Archaeological and Anthropological Sciences _####################_
https://doi.org/10.1007/s12520-021-01327-y

ORIGINAL PAPER

CalcTalus: an online decision support system for the estimation of sex with the calcaneus and talus

Francisco Curate [1] **· João d'Oliveira Coelho** [2] **· Ana Maria Silva** [1,3,4]

Received: 17 February 2021 / Accepted: 24 March 2021

Abstract
The estimation of biological sex is a primary source of information regarding unidentified skeletal individuals in bioarcheological and forensic contexts. This study aims to propose new metric standards for the estimation of sex using variables of the calcaneus and talus. An ancillary goal encompasses the creation of a web-based decision support system for the assessment of sex. Six measurements from the talus and nine from the calcaneus were collected from 180 adult individuals (93 females; 87 males) belonging to the Coimbra Identified Skeletal Collection. Logistic regression (LR), support vector machines (SVM), and a decision-tree algorithm were employed to develop models for sex prediction. Univariable sectioning points generated with a decision-tree algorithm yielded an accuracy under cross-validation from 78.3 to 82.2% with talar measurements, and from 73.6 to 86.4% with calcanei variables. Systematic error ranged from 0.2 to 34.1%. Univariable and multivariable models, produced with LR and SVM, correctly predicted sex in 85.0–91.3% of cases (bias from 0.3 to 4.3%). Obtained cross-validated accuracies obtained with the new models are similar to earlier results on the subject. The performance of multivariable model predictive is substantially superior, hinting the relevance of population-specific standards for sex estimation. The operationalization of these models in a free, user-friendly, web-application—CalcTalus (http://osteomics.com/CalcTalus/)—facilitates the probabilistic assessment of sex, providing performance metrics for the statistical templates.

Keywords Sex estimation · Biological profile · Talus · Calcaneus · Decision support systems

Introduction

Sex estimation is of key significance in the elaboration of the biological profile of anonymous human skeletal remains both in forensic anthropology and bioarchaeology, being a crucial analysis procedure in the preliminary identification of deceased individuals and the reconstruction of the demographic profile of past communities (Spradley and Jantz 2011; Bethard and VanSickle 2020).

Sexual dimorphism is more pronounced in the adult pelvic girdle; therefore, pelvic bones are the most suitable for generating sex estimation methods (İscan and Steyn 2013). In many recovery circumstances, the pelvis is not available for analysis, and other bones have been used to establish the biological sex of unknown skeletal remains, including the teeth, cranium, long bones, and other postcranial bones (Manolis et al. 2009; Spradley and Jantz 2011; Boldsen et al. 2015; Moore et al. 2016; Curate et al. 2017; Inskip et al. 2018; Kazzazi and Kranioti 2018; Gillet et al. 2020; Lescure et al. 2020).

Tarsal bones, and especially the talus and calcaneus, are often recovered in archeological and forensic contexts because sometimes they are partially shielded against taphonomic processes by footwear, and due to their trabecular strength and density (Mays 1998; Saul and Saul 2005; Tuller and Durić 2006). Moreover, the talus and calcaneus show relevant sexual dimorphism (Harris and Case 2012) and have been employed for the estimation of sex of unidentified skeletal remains since the pivotal study by Steele (1976). Steele's method was tailored for European- and African-American populations and, except for one function, requires intact bones

✉ Francisco Curate
franciscocurate@gmail.com; fcurate@uc.pt

1 Research Centre for Anthropology and Health, Department of Life Sciences, University of Coimbra, Coimbra, Portugal

2 Primate Models for Behavioural Evolution Lab, Institute of Human Sciences, University of Oxford, Oxford, UK

3 Centre for Functional Ecology, Department of Life Sciences, University of Coimbra, Coimbra, Portugal

4 University of Lisbon, UNIARQ, Lisbon, Portugal

to be employed. As such, additional population-specific studies further established the relevance of these tarsal bones in the skeletal estimation of sex, also accounting for incomplete and poorly preserved bones (Silva 1995; Bidmos and Dayal 2003; Gualdi-Russo 2007; Robinson et al. 2008; DiMichele and Spradley 2012; Harris and Case 2012; Mahakkanukrauh et al. 2014; Navega et al. 2015; Peckmann et al. 2015; Alonso-Llamazares and Pablos 2019; Sorrentino et al. 2020).

The key purpose of this study was to develop univariable and multivariable models for the estimation of sex from the talus and calcaneus in a sample from the Coimbra Identified Skeletal Collection (Department of Life Sciences, University of Coimbra, Portugal), following the seminal work by Silva (1995) in Portuguese populations. Proposed models were derived from classical and machine learning classifiers and designed for usage in complete/well-preserved bones, or incomplete/fragmentary tali and calcanei. Prediction or classification problems in bioarchaeology or forensic anthropology, such as sex estimation, have been increasingly handled with advanced statistical templates and interactive web applications (Lynch and Stephan 2018; d'Oliveira Coelho and Curate 2019; d'Oliveira Coelho et al. 2020). Hence, an additional objective included the creation of a user-friendly, freely available, web-based decision support system for the classification of sex termed CalcTalus.

Materials and methods

Sample

Data used to generate models for sex estimation with the talus and calcaneus were obtained from 180 adult individuals (93 females and 87 males) curated at the Coimbra Identified Skeletal Collection (CISC, University of Coimbra, Portugal). Ages at death varied between 20 and 89 years (mean = 48.7 years; SD = 18.4). All individuals in the sample were Portuguese nationals (mostly originating from Central Portugal) who were born between 1827 and 1913and perished between 1910 and 1936. The majority proceeded from underprivileged socioeconomic backgrounds and worked in farming, housekeeping, or artisan jobs (Cunha and Wasterlain 2007).

Data collection

Six measurements were collected from the talus and nine from the calcaneus. Both the right and left bones were measured, when available. The measurements include the talus maximum length (TM1), talus width (TM2), talus body height (TM3), talus maximum body height (TM3a), maximum length of the talar trochlea (TM4), maximum width of the talar trochlea (TM5), calcaneus maximum length (CM1), calcaneus length (CM1a), calcaneus load arm width (CM2), calcaneus load arm length (CLAL), calcaneus body height (CM4), calcaneus maximum body height (CMBH), calcaneus body length (CM5), *tuber calcanei* height (CM7), and *tuber calcanei* width (CM8). All measurements were taken according to Martin (1928), except two calcaneal measurements defined by Steele (1976). The talar and calcaneal measurements are described in Appendix 1 and depicted in Figs. 1, 2, 3, and 4. All metric variables were collected by the same observer (author 3) on two different occasions with a digital caliper and measured to the nearest 0.01 mm. All data generated or analyzed during this study are available for other researchers upon reasonable request.

Statistical analysis

Measurements collected in the left and right tali and calcanei were compared, and mean differences between the paired measurements ranged from 0.03 and 0.29 mm in the talus, and 0.02 and 0.70 mm in the calcaneus. The differences between sides are negligible; therefore, the values were averaged and used in the creation of sex prediction models. The human lower limb shows marginal asymmetry, particularly the talus and calcaneus, as already noted by Silva (1995) and Gualdi-Russo (2007).

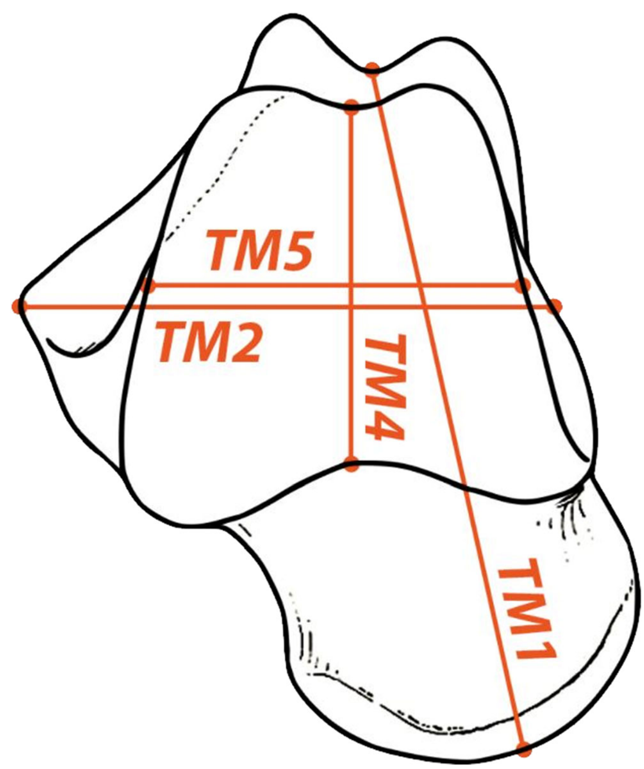

Fig. 1 Variables measured in the talus (superior view); talus maximum length (TM1), talus width (TM2), maximum length of the talar trochlea (TM4), and maximum width of the talar trochlea (TM5)

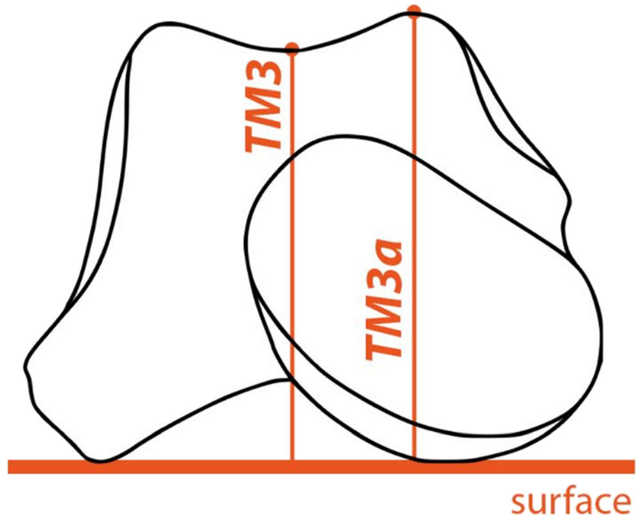

Fig. 2 Variables measured in the talus (anterior view); talus body height (TM3), and talus maximum body height (TM3a)

Descriptive statistics, namely, group means, standard deviation (SD), and 95% confidence intervals (95% CI) for the mean, were estimated for the average of the left and right measurements of the talus and calcaneus. Normal distribution of the variables was evaluated through skewness and kurtosis (Kline 2016), whereas homoscedasticity was evaluated with Levene's test. An independent samples *t*-test was used to assess the null hypothesis that the means of tali and calcanei measurements in males and females were equal. Likewise, the sexual dimorphism index (SDI, Ricklan and Tobias 1986) was estimated:

$$\mathrm{SDI} = \frac{\overline{x}_m - \overline{x}_f}{\overline{x}_m} \times 100,$$

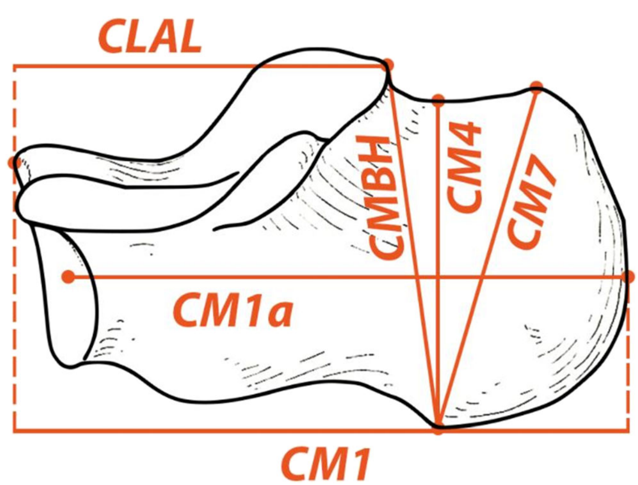

Fig. 3 Variables measured in the talus (lateral view); calcaneus maximum length (CM1), calcaneus length (CM1a), calcaneus load arm length (CLAL), calcaneus body height (CM4), calcaneus maximum body height (CMBH), and tuber calcanei height (CM7)

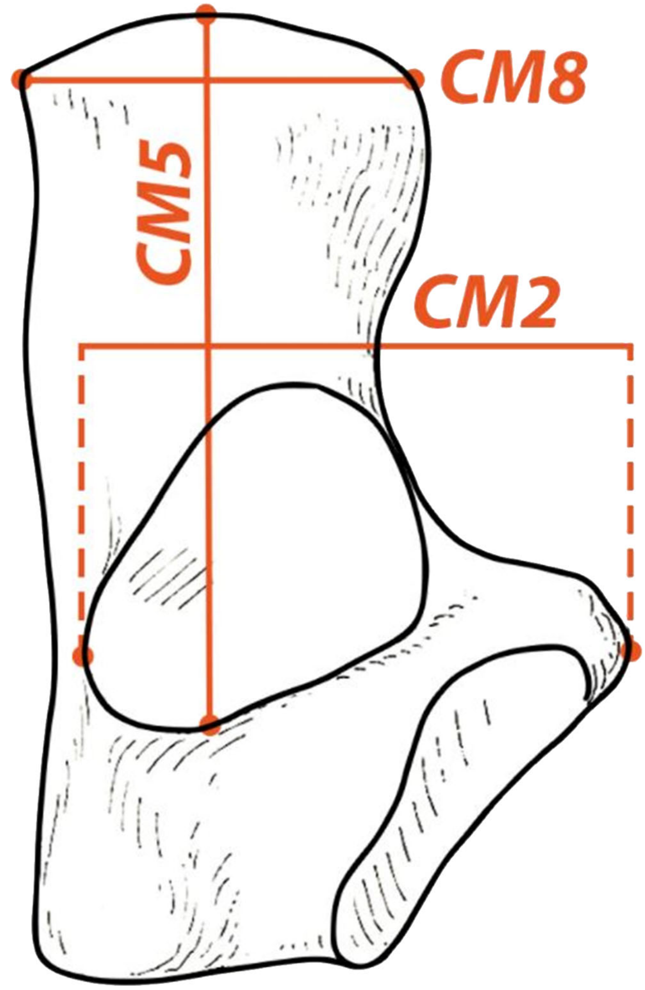

Fig. 4 Variables measured in the talus (superior view); calcaneus load arm width (CM2), calcaneus body length (CM5), and tuber calcanei width (CM8)

where $\overline{x}_m$ and $\overline{x}_f$ are the mean value of any given measurement for males and females, respectively.

Sex prediction models were generated with a classical (logistic regression (LR)) and two machine-learning (support vector machines (SVM) and decision trees (C4.5)) algorithms. LR models the probability of occurrence of two mutually exclusive classes of a dichotomous dependent variable. LR was implemented with stepwise variable selection, and models that optimally separated sexes were designed. Support vector machines are machine-learning classifiers that project a vector of predictors into a higher dimensional plane via linear or non-linear kernel functions. At last, the C4.5 algorithm generates decision trees. It is a straightforward method in machine learning, a descendant of CLS and ID3, that recursively visits individual decision nodes, selecting the optimal possible split. This algorithm uses two heuristic criteria to rank possible tests and select the best split: information gain and the default gain ratio (Quinlan 1993; Wu et al. 2008; Larose and Larose 2015, 2019). Classifier training for all prediction models was achieved with a tenfold cross-validation method to avoid overfitting or underfitting and to ensure that the models will generalize to an independent data set.

To ensure that the new models can be easily applied, tested, and used by bioarcheologists, forensic anthropologists, students, and others, CalcTalus was developed—a simple web-app available online at http://osteomics.com/CalcTalus as part of the Osteomics toolkit for biological anthropologists (d'Oliveira Coelho et al. 2020). The web application includes sex prediction as a probabilistic assessment and provides measures of goodness of fit for the statistical templates. The anthropological descriptions of talar and calcaneal measurements, as well as their graphic illustration, are also presented at the website.

The goodness of fit of the provisional and cross-validated models was evaluated through the overall accuracy (a measure of total agreement between the real and the predicted sex), the proportion of females properly grouped, the proportion of males correctly projected, and Cohen's Kappa (estimates the implementation of a given classifier as if associated to chance only) (Larose and Larose 2015).

All statistical analyses were enacted with R programming language (R Core Team 2021), and WEKA (Waikato Environment for Knowledge Analysis) (Bouckaert et al. 2015). To build CalcTalus, the interactive web application, the caret, and shiny packages were used (Kuhn 2008; Chang et al. 2021).

Results

All metric variables are normally distributed. Summary descriptive statistics for the sample are shown in Table 1. Differences between sexes were statistically significant for all variables of the talus and calcaneus. SDI, an index of sexual dimorphism, ranges from 9.5 to 11.1 for measurements in the talus (maximum length of the talar trochlea and maximum width of the talar trochlea, respectively), and from 8.3 to 11.7 for calcaneal variables (calcaneus body length and calcaneus maximum body height, respectively).

A sectioning point was computed with a decision-tree algorithm for each metric variable of the talus and calcaneus. Performance metrics for each sectioning point are summarized in Table 2. Accuracy under cross-validation varies from 78.3 to 82.2% with talar measurements, and from 73.6 to 86.4% with metric variables of the calcaneus. Bias, or systematic error, herein defined as the absolute difference between the percentage of correctly classified females and the percentage of correctly classified males, ranges from 1.1 to 23.5% in the talus, and from 0.2 to 34.1% in the calcaneus. The variables with the best overall performance under cross-validation were the talus maximum body height (accuracy: 82.2%; bias: 1.1%; kappa: 0.678; sectioning point: 30.0 mm) and the calcaneus load arm length (accuracy: 82.2%; bias: 5.3%; kappa: 0.640; sectioning point: 45.0 mm).

Table 1 Descriptive statistics for the talus and calcaneus (average values of the left and right sides) in both sexes; Coimbra Identified Skeletal Collection

Measurement	Females				Males				SDI	*F*-statistic	*Sig.*
	Mean	SD	95% CI	*N*	Mean	SD	95% CI	*N*			
TM1	48.1	2.3	47.7–48.6	93	53.5	2.7	52.9–54.1	87	10.1	206.708	<0.001
TM2	37.3	2.1	36.9–37.8	93	41.3	2.2	40.9–41.8	87	9.7	157.827	<0.001
TM3	27.7	1.5	27.3–28.0	93	30.7	1.9	30.3–31.1	87	9.8	137.521	<0.001
TM3a	29.0	1.6	28.7–29.3	93	32.1	1.9	31.7–32.5	87	9.7	139.716	<0.001
TM4	29.7	1.7	29.4–30.1	93	32.8	2.0	32.4–33.3	87	9.5	123.801	<0.001
TM5	24.8	1.6	24.5–25.1	93	27.9	1.8	27.5–28.3	87	11.1	151.108	<0.001
CM1	72.0	3.8	71.3–72.8	91	79.2	4.5	78.2–80.1	87	9.1	131.982	<0.001
CM1a	67.2	3.6	66.5–67.9	91	74.2	4.3	73.2–75.1	87	9.4	139.253	<0.001
CM2	37.7	2.2	37.3–38.2	91	41.9	2.3	41.4–42.4	87	10.0	154.858	<0.001
CM4	38.3	2.3	37.8–38.8	90	43.4	3.4	42.6–44.1	87	11.7	135.777	<0.001
CMBH	44.0	2.3	43.5–44.5	91	49.8	3.3	49.1–50.5	86	11.6	189.102	<0.001
CM5	51.6	3.2	51.0–52.3	91	56.3	3.6	55.6–57.1	87	8.3	91.853	<0.001
CM7	41.0	2.9	40.4–41.6	89	46.3	3.8	45.5–47.2	80	11.4	108.182	<0.001
CM8	27.7	1.7	27.4–28.1	88	31.3	2.1	30.9–31.8	85	11.5	151.952	<0.001
CLAL	42.6	2.4	42.1–43.1	91	47.2	2.8	46.6–47.9	87	9.7	131.098	<0.001

SD, standard deviation; 95% CI, 95% confidence interval; SDI, sexual dimorphism index; *sig*., test significance

Table 2 Sectioning points and goodness of fit for the talar and calcaneal measurements

Measurement	Sectioning point	Training sample					Cross-validation				
		Accuracy	% females	% males	Bias	Kappa	Accuracy	% females	% males	Bias	Kappa
TM1	52.25	0.839	1.000	0.667	0.333	0.674	0.811	0.925	0.690	0.235	0.619
TM2	39.75	0.839	0.871	0.805	0.066	0.677	0.817	0.871	0.759	0.112	0.632
TM3	28.50	0.828	0.742	0.920	0.178	0.657	0.811	0.731	0.897	0.166	0.624
TM3a	30.00	0.839	0.817	0.862	0.045	0.678	0.822	0.817	0.828	0.011	0.644
TM4	31.50	0.811	0.914	0.701	0.213	0.619	0.783	0.860	0.701	0.159	0.564
TM5	26.50	0.839	0.903	0.770	0.133	0.676	0.822	0.871	0.770	0.101	0.643
CM1	75.50	0.831	0.857	0.805	0.052	0.662	0.775	0.813	0.736	0.077	0.550
CM1a	70.50	0.837	0.857	0.816	0.041	0.674	0.803	0.769	0.839	0.070	0.607
CM2	40.25	0.820	0.879	0.759	0.120	0.639	0.781	0.780	0.782	0.002	0.562
CM4	42.25	0.819	0.978	0.655	0.323	0.636	0.808	0.967	0.644	0.323	0.614
CMBH	47.75	0.870	0.967	0.767	0.200	0.739	0.864	0.956	0.767	0.189	0.727
CM5	52.75	0.775	0.670	0.885	0.215	0.553	0.736	0.659	0.816	0.157	0.474
CM7	46.00	0.811	0.978	0.625	0.353	0.613	0.805	0.966	0.625	0.341	0.602
CM8	30.25	0.827	0.977	0.671	0.306	0.651	0.798	0.875	0.718	0.157	0.594
CLAL	45.00	0.831	0.868	0.793	0.075	0.662	0.820	0.846	0.793	0.053	0.640

% females, proportion of females correctly classified; % males, proportion of males correctly classified

Univariable and multivariable models for sex estimation were generated with logistic regression and support vector machines (Table 3). The goodness of fit metrics for the suggested models are shown in Table 4. Both a talar (talus maximum length) and a calcaneal (calcaneus maximum body height) variable were used in isolation to estimate sex. The logistic and SVM models that used only the talus maximum length as the predictor variable yield a cross-validated accuracy of 85.0% (bias: 4.3%; kappa: 0.699), while those that employed the calcaneus maximum body height present an accuracy under cross-validation of 85.9% (bias: 4.2%; kappa: 0.717) and 86.0% (bias: 3.0%; kappa: 0.721), respectively.

Multivariable models for the prediction of sex with measurements of the talus and calcaneus show a cross-validated accuracy ranging from 87.2 to 90.8% (logistic regression) and 86.7 to 91.3% (support vector machines). The classification functions that performed better overall include variables from both the talus and the calcaneus: LR_TC3 (accuracy: 90.8%; bias: 0.3%; kappa: 0.815), LR_TC4 (accuracy: 90.8%; bias: 2.6%; kappa: 0.815), SVM_TC4 (accuracy: 91.3%; bias: 1.4%; kappa: 0.827), and SVM_TC5 (accuracy: 91.3%; bias: 3.9%; kappa: 0.825).

All models generated with LR and SVM (see Table 3; keep in mind that sex can be directly estimated with the functions provided in the table) were operationalized in an online responsive site—CalcTalus—available at http://osteomics.com/CalcTalus/ (Fig. 5). The outline of this decision support system application is composed of four main sections: (1) Analysis, (2) Instructions, and (3) About. These are presented as clickable tabs in the top header menu of the application. The first contains all the analytical and predictive functionalities of CalcTalus, which are the following: (a) a sidebar panel menu with all the statistical implementations of the models, including uni- or multivariable inputs and different architectures such as logistic regressions, and support vector machines; (b) a main panel with dynamically generated input menu for metric variables, based on the model selected by the user; and (c) a results section, including relevant summary statistics for each model such as overall accuracy, Cohen's Kappa, and ordered variables by importance.

Besides the "Analysis" section, CalcTalus introduces the "Instructions" tab, which features descriptive and visual guides. This tab serves the purpose of describing calcaneal and talar measurements, and how to properly sample them, with rigorous anatomical descriptions and vectorial drawings of the hindfoot bones in different perspectives. This is crucial to new users or professionals that are not familiarized with the hindfoot metric variables that are used as inputs in the statistical models of CalcTalus. Finally, there is an "About" section with a disclaimer and general information about the web-app, the statistical models, authors, and research.

Discussion

Sexual dimorphism is a foremost source of variation in the human skeleton, with size and shape differences being employed to estimate the biological sex of anonymous

Table 3 Classification functions for univariable and multivariable models

Model	Classification function*
LR_T1	$(1.005\times TM1)-50.925$
LR_T2	$(0.844\times TM1)+(0.498\times TM4)-58.351$
LR_T3	$(0.665\times TM1)+(0.494\times TM4)+(0.471\times TM5)-61.487$
LR_C1	$(0.768\times CMBH)-35.837$
LR_C2	$(0.562\times CMBH)+(0.697\times CM8)-46.718$
LR_TC1	$(0.856\times TM1)+(0.531\times CM8)-58.846$
LR_TC2	$(0.741\times TM1)+(0.374\times TM4)+(0.443\times CM8)-62.136$
LR_TC3	$(0.993\times TM1)+(0.417\times TM4)+(0.581\times CM8)-(0.328\times CM5)-62.676$
LR_TC4	$(0.863\times TM1)+(0.577\times TM4)+(0.824\times CM1a)+(0.533\times CM8)-(1.190\times CM5)-71.346$
LR_TC5	$(0.585\times TM1)+(0.660\times TM4)+(0.753\times CM1a)+(0.403\times CMBH)+(0.504\times CM8)-(1.209\times CM5)-71.346$
SVM_T1	$(0.571\times TM1)-29.000$
SVM_T2	$(0.522\times TM1)+(0.434\times TM4)-40.040$
SVM_T3	$(0.463\times TM1)+(0.359\times TM4)+(0.334\times TM5)-43.360$
SVM_C1	$(0.571\times CMBH)-26.714$
SVM_C2	$(0.290\times CMBH)+(0.522\times CM8)-28.955$
SVM_TC1	$(0.661\times TM1)+(0.383\times CM8)-44.654$
SVM_TC2	$(0.475\times TM1)+(0.276\times TM4)+(0.259\times CM8)-40.189$
SVM_TC3	$(0.648\times TM1)+(0.195\times TM4)+(0.413\times CM8)-(0.195\times CM5)-40.993$
SVM_TC4	$(0.495\times TM1)+(0.296\times TM4)+(0.519\times CM1a)+(0.259\times CM8)-(0.590\times CM5)-46.721$
SVM_TC5	$(0.192\times TM1)+(0.347\times TM4)+(0.457\times CM1a)+(0.274\times CMBH)+(0.396\times CM8)-(0.675\times CM5)-40.829$

*, females classified by negative values and males by positive values; LR, logistic regression; SVM, support vector machines; T, talus; C, calcaneus

skeletal remains in both archeological and medico-legal contexts (Ubelaker and DeGaglia 2017). Although the pelvis is the most dimorphic region of the skeleton, followed by the long bones and cranium (Spradley and Jantz 2011; Brůžek et al. 2017; d'Oliveira Coelho and Curate 2019; Santos et al. 2019), most recovery scenarios result in incomplete and fragmentary remains that hinder the usefulness of standard sex estimation methods (Boldsen et al. 2015). As such, other elements of the skeleton have been employed to generate methods for the assessment of sex, including the talus and calcaneus (Steele 1976; Silva 1995).

The sexual dimorphism index evaluates sexual dimorphism, and results suggest that the maximum width of the talar trochlea and calcaneus maximum body height are the most dimorphic measurements in the talus and calcaneus, respectively. However, the maximum width of the talar trochlea does not perform well in univariable or multivariable models for sex prediction. In fact, SDI is established from a proportion of sample means and does not depict the variance of the samples. Thus, it imparts an unsatisfactory glimpse into the manifestations of sexual dimorphism (Curate et al. 2021).

In general, cross-validated accuracies obtained with the new models are on par with previous literature on the subject (Steele 1976; Silva 1995; Murphy 2002; Bidmos and Dayal 2003; Gualdi-Russo 2007; Lee et al. 2012; Mahakkanukrauh et al. 2014; Peckmann et al. 2015; Alonso-Llamazares and Pablos 2019; Sorrentino et al. 2020; Table 5). A more detailed analysis of the results shows that univariable models usually present good accuracy, but a higher bias—except the sectioning points of TM3a and CLAL, and all the LR and SVM univariable models. Higher bias associated with sectioning points is usually an issue in sex estimation from skeletal remains and also in bones from the hindfoot (Steele 1976; Bidmos and Dayal 2004; Peckmann et al. 2015). Decision trees, like those employed to generate sectioning points, also experience higher bias (Nikita and Nikitas 2020). Univariable models generated with LR and SVM are slightly more accurate and less biased, also allowing for an analytically valued probabilistic assessment of sex (Bartholdy et al. 2020; Klales et al. 2020).

Multivariable models' predictive performances are unquestionably better, as anticipated from previous results in other (Franklin et al. 2014; Curate et al. 2016) and the same skeletal regions (Silva 1995; Mahakkanukrauh et al. 2014; Navega et al. 2015; Peckmann et al. 2015; Alonso-Llamazares and Pablos 2019). In general, SVM-based models performed marginally better—with higher values of accuracy and kappa, but also presenting a modest increase of bias.

CalcTalus (http://osteomics.com/CalcTalus/) derives from the necessity of population-specific standards for sex estimation, particularly those based on the calcaneus and the talus

Table 4 Classification accuracy with the different univariable and multivariable models

Model	Training sample					Cross-validation				
	Accuracy	% females	% males	Bias	Kappa	Accuracy	% females	% males	Bias	Kappa
LR_T1	0.850	0.871	0.828	0.043	0.699	0.850	0.871	0.828	0.043	0.699
LR_T2	0.878	0.871	0.885	0.014	0.756	0.872	0.860	0.885	0.025	0.744
LR_T3	0.894	0.892	0.897	0.005	0.789	0.883	0.871	0.897	0.026	0.767
LR_C1	0.859	0.879	0.837	0.042	0.717	0.859	0.879	0.837	0.042	0.717
LR_C2	0.884	0.886	0.881	0.005	0.767	0.884	0.886	0.881	0.005	0.767
LR_TC1	0.896	0.909	0.882	0.027	0.792	0.896	0.909	0.882	0.027	0.792
LR_TC2	0.908	0.909	0.906	0.003	0.815	0.896	0.898	0.894	0.004	0.792
LR_TC3	0.913	0.909	0.918	0.009	0.827	0.908	0.909	0.906	0.003	0.815
LR_TC4	0.913	0.920	0.906	0.014	0.827	0.908	0.920	0.894	0.026	0.815
LR_TC5	0.924	0.943	0.905	0.038	0.849	0.907	0.920	0.893	0.027	0.814
SVM_T1	0.850	0.871	0.828	0.043	0.699	0.850	0.871	0.828	0.043	0.699
SVM_T2	0.878	0.882	0.874	0.008	0.755	0.867	0.860	0.874	0.014	0.733
SVM_T3	0.878	0.860	0.897	0.037	0.756	0.872	0.849	0.897	0.048	0.745
SVM_C1	0.855	0.852	0.857	0.005	0.709	0.860	0.875	0.845	0.030	0.721
SVM_C2	0.878	0.898	0.857	0.041	0.756	0.872	0.886	0.857	0.029	0.744
SVM_TC1	0.890	0.909	0.871	0.038	0.780	0.890	0.909	0.871	0.038	0.780
SVM_TC2	0.890	0.886	0.894	0.008	0.780	0.890	0.886	0.894	0.008	0.780
SVM_TC3	0.925	0.932	0.918	0.014	0.850	0.902	0.898	0.906	0.008	0.803
SVM_TC4	0.919	0.932	0.906	0.026	0.838	0.913	0.920	0.906	0.014	0.827
SVM_TC5	0.919	0.943	0.893	0.050	0.837	0.913	0.932	0.893	0.039	0.825

% females, proportion of females correctly classified; % males, proportion of males correctly classified; LR, logistic regression; SVM, support vector machines; T, talus; C, calcaneus

(Bidmos and Dayal 2004; Gualdi-Russo 2007), and the increasing requirements for advanced statistical reasoning and predictive modeling under a probabilistic construct (Ousley and Hollinger 2012; Kotěrová et al. 2018; Bartholdy et al. 2020; d'Oliveira Coelho et al. 2020). CalcTalus amplifies the suite of existing web-based applications or packages intended to simplify forensic and bioarcheological methods (Nikita and Lahr 2011; Curate et al. 2016; Gonçalves et al. 2016; Brůžek et al. 2017; Lynch and Stephan 2018; Nikita and Michopoulou 2018; d'Oliveira Coelho and Curate 2019;

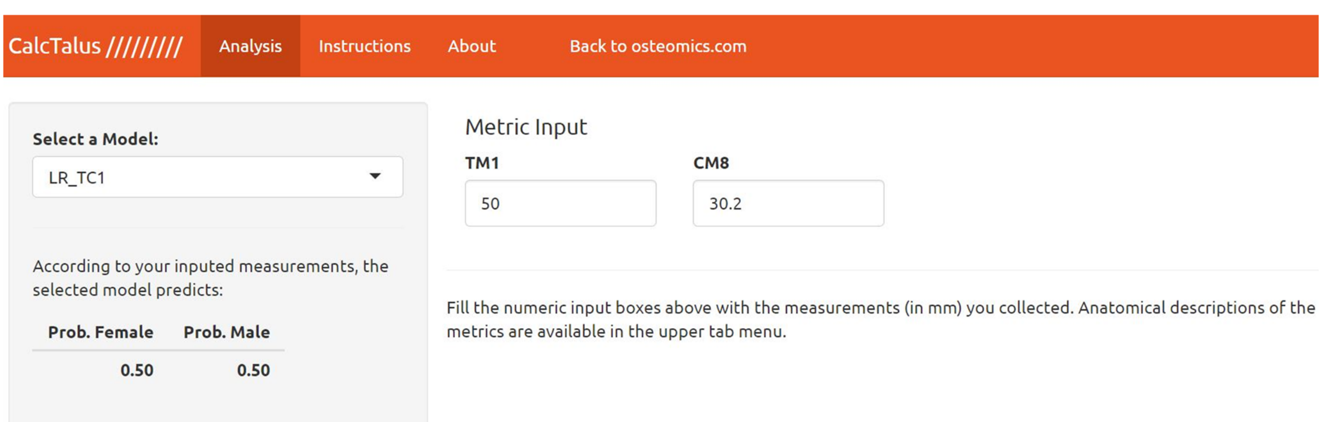

Fig. 5 CalcTalus (http://osteomics.com/CalcTalus) allows the probabilistic estimation of sex in unidentified human skeletons, also featuring the description and illustration of talar and calcaneal variables

Table 5 Accuracy of sex allocation in different studies using osteometric measurements of the talus and/or the calcaneus

Study	Population	Hindfoot bones	Statistical modeling	Accuracy (%)	
Steele (1976)	African- and European-American (*N* = 119)	Talus and calcaneus	LDA	Univariable	---
				Multivariable	79.0–89.0×
Silva (1995)	Portuguese (*N* = 165)	Talus and calcaneus	LDA	Univariable	81.0–82.1×
				Multivariable	82.1–92.9×
Bidmos and Dayal (2003)	South-African (*N* = 120)	Talus	LDA	Univariable	80.0–82.0×
				Multivariable	81.0–88.0×
Gualdi-Russo (2007)	Italian (*N* = 118)	Talus and calcaneus	LDA	Univariable	---
				Multivariable	89.2–95.7×
Lee et al. (2012)	Korean (*N* = 140)	Talus	LDA	Univariable	67.1–82.9*
				Multivariable	72.1–87.1*
Mahakkanukrauh et al. (2014)	Thai (*N* = 252)	Talus	LR	Univariable	79.1–89.8×
				Multivariable	88.0–91.4×
Peckmann et al. (2015)	Greek (*N* = 182)	Talus	LDA	Univariable	65.2–93.4*
				Multivariable	86.7–92.9*
Alonso-Llamazares and Pablos (2019)	African- and European-American (*N* = 164)	Talus and calcaneus	LDA	Univariable	66.1–90.2×
				Multivariable	86.0–96.4×
Current study	Portuguese (*N* = 180)	Talus and calcaneus	LR, SVM, DT	Univariable	78.3–82.2*
				Multivariable	85.0–91.3*

LDA, linear discriminant analysis; LR, logistic regression; SVM, support vector machines; DT, decision trees; ×, resubstitution accuracy; *, cross-validated accuracy

Santos et al. 2019; Santos 2021), providing a free, user-friendly, decision support system for the sex estimation of unidentified skeletal remains. Web-based applications like CalcTalus also function as learning tools, working as interactive and motivating devices in a classroom environment (d'Oliveira Coelho and Curate 2019).

Final remarks

CalcTalus operationalizes a suite of models for sex estimation with measurements of the calcaneus and talus in a free software intended to facilitate data analysis. Generated models, especially multivariable models, are accurate and unbiased, allowing for a probabilistic assessment of sex. In consideration of the population specificity of sexually dimorphic traits of the hindfoot, the newly proposed models should be limited to south-European populations or of south-European descent, and caution is advised when employing the models in other populations. The addition of data from different populations to the database that structures CalcTalus is a future objective.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s12520-021-01327-y.

Code availability Code generated during this study is available for other researchers upon reasonable request.

Funding The Portuguese Foundation for Science and Technology (FCT) is the funding body of the Research Centre for Anthropology and Health (UIDB/00283/2020). FCT also funds João d'Oliveira Coelho through a Ph.D. grant (SFRH/BD/122306/2016).

Data availability All data generated or analyzed during this study are available for other researchers upon reasonable request.

Declarations

Conflict of interest The authors state that they do not have any conflict of interest to declare.

References

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

Bartholdy BP, Sandoval E, Hoogland MLP, Schrader SA (2020) Getting rid of dichotomous sex estimations: why logistic regression should be preferred over discriminant function analysis. J Forensic Sci 65: 1685–1691. https://doi.org/10.1111/1556-4029.14482

Bethard JD, VanSickle C (2020) Applications of sex estimation in paleoanthropology, bioarchaeology, and forensic anthropology. In:

Klales A (ed) Sex estimation of the human skeleton. Academic Press, London, pp 25–34

Bidmos MA, Dayal MR (2003) Sex determination from the talus of South African whites by discriminant function analysis. Am J Forensic Med Pathol 24:322–328. https://doi.org/10.1097/01.paf.0000098507.78553.4a

Bidmos MA, Dayal MR (2004) Further evidence to show population specificity of discriminant function equations for sex determination using the talus of South African blacks. J Forensic Sci 49:1–6. https://doi.org/10.1520/jfs2003431

Boldsen JL, Milner GR, Boldsen SK (2015) Brief communication: sex estimation from modern American humeri and femora, accounting for sample variance structure. Am J Phys Anthropol 158:745–750. https://doi.org/10.1002/ajpa.22812

Bouckaert RR, Frank E, Hall M, et al (2015) WEKA manual for version 3-6-13

Brůžek J, Santos F, Dutailly B, Murail P, Cunha E (2017) Validation and reliability of the sex estimation of the human os coxae using freely available DSP2 software for bioarchaeology and forensic anthropology. Am J Phys Anthropol 164:440–449. https://doi.org/10.1002/ajpa.23282

Chang W, Cheng J, Allaire JJ, et al (2021) shiny: web application framework for R

Cunha E, Wasterlain S (2007) The Coimbra identified osteological collections. In: Grupe G, Peters J (eds) Skeletal series in their socioeconomic context. M. Leidorf, Rahden/Westf., pp 23–33

Curate F, d'Oliveira Coelho J, Gonçalves D et al (2016) A method for sex estimation using the proximal femur. Forensic Sci Int 266:17–23. https://doi.org/10.1016/j.forsciint.2016.06.011

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

Curate F, Mestre F, Garcia SJ (2021) Sex assessment with the radius in Portuguese skeletal populations (late 19th – early to mid 20th centuries). Legal Med 48:101790. https://doi.org/10.1016/j.legalmed.2020.101790

d'Oliveira Coelho J, Curate F (2019) CADOES: an interactive machine-learning approach for sex estimation with the pelvis. Forensic Sci Int 302:109873. https://doi.org/10.1016/j.forsciint.2019.109873

d'Oliveira Coelho J, Curate F, Navega D (2020) Osteomics : decision support systems for forensic anthropologists. In: Obertová Z, Stewart A, Cattaneo C (eds) Statistics and probability in forensic anthropology. Academic Press, London, pp 259–273

DiMichele DL, Spradley MK (2012) Sex estimation in a modern American osteological sample using a discriminant function analysis from the calcaneus. Forensic Sci Int 221:152.e1–152.e5. https://doi.org/10.1016/j.forsciint.2012.03.026

Franklin D, Cardini A, Flavel A, Marks MK (2014) Morphometric analysis of pelvic sexual dimorphism in a contemporary Western Australian population. Int J Legal Med 128:861–872. https://doi.org/10.1007/s00414-014-0999-8

Gillet C, Costa-Mendes L, Rérolle C, Telmon N, Maret D, Savall F (2020) Sex estimation in the cranium and mandible: a multislice computed tomography (MSCT) study using anthropometric and geometric morphometry methods. Int J Legal Med 134:823–832. https://doi.org/10.1007/s00414-019-02203-0

Gonçalves D, d'Oliveira Coelho J, Acosta MA et al (2016) One for all and all for one: linear regression from the mass of individual bones to assess human skeletal mass completeness. Am J Phys Anthropol 160:427–432. https://doi.org/10.1002/ajpa.22979

Gualdi-Russo E (2007) Sex determination from the talus and calcaneus measurements. Forensic Sci Int 171:151–156. https://doi.org/10.1016/j.forsciint.2006.10.014

Harris SM, Case DT (2012) Sexual dimorphism in the tarsal bones: implications for sex determination. J Forensic Sci 57:295–305. https://doi.org/10.1111/j.1556-4029.2011.02004.x

Inskip S, Constantinescu M, Brinkman A, Hoogland M, Sofaer J (2018) The effect of population variation on the accuracy of sex estimates derived from basal occipital discriminant functions. Archaeol Anthropol Sci 10:675–683. https://doi.org/10.1007/s12520-016-0380-6

İscan MY, Steyn M (2013) The human skeleton in forensic medicine. Charles C Thomas, Springfield, IL

Kazzazi SM, Kranioti EF (2018) Sex estimation using cervical dental measurements in an archaeological population from Iran. Archaeol Anthropol Sci 10:439–448. https://doi.org/10.1007/s12520-016-0363-7

Klales AR, Ousley SD, Passalacqua NV (2020) Statistical approaches to sex estimation. In: Klales AR (ed) Sex estimation of the human skeleton. Academic Press, London, pp 203–217

Kline RB (2016) Principles and practices of structural equation modelling. Guilford Press, New York

Kotěrová A, Navega D, Štepanovský M, Buk Z, Brůžek J, Cunha E (2018) Age estimation of adult human remains from hip bones using advanced methods. Forensic Sci Int 287:163–175. https://doi.org/10.1016/j.forsciint.2018.03.047

Kuhn M (2008) Building predictive models in R using the caret package. J Stat Softw 28:1–26

Larose D, Larose C (2015) Data mining and predictive analytics. John Wiley & Sons, Hoboken, NJ

Larose CD, Larose DT (2019) Data science using Python and R. John Wiley & Sons, Inc., Hoboken, NJ

Lee UY, Han SH, Park DK, Kim YS, Kim DI, Chung IH, Chun MH (2012) Sex determination from the talus of Koreans by discriminant function analysis. J Forensic Sci 57:166–171. https://doi.org/10.1111/j.1556-4029.2011.01914.x

Lescure J, Ardevines C, Becerra P, Marrodán MD (2020) New random generalized linear model for sex determination based on cranial measurements. Archaeol Anthropol Sci 12:12. https://doi.org/10.1007/s12520-020-01145-8

Lynch J, Stephan C (2018) Computational tools in forensic anthropology: the value of open-source licensing as a standard. Forensic Anthropol 1:228–243. https://doi.org/10.5744/fa.2018.0025

Mahakkanukrauh P, Praneatpolgrang S, Ruengdit S et al (2014) Sex estimation from the talus in a Thai population. Forensic Sci Int 240:152.e1–152.e8. https://doi.org/10.1016/j.forsciint.2014.04.001

Manolis SK, Eliopoulos C, Koilias CG, Fox SC (2009) Sex determination using metacarpal biometric data from the Athens Collection. Forensic Sci Int 193:130.e1–130.e6. https://doi.org/10.1016/j.forsciint.2009.09.015

Mays S (1998) The archaeology of human bones. Routledge, London

Moore MK, DiGangi EA, Niño Ruíz FP et al (2016) Metric sex estimation from the postcranial skeleton for the Colombian population. Forensic Sci Int 262:286.e1–286.e8. https://doi.org/10.1016/j.forsciint.2016.02.018

Murphy AMC (2002) The calcaneus: sex assessment of prehistoric New Zealand Polynesian skeletal remains. Forensic Sci Int 129:205–208. https://doi.org/10.1016/S0379-0738(02)00301-8

Navega D, Vicente R, Vieira DN, Ross AH, Cunha E (2015) Sex estimation from the tarsal bones in a Portuguese sample: a machine learning approach. Int J Legal Med 129:651–659. https://doi.org/10.1007/s00414-014-1070-5

Nikita E, Lahr MM (2011) Simple algorithms for the estimation of the initial number of individuals in commingled skeletal remains. Am J Phys Anthropol 146:629–636. https://doi.org/10.1002/ajpa.21624

Nikita E, Michopoulou E (2018) A quantitative approach for sex estimation based on cranial morphology. Am J Phys Anthropol 165:507–517. https://doi.org/10.1002/ajpa.23376

Nikita E, Nikitas P (2020) Data mining and decision trees. In: Obertová Z, Stewart A, Cattaneo C (eds) Statistics and probability in forensic anthropology. Academic Press, London, pp 87–105

Ousley SD, Hollinger RE (2012) The pervasiveness of Daubert. In: Dirkmaat DC (ed) A companion to forensic anthropology. Blackwell Publishing Ltd, Chichester, pp 654–665

Peckmann TR, Orr K, Meek S, Manolis SK (2015) Sex determination from the talus in a contemporary Greek population using discriminant function analysis. J Forensic Legal Med 33:14–19. https://doi.org/10.1016/j.jflm.2015.03.011

Quinlan JR (1993) C4.5: Programs for machine learning. Morgan Kaufmann Publishers, San Francisco, CA

R Core Team (2021) A language and environment for statistical computing. R Found. Stat. Comput. 2:https://www.R-project.org

Ricklan DE, Tobias PV (1986) Unusually low sexual dimorphism of endocranial capacity in a Zulu cranial series. Am J Phys Anthropol 71:285–293. https://doi.org/10.1002/ajpa.1330710304

Robinson C, Eisma R, Morgan B, Jeffery A, Graham EAM, Black S, Rutty GN (2008) Anthropological measurement of lower limb and foot bones using multi-detector computed tomography. J Forensic Sci 53:1289–1295. https://doi.org/10.1111/j.1556-4029.2008.00875.x

Santos F (2021) rdss: an R package to facilitate the use of Murail et al.'s (1999) approach of sex estimation in past populations. Int J Osteoarchaeol. https://doi.org/10.1002/oa.2957

Santos F, Guyomarc'h P, Rmoutilova R, Bruzek J (2019) A method of sexing the human os coxae based on logistic regressions and Bruzek's nonmetric traits. Am J Phys Anthropol 169:435–447. https://doi.org/10.1002/ajpa.23855

Saul JM, Saul F (2005) "The game is afoot": feet help solve forensic puzzles in the United States and overseas. In: Rich J, Dean DE, Powers RH (eds) Forensic medicine of the lower extremity : human identification and trauma analysis of the thigh, leg, and foot. Human Press, Tottowa, pp 359–374

Silva AM (1995) Sex assessment using the calcaneus and talus. Antropol Port 13:107–119

Sorrentino R, Belcastro MG, Figus C, Stephens NB, Turley K, Harcourt-Smith W, Ryan TM, Benazzi S (2020) Exploring sexual dimorphism of the modern human talus through geometric morphometric methods. PLoS One 15:1–17. https://doi.org/10.1371/journal.pone.0229255

Spradley MK, Jantz RL (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

Steele DG (1976) The estimation of sex on the basis of the talus and calcaneus. Am J Phys Anthropol 45:581–588. https://doi.org/10.1002/ajpa.1330450323

Tuller H, Durić M (2006) Keeping the pieces together: comparison of mass grave excavation methodology. Forensic Sci Int 156:192–200. https://doi.org/10.1016/j.forsciint.2004.12.033

Ubelaker DH, DeGaglia CM (2017) Population variation in skeletal sexual dimorphism. Forensic Sci Int 278:407.e1–407.e7. https://doi.org/10.1016/j.forsciint.2017.06.012

Wu X, Kumar V, Quinlan JR et al (2008) Top 10 algorithms in data mining. Knowl Inf Syst 14:1–37. https://doi.org/10.1007/s10115-007-0114-2

Martin R (1928) Lehrbuch der Anthropologie in Systematischer Darstellung mit Besonderer Berücksichtigung der anthropologischen Methoden für Studierende, Artze und Forschungsreisende, Kraniologie, Osteologie, vol. 2. Gustav Fischer, Jena

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.