that only a small part of the Syracuse collection could be used in this study, thus constituting a relatively small sample, the anteroposterior diameter of the radius appears to be a good indicator of sexual dimorphism. Indeed, numerous studies have highlighted the importance of this parameter, which sometimes offers the best accuracy (Waghmare *et al.*, 2011).

These results show that methods originally designed for specific populations can sometimes be applied to other population samples, provided they have been evaluated on a documented reference population. Because of the relatively limited sample size of the Syracuse collection in this study, the results need to be taken with caution, but it would appear that the univariate methods of sex estimation could be consistent in the estimation of the sex of an Italian population. These results suggest the possibility of applying the methods tested in the case of Italian archaeological remains. Moreover, the strength of this study lies in its global approach, as it studied sexual dimorphism on the Syracusan Italian collection using measurements from several bones, offering a broader perspective on sexual dimorphism on an Italian reference sample.

Appendix

METHOD	SAMPLE ORIGIN	SAMPLE SIZE	EQUATIONS	CUT-OFF POINT
Waghmare *et al.*, 2011 (Indian population)	Bone bank of Department of Anatomy, Government Medical College Aurangabad (Maharashtra)	Total : 198 M: 133 F: 65	Sex = $-11.944+1.093 \times RAPD$	F < -0.00003 < M
Duangto *et al.*, 2020 (Thai population)	Forensic Osteology Research Center, Chiang Mai University, Chiang Mai, Thailand	Total: 228 M: 114 F: 114	Sex = $(1.385 \times RAPD) - 15.250$	F < 0 < M
Mokoena *et al.*, 2019 (Mixed ancestry South Africans)	Raymond A. Dart Collection of Human Skeletons & UCT Human Skeletal Collection	Total: 111	Sex = $-9.873 + 0.915 \times RAPD$	F < -0.128 < M

Annex. 1. *Summary table of the sex estimation methods used for the univariate analysis. M= Male; F= Female; RAPD= anteroposterior diameter of the radius.*

METHOD	SAMPLE ORIGIN	SAMPLE SIZE	EQUATIONS	CUT-OFF POINT
Mokoena *et al.*, 2019 (Mixed ancestry South Africans)	Raymond A. Dart Collection of Human Skeletons & UCT Human Skeletal Collection	Total: 111	Sex = -18.309 + 0.056 × RML + 0.497 × RAPD	- F < -0.408< M
Tise *et al.*, 2013 (Hispanic)	Pima county Office of the Medical Examiner & Forensic Anthropology Data Bank	Total: 142 Males: 114 Females: 28	Sex = (0.1331× RML) + (1.06951 × RAPD) + (-42.7206)	F < 0 < M

Annex. 2. *Summary table of the sex estimation methods used for the multivariate analysis. M= Male; F= Female; RAPD= anteroposterior diameter of the radius; RML= maximum length of the radius.*

LIST	VARIABLES	WILKS' LAMBDA	P-VALUE
1	LRAPD + LTPB	0.416574217475033	4.23118926896267e-05
2	LRML + LRAPD	0.422134154665651	4.92811302234866e-05
3	LRAPD + LFAB	0.429915961857321	6.08010463300409e-05
4	LRAPD + LFML	0.43211305214135	6.44718188090134e-05
5	LRAPD + LFBL	0.434813836093893	6.92609724352037e-05
6	LHHD + LRAPD	0.436500621005562	7.24145702973694e-05
7	LRAPD + LTMLD	0.436719215455577	7.28327088328607e-05
8	LHML + LRAPD	0.438416630915483	7.61554053104749e-05
9	LHMLD + LRAPD	0.445557385000292	9.17047715719127e-05
10	LRAPD + LFEB	0.445753952230908	9.21711123339619e-05
11	LRAPD + LFAPD	0.44584479185173	9.2387353053677e-05
12	LRMLD + LRAPD	0.447565939805548	9.65730243365016e-05
13	LRAPD + LFMLD	0.448173126163709	9.80904734018233e-05
14	LRAPD + LFHD	0.450715810928001	0.000104684437053651
15	LRAPD + LTML	0.452209545161818	0.000108744384349642
16	LRAPD + LTAPD	0.453399876843938	0.000112082052401391
17	LHMLD + LTPB	0.476759730471814	0.00019973078282879
18	LHMLD + LRML	0.479203383826205	0.000211825663371805
19	LHAPD + LTPB	0.498213340832684	0.000331342050113896
20	LHAPD + LRML	0.503358553583923	0.000372898562636397
21	LHMLD + LFAB	0.504616414551028	0.000383756544559665

22	LHEB + LHAPD	0.507885040663071	0.00041333515467028
23	LRML + LTPB	0.51225725330258	0.000456155945525168
24	LHEB + LTPB	0.514108213756045	0.000475474509358237
25	LHEB + LHMLD	0.514559647575509	0.000480298069161882
26	LRML + LFAPD	0.51620839260278	0.000498296919821073
27	LFAPD + LTPB	0.518639446424804	0.000525961273235348
28	LHAPD + LFAB	0.518716729260921	0.000526863278128411
29	LHHD + LRML	0.519660740696128	0.000537995870360186
30	LHMLD + LFEB	0.522929470915734	0.000578223661313564
31	LRML + LFEB	0.524361096550914	0.000596692150111429
32	LHEB + LTMLD	0.525845332266694	0.000616406581070595
33	LHEB + LHHD	0.52734639625658	0.000636947736784371
34	LHEB + LFEB	0.529090974312602	0.000661605428279052
35	LHEB + LFAB	0.533682979179436	0.000730732709886633
36	LRML + LFAB	0.53465432996176	0.000746174693958934
37	LTPB + LTAPD	0.536967665760093	0.000784157951803467
38	LHMLD + LHAPD	0.537244026513062	0.000788811705777876
39	LHAPD + LFEB	0.538144088781526	0.000804143586833117
40	LRML + LFBL	0.539289527373065	0.000824048639589933
41	LHEB + LFAPD	0.540420255839304	0.000844138319335063
42	LRML + LFML	0.541396254362048	0.000861837402794708
43	LHMLD + LTAPD	0.541444512332287	0.000862721256174933
44	LHAPD + LTAPD	0.546875716688662	0.000967652347268978
45	LRML + LTAPD	0.547750395194615	0.000985600794846697
46	LRML + LRMLD	0.548204831876511	0.000995045361521109

Annex. 3. *Detailed list of all significant combinations of two variables ($p < 0.001$).*

Bibliographical references

Auerbach, B.M. 2014. *Goldman Osteometric Data Set*. Internet Edition: http://web.utk.edu/~auerbach/GOLD.htm.

Bruzek, J. 2002. A method for visual determination of sex, using the human hip bone, *American Journal of Physical Anthropology*, 117(2): 157-168.

Capitaneanu, C., Willems, G., Thevissen, P. 2017. A systematic review of odontological sex estimation methods, *Journal of Forensic Odontostomatology*, 35(2): 1-19.

Cohen, J. 1988. *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cox, S.L., Nicklisch, N., Francken, M., Wahl, J., Meller, H., Haak, W., Alt, K.W., Rosenstock, E., Mathieson, I. 2023. Socio-cultural practices may have affected sexual dimorphism in stature in Early Neolithic Europe, *Nature Human Behaviour*, 8: 243-255.

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.

Curate, F., Coelho, J., Ferreira, M.T., Navega, D., Gonçalves, D., Cunha, E. 2020. Sex assessment with the radius in Portuguese skeletal populations (late 19th to early 21st centuries), *Legal Medicine*, 47: 101769.

Duangto, P., Mahakkanukrauh, P. 2020. Sex estimation from upper limb bones in a Thai population, *Anatomy & Cell Biology*, 53(1): 36-43.

Frutos, L.R. 2005. Metric determination of sex from the humerus in a Guatemalan forensic sample, *Forensic Science International*, 147(2-3): 153-157.

İşcan, M.Y., Steyn, M. 2013. *The Human Skeleton in Forensic Medicine* (3rd ed.). Springfield, IL: Charles C Thomas Publisher.

Kranioti, E.F., Apostol, M.A. 2014. Sexual dimorphism of the tibia in contemporary Greeks, Italians, and Spanish: forensic implications, *International Journal of Legal Medicine*, 129: 357-363.

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.

Martin, R. 1928. *Lehrbuch der Anthropologie in systematischer Darstellung. Mit besonderer Berücksichtigung der anthropologischen Methoden für Studierende, Ärzte und Forschungsreisende*. Kraniologie, Osteologie. Jena: Gustav Fischer.

Mokoena, P., Billings, B.K., Gibbon, V., Bidmos, M.A., Mazengenya, P. 2019. Development of discriminant functions to estimate sex in upper limb bones for mixed ancestry South Africans, *Science & Justice*, 59(6): 660-666.

Phenice, T.W. 1969. A newly developed visual method of sexing the os pubis, *American Journal of Physical Anthropology*, 30(2): 297-301.

Ruff, C.B. 1994. Morphological adaptation to climate in modern and fossil hominids, *American Journal of Physical Anthropology*, 37(S19): 65-107.

Sathyanarayanan, S., Tantri, B.R. 2024. Confusion matrix-based performance evaluation metrics, *African Journal of Biomedical Research*, 27(4, Suppl.): 25-33.

Smal, S.A. 2020. Assessing the sexual dimorphism of the upper limb in a Free State skeletal population using univariate measurements. Master's thesis, University of the Free State, South Africa.

Spradley, M.K., Jantz, R.L. 2011. Sex estimation in forensic anthropology: Skull versus postcranial elements, *Journal of Forensic Sciences*, 56(2): 289-296.

Tise, M.L., Spradley, M.K., Anderson, B.E. 2012. Postcranial sex estimation of individuals considered Hispanic, *Journal of Forensic Sciences*, 58(S1).

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, *Anthropological Review*, 80(3): 251-263.

Ujaddughe, O.M., Haberfeld, J., Bidmos, M.A., Olateju, O.I. 2025. Evaluation of standards for sex estimation using measurements obtained from reconstructed computed tomography images of the femur of contemporary Black South Africans, *International Journal of Legal Medicine*, 139: 1409-1422.

Vassallo, S., Davies, C., Biehler-Gomez, L. 2021. Sex estimation using scapular measurements: discriminant function analysis in a contemporary Australian

population, *Australian Journal of Forensic Sciences*, 53(6): 785-798.
Waghmare, J., Deshmukh, P., Waghmare, P. 2011. Determination of sex from the shaft and tuberosity of radius: A multivariate discriminant function analysis, *Annals of Biological Research*.