Sex Estimation in Terms of Inclination and Alsberg in Proximal Femur by using Machine Learning Algorithms

S. Sertel Meyvaci[1,*], Handan Ankarali[2] and Mustafa Hizal[3]

[1]*Bolu Abant Izzet Baysal University, Faculty of Medicine, Department of Anatomy, Bolu, Turkiye*

[2]*Istanbul Medeniyet University, Faculty of Medicine, Department of Biostatistics and Medical Informatics, Istanbul, Turkiye*

[3]*Bolu Abant Izzet Baysal University, Faculty of Medicine, Department of Radiology, Bolu, Turkiye*

Abstract: The fact that the femur has a solid structure ensures that the integrity of the bone is preserved, making it favorable in the sex determination process. In this study, it was aimed to predict sex by using angular variables of femur from computed tomography (CT) images and machine learning algorithms. In this study, a total of 4 angular measurements, including the femoral inclination angle (FIA) and Alsberg angle (FAA) of the proximal femur on both sides, were evaluated on CT images of 88 female and 92 male adults. Logistic regression (LR) and classification and regression tree (CART) machine learning algorithms were used for sex diagnosis. 5-fold cross-validation method was used in the training and testing processes of the models. Model performances were evaluated with area under the ROC curve, Precision and Recall statistics. Of the 4 angle measurements evaluated, only the right side FIA mean was significantly higher in women (p=0.042). The sex diagnosis success of the LR model and the CART algorithm were found to be similar (p values 0.014 and 0.017, respectively). When the success criteria of each algorithm were examined, we saw that although sex estimation was significant, (Acc 0.61, Acc 0.60, respectively) they were not very successful. We found that the machine learning algorithms applied to the variables of proximal femur angle parameters gave low accuracy of sex and the effect of both models on sex estimation was similar.

Keywords: Artificial neural networks, Upper femur angles, Sex assessment, Forensic anthropology, Femoral neck-shaft angle, Collo-diaphyseal angle.

INTRODUCTION

Sex identification of human skeletal remains is important in establishing the biological profile of an unidentified individual. Accurately estimating sex of skeletal remains is based on the presence of pelvic bones in forensic and bioarchaeological environments and their well-preserved nature [1]. In cases where pelvic bones cannot be used, methods developed from other skeletal components, including the femur, have been needed for sex assessment. The femur, whose main function is to bear weight and to provide stability to walk, is the longest, heaviest and strongest bone in the human body [2, 3]. Femoral inclination angle (FIA) is important in the evaluation of hip anatomy and useful for operation planning [4]. FIA, participates in the adult hip joint structure and also changes with age [5].

Studies evaluating sex determination in the femur have examined various anatomical structures of the proximal and distal femur [1, 6]. In a study examining FIA in the proximal femur, a difference between both sexes was reported [7]. In another study, although FIA and femoral Alsberg angle (FAA) were evaluated according to age and sex, determination of sex was not assessed [5].

*Address correspondence to this author at the Department of Anatomy, Golkoy Campus, Bolu Abant Izzet Baysal University, Faculty of Medicine, Bolu, Turkiye; Tel: +90374 253 46 56; E-mail: sedasertelmeyvaci@gmail.com

Since plain radiographs are two-dimensional, radiographic assessment of femoral rotation is difficult [8], and accurate measurement of bone geometry is better achieved using 3-dimensional computed tomography (CT) [9].

The aim of this study is to estimate sex with the help of FIA and FAA parameters in the proximal femur using current methods such as machine learning algorithms. However, due to importance of these angular parameters in preoperative planning in the clinical field, our preliminary results present sex-related differences.

MATERIAL AND METHOD

In this study, a total of 4 angular parameters, bilateral FIA and FAA, of the femur of 180 individuals (88 females, 92 males) aged between 20-83 years who applied to the Department of Radiology, Faculty of Medicine, Bolu Abant Izzet Baysal University for various reasons to have an abdominal CT scan and did not have any pathology or previous surgery that would affect the proximal femur morphometry, were evaluated retrospectively. Age information of the cases was determined from the images obtained. Visuals of the measured variables are given in Figure **1**. This study, ethical approval was received from Bolu Abant Izzet Baysal University Clinical Research Ethics Committee Approval with decision number 2022/256.

E-ISSN: 1929-6029/25

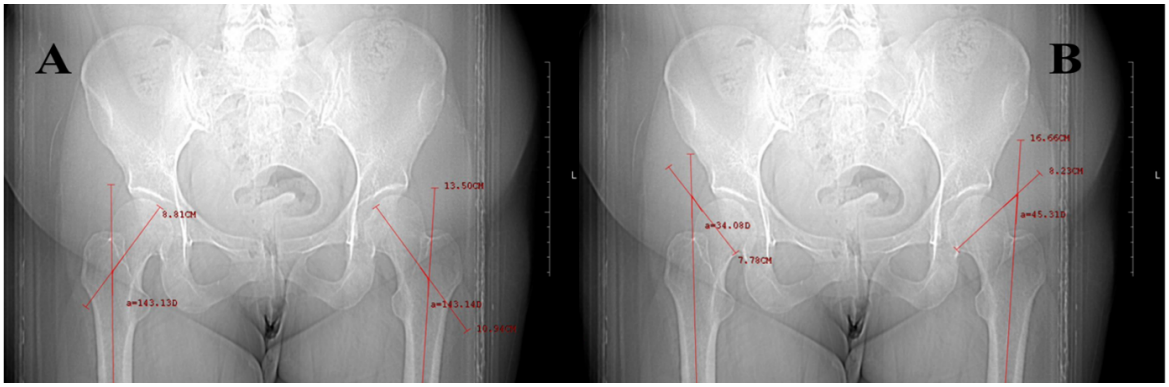

Figure 1: Measurement of proximal femoral angular parameters on CT image.

Definition of Femoral Angular Measurement Parameters and Locations

A. Inclination Angle (FIA)

Angle formed at the intersection of the anatomical axis of the femur with the axis of the collum femoris and the opening facing inward.

B. Alsberg Angle (FAA)

Angle formed at the intersection of the epiphyseal line at the junction of the collum femoris and the caput femoris and the anatomical axis of the femur and the opening facing inward (Figure **1**).

Statistical Analysis

Descriptive values of the measured angles were computed as mean, standard deviation, median and first and third quartiles. Normality assumption was checked with the Shapiro-Wilk test and it was determined that they did not have normal distribution in both women and men. In the first step, the sexes were compared with the Mann-Whitney U test in terms of the measured angles. In the second step, binary logistic regression (LR) and classification and regression tree (CART) algorithms were used to classification of sex. Five-fold cross-validation method was used in the training and testing processes of the models. Model performances were evaluated with area under the ROC curve, precision and recall statistics. We accepted as $p<0.05$ for the statistical significance level and SPSS (ver. 23) program was used in the calculations.

RESULTS

Descriptive statistics of FIA and FAA in the proximal femur according to sex are presented in Table **1**. The average age of males was 52.46±14.02 and the average age of females was 50.98±13.39, and there was no significant difference in age between sexes (p=0.470). When Table **1** is examined, only the right femur inclination angle (RFIA) average was found to be significantly greater in females as a result of univariate evaluation, while there was no significant difference between sexes in terms of the other 3 angle measurements, namely right femur Alsberg angle (RFAA), left femur inclination angle (LFIA), and left femur Alsberg angle (LFAA). According to this result, the 4 angles consisting of the right and left measurements were re-examined with multivariate

Table 1: Difference in Proximal Femur Inclination and Alsberg Angles According to Sex

	SEX												P*
	Male						Female						
	N	Mean	SD	Percentiles			N	Mean	SD	Percentiles			
				25th	Median	75th				25th	Median	75th	
Age	92	52.46	14.015	43.00	55.00	63.75	88	50.98	13.393	42.25	51.0	61.0	0.470
RFIA	92	134.72	5.115	132.00	134.00	138.00	88	136.08	3.671	133.25	136.0	139.0	0.042
LFIA	92	134.99	3.743	132.25	134.50	138.00	88	135.40	3.337	133.0	135.0	138.0	0.441
RFAA	92	50.53	4.573	47.25	50.50	54.00	88	51.34	4.775	47.25	51.5	55.0	0.248
LFAA	92	50.35	4.546	47.00	50.00	54.00	88	50.91	4.570	48.0	51.0	55.0	0.410

*: Mann-Whitney U test.

classification models and it was seen that only RFIA could distinguish sexs significantly. LR and CART results were given from these model results. However, since only the RFIA measurement was found to be significant in sex distinguishing in the models in question, it is seen that the results are not different from the univariate evaluation.

Classification of Males and Females using FIA and FAA Measurements on the Proximal Femur: Results of Machine Learning Algorithms

In our study, the classification success of 2 different machine learning algorithms for sex estimation with the help of FIA and FAA measurements in the proximal femur was comparatively examined. For this purpose, LR and CART were used. The confusion matrix showing the compatibility between the estimated values of the algorithms used in the study and the actual results is given in Table **2**. The criteria presented in Table **3** were evaluated as the criterion of diagnostic success. Only RFIA was included in the LR model from the 4 angle measurements and it was determined that it had a significant effect on sex estimation (p=0.047).

When the diagnostic tree of the CART algorithm is examined, 9 out of 10 individuals with a RFIA of less than 128.5 were male. Of the 99 individuals with a RFIA between 128.5 and 136.5, 54 were male and the rest were female. Of the 71 individuals with a RFIA greater than 136.5, 42 were female. According to this result, the classification success of the CART algorithm is high, especially in cases with a RFIA of less than 128.5 and a RFIA greater than 136.5.

Table **2** shows the correct and incorrect classification of 92 male and 88 female individuals included in the study to their own sex.

The success of the algorithms in distinguishing female sex was defined by sensitivity (Sen) and the success in distinguishing male sex was defined by specificity (Spe). In addition, the general determination success of the algorithms were explained by the area under the ROC curve (AUC) and the sex diagnostic success of the LR model and the CART algorithm was found to be similar (p values 0.014 and 0.017, respectively). In both algorithms, only RFIA was included in the model as in the univariate analysis and they gave similar results.

When the success criteria of the algorithms were examined (Table **3**), although their success in distinguishing sexes is found to be significant, it was not very successful. When the success of the algorithms used in sex estimation was evaluated with the ROC curve, the graph given in Figure **2** was obtained. Each curve represents the success of an algorithm. The AUC values in Table **3** correspond to the area under these curves. In addition, the values on the curve are the cut-off values of the probabilities in sex estimation obtained from the algorithms.

DISCUSSION

Due to the better classification provided by deep learning based on radiological methods, it has attracted significant attention in forensic anthropology. Furthermore, it has been stated that deep learning can

Table 2: Number of Individuals Classified Correctly and Incorrectly according to Two Different Algorithms

		Predicted from Model					
		LR*			CART*		
Actual	Sex	*Male*	*Female*	Total	*Male*	*Female*	Total
	Male	55	37	92	63	29	92
	Female	42	46	88	46	42	88
	Total	97	83	180	109	71	180

*: **LR:** Logistic regression, **CART:** Classification and regression tree.

Table 3: Success Criteria of Algorithms in Sex Estimation

Model	AUC±SE	Precision (PPV)*	Recall (sensitivity)	Specifity	FP*	FN*
LR	0.606±0.042	0.554	0.523	0.598	0.402	0.477
CART	0.603±0.042	0.592	0.477	0.685	0.315	0.592

: FN:** False negatif**; FP:** False pozitif**; PPV:** Pozitive predictive value**;** Pozitive group is "Female***", Negative group is "**Male**".

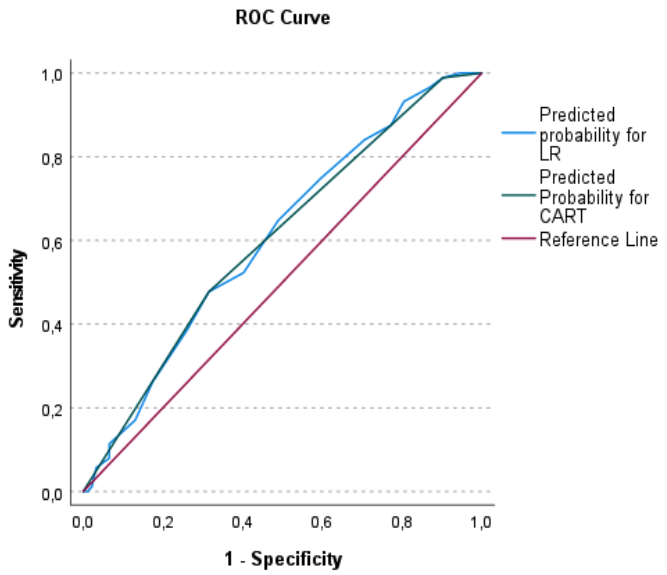

Figure 2: ROC curves of the algorithms.

be applied for fast and accurate sex estimation from skeletal remains [10]. Deep learning method not only improves the diagnosis and treatment of diseases related to the proximal femur in clinical settings, but also contributes to the development of forensic science due to its applicability in sex determination studies [11-13]. Sex estimation from skeletal remains is one of the most important issues in forensic anthropology. Traditionally, the cranium and pelvis bones, which are more sexually dimorphic, are used for accurate sex estimation. The cranium is less reliable than long bones, as several studies have demonstrated that long bones provide better accuracy for sex estimation when the pelvis is unavailable [14]. The fact that the femur, which has been proven to accurately predict sex, has a solid structure also provides an advantage over other bones in the sex determination process, as it will preserve the integrity of this bone [15, 16].

Many researchers have obtained data by measuring various parameters on the femur and used this data in sex determination. In addition to traditional methods used in sex determination, a current method, deep learning detection and classification, was used to examine the proximal femur for sex estimation using digital pelvis radiography. In a study, which presents a computer-aided method, an automatic proximal femur detection based on deep learning (YOLOv5) was performed and it was reported that the accuracy was characterized by the intersection point on the union (IOU). The IOU average was 0.92 and varied between 0.73 and 0.992, indicating that YOLOv5 performed well in sex determination of the proximal femur. The sex estimation accuracy for the Chinese Han population test dataset was found to be 94.6%. The sex estimation accuracy for the white population samples was 82.9%. The accuracy of deep machine learning convolutional neural networks (CNN) tested on the Chinese population was reported to be significantly higher than that tested on the Caucasian population ($p<0.001$) [17].

A comparative study of the performance of supervised learning algorithms in proximal femur sex prediction in contemporary Egyptian and Turkish samples compared nine machine learning algorithms including linear discriminant analysis (LDA), quadratic discriminant analysis (QDA), support vector machine (SVM), Decision Tree (DT), Gaussian process (GPC), Naïve Bayesian (NBC), K-Nearest Neighbor (KNN), Random Forest (RFM) and Adaptive boosting (Adaboost). They reported that the raw models RFM, LDA, QDA and SVM performed best without calibration techniques for posterior probabilities >0.95, while LDA and QDA remained the fastest and simplest classifiers. Random forest (RFM) was found to be the best supervised learning method for sex estimation with an accuracy rate of 93-94% [18].

The femur was evaluated in a study examining linear discriminant analysis (LDA), penalized logistic regression (PLR), random forest (RF), support vector machine (SVM), and artificial neural network (ANN) classification models for sex estimation based on 18 combined measurements of long bones (radius, humerus, femur, and tibia) using machine learning classifiers. In cross-validation and sample testing, it was observed that the accuracy rate of the isolated femur bone for the reference sample varied between 85.4% and 89.4%, and all used algorithms had similar results [16].

In our study, purpose was to estimate sex with machine learning algorithms using the measurement values of FIA and FAA parameters. We found that there was no significant difference between the sexes in terms of 3 angle measurements, LFIA, RFAA, LFAA, and only the RFIA average was found to be significantly larger in women ($p=0.042$).

As a result of our comparative examination of the classification success of 2 different machine learning algorithms used in sex estimation, only the RFIA angle was included in the LR model out of 4 angle measurements and it had a significant effect on sex estimation ($p=0.047$). LR algorithm showed 0.61 Acc,

0.598 Spe, 0.523 Sen and 0.554 PPV values whilst, in the CART algorithm, we found 0.60 Acc, 0.685 Spe, 0.477, Sen and 0.592 PPV values. When the diagnosis tree of this algorithm was examined, it was seen that the classification success was high, especially in cases where RFIA was less than 128.5 and RFIA was greater than 136.5. Although the sex estimation success of the LR model (p=0.014) and the CART algorithm (p=0.017) was found to be similar and significant, it was seen that they could not distinguish sexes very successfully. The 61% accuracy rate obtained in our study is low compared to other studies, which negatively affects its potential to be used alone in the forensic identification process. When the results were compared, the highest accuracy rate was reported as 94.6% [17]. We think that the reason for the different accuracy rates in the results of the studies may be due to the differences in the parameters and algorithms of the femur used, as well as ethnic origin.

In our study, we investigated the potential of angular measurements of the proximal femur in sex estimation in a Turkish population using an up-to-date method, machine learning algorithms, and we found that some angular parameters of the proximal femur were dimorphic in sex estimation, but they did not have sufficient potential to be used alone in sex determination.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

This study was conducted in accordance with the ethical guidelines of Bolu Abant Izzet Baysal University and approved by the Bolu Abant Izzet Baysal University Clinical Research Ethics Committee (Approval Number: 2022/256).

Informed consent includes appropriate statements.

CONSENT FOR PUBLICATION

Not applicable.

FUNDING

No financial support has been received from any institution or organization.

CONFLICT OF INTEREST

No conflict of interest.

REFERENCES

[1] Curate F, Coelho J, Gonçalves D, Coelho C, Ferreira MT, Navega D, Cunha E. A method for sex estimation using the proximal femur. Forensic Sci Int 2016; 266: 579.e1-579.e7. https://doi.org/10.1016/j.forsciint.2016.06.011

[2] Wagner FV, Negrão JR, Campos J, Ward SR, Haghighi P, Trudell DJ, Resnick D. Capsular ligaments of the hip: anatomic, histologic, and positional study in cadaveric specimens with MR arthrography. Radiology 2012; 263(1): 189-98. https://doi.org/10.1148/radiol.12111320

[3] Reynolds A. The fractured femur. Radiol Technol 2013; 84(3): 273-91; quiz pp. 292-4.

[4] Boese CK, Dargel J, Oppermann J, Eysel P, Scheyerer MJ, Bredow J, Lechler P. The femoral neck-shaft angle on plain radiographs: a systematic review. Skeletal Radiol 2016; 45(1): 19-28. https://doi.org/10.1007/s00256-015-2236-z

[5] Sertel Meyvaci S, Kaya Y. Investigation of the femoral inclination and Alsberg angles according to age and gender in adults. Anatomy 2020; 14(3). https://doi.org/10.2399/ana.20.827122

[6] van den Heever DJ, Scheffer C, Erasmus P, Dillon E. Classification of gender and race in the distal femur using self organising maps. Knee 2012; 19(4): 488-492. https://doi.org/10.1016/j.knee.2011.06.009

[7] Tuck SP, Pearce MS, Rawlings DJ, Birrell FN, Parker L, Francis RM. Differences in bone mineral density and geometry in men and women: the Newcastle Thousand Families Study at 50 years old. Br J Radiol 2005; 78: 493-498. https://doi.org/10.1259/bjr/42380498

[8] Blankstein M, Schottel P, Brinker M, Krieg J. Limb Rotation in Lower Extremity Fractures. Instr Course Lect 2024; 73: 901-918.

[9] Carballido-Gamio J, Nicolella DP. Computational anatomy in the study of bone structure. Curr Osteoporos Rep 2013; 11(3): 237-45. https://doi.org/10.1007/s11914-013-0148-1

[10] Cao Y, Ma Y, Vieira DN, Guo Y, Wang Y, Deng K, Chen Y, Zhang J, Qin Z, Chen F, Huang P, Zhang J. A potential method for sex estimation of human skeletons using deep learning and three-dimensional surface scanning. Int J Legal Med 2021; 135(6): 2409-2421. https://doi.org/10.1007/s00414-021-02675-z

[11] Deng Y, Wang L, Zhao C, Tang S, Cheng X, Deng HW, Zhou W. A deep learning-based approach to automatic proximal femur segmentation in quantitative CT images. Med Biol Eng Comput 2022; 60(5): 1417-1429. https://doi.org/10.1007/s11517-022-02529-9

[12] Jang SJ, Alpaugh K, Kunze KN, Li TY, Mayman DJ, Vigdorchik JM, Jerabek SA, Gausden EB, Sculco PK. Deep-Learning Automation of Preoperative Radiographic Parameters Associated With Early Periprosthetic Femur Fracture After Total Hip Arthroplasty. J Arthroplasty 2024; 39(5): 1191-1198.e2. https://doi.org/10.1016/j.arth.2023.11.021

[13] Homma Y, Zhuang X, Ohtsu H, Ishii S, Shirogane Y, Hayashi K, Watari T, Baba T, Ishijima M. Highly accurate acoustical prediction using support vector machine algorithm for post-operative subsidence after cementless total hip arthroplasty. Int Orthop 2023; 47(1): 187-192. https://doi.org/10.1007/s00264-022-05641-5

[14] Spradley MK, Jantz RL. Sex estimation in forensic anthropology: skull versus postcranial elements. Journal of Forensic Sciences 2011; 56(2): 289-296. https://doi.org/10.1111/j.1556-4029.2010.01635.x

[15] Maijanen H, Junno JA, Keisu A, Niinimäki J, Lehenkari P, Oura P. Sex estimation from knee breadth dimensions in a Finnish population. Leg Med (Tokyo) 2021; 51: 101873. https://doi.org/10.1016/j.legalmed.2021.101873

[16] Knecht S, Santos F, Ardagna Y, Alunni V, Adalian P, Nogueira L. Sex estimation from long bones: a machine learning approach. Int J Legal Med 2023; 137(6): 1887-1895. https://doi.org/10.1007/s00414-023-03072-4

[17] Li Y, Niu C, Wang J, Xu Y, Dai H, Xiong T, Yu D, Guo H, Liang W, Deng Z, Lv J, Zhang L. A fully automated sex estimation for proximal femur X-ray images through deep learning detection and classification. Leg Med (Tokyo) 2022; 57: 102056. https://doi.org/10.1016/j.legalmed.2022.102056

[18] Attia MAH, Attia MH, Tarek Farghaly Y, Ahmed El-Sayed Abulnoor B, Curate F. Performance of the supervised learning algorithms in sex estimation of the proximal femur: A comparative study in contemporary Egyptian and Turkish samples. Sci Justice 2022; 62(3): 288-309. https://doi.org/10.1016/j.scijus.2022.03.003