Dear author,

Please note that changes made in the online proofing system will be added to the article before publication but are not reflected in this PDF.

We also ask that this file not be used for submitting corrections.

Forensic Science [International](http://dx.doi.org/10.1016/j.forsciint.2017.02.035) xxx (2016) xxx–xxx



# Forensic Science International



journal homepage: <www.elsevier.com/locate/forsciint>

# Sex estimation with the total area of the proximal femur: A <sup>2</sup> densitometric approach

# <sup>3</sup> � Francisco Curate<sup>a,b,c,\*</sup>, Anabela Albuquerque<sup>d</sup>, Izilda Ferreira<sup>d</sup>, Eugénia Cunha<sup>b,e</sup>

<sup>4</sup> a Research Centre for Anthropology and Health, Department of Life Sciences, University of Coimbra, Coimbra, Portugal<br><sup>5</sup> Laboratory of Forensic Anthropology, Department of Life Sciences, University of Coimbra, Coimbra,

<sup>7</sup> <sup>d</sup> The Coimbra Hospital and University Centre, Coimbra, Portugal<br><sup>8</sup> e Centre for Functional Ecology, Department of Life Sciences, University of Coimbra, Coimbra, Portugal

### A R T I C L E I N F O

Article history: Received 19 December 2016 Received in revised form 21 February 2017 Accepted 28 February 2017 Available online xxx

Keywords: Forensic anthropology population data Human identification Biological profile Sex diagnosis Bone densitometry

### A B S T R A C T

The estimation of sex is a central step to establish the biological profile of an anonymous skeletal individual. Imaging techniques, including bone densitometry, have been used to evaluate sex in remains incompletely skeletonized. In this paper, we present a technique for sex estimation using the total area (TA) of the proximal femur, a two-dimensional areal measurement determined through densitometry. TA was acquired from a training sample (112 females; 112 males) from the Coimbra Identified Skeletal Collection (University of Coimbra, Portugal). Logistic regression (LR), linear discriminant analysis (LDA), reduce error pruning trees (REPTree), and classification and regression trees (CART) were employed in order to obtain models that could predict sex in unidentified skeletal remains. Under cross-validation, the proposed models correctly estimated sex in 90.2–92.0% of cases (bias ranging from 1.8% to 4.5%). The models were evaluated in an independent test sample (30 females; 30 males) from the 21st Century Identified Skeletal Collection (University of Coimbra, Portugal), with a sex allocation accuracy ranging from 90.0% to 91.7% (bias from 3.3% to 10.0%). Overall, data mining classifiers, especially the REPTree, performed better than the traditional classifiers (LR and LDA), maximizing overall accuracy and minimizing bias. This study emphasizes the significant value of bone densitometry to estimate sex in cadaveric remains in diverse states of preservation and completeness, even human remains with soft tissues.

© 2017 Elsevier B.V. All rights reserved.

## <sup>9</sup> 1. Introduction

 $10 \tQ3$  The assessment of biological sex constitutes a focal research  $11$  demand in the forensic evanination of burnar eligibility remains <sup>11</sup> demand in the forensic examination of human skeletal remains,<br><sup>12</sup> with additional parameters of the higherical profile (e.g. stature or <sup>12</sup> with additional parameters of the biological profile (*e.g.*, stature or  $\frac{13}{2}$  are  $\frac{13}{2}$  superlative  $13$  age) typically estimated as sex-specific [1,2]. Superlative  $14$  approaches for the sound estimation of unknown skeletal <sup>14</sup> approaches for the sexual estimation of unknown skeletal  $^{15}$  individuals usually denend on the recovery and analysis of well- $^{15}$  individuals usually depend on the recovery and analysis of well-<br> $^{16}$  recogned polyic bones  $^{11}$ ,  $^{21}$ , likewise, the cranium and long <sup>16</sup> preserved pelvic bones  $[1-3]$ . Likewise, the cranium and long<br><sup>17</sup> bones have been employed to accurately assess sox in human <sup>17</sup> bones have been employed to accurately assess sex in human  $\frac{18}{18}$  closeled to a straight in the longest and as a rule the <sup>18</sup> skeletal remains  $[3-6]$ . The femur is the longest and, as a rule, the long commonly recovered in both <sup>19</sup> strongest skeletal element, being commonly recovered in both  $\frac{20}{\pi}$  for the strong strong and archeological contexts [5]. As such it is not surprising forensic and archeological contexts  $[5]$ . As such, it is not surprising

<http://dx.doi.org/10.1016/j.forsciint.2017.02.035> 0379-0738/© 2017 Elsevier B.V. All rights reserved. that, alongside the cranium and pelvis, the femur has received  $^{21}$ <br>meet of the stration in studies of sexual dimensions with sexual  $^{22}$ most of the attention in studies of sexual dimorphism, with several  $^{22}$ <br>dimonsions of the formula angles of for the prodiction of sexual  $^{23}$ dimensions of the femur employed for the prediction of sex in  $\frac{23}{94.24}$ skeletal remains  $[4,6-10]$ . Q4 24

In forensic settings, sex estimation is usually performed in fully  $25$ <br>platenized, bodies, with the support of standard osteometric  $26$ skeletonized bodies with the support of standard osteometric  $\frac{26}{3}$ techniques, but periodically forensic identification of unknown  $^{27}$ <br>individuals negatives the study of incomplete pertially fleshed ex.  $^{28}$ individuals requires the study of incomplete, partially fleshed or  $^{28}$ <br>charged remains [11.12]. Modical imaging techniques can be used charred remains [11,12]. Medical imaging techniques can be used  $^{29}$ <br>to observe remains not completely ekolotonized in which ckolotal  $^{30}$ to observe remains not completely skeletonized in which skeletal  $30$ <br>preparation (e.g. maceration) is impractical or even unreasonable  $31$ preparation (*e.g.*, maceration) is impractical, or even unreasonable  $31$ <br>from a social or cultural standnoint. Accordingly imaging  $32$ from a social or cultural standpoint. Accordingly, imaging  $32$ <br>techniques such as computer tomography or projectional radios techniques, such as computer tomography or projectional radiog- $\frac{33}{2}$ <br>raphy have been extensively used to address the estimation of sex  $\frac{34}{2}$ raphy, have been extensively used to address the estimation of sex  $34$ <br>in crapial and posteranial bones [12-18] including the femur  $35$ in cranial and postcranial bones  $[12-18]$ , including the femur  $35$ <br> $[1119.20]$  $[11,19,20]$ .<br>
Dual Y Fau absorptiometry (DYA) or hope densite metry is an  $37$ 

Dual X-ray absorptiometry (DXA), or bone densitometry, is an  $^{37}$ <br>plication of low operay prejectional rediscrepty, concretive  $^{38}$ application of low energy projectional radiography, generally

Q2 \* Corresponding author. E-mail address: [fcurate@uc.pt](mailto:fcurate@uc.pt) (F. Curate).

 $39$  recognized as the gold-standard technique to evaluate bone<br> $40$  existent during  $(D400)$  and discusses actores with  $(D400)$  Given 40 mineral density (BMD) and diagnose osteoporosis  $[21,22]$ . Given  $41$  that  $\mathbb{N}$ A is a true dimensional seem and have density segment be 41 that DXA is a two-dimensional scan, real bone density cannot be  $\frac{42}{100}$  determined instead bone mineral content (DMC in mann) in a 42 determined; instead, bone mineral content (BMC, in grams) in a<br>43 given projected area (in  $cm<sup>2</sup>$ ) is moasured. Areal BMD is thus  $43$  given projected area (in cm<sup>2</sup>) is measured. Areal BMD is thus 44 determined by dividing the BMC by area. DXA has been<br>45 infrastrontly poplied in the foreneic sciences although it can be <sup>45</sup> infrequently applied in the forensic sciences, although it can be exploited to estimate say age at death and appearty  $\frac{10.22 \times 261}{10.22 \times 261}$ <sup>46</sup> exploited to estimate sex, age at death and ancestry  $[10,23-26]$ . 47 Some advantages of DXA application in the forensic sciences are<br> $\frac{48}{2}$  summarized by Mbostlay [33]  $\frac{48}{49}$  summarized by Wheatley [23].

<sup>49</sup> The main purpose of this study is to generate and test models for  $\frac{50}{2}$  the argument for the anglicition of sex head on the tatel area of the argument for  $\mu$ <sup>50</sup> the prediction of sex based on the total area of the proximal femur, a<br> $\frac{51}{2}$  two dimensional areal measurement performed with DYA. Also the  $51$  two-dimensional areal measurement performed with DXA. Also, the performance of classical classifiers, such as logistic regression and  $52$  performance of classical classifiers, such as logistic regression and<br> $53$  Eisher's linear discriminant analysis, which have been extensively  $53$  Fisher's linear discriminant analysis, which have been extensively<br> $54$  used for classification of problems where the dependent variable is  $54$  used for classification of problems where the dependent variable is  $55$  displatements is compared with that of classification and regression  $55$  dichotomous, is compared with that of classification and regression <sup>56</sup> trees and reduce error pruning trees, which are non-parametric decision tree logging to the investor decision tree learning techniques.

# <sup>58</sup> 2. Materials and methods

 $59$  The samples used in this study were obtained from two  $50$ <sup>60</sup> Portuguese Identified Skeletal Collections [27,28]. A training set  $\frac{61}{2}$  from the Collection Islands Collection (CISC University of <sup>61</sup> from the Coimbra Identified Skeletal Collection (CISC, University of  $^{62}$  Coimbra Dettural), comprising 224 individuals (112 females and 62 Coimbra, Portugal), comprising 224 individuals (112 females and  $63$  112 males), was used to fit the models for sex estimation  $^{63}$  112 males), was used to fit the models for sex estimation.<br> $^{64}$  ladividual ages at doath ranged from 20 to 06 years. Dates of doath  $^{64}$  Individual ages at death ranged from 20 to 96 years. Dates of death  $^{65}$  $^{65}$  spanned from 1910 to 1936. A second sample, from the 21st<br> $^{66}$  Septem: Identified Skalatel Callegian (ISC) W. University of <sup>66</sup> Century Identified Skeletal Collection (ISC/XXI, University of  $\frac{67}{100}$  Coimbre, Deptyred), included CO individuals (20 females and  $^{67}$  Coimbra, Portugal), included 60 individuals (30 females and  $^{68}$  $^{68}$  30 males) and was employed to test the predictive value of the  $^{69}$  $^{69}$  models generated in the CISC sample: this is the testing, or holdout,<br> $^{70}$  cample, All individuals died between 1005 and 2001, Age at death  $^{70}$  sample. All individuals died between 1995 and 2001. Age at death  $^{71}$  ranged from 22 to 07 years old. Only individuals with at least one  $^{71}$  ranged from 33 to 97 years old. Only individuals with at least one<br> $^{72}$  famur, showing no macroscopical signs of post depositional  $^{72}$  femur showing no macroscopical signs of post-depositional  $^{73}$  change and lacking significant pathological modifications were  $^{73}$  change and lacking significant pathological modifications were<br> $^{74}$  included in the samples  $^{74}$  included in the samples.<br> $^{75}$  In the domain of den

<sup>75</sup> In the domain of densitometry, the proximal femur has been<br><sup>76</sup> partitioned into distinctive regions of interest. The total area (TA <sup>76</sup> partitioned into distinctive regions of interest. The total area (TA,  $\text{cm}^2$ ) of the proximal femig (also known in the medical literature  $77 \text{ cm}^2$ ) of the proximal femur (also known in the medical literature  $^{78}$  as total area of the hip) is the sum of three individual areas:<br> $^{79}$  fomoral poek treebanteric region and intertreebanteric provincial femoral neck, trochanteric region, and intertrochanteric/proximal

<b>SAN WARRANT</b> r. ĉ
e 一

Fig. 1. The total area  $(cm<sup>2</sup>)$  of the proximal femur (gray color).

diaphysis regions (Fig. 1) [21,22]. A femur from each individual (as a rule, the bone from the left side) was scanned with a Hologic QDR-4500A densitometer (Hologic, Inc., Bedford, MA) at the Nuclear Medicine Unit (Coimbra Hospital and University Centre, Portugal) and the computer produced the above designated semi-automated regions of interest (if required the technologist made minor adjustments) and the area ( $cm<sup>2</sup>$ ) for each region is calculated.  $86$ Subsequently TA was automatically determined by the densitometer's software (Fig. 2). Femora were placed in anteroposterior position; with the femoral neck parallel to the plane of the scanner; in a low-density cardboard container with 10 cm depth of dry rice acting as a surrogate for soft tissue (soft tissues and bone marrow slightly influence the reading of bone mineral content but not TA). Fifty femora were scanned in two different days to check repeatability of the DXA measurements. The magnitude of the  $\frac{94}{12}$ intraobserver error was assessed with the relative technical error





Total BMD CV 1.0% WHO Classification: Osteoporosis Fracture Risk: High

Fig. 2. Results summary for a DXA scanning (CISC, female, 80 years old). In this example, TA is 43.24 which is the sum of three different areas: neck, trochanteric and intertrochanteric.

Please cite this article in press as: F. Curate, et al., Sex estimation with the total area of the proximal femur: A densitometric approach, Forensic Sci. Int. (2017), <http://dx.doi.org/10.1016/j.forsciint.2017.02.035>

# **DXA Results Summary:**

Table 1

Descriptive statistics for TA (cm<sup>2</sup>) in both sexes; Coimbra Identified Skeletal Collection (CISC), 21st Century Identified Skeletal Collection (ISC/XXI) and pooled samples.



SD: standard deviation; 95% CI: 95% confidence interval.

96 of measurement (rTEM) [29] and it was very low (rTEM = 0.42),<br>97 survey that the negliging of the female use neithermod 97 suggesting that the positioning of the femur was performed<br>98 supprensiately physiological length of the femur was obtained appropriately. Physiological length of the femur was obtained <sup>99</sup> following Martin [30].

100 Descriptive statistics are presented as group means, standard<br>101 deviation (SD) and 05% sonfidence intervals (05% CU for the mean 101 deviation (SD) and 95% confidence intervals (95% CI) for the mean.<br>102 Mermality of the data was assessed through elements and <sup>102</sup> Normality of the data was assessed through skewness and<br><sup>103</sup> Lutteria and homographicity with a Layona's test [21] A t test  $103$  kurtosis, and homoscedasticity with a Levene's test [31]. A t-test  $104$  (independent camples), we used to evaluate the pull  $104$  (independent samples) was used to evaluate the null  $105$  hypothesis that TA mean in males and females was equal. To  $105$  hypothesis that TA mean in males and females was equal. To  $106$  assess sexual dimension the ensuing indicator was employed  $\frac{106}{107}$  assess sexual dimorphism, the ensuing indicator was employed  $[32]$ :

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

109 where  $\bar{x}_m$  and  $\bar{x}_f$  are the mean TA values for males and females,<br>110 respectively respectively.

111 The models for the mathematical prediction of sex were<br>112  $\sigma$  concented through linear discriminant analysis (LDA) logistic <sup>112</sup> generated through linear discriminant analysis (LDA), logistic  $\frac{113}{2}$  generation (LD), classification and regression trees (CAPT), and 113 regression (LR), classification and regression trees (CART), and reduce error pruning trees (PERTree), LDA is the eldest election 114 reduce error pruning trees (REPTree). LDA is the oldest classifier  $115$  retill in use and is founded upon the potion of identifying a linear  $115$  still in use and is founded upon the notion of identifying a linear<br> $116$  combination of prodictor unriples that optimally congrates  $116$  combination of predictor variables that optimally separates  $117$  mutually syclusive groups. Discriminant analysis than creates a  $117$  mutually exclusive groups. Discriminant analysis then creates a<br> $118$  discriminant function that parsimoniously enitomizes the differ- $118$  discriminant function that parsimoniously epitomizes the differ-<br> $119$  apces, between groups and classifies now individuals with  $\frac{119}{120}$  ences between groups and classifies new individuals with  $\frac{120}{120}$  unknown group momborship  $\frac{122}{120}$  Logistic regression is a  $120$  unknown group membership [33]. Logistic regression is a  $121$  non-parametric statistical modeling approach that can be used  $121$  non-parametric statistical modeling approach that can be used<br> $122$  to describe the relationship of one or more independent variables  $122$  to describe the relationship of one or more independent variables<br> $123$  to a disbetement dependent variable [24]. Classification and  $123$  to a dichotomous dependent variable [34]. Classification and regression trees are binary requiring classifiers that generate  $124$  regression trees are binary recursive classifiers that generate  $125$  biomarchical decision trees by partitioning data among classes of <sup>125</sup> hierarchical decision trees by partitioning data among classes of<br><sup>126</sup> **Q5** the criterion at a given node resulting from an "iffthen" rule <sup>126</sup> **the criterion at a given node, resulting from an "if/then" rule**<br><sup>127</sup> directed to a set of predictors [35,36]. Beduce error pruning trees is <sup>127</sup> directed to a set of predictors  $[35,36]$ . Reduce error pruning trees is<br><sup>128</sup> the simplest method in decision tree pruning and is founded on the <sup>128</sup> the simplest method in decision tree pruning and is founded on the  $\frac{129}{12}$  principle of computing the information gain with entropy and  $129$  principle of computing the information gain with entropy and  $130$  minimizing the error that ensues from variance [36.37]. For general <sup>130</sup> minimizing the error that ensues from variance [36,37]. For general<br><sup>131</sup> maybe to the LIP CAPT and PEPTree see, for example Maroco 131 reviews of LDA, LR, CART and REPTree see, for example, Maroco  $132$  ret al. 1331 Hospital II 1341 Mu et al. 1351 and Cupta et al. 1361 In <sup>132</sup> et al. [33], Hosmer et al. [34], Wu et al. [35], and Gupta et al. [36]. In<br><sup>133</sup> erder to avoid overfitting and to insure that the results are order to avoid overfitting and to insure that the results are 134 generalizable to an independent data set, a 10-fold cross-validation 135 approach was followed to train the classifiers.

136 The performance of the provisional and cross-validated models<br>137 as well as the discriminative power of the models in the testing  $137$  – as well as the discriminative power of the models in the testing<br> $138$  distance was ovaluated through overall accuracy (a moasure of <sup>138</sup> dataset – was evaluated through overall accuracy (a measure of  $\frac{139}{25}$  accessment between the decumented and the predicted sex)  $139$  agreement between the documented and the predicted sex),<br> $140$  consitivity (the prepartion of males that were correctly recog  $140$  sensitivity (the proportion of males that were correctly recog-<br> $141$  nized) specificity (the proportion of females that were properly <sup>141</sup> nized), specificity (the proportion of females that were properly  $\frac{142}{142}$  predicted). Cohen's Kappa (also a measure of total agreement but <sup>142</sup> predicted), Cohen's Kappa (also a measure of total agreement but  $^{143}$  adjusting for those that occur by chance alone) and Area Under the <sup>143</sup> adjusting for those that occur by chance alone) and Area Under the  $\frac{144}{2}$  Receiver Operating Characteristic Curve (AUC) Receiver Operating Characteristic Curve (AUC).

145 **All analyses were performed with R programming language**<br>146 **All and Waikato Environment for Knowledge Analysis [40]** [38,39] and Waikato Environment for Knowledge Analysis [40].

## **147** 3. Results **147**

Descriptive statistics for the Coimbra Identified Skeletal 148<br>Ilection and the 21st Century Identified Skeletal Collection 149 149<br>
149 complex are summarized in Table 1. The total area of the provincial<br>
150 samples are summarized in Table 1. The total area of the proximal  $150$ <br>formula is statistically different between seves both in the training  $151$ femur is statistically different between sexes both in the training  $151$ <br>(t) 20.907;  $d_f = 222$ ;  $p_f = 0.001$ ) and the testing samples (t; 152 (t:  $-20.907$ ; df=222; p < 0.001) and the testing samples (t:  $152$  $-11.666$ ; df = 58; p < 0.001). Kernel density plots show the distribution of TA values per sex (Figs. 3 and 4). TA is  $23.0\%$  and  $154$ <br>21.0% larger in males in the CISC and ISC/VVI samples, respectively.  $155$ 21.0% larger in males in the CISC and ISC/XXI samples, respectively.  $155$ <br>The tatal area of the provincil formula products in a strengty.  $156$ The total area of the proximal femur is moderately to strongly  $156$ <br>correlated with femoral physiological length in both camples and  $157$ correlated with femoral physiological length in both samples and  $157$ <br>savec (CISC: Degreen's TA\*EDI<br> $158$ sexes (CISC: Pearson's TA\*FPL<sub>females</sub>: 0.578; p < 0.001/Pearson's 158<br>TA\*FPL<sub>males</sub>: 0.559; p < 0.001 | ISC/XXI: Pearson's TA\*FPL<sub>females</sub>: 159<br>0.775; p < 0.001/Pearson's TA\*FPL : 0.527; p < 0.001) but it is 1.60 <sup>160</sup> 0.725; p < 0.001/Pearson's TA\*FPLmales: 0.537; p < 0.001) but it is not correlated with age at death (CISC: Pearson's TA\*age<sub>females</sub>: <sup>161</sup><br>0.170: p = 0.073/Pearson's TA\*age : 0.116: p = 0.222 | ISC/XXI. 162 <sup>162</sup> 0.170; p = 0.073/Pearson's TA\*agemales: 0.116; p = 0.222 | ISC/XXI: Pearson's TA\*age<sub>females</sub>:  $-0.195$ ; p=0.303/Pearson's TA\*age<sub>males</sub>:  $163$ 0.253;  $p = 0.177$ ). 164<br>The logistic regression model is summarized in Table 2, It is 165

The logistic regression model is summarized in Table 2. It is the lost of the local time in the summarized in Table 2. It is the local time field that the encyclopediate in the meaning of the local time in the meaning of t defined by the ensuing equation (females classified with negative  $166$ <br>values males classified with positive values): values, males classified with positive values):

$$
Sex = 0.800 * TA - 30.498
$$

The sex was correctly predicted in 92.0% of all individuals  $168$ <br>naitivity 01.1% enough situated and (sensitivity: 91.1%; specificity: 92.9%), with a significant discrimi-<br>nant-capability in both the provisional and cross validation nant capability in both the provisional and cross-validation  $170$ <br>models In the holdout sample (ISC/YYI) sex was accurately models. In the holdout sample (ISC/XXI), sex was accurately  $111$ <br>estimated in 01.7% of the cases. The model appropriately identified estimated in 91.7% of the cases. The model appropriately identified  $172$ <br>06.7% of familes and 96.7% of males (Table 2) 96.7% of females and 86.7% of males (Table 3).

Box's M was used to test the equality of the variance–covariance  $174$ <br>trises (Boy's M+2.457+ p=0.117) Lipear discriminant applyeis  $175$ matrices (Box's M: 2.467; p=0.117). Linear discriminant analysis  $175$ <br>produced a single discriminant function with a qutoff point squal  $176$ produced a single discriminant function with a cutoff point equal  $176$ <br>to zero (seems above zero glassified as males and below zero as  $177$ to zero (scores above zero classified as males and below zero as  $177$ <br> $fomalso$ : females):

# $Sex = 0.279 * TA - 10.738$

In both the provisional and cross-validation models, sex was  $179$ <br>rectly estimated in 90.6% of individuals (sepsitivity: 88.4%; 180 correctly estimated in 90.6% of individuals (sensitivity: 88.4%;  $^{180}$ <br>specificity: 92.9%) In the testing sample sex was correctly  $^{181}$ specificity: 92.9%). In the testing sample, sex was correctly  $181$ <br>assessed in 91.7% of the individuals (sensitivity: 86.7%; specificity;  $182$ assessed in 91.7% of the individuals (sensitivity: 86.7%; specificity:  $\frac{182}{183}$ 183<br>The CAPT decision tree is uttaily simple and straightforward 184

The CART decision tree is utterly simple and straightforward,  $184$ <br>d provided a sectioning point of 27.21 in which  $T_A$  < 27.21 – and provided a sectioning point of 37.31, in which  $TA < 37.31 =$   $185$ <br>EEMALE, and  $TA > 27.31 = MALE$ . The decision rule correctly, FEMALE, and  $TA \geq 37.31 = MALE$ . The decision rule correctly  $186$ <br>classified 02.2% of all individuals in the provisional model with classified 93.3% of all individuals in the provisional model, with  $187$ <br>a consitivity of 05.5% and a specificity of 01.1%. In the cross a sensitivity of 95.5% and a specificity of 91.1%. In the cross-<br>unlideted model, suggell assument use 00.2% (sensitivity 03.0%, 189 validated model, overall accuracy was  $90.2\%$  (sensitivity:  $92.0\%$ ;  $189$ <br>specificity:  $89.4\%$ ) In the testing sample, overall assumes reached specificity: 88.4%). In the testing sample, overall accuracy reached  $190$ <br>00.0%, with 02.2% males, and 86.7% fomales correctly assigned 90.0%, with 93.3% males and 86.7% females correctly assigned  $\frac{191}{192}$ <sup>192</sup> (Table 3).

The reduced error pruning tree classifier provided a sectioning  $193$ <br>interference in which  $TA = 2777$   $\times$  FMALE and  $TA \ge 2777$   $194$ point of 37.77, in which TA < 37.77 = FEMALE, and TA  $\geq$  37.77 =  $194$ <br>MALE Overall accuracy was 93.9% (with the same sensitivity and  $195$ MALE. Overall accuracy was 92.9% (with the same sensitivity and



Fig. 3. Kernel density distribution of TA  $\rm (cm^2)$  by sex (CISC sample).

<sup>196</sup> specificity) in the provisional model, and 90.6% (sensitivity:<br><sup>197</sup> and  $2\%$  specificity: 80.2%) in the grees validated model. In the ISC <sup>197</sup> 92.0%; specificity: 89.3%) in the cross-validated model. In the ISC/<br> $\frac{198}{198}$  WH haldwe segments  $\frac{0.11\%}{0.5}$  of all individuals were segmented <sup>198</sup> XXI holdout sample, 91.1% of all individuals were correctly<br> $\frac{199}{2}$  elassified with 00.0% famales and 02.2% males properly allocated <sup>199</sup> classified, with 90.0% females and 93.3% males properly allocated  $\frac{200}{(3.516 \times 2)}$  $(Table 3).$ 

# <sup>201</sup> 4. Discussion

<sup>202</sup> Sexual dimorphism in the human skeleton has been classically<br> $^{203}$  investigated in the politic enginement long bangs. In assess of  $203$  investigated in the pelvis, cranium and long bones. In cases of  $204$  commingled, easttored, fractional, and/or, fractional buman  $^{204}$  commingled, scattered, fractional and/or fragmented human  $^{205}$  cludatal remains the polyie is not always available for forensic  $205$  skeletal remains, the pelvis is not always available for forensic  $206$  analysis as a set of the dimensional elements including 206 analysis. As such, other dimorphic skeletal elements – including  $207$  the formula  $2.4 \times 10^{-3}$  are widely used in sex determination. Because <sup>207</sup> the femur  $[2,4,6]$  – are widely used in sex determination. Research  $\frac{208}{\pi}$  in forensic, anthropology typically involves the analysis of  $^{208}$  in forensic anthropology typically involves the analysis of  $^{209}$  codoveric remains in different states of preservation and com- $209$  cadaveric remains in different states of preservation and com-<br> $210$  plateness including human remains with or without soft tissues <sup>210</sup> pleteness, including human remains with or without soft tissues.<br> $^{211}$  lmaging anneaehos for the assessment of foatures related with the  $^{211}$  Imaging approaches for the assessment of features related with the  $^{212}$  higherical profile should be professed in gases when skeletal <sup>212</sup> biological profile should be preferred in cases when skeletal<br><sup>213</sup> preparation is socially offensive or simply not viable [6.11.12.41]. In 213 preparation is socially offensive or simply not viable [6,11,12,41]. In  $214$  cuch cases DYA is a suitable technique to estimate sex [10.22.24]. such cases, DXA is a suitable technique to estimate sex [10,23,24], and purportedly age at death and ancestry  $[10,24-26]$  -even in the  $215$ <br>axes of recovery of a single formulation case of recovery of a single femur.  $^{216}$ <br>The channel caused dimensions of the tatal area of the  $^{217}$ 

The observed sexual dimorphism of the total area of the  $^{217}$ <br>wimal formus in both the training (CISC) and testing samples  $^{218}$ proximal femur in both the training (CISC) and testing samples  $^{218}$ <br>(ISC) 219 as assessed through DYA, was in agreement with the  $^{219}$ (ISC/XXI), as assessed through DXA, was in agreement with the  $^{219}$ <br>results established in opidemial studies [42,42]. TA subjects  $^{220}$ results established in epidemiological studies  $[42,43]$ . TA exhibits  $220$ a slight variation with ancestry; notwithstanding, differences  $^{221}$ <br>between exuse are large (since 10 cm<sup>2</sup>) and experient within env.  $^{222}$ between sexes are large (circa 10 cm<sup>2</sup>) and consistent within any  $222$ population (>20% variation between sexes) [42]. Sexual differences  $^{223}$ <br>in hence sine are established early in life neesibly sure in wtem by  $^{224}$ in bone size are established early in life, possibly even in utero, but  $224$ <br>are more poticeable after puberty [44.45]. For example, periodical  $225$ are more noticeable after puberty  $[44,45]$ . For example, periosteal  $225$ <br>grouth which expands have diameter accelerates during puberty  $226$ growth, which expands bone diameter, accelerates during puberty  $226$ <br>in males: while explicit completion of longitudinal growth and  $227$ in males; while earlier completion of longitudinal growth and  $^{227}$ <br>inhibition of periodical appecition produces smaller bones in  $^{228}$ inhibition of periosteal apposition produces smaller bones in  $\frac{228}{229}$ <br>fomales [45,46] Pape grouth and size is influenced by genetic and  $\frac{229}{2}$ females [45,46]. Bone growth and size is influenced by genetic and  $229$ <br>hormonal factors mechanical loading and nutrition among  $230$ hormonal factors, mechanical loading and nutrition, among  $^{230}$ <br>others and it is probable that the ensuing effect on hope sine  $^{231}$ others, and it is probable that the ensuing effect on bone size  $\frac{231}{232}$ may be sex-specific [46–49]. The structural phenotype of the  $^{232}$ <br>provimal femur in particular shows bigh heritability [48,50] also  $^{233}$ proximal femur, in particular, shows high heritability [48,50], also  $233$ <br>conforming to Wolff's law and Harold Erost's mechanostat model  $234$ conforming to Wolff's law and Harold Frost's mechanostat model  $^{234}$ <br>[51.52] The moderate to strong association of TA with femoral  $^{235}$ [51,52]. The moderate to strong association of TA with femoral  $^{235}$ <br>physiological length, suggests, that, sex, dimorphism in the  $^{236}$ <sup>236</sup> physiological length suggests that sex dimorphism in the

### Table 2

Logistic regression model fitting for the training sample (CISC).



TA: total area (cm<sup>2</sup>);  $\beta$ : the coefficient for the constant in the null model; SE: standard error; Wald: Wald chi-square test; Exp ( $\beta$ ): exponentiation of the  $\beta$  coefficient.

F. Curate et al. / Forensic Science International xxx (2016) xxx–xxx 5



Fig. 4. Kernel density distribution of TA  $(cm^2)$  by sex (ISC/XXI sample).

237 expression of TA has a size effect component. BMD declines during<br>238 aging in all populations particularly in females [25] but bone area <sup>238</sup> aging in all populations, particularly in females  $[25]$ , but bone area<br><sup>239</sup> tends to remain constant or increase marginally with age in adults 239 tends to remain constant or increase marginally with age in adults  $\frac{239}{1421}$  Even in the latter case, area increases much less than the  $\frac{240}{421}$  [42]. Even in the latter case, area increases much less than the  $\frac{241}{4}$  degree of some dimension. In the observed samples Taure pot  $^{241}$  degree of sexual dimorphism. In the observed samples, TA was not  $^{242}$  $242$  associated with age at death.<br> $243$  Sex association with the t

<sup>243</sup> Sex assessment with the total area of the proximal femur in  $\frac{244}{12}$  by proximal femur in the human skeletal remains shows high overall accuracy in the

### Table 3

Classification accuracy with the different classifiers.

cross-validated models (always exceeding  $90\%$ ), with an effective  $^{245}$ <br>performance, independently of the classifier used to create the  $^{245}$ performance, independently of the classifier used to create the  $^{246}$ <br>classification models. The allocation accuracy in a holdout sample  $^{247}$ classification models. The allocation accuracy in a holdout sample  $247$ <br>not used to develop the models was also very high suggesting that  $248$ not used to develop the models was also very high, suggesting that  $^{248}$ <br>the results are generalizable to independent datasets. Notwith  $^{249}$ the results are generalizable to independent datasets. Notwith- $^{249}$ <br>standing classification bias (the difference between properly  $^{250}$ standing, classification bias (the difference between properly  $^{250}$ <br>classified females and males) with the traditional classifiers (LR  $^{251}$ classified females and males) with the traditional classifiers (LR  $^{251}$ <br>and JDA with 13.3% of misclassified females and only 3.3%  $^{252}$ <sup>252</sup> and LDA, with 13.3% of misclassified females and only 3.3%



LR: logistic regression; LDA: linear discriminant analysis; CART: classification and regression trees; REPTree: reduce error pruning trees; AUC: area under the receiver operating characteristic curve.

<sup>253</sup> misclassified males) and the CART algorithm  $(6.7\%$  misclassified<br><sup>254</sup> famales and 12.3% misclassified males) was problematic in the  $^{254}$  females and 13.3% misclassified males) was problematic in the  $^{255}$  $^{255}$  testing sample.

<sup>256</sup> Sex specific accuracy is probably related with secular change in  $257$  hand dimensions [52,54], usually inducing a bigber proportion of <sup>257</sup> bone dimensions [53,54], usually inducing a higher proportion of  $^{258}$  micclassified famales when a model fitted in a chronologically <sup>258</sup> misclassified females when a model fitted in a chronologically<br><sup>259</sup> older sample is used to estimate sex. The training sample (CISC) is <sup>259</sup> older sample is used to estimate sex. The training sample (CISC) is,<br> $\frac{260}{20}$  on average composed by individuals that were been much earlier. <sup>260</sup> on average, composed by individuals that were born much earlier<br> $\frac{261}{\pi}$  than individuals in the testing sample (ISC/YYI) with other  $^{261}$  than individuals in the testing sample (ISC/XXI) – with other  $^{262}$  relevant differences between samples including socioeconomic  $\frac{262}{100}$  relevant differences between samples, including socioeconomic<br> $\frac{263}{1000}$  status and mortality pattern – but the magnitude of sexual  $^{263}$  status and mortality pattern – but the magnitude of sexual  $^{264}$  dimorphism in the total area of the proximal femur is very similar  $^{264}$  dimorphism in the total area of the proximal femur is very similar<br> $^{265}$  in both samples. This is also relevant for the assessment of this  $^{265}$  in both samples. This is also relevant for the assessment of this  $^{266}$  method in samples of non-Portuguese origin Besides the problem  $^{266}$  method in samples of non-Portuguese origin. Besides the problem  $^{267}$  of secular change the selection of the statistical model also seems <sup>267</sup> of secular change, the selection of the statistical model also seems<br><sup>268</sup> critical to lower error rate and bias [23,55]. In fact, the decision rule <sup>268</sup> critical to lower error rate and bias [33,55]. In fact, the decision rule<br>269 consider hy the PEPTree classifier maximized the querell acquirant <sup>269</sup> provided by the REPTree classifier maximized the overall accuracy  $270$  and it is proposition bigger in the overall accuracy  $270$ <sup>270</sup> while improving bias: misclassification difference between sexes<br><sup>271</sup> in the holdout sample was lower than the recommended  $F^{\alpha}$  $271$  in the holdout sample was lower than the recommended 5%<br> $272$  throshold [12]  $\frac{272}{273}$  threshold [12].

<sup>273</sup> Classical statistical techniques, such as LR and LDA, have been<br><sup>274</sup> videly used to assess soy in forensis, contexts, <sup>[16</sup>] <sup>274</sup> widely used to assess sex in forensic contexts  $[1,6 275$  15,18,19,32,56,57], but the promising performance of data mining<br> $276$  and the suith sharifland like summer vector are things and the  $^{276}$  methods, with classifiers like support vector machines, random<br> $^{277}$  fancts an elastification trace has led to a resent research annot in  $277$  forests or classification trees, has led to a recent research appeal in  $278$  their application to classification problems in forencie anthropol <sup>278</sup> their application to classification problems in forensic anthropol-<br><sup>279</sup> agus 16.52.55.58.601. Pecults are conflicting about classification  $^{279}$  ogy [6,53,55,58–60]. Results are conflicting about classification  $^{280}$  accuracy of data mining classifications as compared to traditional  $^{280}$  accuracy of data mining classifiers as compared to traditional  $^{281}$  methods  $^{10}$   $^{80}$   $^{82}$   $^{82}$   $^{89}$  uith the classifiers' performance  $^{281}$  methods [e.g., Refs. 53,58] with the classifiers' performance<br> $^{282}$  affected by the different arrangements of predictors data <sup>282</sup> affected by the different arrangements of predictors, data<br><sup>283</sup> assumptions parameters' tuning and sample sizes [33] In general <sup>283</sup> assumptions, parameters' tuning and sample sizes [33]. In general,  $\frac{284}{100}$  curves the show that both traditional and decision tree  $^{284}$  our results show that both traditional and decision tree<br> $^{285}$  learning techniques perform very well under cross-validation  $^{285}$  learning techniques perform very well under cross-validation<br> $^{286}$  but except for the PEPTree algorithm, the models display <sup>286</sup> but, except for the REPTree algorithm, the models display<br><sup>287</sup> unbalanced classification efficiency in the testing sample  $287$  unbalanced classification efficiency in the testing sample.<br> $288$  Overall correct classification in this study is compara

 $288$  Overall correct classification in this study is comparable to  $289$  other seemingly bigbly accurate methods including techniques <sup>289</sup> other seemingly highly accurate methods, including techniques<br><sup>290</sup> using the polyic region [8,61,62], the cranium [58,63], and different <sup>290</sup> using the pelvic region  $[8,61,62]$ , the cranium  $[58,63]$ , and different  $291$  long bones  $[18, 0.11, 2.50]$ . The bigh overall accuracy and low bias <sup>291</sup> long bones [1,8,9,11,23,59]. The high overall accuracy and low bias<br><sup>292</sup> obtained in the testing sample with the REPTree model is  $292$  obtained in the testing sample with the REPTree model is  $293$  particularly relevant since for many published models only  $293$  particularly relevant, since for many published models only  $294$  requbitivition and gross validation acquired varies are reported <sup>294</sup> resubstitution and cross-validation accuracy rates are reported<br> $295$   $1221$  Quanfitting is often a sensequence in the first  $222$  while  $^{295}$  [32]. Overfitting is often a consequence in the first case, while<br> $^{296}$  access which is usually extincted will also be likely and itina <sup>296</sup> cross-validation usually estimates well only the likely prediction<br><sup>297</sup> cause 1041, As such as were only the segment to assess the <sup>297</sup> error [64]. As such, a more valuable approach to assess the  $\frac{298}{298}$  expecting agree of a classification model is to use an  $^{298}$  generalization error of a classificatory model is to use an  $^{299}$  independent dataset independent dataset.

# <sup>300</sup> 5. Conclusions

 $301$  The new models for the estimation of sex based on the total area<br> $302$  of the provincil formula model performed with DYA  $302$  of the proximal femur, a measurement performed with DXA,  $303$  display great assumely both in great alleged and in an  $303$  display great accuracy both in cross-validation and in an  $304$  independent cample. The model based in a fact decision tree  $304$  independent sample. The model based in a fast decision tree<br> $305$  learning algorithm (PEDTree) reduces hiss in the holdout sample to  $\frac{305}{200}$  learning algorithm (REPTree) reduces bias in the holdout sample to  $\frac{306}{200}$  appropriate lovels. The proposed models should endure additional  $\frac{306}{207}$  appropriate levels. The proposed models should endure additional  $\frac{307}{207}$  validation in independent skeletal remains (particularly of  $307$  validation in independent skeletal remains (particularly of  $308$  non-Portuguese origin) to substantiate their reliability in forensic  $\frac{308}{200}$  non-Portuguese origin) to substantiate their reliability in forensic  $\frac{309}{200}$  and/or bioarcheological contexts and/or bioarcheological contexts.

# <sup>310</sup> Conflict of interest

- 
- <sup>311</sup> The authors declare that they have no conflict of interest.

# Acknowledgments 312

The authors are grateful to the Fundação para a Ciência e  $\frac{8 \times 313}{314}$ Tecnologia (grant number  $\#$  SFRH/BPD/74015/2010) and to the  $^{314}$ <br>315 anonymous referees.

## **References** 316

- [1] M.K. Spradley, R.L. Jantz, Sex estimation in forensic anthropology: skull versus <sup>317</sup> postcranial elements, J. Forensic Sci. <sup>56</sup> (2011) <sup>289</sup>–296, doi[:http://dx.doi.org/](http://dx.doi.org/10.1111/j.1556-4029.2010.01635.x) <sup>318</sup> [10.1111/j.1556-4029.2010.01635.x](http://dx.doi.org/10.1111/j.1556-4029.2010.01635.x).
- [2] J. Bruzek, P. Murail, Methodology and reliability of sex diagnosis from the 319 skeleton, in: A. Schmitt, E. Cunha, J. Pinheiro (Eds.), Forensic Anthropology and<br>skeleton, in: A. Schmitt, E. Cunha, J. Pinheiro (Eds.), Forensic Anthropology and<br>320 Medicine: Complementary Sciences from Recovery to Cause of Death, Humana  $\frac{320}{321}$ Press, New Jersey, 2006, pp. 225-242.
- [3] A.M. Christensen, N.V. Passalacqua, E.J. Bartelink, Forensic Anthropology: <sup>322</sup> Current Methods and Practice, Academic Press, San Diego, CA, 2014.
- [4] T.D. Stewart, Essentials of Forensic Anthropology, Charles <sup>C</sup> Thomas, <sup>323</sup> Springfield, IL, 1979.
- [5] T.D. White, M.T. Black, P.A. Folkens, Human Osteology, Academic Press, San 324 Diego, CA, 2012.
- [6] K. Krishan, P.M. Chatterjee, T. Kanchan, S. Kaur, N. Baryah, R.K. Singh, A review 825 Sex estimation techniques during examination of skeletal remains in the sex estimation techniques during examination of skeletal remains in the forecast forecast of skeletal remains in the sex estimate of  $\frac{326}{\text{m}}$ forensic anthropology casework, Forensic Sci. Int. 261 (2016) 165.e1–165.e8,  $327$ <br>doi:[http://dx.doi.org/10.1016/j.forsciint.2016.02.007.](http://dx.doi.org/10.1016/j.forsciint.2016.02.007)
- [7] F. Curate, J. Coelho, D. Gonçalves, C. Coelho, M.T. Ferreira, D. Navega, E. Cunha, A method for sex estimation using the proximal femur, Forensic Sci. Int. 266<br>for sex estimation using the proximal femur, Forensic Sci. Int. 266<br> $329$ <sup>329</sup> (2016) 579.e1–579.e7, doi:[http://dx.doi.org/10.1016/j.forsciint.2016.06.011.](http://dx.doi.org/10.1016/j.forsciint.2016.06.011)
- [8] J. Albanese, G. Eklics, A. Tuck, A metric method for sex determination using the <sup>330</sup> proximal femur and fragmentary hipbone, J. Forensic Sci. <sup>53</sup> (2008) <sup>331</sup> <sup>1283</sup>–1288, doi[:http://dx.doi.org/10.1111/j.1556-4029.2008.00855.x](http://dx.doi.org/10.1111/j.1556-4029.2008.00855.x).
- [9] V. Alunni-Perret, P. Staccini, G. Quatrehomme, Sex determination from the 332 v. Alumn-Ferret, F. Statem, G. Quatteriorme, Sex determination from the<br>distal part of the femur in a French contemporary population, Forensic Sci. Int. 333<br>175 (2000) 112, 117 doi: http://dv. doi.org/10.1016/i forecij <sup>333</sup> <sup>175</sup> (2008) <sup>113</sup>–117, doi:[http://dx.doi.org/10.1016/j.forsciint.2007.05.018.](http://dx.doi.org/10.1016/j.forsciint.2007.05.018)
- [10] R.A. Meeusen, A.M. Christensen, J.T. Hefner, The use of femoral neck axis length <sup>334</sup> to estimate sex and ancestry, J. Forensic Sci. <sup>60</sup> (2015) <sup>1300</sup>–1304, doi:[http://](http://dx.doi.org/10.1111/1556-4029.12820) <sup>335</sup> [dx.doi.org/10.1111/1556-4029.12820](http://dx.doi.org/10.1111/1556-4029.12820).
- [11] E.F. Kranioti, N. Vorniotakis, C. Galiatsou, Sex identification and software  $\frac{336}{480}$  development using digital femoral head radiographs, Forensic Sci. Int. 189  $\frac{336}{42009}$  development using digital femoral head radiographs, Forensic Sci. Int. 189  $\frac{337}{42009}$ (2009) 113.e1-113.e7, doi:<http://dx.doi.org/10.1016/j.forsciint.2009.04.014>.<br>[12] R. DeSilva, A. Flavel, D. Franklin, Estimation of sex from the metric assessment
- [12] R. DeSilva, A. Flavel, D. Franklin, Estimation of sex from the metric assessment <sup>338</sup> of digital hand radiographs in <sup>a</sup> Western Australian population, Forensic Sci. <sup>339</sup> Int. <sup>244</sup> (2014) 314.e1–314.e7, doi[:http://dx.doi.org/10.1016/j.for-](http://dx.doi.org/10.1016/j.forsciint.2014.08.019) <sup>340</sup> [sciint.2014.08.019](http://dx.doi.org/10.1016/j.forsciint.2014.08.019).
- [13] P.J. Macaluso Jr., J. Lucena, Estimation of sex from sternal dimensions derived <sup>341</sup> from chest plate radiographs in contemporary Spaniards, Int. J. Leg. Med. <sup>341</sup><br>from chest plate radiographs in contemporary Spaniards, Int. J. Leg. Med. <sup>342</sup> <sup>342</sup> (2014) <sup>389</sup>–395, doi[:http://dx.doi.org/10.1007/s00414-013-0910-z](http://dx.doi.org/10.1007/s00414-013-0910-z).
- [14] C. Robinson, R. Eisma, B. Morgan, A. Jeffery, E.A.M. Graham, S. Black, G.N. Rutty, <sup>343</sup> Anthropological measurement of lower limb and foot bones using multidetector computed tomography, J. Forensic Sci. 53 (2008) 1289–1295, doi:<br>  $344$ <br>  $345$ http://dx.doi.org/10.1111/j.1556-4029.2008.00875.x
- [15] M. López-Alcaraz, P.M. Garamendi, I. Alemán, L.M. Botella, Image analysis of 1. Super-Finder of sex determination in a computed tomography sample, pubic bone for sex determination in a computed tomography sample, 347 Int. J. Leg. Med. 127 (2013) 1145–1155, doi:[http://dx.doi.org/10.1007/s00414-](http://dx.doi.org/10.1007/s00414-013-0900-1) 348
- [16] S.U. Ramadan, N. Türkmen, N.A. Dolgun, D. Gökharman, R.G. Menezes, M.  $\frac{349}{1000}$  Kacar, Sex determination from measurements of the sternum and  $\frac{349}{10000}$  from  $\frac{349}{10000}$ fourth rib using multislice computed tomography of the chest, Forensic Sci. Int.  $\frac{350}{351}$ <sup>351</sup> <sup>197</sup> (2010) <sup>3</sup>–7, doi[:http://dx.doi.org/10.1016/j.forsciint.2009.12.049.](http://dx.doi.org/10.1016/j.forsciint.2009.12.049)
- [17] D. Ilgüy, M. Ilgüy, N. Ersan, S. Dölekoğlu, E. Fişekçioğlu, Measurements of the 352 b. ngay, w. aigay, w. zisan, b. bolckogia, b. 1336, b. 1336, b. externed magnum and mandible in relation to sex using CBCT, J. Forensic Sci. 353<br>53 (2014) 601–605 doi: http://dx. doi. pre/10.1111/1556-4029.12376 <sup>353</sup> <sup>59</sup> (2014) <sup>601</sup>–605, doi:[http://dx.doi.org/10.1111/1556-4029.12376.](http://dx.doi.org/10.1111/1556-4029.12376)
- [18] C.R. Torwalt, R.D. Hoppa, A test of sex determination from measurements of <sup>354</sup> chest radiographs, J. Forensic Sci. <sup>50</sup> (2005) <sup>785</sup>–790.
- [19] A. Harma, H.M. Karakas, Determination of sex from the femur in Anatolian <sup>355</sup> Caucasians: <sup>a</sup> digital radiological study, J. Forensic Sci. <sup>14</sup> (2007) <sup>190</sup>–194, doi: 356 <http://dx.doi.org/10.1016/j.jcfm.2006.05.008>.
- [20] A. Mitra, A.P. Vida, R.N. Ali, M. Farzaneh, V.F. Maryam, Y. Vahid, Sexing based on measurements of the femoral head parameters on pelvic radiographs,<br>measurements of the femoral head parameters on pelvic radiographs,<br>J. Forensic Leg. Med. 23 (2014) 70–75, doi[:http://dx.doi.org/10.1016/j.](http://dx.doi.org/10.1016/j.jflm.2014.01.004) 359<br>iflm.2014.01
- [21] S.L. Bonnick, L.A. Lewis, Bone Densitometry for Technologists, Springer, New 360 York, 2013.
- [22] F. Curate, Osteoporosis and paleopathology: a review, J. Anthropol. Sci. 92 <sup>361</sup> (2014) <sup>119</sup>–146, doi[:http://dx.doi.org/10.4436/JASS.92003](http://dx.doi.org/10.4436/JASS.92003).
- [23] B.P. Wheatley, An evaluation of sex and body weight determination from the  $\frac{362}{262}$  proximal femur using DXA technology and its potential for forensic  $\frac{362}{262}$ anthropology, Forensic Sci. Int. 147 (2005) 141–145, doi[:http://dx.doi.org/](http://dx.doi.org/10.1016/j.forsciint.2004.09.076) 364<br>[10.1016/j.forsciint.2004.09.076](http://dx.doi.org/10.1016/j.forsciint.2004.09.076).
- [24] R. Fernández Castillo, M.C. López Ruiz, Assessment of age and sex by means of 16. Extra bone densition, M.C. Lopez Kuiz, Assessment of age and sex by means of DXA bone densition in forensic Sci. Int. 365 <sup>366</sup> <sup>209</sup> (2011) <sup>53</sup>–58, doi:<http://dx.doi.org/10.1016/j.forsciint.2010.12.008>.

- [25] F. Curate, A. Albuquerque, E.M. Cunha, Age at death estimation using bone 1231 1. Curace, A. Abduquerque, E.M. Curina, Age at death estimation dsing bond<br>densitometry: testing the Fernández Castillo and López Ruiz method in two<br>documented skeletal samples from Portugal, Forensic Sci. Int. 226 (2 <sup>369</sup> e1–296.e6, doi:<http://dx.doi.org/10.1016/j.forsciint.2012.12.002>.
- [26] A.M. Christensen, W.D. Leslie, S. Baim, Ancestral differences in femoral neck 370 axis length: possible implications for forensic anthropological analyses,<br>axis length: possible implications for forensic anthropological analyses,<br>Forensic Sci. Int. 236 (2014) 193.e1–193.e4, doi:http://dx.doi.org/10.
- [27] E. Cunha, S. Wasterlain, The Coimbra identified osteological collections, in: G. <sup>373</sup> Grupe, J. Peters (Eds.), Skeletal Series and Their Socio-Economic Context, <sup>374</sup> Verlag Marie Leidorf GmbH, Rahden/Westf, 2007, pp. <sup>23</sup>–33.
- [28] M.T. Ferreira, R. Vicente, D. Navega, D. Gonçalves, F. Curate, E. Cunha, <sup>A</sup> new <sup>375</sup> forensic collection housed at the University of Coimbra, Portugal: the 21st 376 century identified skeletal collection, Forensic Sci. Int. 245 (2014)  $202.e1-202.e5$ , doi:[http://dx.doi.org/10.1016/j.forsciint.2014.09.021.](http://dx.doi.org/10.1016/j.forsciint.2014.09.021)
- [29] S. Ulijaszek, D. Kerr, Anthropometric measurement error and the assessment <sup>378</sup> of nutritional status, Br. J. Nutr. <sup>82</sup> (1999) <sup>165</sup>–177.
- [30] R. Martin, Lehrbuch der Anthropologie in Systematischer Darstellung mit 379 Besonderer Berücksichtigung der anthropologischen Methoden für Studier-<br>380 Besonderer Berücksichtigung der anthropologischen Methoden für Studier-380 ende, Ärtze und Forschungsreisende, Kraniologie, Osteologie, vol. 2, Gustav Fischer, Jena, 1928.
- [31] R.B. Kline, Principles and Practice of Structural Equation Modeling, The 382 Guilford Press, New York, 2010.
- [32] I. Gama, D. Navega, E. Cunha, Sex estimation using the second cervical  $383$  vertebra: a morphometric analysis in a documented Portuguese skeletal<br>384 vertebra: a morphometric analysis in a documented Portuguese skeletal<br>384 <sup>384</sup> sample, Int. J. Leg. Med. <sup>129</sup> (2015) <sup>365</sup>–372, doi[:http://dx.doi.org/10.1007/](http://dx.doi.org/10.1007/s00414-014-1083-0) <sup>385</sup> [s00414-014-1083-0.](http://dx.doi.org/10.1007/s00414-014-1083-0)
- [33] J. Maroco, D. Silva, A. Rodrigues, M. Guerreiro, I. Santana, A. de Mendonça, Data 386 5386 mining methods in the prediction of Dementia: a real-data comparison of the prediction of the same space 387 accuracy, sensitivity and specificity of linear discriminant analysis, logistic<br>388 regression, neural networks, support vector machines, classification trees and<br>389 regression, figure and file and the support vector <sup>389</sup> random forests, BMC Res. Notes <sup>4</sup> (2011) 299, doi:[http://dx.doi.org/10.1186/](http://dx.doi.org/10.1186/1756-0500-4-299) <sup>390</sup> [1756-0500-4-299](http://dx.doi.org/10.1186/1756-0500-4-299).
- [34] D.W. Hosmer, S. Lemeshow, R.X. Sturdivant, Applied Logistic Regression, John <sup>391</sup> Wiley & Sons, Inc, Hoboken, New Jersey, 2013.
- [35] X. Wu, V. Kumar, J.R. Quinlan, J. Ghosh, Q. Yang, H. Motoda, G.J. Mclachlan, A. 392 ISBN 0.1301 Ng, B. Liu, P.S. Yu, Z. Liu, P. David, J.H. Dan, Top 10 algorithms in data<br>393 minima Knowl Inf Syst 14 (2008) 1–37 doi: http://dx.doi.org/10.1007/c10115– 393 mining, Knowl. Inf. Syst. 14 (2008) 1–37, doi:[http://dx.doi.org/10.1007/s10115-](http://dx.doi.org/10.1007/s10115-007-0114-2)<br>394 [007-0114-2](http://dx.doi.org/10.1007/s10115-007-0114-2).
- [36] D.L. Gupta, A.K. Malviya, S. Singh, Performance analysis of classification tree<br>learning algorithms, Int. J. Comput. Appl. 55 (2012) 39–44.<br>**Q7** [37] LR. Quiplan, Simplifying decision trees, Int. J. Hum, Comput. Stud.
- **U** [37] J.R. Quinlan, Simplifying decision trees, Int. J. Hum. Comput. Stud. 51 (1999) 497–510.
- [38] R Development Core Team, R. A language and environment for statistical 397 computing, R Foundation for Statistical Computing, Vienna, Austria, 2016.<br>398 computing, Vienna, Austria, 2016. [http://www.R-project.org/.](http://www.R-project.org/)
- [39] W. Chang, H. Wickham, ggvis. Interactive Grammar of Graphics. <sup>R</sup> Package <sup>399</sup> Version 0.4.2, (2016) . [http://CRAN.R-project.org/package=ggvis.](http://CRAN.R-project.org/package=ggvis)
- [40] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, I.H. Witten, The 400 (40) W. Hall, L. Halls, G. Hollics, D. Hallmiger, F. Reuterham, H. Witten, The<br>WEKA data mining software: an update, SIGKDD Explor. Newsl. 11 (2009)<br>10–18.<br>[41] C. Carneiro, F. Curate, P. Borralho, E. Cunha, Radiograph
- [41] C. Carneiro, F. Curate, P. Borralho, E. Cunha, Radiographic fetal osteometry: <sup>402</sup> approach on age estimation for the Portuguese population, Forensic Sci. Int. 403 231 (2013) 397.e1-397.e5, doi:[http://dx.doi.org/10.1016/j.forsciint.2013.05.](http://dx.doi.org/10.1016/j.forsciint.2013.05.039)<br>404 [039.](http://dx.doi.org/10.1016/j.forsciint.2013.05.039)<br>42 A.C. Looker, H.W. Wahner, W.L. Dunn, M.S. Calvo, T.B. Harris, S.P. Heyse,
- [42] A.C. Looker, H.W. Wahner, W.L. Dunn, M.S. Calvo, T.B. Harris, S.P. Heyse, <sup>405</sup> J.C.C. Johnston, R.L. Lindsay, Proximal femur bone mineral levels of US 406 adults, Osteoporos. Int. 5 (1995) 389-409, doi:[http://dx.doi.org/10.1007/](http://dx.doi.org/10.1007/BF01622262)<br>407 BE01622262
- [43] A.C. Looker, L. Borrud, J. Hughes, B. Fan, J. Shepherd, L.J. Melton III, Lumbar 408 (45) A.C. Looker, L. Borrud, J. Hughes, B. Fair, J. Shepherd, L.J. Menton in, Lumbar<br>spine and proximal femur bone mineral density, bone mineral content, and<br>409 409 bone area: United States, 2005–2008, Vital Health Stat. 11 (2012) 1–132<br>410 **bulgational Contor for Health Statistics** National Center for Health Statistics.
- [44] E. Seeman, From density to structure: growing up and growing old on the 411 surfaces of bone, J. Bone Miner. Res. 12 (1997) 509–521, doi:[http://dx.doi.org/](http://dx.doi.org/10.1359/jbmr.1997.12.4.509)<br>[10.1359/jbmr.1997.12.4.509.](http://dx.doi.org/10.1359/jbmr.1997.12.4.509)
- [45] E. Seeman, Structural basis of growth-related gain and age-related loss of bone <sup>413</sup> strength, Rheumathology (Oxford) <sup>47</sup> (2008) iv2–iv8, doi:[http://dx.doi.org/](http://dx.doi.org/10.1093/rheumatology/ken177) <sup>414</sup> [10.1093/rheumatology/ken177.](http://dx.doi.org/10.1093/rheumatology/ken177)
- [46] J.W. Nieves, C. Formica, J. Ruffing, M. Zion, P. Garrett, R. Lindsay, F. Cosman, Males have larger skeletal size and bone mass than females, despite<br>comparable body size, J. Bone Miner. Res. 20 (2005) 529–535, doi.http://dx.  $416$ <br>coi.org/10.1359/JBMR.041005.<br>[47] M.C.H. Van der Meulen, M.W. Ashford,
- Determinants of femoral geometry and structure during adolescent growth, J.<br>Determinants of femoral geometry and structure during adolescent growth, J.<br>Orthop. Res. 14 (1998) 22–29, doi:http://dx.doi.org/10.1002/jor.11001
- [48] J. Gregory, R. Aspden, Femoral geometry as <sup>a</sup> risk factor for osteoporotic hip <sup>420</sup> fracture in men and women, Med. Eng. Phys. <sup>30</sup> (2008) <sup>1275</sup>–1286, doi:[http://](http://dx.doi.org/10.1016/j.medengphy.2008.09.002) <sup>421</sup> [dx.doi.org/10.1016/j.medengphy.2008.09.002](http://dx.doi.org/10.1016/j.medengphy.2008.09.002).
- [49] V. Gilsanz, A. Kovanlikaya, G. Costin, T.F. Roe, J. Sayre, F. Kaufman, Differential effect of gender on the sizes of the bones in the axial and appendicular 422 skeletons, J. Clin. Endocrinol. Metab. 82 (1997) 1603–1607, doi[:http://dx.doi.](http://dx.doi.org/10.1210/jcem.82.5.3942) 424 [org/10.1210/jcem.82.5.3942](http://dx.doi.org/10.1210/jcem.82.5.3942).
- [50] D.L. Koller, G. Liu, M.J. Econs, S.L. Hui, P.A. Morin, G. Joslyn, L.A. Rodriguez, P.M. 125 Conneally, J.C. Christian, C.C. Johnston Jr., T. Foroud, M. Peacock, Genome<br>Conneally, J.C. Christian, C.C. Johnston Jr., T. Foroud, M. Peacock, Genome<br>Screen for quantitative trait loci, underlying normal variation in screen for quantitative trait loci underlying normal variation in femoral  $426$ <br>structure LPone Miner Pes 16(2011) 025, 001 doi: http://dx.doi.org/10.1250/ 427 structure, J. Bone Miner. Res. 16 (2011) 985–991, doi[:http://dx.doi.org/10.1359/](http://dx.doi.org/10.1359/jbmr.2001.16.6.985) 427<br>[jbmr.2001.16.6.985.](http://dx.doi.org/10.1359/jbmr.2001.16.6.985)<br>[51] L.J. Melton III, T.J. Beck, S. Amin, S. Khosla, S.J. Achenbach, A.L. Oberg, B.L. Riggs,
- L.J. Metion in, i.j. beck, S. Annin, S. Knosid, S.J. Achembach, A.L. Oberg, B.L. Kiggs, 429<br>Contributions of bone density and structure to fracture risk assessment in men<br>and women, Osteoporos. Int. 16 (2005) 460–467, doi
- [52] H.M. Frost, Bone'<sup>s</sup> mechanostat: <sup>a</sup> <sup>2003</sup> update, Anat. Rec. 275A (2003) <sup>432</sup> <sup>1081</sup>–1101, doi[:http://dx.doi.org/10.1002/ar.a.10119](http://dx.doi.org/10.1002/ar.a.10119).
- [53] D. Navega, R. Vicente, D.N. Vieira, A.H. Ross, E. Cunha, Sex estimation from the  $\mu$ , Savega, R. Vietne, D.N. Viena, A.H. Ross, E. Culma, Sex estimation from the learning approach, Int. J. Leg. 433<br>Med 120 (2015) CF1, CF0, doi: between the learning approach, Int. J. Leg. 434 Med. 129 (2015) 651–659, doi:[http://dx.doi.org/10.1007/s00414-014-1070-5.](http://dx.doi.org/10.1007/s00414-014-1070-5)<br>[54] A.H. Ross, D.H. Ubelaker, E.H. Kimmerle, Implications of dimorphism,
- Fri. No., performance, the thermal matrices, the propulation variation, and secular change in estimating population affinity  $435$ <br>in the Iberian Peninsula, Forensic Sci. Int. 206 (2011) 214.e1–214.e5, doi:[http://](http://dx.doi.org/10.1016/j.forsciint.2011.01.003)  $436$ <br>d
- [55] D.G. Mcbride, M.J. Dietz, M.T. Vennemeyer, R.M. Dg, D. Mj, V. Mt, Bootstrap methods for sex determination from the Os Coxae using the ID3 Algorithm, 438<br>methods for sex determination from the Os Coxae using the ID3 Algorithm, 439<br>J. Forensic Sci. 46 (2001) 427–431.
- [56] R. DiBennardo, J.V. Taylor, Classification and misclassification in sexing the 440<br>black femur by discriminant function analysis, Am. J. Phys. Anthropol. 58<br>(1092) 145, 151, deithery (levels) regulation analysis, 1220500006 <sup>441</sup> (1982) <sup>145</sup>–151, doi[:http://dx.doi.org/10.1002/ajpa.1330580206](http://dx.doi.org/10.1002/ajpa.1330580206).
- [57] A.B. Acharya, S. Prabhu, M.V. Muddapur, Odontometric sex assessment from <sup>442</sup> logistic regression analysis, Int. J. Leg. Med. <sup>125</sup> (2011) <sup>199</sup>–204, doi[:http://dx.](http://dx.doi.org/10.1007/s00414-010-0417-9) <sup>443</sup> [doi.org/10.1007/s00414-010-0417-9.](http://dx.doi.org/10.1007/s00414-010-0417-9)
- [58] F. Santos, P. Guyomarc'h, J. Bruzek, Statistical sex determination from 1. Salitos, 1. Sulgoniac II, 3. Braziski, Statistical Sex decernimiation from<br>
craniometrics: comparison of linear discriminant analysis, logistic regression,<br>
445<br>
and support vector machines, Forensic Sci. Int. 245 (2014 <http://dx.doi.org/10.1016/j.forsciint.2014.10.010>.<br>[59] P. du Jardin, J. Ponsaille, V. Alunni-Perret, G. Quatrehomme, A comparison
- F. du Jardin, J. Ponsaille, V. Alumni-Perret, G. Quatrehomme, A comparison 447<br>between neural network and other metric methods to determine sex from the 447<br>determine sex from the sex from the sex from the 448 upper femur in a modern French population, Forensic Sci. Int. 192 (2009) 127.  $448$ <br>e1. 127.06. doi: http://dy. doi. arg/10.1016/i foregiint. 2000.07.014. 449 <sup>449</sup> e1–127.e6, doi:[http://dx.doi.org/10.1016/j.forsciint.2009.07.014.](http://dx.doi.org/10.1016/j.forsciint.2009.07.014)
- [60] M. Mahfouz, A. Badawi, B. Merkl, E.E. Abdel, E. Pritchard, K. Kesler, M. Moore, R.  $\frac{1}{2}$  Jantz, Patella sex determination by 3D statistical shape models and<br>pantz, L. Jantz, Patella sex determination by 3D statistical shape models and<br>poplicial shape models and  $\frac{450}{451}$ nonlinear classifiers, Forensic Sci. Int. 173 (2007) 161–170, doi[:http://dx.doi.](http://dx.doi.org/10.1016/j.forsciint.2007.02.024) 452<br>[org/10.1016/j.forsciint.2007.02.024](http://dx.doi.org/10.1016/j.forsciint.2007.02.024).
- [61] J.Bruzek,Amethod for visualdeterminationof sex,usingthehumanhipbone,Am. J. <sup>453</sup> Phys. Anthropol. <sup>117</sup> (2002) <sup>157</sup>–168, doi:[http://dx.doi.org/10.1002/ajpa.10012.](http://dx.doi.org/10.1002/ajpa.10012)
- [62] A. Clavero, M. Salicrú, D. Turbón, Sex prediction from the femur and hip bone  $\frac{454}{1000}$  a sample of CT images from a Spanish population, Int. J. Leg. Med. 129  $\frac{454}{1000}$ <sup>455</sup> (2015) <sup>373</sup>–383, doi:<http://dx.doi.org/10.1007/s00414-014-1069-y>.
- [63] B.A. Williams, T.L. Rogers, Evaluating the accuracy and precision of cranial morphological traits for sex determination, J. Forensic Sci. 51 (2006) 729–735, 456 doi[:http://dx.doi.org/10.1111/j.1556-4029.2006.00177.x](http://dx.doi.org/10.1111/j.1556-4029.2006.00177.x). 457
- [64] T. Hastie, R. Tibshirani, J. Friedman, The Elements of Statistical Learning: Data 1. Hasue, K. Hoshirani, J. Friedman, The Elements of Statistical Learning. Data 458<br>Mining, Inference and Prediction, Springer, New York, 2009.