PMAC
PRINCE MAHIDOL AWARD CONFERENCE 2025

E-ISSN 2228-8082
Volume 77, Number 12, December 2025

Siriraj Medical Journal
is deeply sadden by the passing of

**Her Majesty Queen Sirikit
The Queen Mother**

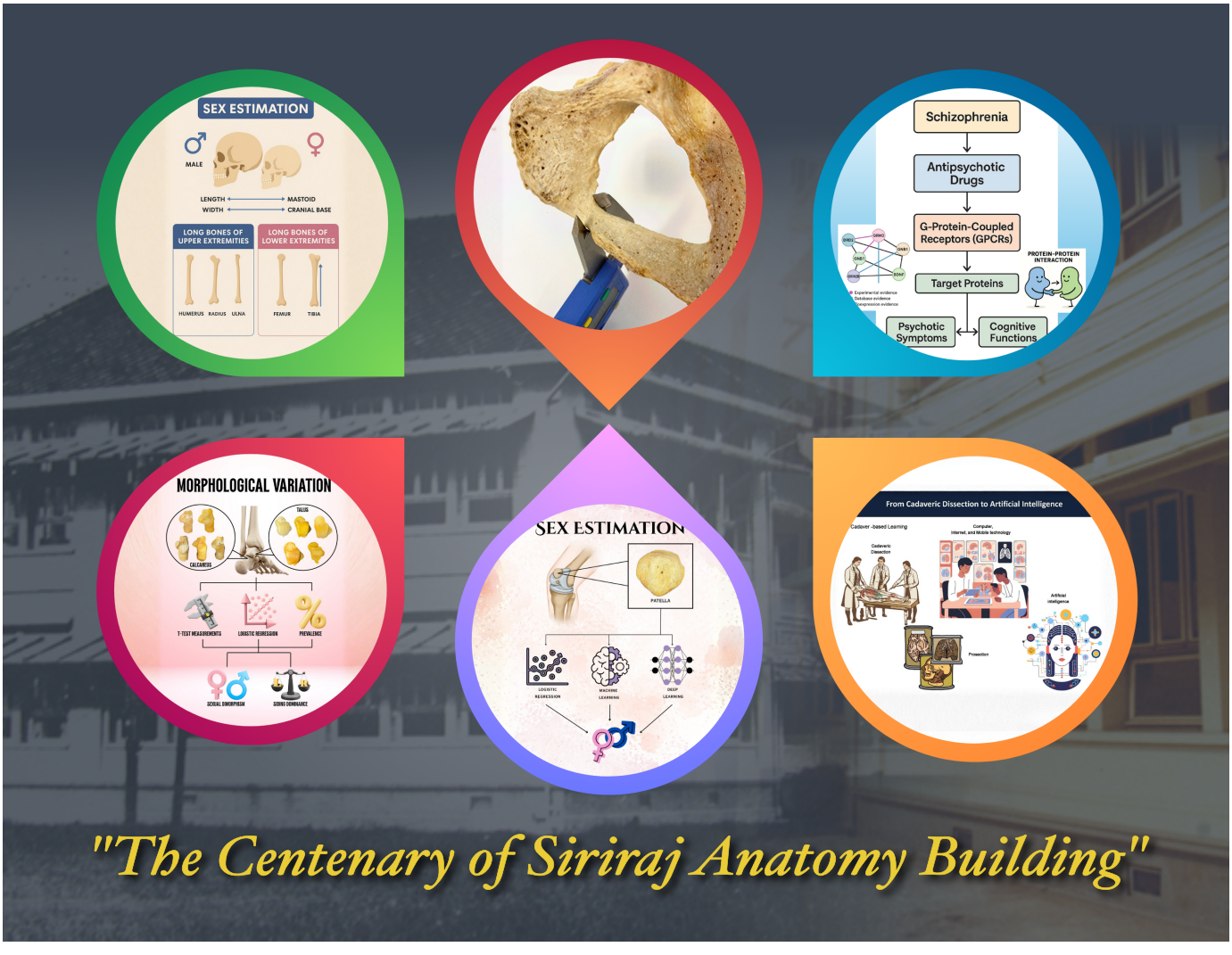
We pay tribute to Her Majesty's
lifelong devotion to the nation
and the people of Thailand

SMJ

Siriraj Medical Journal

The world-leading biomedical science of Thailand

ORIGINAL ARTICLE REVIEW ARTICLE MONTHLY



"The Centenary of Siriraj Anatomy Building"

Indexed by

Scopus® DOAJ EBSCO

https://he02.tci-thaijo.org/index.php/sirirajmedj/index
E-mail: sijournal92@gmail.com

SMJ
Siriraj Medical Journal
The world-leading biomedical science of Thailand
Volume 77 Number 12
December 2025

ORIGINAL ARTICLE

REVIEW ARTICLE

SIRIRAJ MEDICAL JOURNAL

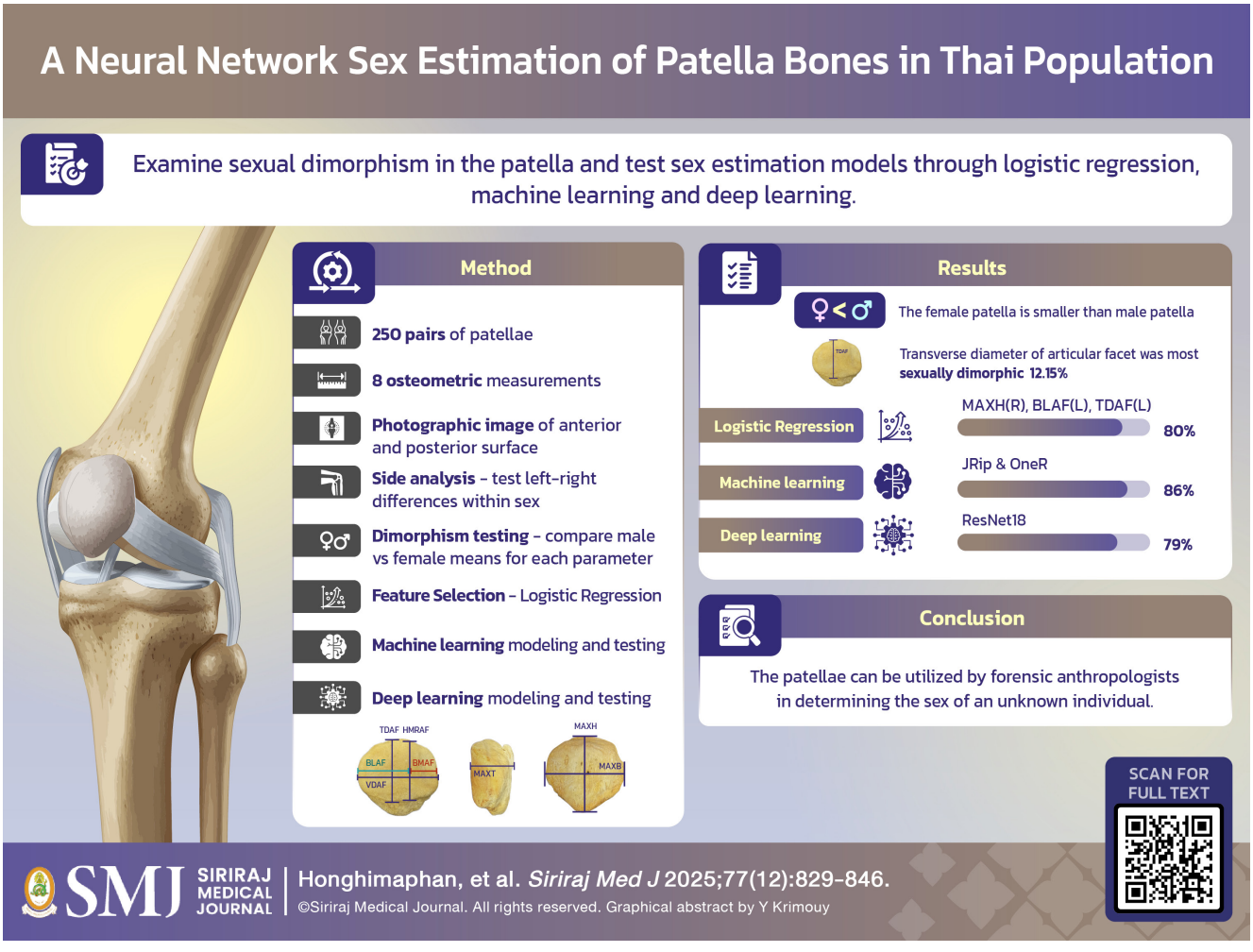
https://he02.tci-thaijo.org/index.php/sirirajmedj/index

Office: His Majesty the King's 80th Birthday Anniversary 5th December 2007 Building (SIMR), 2nd Fl., Room No.207
Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok 10700, Thailand
E-mail: sijournal92@gmail.com

The Efficacy of Deep Learning Model on Sex Estimation in Patellae for A Thai Population

Saranya Honghimaphan, B.Sc.[1], Patara Rattanachet, Ph.D.[1], Natipong Chatthai, M.Sc.[1], Parawee Jitrabeab, M.Sc.[1], Sittiporn Ruengdit, Ph.D.[2], Robert W.Mann, Ph.D.[3], Napakorn Sangchay, M.D., Ph.D.[1,*]

[1]Siriraj Anatomical and Anthropological Bone Research Centre (Virapan Davivongs), Department of Anatomy, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand, [2]Department of Forensic Medicine, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand, [3]John A. Burns School of Medicine, University of Hawai'i, Honolulu, Hawai'i, USA.



*Corresponding author: Napakorn Sangchay
E-mail: napakorn.sac@mahidol.ac.th
Received 13 August 2025 Revised 10 October 2025 Accepted 18 October 2025
ORCID ID:http://orcid.org/0000-0001-9250-3702
https://doi.org/10.33192/smj.v77i12.277048

ABSTRACT

Objective: This study investigates the sexual dimorphism in the patellae to develop sex estimation equations and determine the accuracy using machine learning (ML) and deep learning (DL) classification models.

Materials and Methods: The sample of 250 pairs of patellae from the Siriraj Anatomical and Anthropological Bone Research Center (Si-AABRC) were measured for eight parameters. The data were statistically analysed using logistic regression model alongside, ML and DL were used to predict the best classifiers in sex classification. Rather than traditional radiographic images, this paper tries a novel integration of photographic images.

Results: The average values for each parameter were significantly larger in males than females ($p < 0.05$), suggesting the presence of sexual dimorphism within the patellae between each sex. The most dimorphic parameter was Transverse Diameter of Articular Facet (TDAF). The parameters in females showed no significant difference between left and right except for Breadth of the Medial Articular Facet (BMAF). However, in males a significant difference was observed for Maximum Height (MAXH), Transverse Diameter of Articular Facet (TDAF) and Breadth of the Lateral Articular Facet (BLAF). The logistic regression equation generated included the following parameters: MAXH (R), BLAF (L), and TDAF (L). The overall accuracy obtained for different sex estimation models ranged from 80%, 80% to 86% and 49.7% to 79.2% using logistic regression, ML and DL, respectively.

Conclusion: The patellae can be utilized by forensic anthropologists in determining the sex of an unknown individual.

Keywords: Patellae; sexual dimorphism; logistic regression; machine learning; deep learning; Thai population (Siriraj Med J 2025; 77: 829-846)

INTRODUCTION

Determining the identity of an unknown deceased is critical during a medicolegal investigation.[1] A medical examiner and a forensic anthropologist work within the legal system to provide evidence and facilitate the identification of the deceased. This process ensures that justice is served to the individual and their family by resolving ambiguous questions, providing closure, and supplying accurate information for litigation purposes.[2]

In many instances, especially cold cases, the remains discovered are skeletonized. Hence, forensic anthropologists analyze the skeletal remains to establish a biological profile. The profile comprises sex, age, stature, and ancestry.[3] Sex estimation is the initial step in the biological identification of the unknown as sex influences the other three parameters.[4] Moreover, determining the sex minimizes the pool of candidates significantly by 50%.[5]

The human skeletal system is composed of several types of bones that can be categorized into long bones, flat bones, short bones, irregular bones and sesamoid bones. However, of all the bones in the human body, the bone that serve as the primary choice for anthropological sex estimation are a combination of irregular and flat bones; the pelvic (innominate) bone and the skull. The pelvic bone is primarily chosen for sex estimation as distinct morphological features can be witnessed between the sexes.[6,7] Furthermore, the pelvis has been known to exhibit an accuracy of up to 95% in determining the sex.[8] Nonetheless, in the absence of the pelvic bone, the skull serves as the secondary choice in sex identification. This is because the accuracy has been observed to be approximately 92%.[9] However, when unknown sets of remain are discovered, it is not always that a full set of skeletons is available. The incomplete nature of remains can be attributed to reasons such as animal scavenging, trauma, and the efficiency in retrieving and recovering the remains. Thus, in the context of forensic anthropology, different skeletal parts and bones are required to develop different methodologies in assessing the sex. One such bone is the patella.

The patella is the biggest sesamoid bone that is triangular in shape and possesses three borders – superior, medial and lateral, two surfaces – anterior and posterior, a base, and an apex.[10] The patella is located within the patellofemoral groove of a femur and is embedded and attached to the quadricep femoris tendon and the patellar ligament.[11] The posterior surface of the patella articulates with the femur forming an anterior articular surface of the knee joint that functions in protecting against trauma.[2] Furthermore, the patella is also involved in the extensor mechanism of the quadricep muscle by acting as a fulcrum to increase the efficiency in the extension and flexion of the knee joint.[12] Therefore, due to its morphology, location and function, the patella has been regarded as highly unsusceptible to post-mortem or taphonomy changes.[13]

In the past, patellae have been used in forensic cases to help identify unknown individuals. For example, in a 2008 case from Goias, Brazil,[14] an unidentifiable decomposed body was discovered in a forest. Due to the severe state of decomposition, a fingerprint comparison was impossible. Alternatively, a positive identification was possible by comparing the patellae retrieved post-mortem to a radiographic image taken antemortem during knee surgery. The patella proved to be useful as the prior left knee surgery involving metallic surgical wires provided the evidence in the form of two channels in the patella that had exact dimensions to the surgical wires.[14] Hence, this case showed that when a body is recovered, the conditions are not always favourable for identification. Several different parts of the skeleton should be used to determine the sex. In this instance, the patella has a location and position that allows it to resist post-mortem changes and is a useful tool in identifying the unknown. Thus, illustrates the importance of utilizing images to aid in distinguishing sex and identifying the individual

Previously, numerous studies have assessed the sexual dimorphism of the patella. However, parameters regarded as the best predictors in sex estimation for both cranial and post cranial elements vary across populations, highlighting the importance of sexual dimorphism for each bone in specific population groups. While some studies have focused on morphometric measurments, others have incorporated machine learning for sex estimation.[7,12] Nonetheless, many of these studies are limited by unequal[15,16] or small sample sizes[17,18], focus only on one side of the patella[17] or restrict their method by primarily relying on discriminant function analysis for sex estimation.[19] More recently, Cavlak et al[20], have applied the framework of convolutional neural network to the MRI slices of patella, demonstrating the use of deep learning in sex estimation. As a result, building on these knowledge and foundation, this study aims to investigate sexual dimorphism using two different approaches; morphometric measurement inspected through logistic regression and machine learning, and photographic images evaluated through deep learning. In addition to addressing the comparative analysis, this studies aims to address limitation by using equal sample size and incorporating additional parameters. The contribution of this research is also in extending the image based analysis using photographs for a specific population. Therefore, it allows for a robust research by comparing traditional and advanced classification tools. This is vital for forensic anthropologists to determine the best method for estimating sex from the patella.

MATERIALS AND METHODS

Sample selection

This study analysed a total of 250 pairs of patellae in the Siriraj Anatomical and Anthropological Research Center's (Si-AABRC) bone collection within the Department of Anatomy, Faculty of Medicine at Siriraj Hospital, Mahidol University. The sample consisted of an equal number of pairs of each sex, with an average age of 51 years. As both the left and right sides of the patella were used for the analysis, any remains without both patellae were excluded. Furthermore, patellae presented with fracture, erosion, trauma and extreme osteophytic formation were also excluded.

Parametric measurement

In this study, eight parameters in the patella were measured using a digital vernier calliper. The parameters were developed and adapted from the study conducted by Dayal & Bidmos[19,21] assessing the sexual dimorphism of the patella in the South African population as well as the study assessing the sexual dimorphism of the patella in the Southern Part of Andhra Pradesh.[17] The parameters shown in Fig 1. include maximum height (MAXH), maximum breadth (MAXB), maximum thickness (MAXT), height of the median ridge of the articular facet (HMRAF), breadth of lateral articular facet (BLAF), breadth of medial articular facet (BMAF), transverse diameter of articular facet (TDAF), and vertical diameter of the articular facet (VDAF). Detailed descriptions of the parameters are presented in Table 1. A patellar index (PI) was calculated to obtain a ratio between the MAXH and MAXL using the formula described below.

$$\frac{MAXH}{MAXB} \times 100$$

Statistical analysis

The statistical analysis for the osteometric assessment of sexual dimorphism was conducted using Jamovi Version 2.5 - Sydney, Australia.[22] The level of significance for each test was $\alpha = 0.05$. Descriptive statistics comprising mean and standard deviation for all the parameters were calculated. The sample data were subjected to a Shapiro-Wilk normality test. If the data did not violate the assumption of normality, the difference between the left and right side of the patellae within each sex was assessed using independent t-test to determine the sexual dimorphism. However, if the data violated the aforementioned assumption, Welch's *t*-test and Mann–Whitney U test were used, respectively. The percentage of sexual dimorphism was calculated using the formula

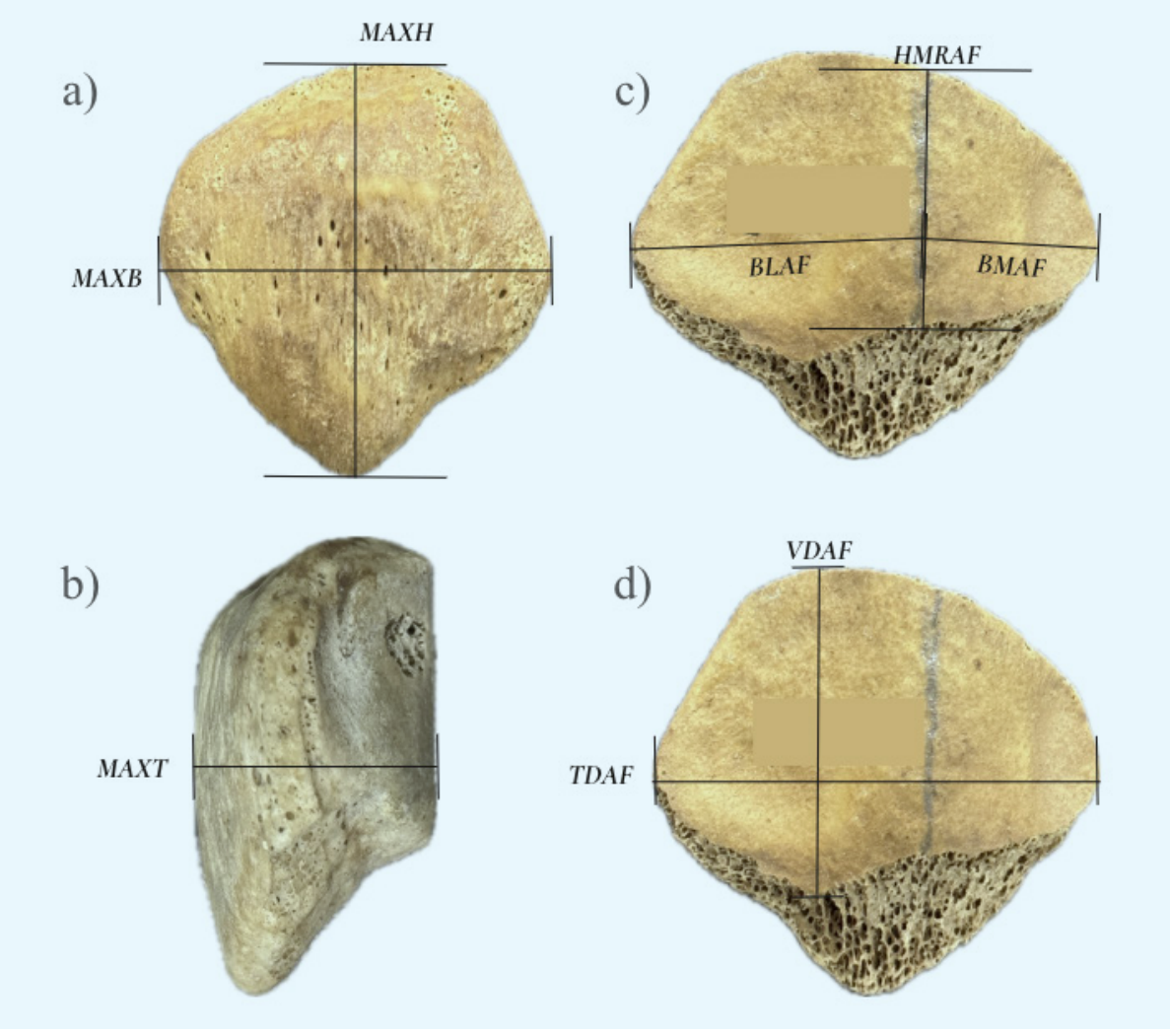


Fig 1. a) Patella measurements for the eight parameters. MAXH - Maximum Height, MAXB - Maximum Breadth, b) MAXT - Maximum Thickness, c) HMRAF - Height of median ridge of the articular facet, BLAF - Breadth of lateral articular facet, BMAF - Breadth of medial articular facet, d) VDAF - Vertical diameter of the articular facet, TDAF - Transverse diameter of the articular facet.

obtained from a study on sexual dimorphism of the root lengths in Tamil population.[23]

$$\frac{\overline{x}_{male} - \overline{x}_{female}}{\overline{x}_{male}} \times 100$$

To assess the difference between the left and the right side of the patellae within each sex, the Wilcoxon Signed Rank (for data violating the normality) and paired t-test (for data following normal distribution) were used. Moreover, the mean difference was also calculated. With the presence of multiple parameters being analyzed, a Bonferroni and Benjamini–Hochberg test was also conducted and reported.

Reliability analysis by interobserver and intraobserver

In determining the reliability of the parameters, 20% of the samples (25 pairs of males and 25 pairs of females) were randomly selected for interobserver and intraobserver evaluations. The intraobserver and interobserver measurements were collected two weeks apart. The Intraclass Correlation Coefficient (ICC) was calculated to discover the degree of correlation and reliability between two observers.[24]

Following the guidelines by Koo and Yi (2016)[25] and Shrout and Fleiss (1979)[26], intraobserver reliability was assessed using ICC(3,1), a two-way mixed-effects model and interobserver reliability was analyzed using (ICC(2,k), a two-way random-effects model. Both the upper and lower confidence intervals was reported. The ICC score were interpreted as follows: ICC < 0.50 is considered poor, 0.50 – 0.75 is moderate, 0.75 – 0.90 is good and >0.90 is considered excellent.

Logistic regression analysis

In the next step, a logistic regression equation was

TABLE 1. Description of the parameters measured on the patellae.

Variable	Acronym	Description
Maximum Height	MAXH	The maximum linear distance between the superior point of the apex to the inferior point at the base of the patellae
Maximum Breadth	MAXB	The maximum linear distance between the medial border to the lateral border of the patellae
Maximum Thickness	MAXT	The maximum linear distance between the anterior surface and the posterior surface of the patellae.
Height of the Median Ridge of the Articular Facet	HMRAF	Distance between the superior point of the median ridge to the inferior point of the median ridge on the posterior surface of the patella
Breadth of Medial Articular Facet	BMAF	Distance between the median ridge and the medial border of the articular facet on the posterior aspect of the patella.
Breadth of Lateral Articular Facet	BLAF	Distance between the median ridge and the lateral border of the articular facet on the posterior aspect of the patella.
Transverse Diameter of Articular Facet	TDAF	Maximum linear distance between the medial border of the articular facet and lateral border of the articular facet on the posterior aspect of the patella
Vertical Diameter of Articular Facet	VDAF	Maximum linear distance between the most superior point of the articular facet and the most inferior point of the articular facet on the posterior aspect of the patella

formulated using R Studio (Version 2024.4.2.764.1).[27] The model was developed using the generalized linear model function – glm(). The output of this function represents a constant (a), a coefficient (b) and the measurement value of the predicting parameter (x) in the form of $a + b \times x$.[28] This model is then evaluated for the log of odd ratio (logit) and a probability of classification equation (P).

$$\frac{1}{1+e^{-(a+b1.x1+b2.x2.....)}}$$

The equation for the probability of classification predicts the outcome to be either male or female, where $P > 0.5$ classifies the measurement to be male and $P \leq 0.5$ classifies the measurement to be female.[29] A logit equation for each parameter and each side was developed. However, as there are multiple predictors, a stepwise model selection function was performed to obtain the parameter that provides the best efficacy for sex estimation based on an AIC, which is a computational value that depicts the efficacy of the model.[28] The model that presented with the lowest AIC and highest Area Under the Curve (AUC) was selected. A ROC Curve as well as a Calibration plot was generated. For the model that provides the best efficacy, a Hosmer–Lemeshow χ^2 alongside p-value and a Brier score was calculated. The model was then tested using a blind test where 100 additional pairs of the patellae were measured and the values were put in the model to determine the accuracy.

Machine learning

Machine Learning analyses were performed using MATLAB version 9.14 - 2023a[30] and Waikato Environment for Knowledge Analysis software (WEKA 3.8.6).[31] To prevent the subject-level leakage, each row in the dataset represented one subject containing the measurement of both left and right patella.

For analysis in WEKA, 250 subject pairs were randomly assigned to 10 groups, as the models were evaluated using 10-fold cross validation. During the training phase, the models were cross-validated using a separate test dataset selected through the 'Supplied Test Set' option. In each fold models were trained on 90% of the dataset and validated on the 10% of the dataset. The mean and standard deviation of the performance metrics were calculated across 10 folds.

For analysis in MATLAB, the response variable was set at "Sex". The model training was then conducted

using the Classification Learner application with 10 fold cross validation. The mean and standard deviation of the performance metrics were reported for models exhibiting more than 80% of accuracy.

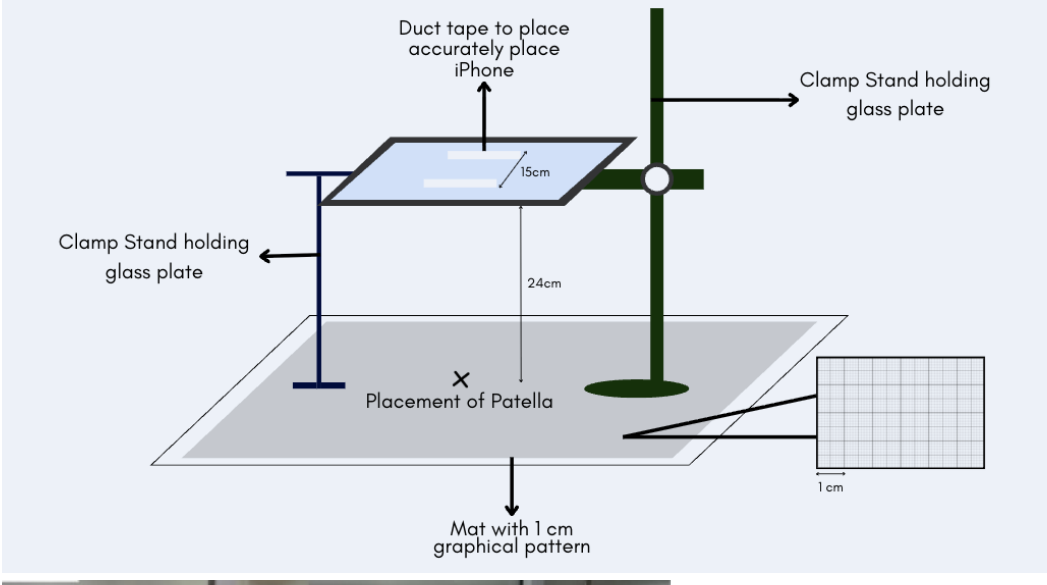
Deep learning

For the deep learning approach, 1000 images [left, right, anterior and posterior] were photographed using an iPhone 15 Pro at 3x magnification placed on an instrument (Fig 2) at a height of 24 cm. Each patella was secured with a clay to ensure a levelled orientation allowing the anterior surface and the posterior surface of the patella to be photographed (Fig 3). The deep learning assessment was done to only classify using anterior surface and only classify using posterior surface

Images in the dataset of the left and right patella from the same subject were assigned exclusively to either training set ("Training → Male/Female") or the validation test ("Test → Male/Female") for both anterior images and posterior images, ensuring a subject-level split. Pretrained models like Visual Geometry Group (VGG19), MobileNet V2 and Deep Residual Learning (ResNet18) was utilized. The input images were resized to 224 pixels x 224 pixels with single precision and processed to 3 channel RGB. An augmentation was applied including random rotation (±10°), translation (±10 pixels), and scaling (0.9–1.1) in both axes.

Following this a 5-fold cross validation was utilized on the training dataset where each were set for 20 epochs, initial learning rate = 1×10^{-4} and batch size of 32. The sample was evaluated on a test set. Performance metrics was evaluated; accuracy, sensitivity and specificity.

a)



b)

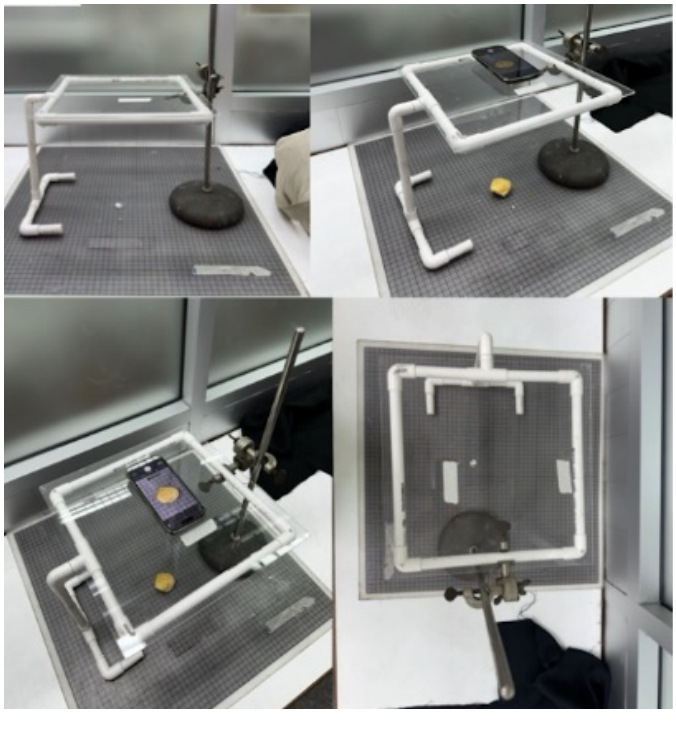


Fig 2. a) A detailed schematic created using the Canva software showcasing the instrument used in taking images of the patellae. b) Images showcasing the set-up used for photographing the patella.

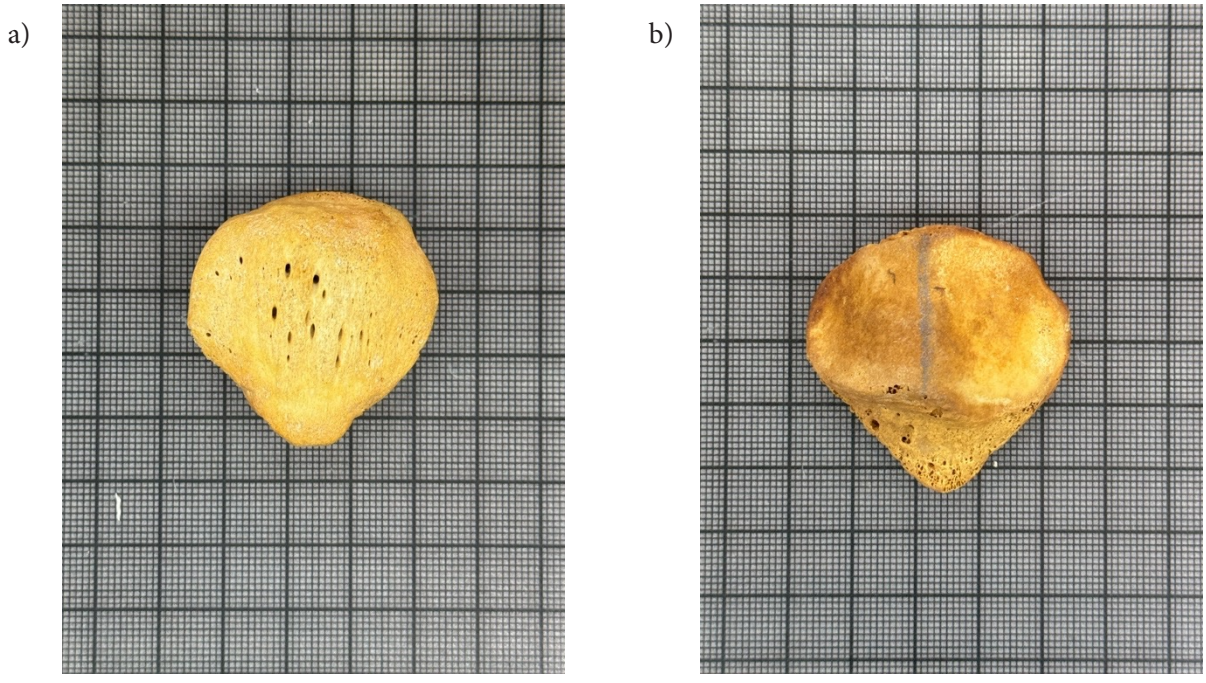


Fig 3. a) Representative image of the anterior view of the patella. b) Representative image of the posterior view of the patella

RESULTS

Sexual dimorphism assessment and descriptive statistics

Descriptive statistics consisting of mean and standard deviation (SD) were determined for all eight parameters for both sexes and sides. Comparing each parameter between males and females using parametric and non-parametric test, a $p < 0.001$ which is less than $\alpha = 0.05$ was obtained. Seen in Table 2, the percentage of sexual dimorphism for each parameter is positive. The highest percentage and lowest percentage of sexual dimorphism were observed in TDAF (Left) at 12.15% and HMRAF (Right) at 6.72% respectively. Furthermore, the average PI on the left side of male was 95.97 and female was 95.82. The average patellar PI on the right side of both male and female were slightly smaller at 95.23 and 95.42 respectively. This is concurrent with the Bonferroni and Benjamini–Hochberg test where all parameters were significant. ($q = 0.002$) (Table 3).

Difference in patellae side assessment

Seen in Table 3, the results assessing the difference in the patellar side for each parameter for both males and females are summarized. All the parameters in females only the BMAF had a significant difference ($p < 0.001$), whereas in males MAXH ($p < 0.001$), TDAF ($p = 0.041$), and BLAF ($p = 0.012$) had a significant difference. However, from performing the Bonferroni and Benjamini–Hochberg only MAXH ($q = 0.002$) and BLAF ($q = 0.020$) showed significant difference within males and for females only BMAF ($q = 0.002$) (Table 3).

Reliability analysis by interobserver and intraobserver

The ICC values for intraobserver and interobserver measurements on 50 randomly selected samples are summarized in Table 4. The intraobserver values ranged from 0.842 to 0.997 and the interobserver values ranged from 0.502 to 0.995.

Regarding the both the reliability analysis, it was found that the highest ICC value belonged to the MAXB in both males and females. For the intraobserver value, the highest ICC was 0.995 for male on the left side and 0.997 for female on the right side. For the interobserver value, the highest ICC was 0.995 for male on the right side and 0.994 for both sides in female.

However, the parameter that proved to show the lowest reliability in interobserver measurement was the VDAF for both male and female as well as for both sides with values ranging from 0.502 to 0.706. The parameters that showcased the highest reliability for both male and female include MAXH, MAXB, and MAXT.

Logistic regression

The logistic regression analysis as seen in Table 5 indicates TDAF–transverse diameter of the articular facet on the left side as the most effective parameter in sex determination. This is because the model for TDAF produced the lowest AIC score of 212.20 and the highest AUC of 0.8941 (Table 5). Additionally, the ROC curve presented in Fig 5 visually demonstrates differences between the parameters, highlighting TDAF on the left side taking precedence in sex estimation.

TABLE 2. Descriptive statistics and percentage of sexual dimorphism of each parameter for both the sexes and both the sides for the patellae.

Parameters and Side	Male (n=125) Mean	Male (n=125) SD	Female (n=125) Mean	Female (n=125) SD	% Sexual Dimorphism	P Value	Benjamini Hochberg P (All)	Benjamini Hochberg P (Within)	Bonferroni P (All)	Bonferroni P (Within)
MAXH (L)[a]	42.2	3.41	37.3	2.67	11.61%	<0.001	0.00178	0.001	0.032	0.008
MAXH (R)[a]	41.7	3.27	37.2	2.61	10.79%	<0.001	0.00178	0.001	0.032	0.008
MAXB (L)	44.0	3.18	39.0	3.01	11.36%	<0.001	0.00178	0.001	0.032	0.008
MAXB (R)	43.9	3.20	39.1	3.09	10.93%	<0.001	0.00178	0.001	0.032	0.008
MAXT (L)[b]	20.8	1.80	18.7	1.49	10.10%	<0.001	0.00178	0.001	0.032	0.008
MAXT (R)[b]	20.7	1.80	18.6	1.49	10.14%	<0.001	0.00178	0.001	0.032	0.008
HMRAF (L)	26.7	2.91	24.8	2.37	7.12%	<0.001	0.00178	0.001	0.032	0.008
HMRAF (R)	26.8	2.79	25.0	2.41	6.72%	<0.001	0.00178	0.001	0.032	0.008
BMAF (L)[b]	20.4	2.15	18.1	1.85	11.27%	<0.001	0.00178	0.001	0.032	0.008
BMAF (R)	20.6	1.89	18.6	1.88	9.71%	<0.001	0.00178	0.001	0.032	0.008
BLAF (L)	27.0	2.04	23.9	1.90	11.48%	<0.001	0.00178	0.001	0.032	0.008
BLAF (R)[b]	26.7	2.17	23.8	1.96	10.86%	<0.001	0.00178	0.001	0.032	0.008
TDAF (L)	42.8	3.09	37.6	2.97	12.15%	<0.001	0.00178	0.001	0.032	0.008
TDAF (R)	42.5	3.14	37.6	2.97	11.53%	<0.001	0.00178	0.001	0.032	0.008
VDAF (L)[b]	31.6	2.74	29.1	2.46	7.91%	<0.001	0.00178	0.001	0.032	0.008
VDAF (R)	31.5	2.35	29.0	2.27	7.94%	<0.001	0.00178	0.001	0.032	0.008

n = Sample size (pair of patellae)
a = Levene's test (Homogeneity of variance test): violation of the assumption of equal variances. Welch's t-test.
b = Shapiro Wilk (Normality Test): violation of the assumption of normality. Mann-Whitney U t-test.
Abbreviations: MAXH = Maximum Height, MAXB = Maximum Breadth, MAXT = Maximum Thickness, HMRAF = Height of the Median Ridge of the Articular Facet, BMAF = Breadth of Medial Articular Facet, BLAF = Breadth of Lateral Articular Facet, TDAF = Transverse Diameter of Articular Facet, VDAF = Vertical Diameter of Articular Facet.

TABLE 3. Comparison between the left and right sides for each parameter, separately for males and females.

Parameters	Male (n=125) Mean Difference	Male (n=125) P Value	Male (n=125) Benjamini Hochberg P (All)	Male (n=125) Benjamini Hochberg P (Within)	Male (n=125) Bonferroni P (All)	Male (n=125) Bonferroni P (Within)	Female (n=125) Mean Difference	Female (n=125) P Value	Female (n=125) Benjamini Hochberg P (All)	Female (n=125) Benjamini Hochberg P (Within)	Female (n=125) Bonferroni P (All)	Female (n=125) Bonferroni P (Within)
MAXH (L&R)	0.5100[a]	<0.001[a]	0.002	0.008	0.032	0.008	0.1128	0.207	0.276	0.552	1.000	1.000
MAXB (L&R)	0.1329	0.24	0.307	0.320	1.000	1.000	-0.07	0.862	0.890	0.862	1.000	1.000
MAXT (L&R)	0.0400[a]	0.400[a]	0.457	0.457	1.000	1.000	0.01	0.591	0.652	0.788	1.000	1.000
HMRAF(L&R)	-0.2000[a]	0.064[a]	0.098	0.128	1.000	0.512	-0.1400[a]	0.113[a]	0.157	0.452	1.000	0.904
BMAF (L&R)	-0.1700[a]	0.098[a]	0.143	0.157	1.000	0.784	-0.4501[a]	<0.001[a]	0.002	0.008	0.032	0.008
BLAF (L&R)	0.2650[a]	0.012[a]	0.020	0.048	0.384	0.096	0.0549[a]	0.388[a]	0.457	0.621	1.000	1.000
TDAF (L&R)	0.2600[a]	0.041[a]	0.066	0.109	1.000	0.328	-0.0450[a]	0.824[a]	0.879	0.862	1.000	1.000
VDAF (L&R)	0.0150[a]	0.890[a]	0.890	0.890	1.000	1.000	0.0500[a]	0.333[a]	0.410	0.621	1.000	1.000

L = Left and R = Right.
n = Sample size (pair of patellae)
a = Shapiro Wilk (Normality Test): violation of the assumption of normality. Wilcoxon signed-rank test
Abbreviations: MAXH = Maximum Height, MAXB = Maximum Breadth, MAXT = Maximum Thickness, HMRAF = Height of the Median Ridge of the Articular Facet, BMAF = Breadth of Medial Articular Facet, BLAF = Breadth of Lateral Articular Facet, TDAF = Transverse Diameter of Articular Facet, VDAF = Vertical Diameter of Articular Facet.

TABLE 4. Intraclass Correlation Coefficient (ICC) and descriptive analysis for the Intraobserver and Interobserver measurement.

Parameters	Side	Male					Female				
		Mean	SD	ICC[a]	Lower CI	Upper CI	Mean	SD	ICC[a]	Lower CI	Upper CI
Intraobserver (n=25) ICC(3,1) [95% CI]											
MAXH	L	42.5	4.31	0.981	0.962	0.990	37.1	2.17	0.994	0.989	0.997
	R	42.1	4.41	0.984	0.969	0.992	37.3	2.13	0.993	0.986	0.996
MAXB	L	43.8	3.92	**0.995**	0.990	0.998	39.0	2.89	0.996	0.993	0.998
	R	43.9	3.97	0.993	0.987	0.997	39.3	2.73	**0.997**	0.994	0.998
MAXT	L	20.3	1.93	0.986	0.973	0.993	18.5	1.30	0.982	0.965	0.991
	R	20.1	1.77	0.992	0.984	0.996	18.6	1.28	0.957	0.917	0.978
HMRAF	L	26.6	3.22	0.916	0.840	0.957	24.8	2.24	0.873	0.766	0.933
	R	27.2	2.94	0.951	0.906	0.975	24.7	2.42	0.903	0.819	0.949
BMAF	L	20.8	1.77	0.861	0.741	0.927	18.3	1.78	0.894	0.802	0.944
	R	20.9	1.83	0.906	0.821	0.951	19.1	1.58	0.852	0.729	0.921
BLAF	L	26.5	2.26	0.927	0.859	0.962	23.9	1.77	0.863	0.749	0.928
	R	26.4	2.78	0.960	0.922	0.979	23.8	1.79	0.842	0.713	0.916
TDAF	L	42.7	3.74	0.982	0.965	0.991	37.6	2.62	0.973	0.948	0.986
	R	42.5	3.69	0.980	0.961	0.990	37.7	2.36	0.938	0.881	0.968
VDAF	L	31.0	2.65	0.967	0.935	0.983	29.2	1.89	0.965	0.933	0.982
	R	31.7	2.58	0.969	0.940	0.984	29.1	1.98	0.967	0.937	0.983
Interobserver (n=25) ICC(2,k) [95% CI]											
MAXH	L	42.0	4.36	0.984	0.967	0.992	36.4	2.16	0.948	0.605	0.983
	R	41.3	4.63	0.984	0.955	0.993	36.4	2.22	0.926	0.391	0.976
MAXB	L	43.5	3.76	0.992	0.977	0.996	38.8	2.85	**0.994**	0.981	0.997
	R	43.7	3.90	**0.995**	0.988	0.998	39.1	2.60	**0.994**	0.973	0.998
MAXT	L	19.9	1.91	0.973	0.830	0.991	18.3	1.31	0.980	0.895	0.992
	R	19.8	1.81	0.970	0.890	0.988	18.4	1.38	0.956	0.786	0.983
HMRAF	L	25.3	3.46	0.867	0.462	0.949	23.6	2.05	0.781	0.128	0.917
	R	25.4	3.26	0.876	0.036	0.963	23.6	2.29	0.786	0.220	0.915
BMAF	L	19.6	1.91	0.874	0.499	0.951	17.4	1.75	0.898	0.406	0.964
	R	19.5	1.71	0.791	0.158	0.921	18.2	1.57	0.820	0.383	0.927
BLAF	L	25.9	2.54	0.911	0.588	0.967	23.6	1.92	0.889	0.472	0.959
	R	25.8	2.73	0.949	0.563	0.984	23.6	1.91	0.911	0.819	0.955
TDAF	L	41.4	3.70	0.951	0.571	0.984	36.5	2.57	0.910	0.110	0.973
	R	41.0	3.49	0.923	0.262	0.976	37.1	2.69	0.931	0.411	0.978
VDAF	L	28.5	2.35	0.706	-0.112	0.898	26.6	1.90	0.502	-0.142	0.791
	R	28.8	2.23	0.665	-0.125	0.879	26.6	2.0	0.586	-0.140	0.839

a = Intraclass Correlation Coefficient for the comparison between the intraobserver measurement and original as well as, interobserver measurement and original

Abbreviations: MAXH = Maximum Height, MAXB = Maximum Breadth, MAXT = Maximum Thickness, HMRAF = Height of the Median Ridge of the Articular Facet, BMAF = Breadth of Medial Articular Facet, BLAF = Breadth of Lateral Articular Facet, TDAF = Transverse Diameter of Articular Facet, VDAF = Vertical Diameter of Articular Facet.

TABLE 5. Logistics Regression Model for single and combination of parameters for sex estimation.

Single Logistic Regression					
Parameters	**Side**	**Logit**	**AIC**	**Odds Ratio (95% CI)**	**AUC**
MAXH	L	-19.993 + 0.505 ×MAXHL	230.67	1.66 (1.48 – 1.89)	0.8652
	R	-20.120 + 0.512×MAXHR	236.92	1.67 (1.48 – 1.91)	0.8607
MAXB	L	-21.415 + 0.516×MAXBL	227.33	1.68 (1.49 – 1.92)	0.8786
	R	-19.701 + 0.475×MAXBR	237.55	1.61 (1.44 – 1.83)	0.8634
MAXT	L	-15.894 + 0.809×MAXTL	263.66	2.25 (1.83 – 2.83)	0.8210
	R	-15.497 + 0.791×MAXTR	267.27	2.21 (1.80 – 2.77)	0.8190
HMRAF	L	-7.190 + 0.280×HMRAFL	319.27	1.32 (1.19 – 1.48)	0.6977
	R	-7.030 + 0.272×HMRAFR	321.28	1.31 (1.18 – 1.47)	0.6921
BMAF	L	-11.445 + 0.594×BMAFL	276.02	1.81 (1.55 – 2.16)	0.8062
	R	-11.089 + 0.566×BMAFR	286.98	1.76 (1.51 – 2.09)	0.7835
BLAF	L	-20.209 + 0.795×BLAFL	232.64	2.22 (1.84 – 2.74)	0.8709
	R	-17.549 + 0.697×BLAFR	250.69	2.01 (1.70 – 2.43)	0.8458
TDAF	L	-22.917 + 0.571×TDAFL	212.20	1.77 (1.56 – 2.05)	0.8941
	R	-20.223 + 0.505×TDAFR	229.50	1.66 (1.48 – 1.89)	0.8767
VDAF	L	-12.600 + 0.416×VDAFL	296.40	1.52 (1.34 – 1.74)	0.7694
	R	-13.677 + 0.453×VDAFR	288.44	1.57 (1.39 – 1.81)	0.7784

Best Logistic Regression Model						
Predictor	**Coefficient**	**Std. Error**	**P-value**	**Odds Ratio (95% CI)**	**AIC**	**AUC**
Intercept	-26.87	3.30	<0.001	2.14×10^{-12} (1.92×10^{15} – 8.50×10^{10})	205	0.907 (95% CI: 0.8685-0.945)
MAXH.R	0.219	0082	0.007	1.245 (1.065 – 1.469)		
BLAF.L	0.281	0.149	0.059	1.325 (0.993 – 1.783)		
TDAF.L	0.277	0.110	0.012	1.319 (1.070 – 1.652)		

Threshold Metrics of the Best Logistic Regression Model				
Sensitivity	**Specificity**	**Precision (PPV)**	**NPV**	**Accuracy (95% CI)**
0.864	0.808	0.818	0.856	0.836 (0.7842 - 0.8797)

Blind Test (n=100) on the Best Logistic Regression Model				
0.76	0.84	0.79	0.81	0.80

Confusion Matrix - Logistic Regression Equation

		Ground Truth	
		M	F
Predicted	M	108	24
	F	17	101

Confusion Matrix - Blind Test

		Ground Truth	
		M	F
Predicted	M	34	9
	F	11	46

Fig 4. A confusion matrix representing the performance evaluation of the best model obtained through logistic regression analysis.

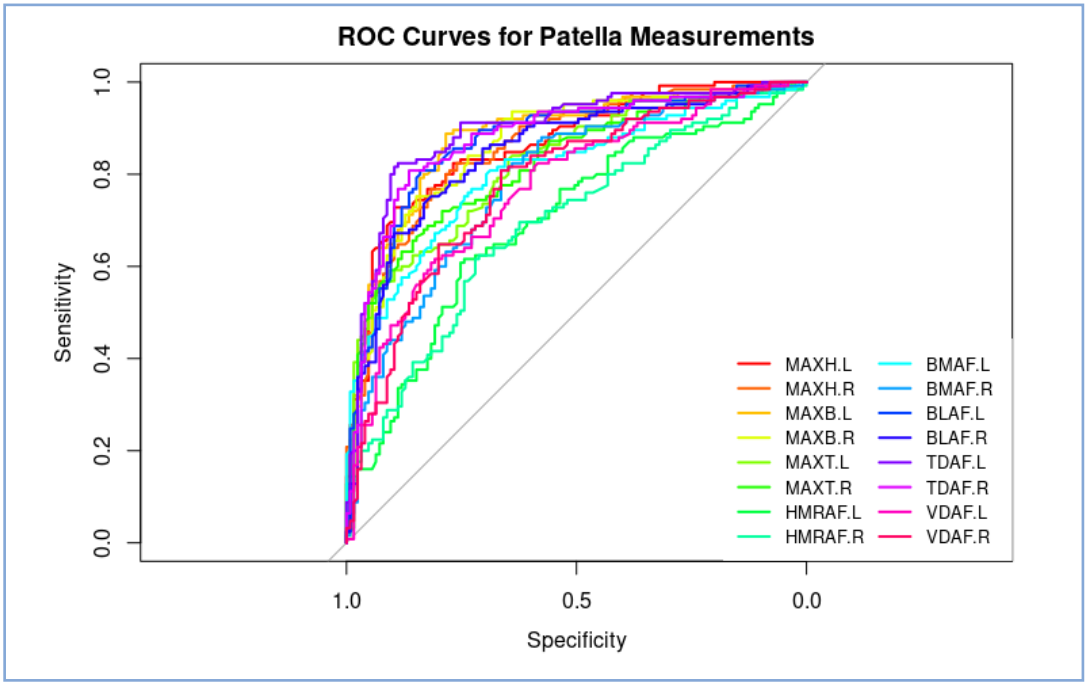


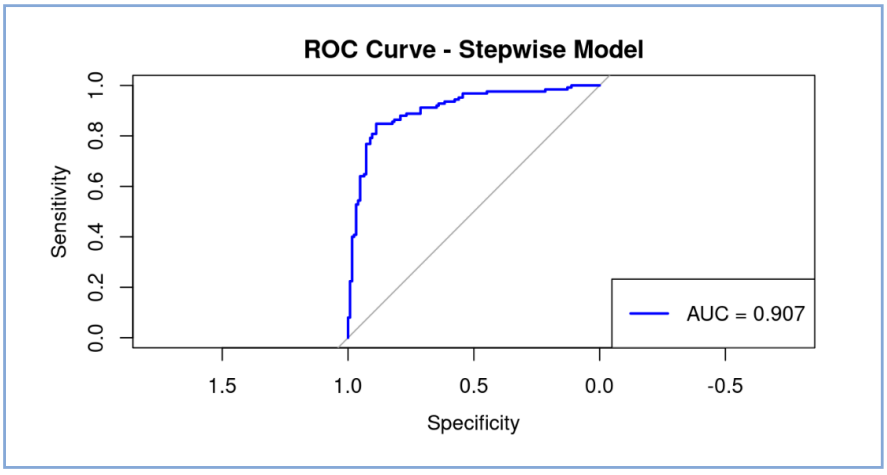
Fig 5. Receiver operating characteristic curve (ROC curves) for the all the univariable measured on the patella.

Moreover through stepwise logistic regression, the most notable model for sex estimation incorporated the parameters MAXH(R), BLAF(L), TDAF(L), formulating the equation: -26.87+(TDAF.L×0.277)+(MAXH.R×0.219)+(BLAF.L×0.281) (Table 5). However, within these three the parameter that contributes most to the prediction is MAXH(R) and TDAF(L) as their p values are statistically significant at 0.007 and 0.012 respectively. The analysis resulted in an AIC score of 205 and an AUC of 0.907 (95% CI: 0.8685 - 0.945) (Fig 6). The performance metrics revealed a sensitivity of 0.864, specificity of 0.808, a precision of 0.818, NPV of 0.856 and accuracy of 0.836 (95% CI: 0.7842 - 0.8797) (Table 5). A Hosmer–Lemeshow test revealed X^2 of 8.2406 with a p-value of 0.4103 – an appropriate goodness of fit. In addition, the Brier score revealed a low prediction error with a value of 0.12.

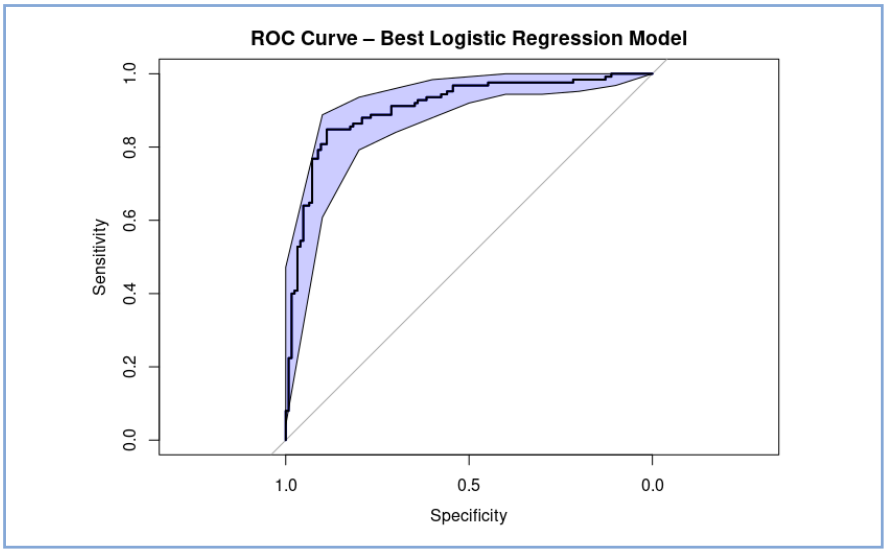
Machine learning and deep learning

Different models were obtained from using WEKA software and MATLAB software that exhibited the accuracy of 80% or more (Table 6). The models in MATLAB that produced the highest were Coarse Tree and Coarse

a)



b)



c)

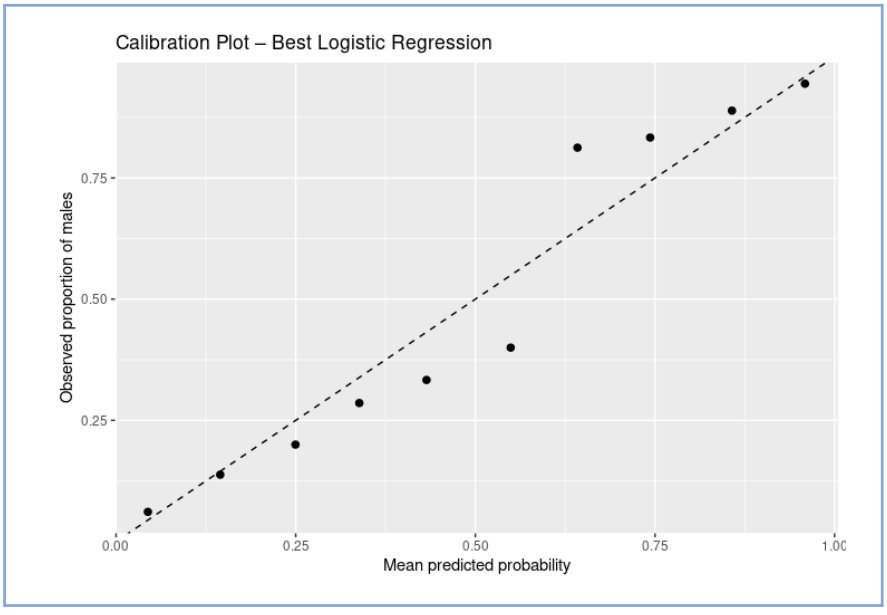


Fig 6. a) Receiver operating characteristic curve (ROC curves) for the equation that best determines sex in patella b) Receiver operating characteristic curve (ROC curves) with confidence band. c) Calibration plot for the equation that best determines sex in patella: TDAF.L + MAXH.R + BLAF.L.

TABLE 6. Comparing the top ten different machine learning models from WEKA and MATLAB to determine the model with the highest accuracy in sex estimation.

Model	Accuracy (%)	TP Rate Recall Sensitivity	FP Rate	Precision PPV	F-measure	Specificity
WEKA						
Bayes Net	83.60±0.06	0.85±0.06	0.17±0.11	0.84±0.10	0.84 ±0.06	0.83±0.11
Naives Bayes	83.2±0.07	0.83±0.09	0.17±0.10	0.84±0.09	0.83±0.07	0.83±0.10
Logistic	81.2±0.08	0.82±0.10	0.19±0.12	0.82±0.11	0.81±0.08	0.81±0.12
SGD	82±0.07	0.82±0.09	0.18±0.12	0.83±0.11	0.82±0.07	0.82±0.12
Simple Logistic	82.4±0.09	0.85±0.07	0.2±0.14	0.82±0.12	0.83±0.08	0.8±0.14
SMO	84.4±0.05	0.84±0.07	0.15±0.10	0.86±0.09	0.84±0.04	0.85±0.10
LWL	84.4±0.08	0.8±0.10	0.11±0.09	0.88±0.09	0.84±0.09	0.89±0.09
Decision Tree	82.8±0.08	0.83±0.10	0.17±0.09	0.83±0.08	0.83±0.08	0.83±0.09
JRip	86±0.09	0.85±0.10	0.13±0.12	0.88±0.11	0.86±0.09	0.87±0.12
OneR	86±0.08	0.87±0.11	0.15±0.12	0.86±0.08	0.86±0.07	0.85±0.12
PART	82.4±0.08	0.78±0.07	0.13±0.12	0.87±0.11	0.82±0.08	0.87±0.12
Decision Stump	84.4±0.08	0.8±0.10	0.11±0.09	0.88±0.09	0.84±0.09	0.89±0.09
Hoeffding Tree	83.2±0.07	0.82±0.09	0.16±0.10	0.84±0.09	0.83±0.07	0.84±0.10
LMT	82.8±0.08	0.83±0.09	0.18±0.11	0.83±0.10	0.83±0.08	0.83±0.11
Random Forest	82±0.07	0.8±0.08	0.16±0.11	0.84±0.10	0.82±0.07	0.84±0.11
REP Tree	83.2±0.08	0.82±0.09	0.16±0.12	0.85±0.10	0.83±0.07	0.84±0.12
AdaBoostM1	84±0.08	0.81±0.09	0.13±0.08	0.86±0.08	0.83±0.08	0.87±0.08
Attribute Selected Classifier	83.6±0.06	0.79±0.06	0.12±0.10	0.87±0.10	0.83±0.07	0.88±0.10
Bagging	84.4±0.05	0.83±0.06	0.14±0.07	0.86±0.07	0.84±0.05	0.86±0.07
Classification via Regression	84.4±0.08	0.8±0.09	0.11±0.09	0.88±0.09	0.84±0.08	0.89±0.09
Filtered Classifier	82.4±0.06	0.8±0.06	0.15±0.08	0.84±0.08	0.82±0.06	0.85±0.08
Iterative Classifier Optimizer	82.8±0.06	0.79±0.08	0.13±0.09	0.86±0.09	0.82±0.07	0.87±0.09
Logit Boost	82.8±0.06	0.79±0.08	0.14±0.07	0.86±0.07	0.82±0.07	0.86±0.07
Random Subpace	82.8±0.08	0.82±0.08	0.16±0.13	0.85±0.12	0.83±0.08	0.84±0.13
MATLAB						
Coarse Tree	83.7±0.02	0.81±0.01	0.13±0.02	0.86±0.02	0.83±0.02	0.87±0.02
Linear Discriminant	81.7±0.01	0.81±0.02	0.18±0.01	0.82	0.82±0.0	0.82±0.01
Binary GLM LR	80.9	0.81±0.01	0.19±0.01	0.81±0.01	0.81±0.	0.81±0.01
Efficient LR	81.2±0.01	0.82±0.01	0.19	0.81	0.81±0.01	0.81
Gaussian Naive Bayes	83.2±0.01	0.82±0.01	0.16	0.84±0.01	0.83±0.01	0.84±0.01
Kernel Naive Bayes	83.3±0.01	0.83	0.16±0.01	0.84±0.01	0.83±0.01	0.84±0.01
Linear SVM	82.7±0.01	0.83±0.01	0.18±0.01	0.82±0.01	0.83±0.01	0.82±0.01
Medium Gaussian SVM	81.9±0.01	0.80	0.16±0.01	0.83±0.01	0.81±0.01	0.84±0.01
Coarse Gaussian SVM	83.7	0.84±0.01	0.17±0.01	0.83±0.01	0.84±0.01	0.83±0.01
Medium KNN	80.7±0.01	0.78±0.01	0.17±0.01	0.82±0.01	0.80±0.01	0.83±0.01
Coarse KNN	82.4±0.01	0.80±0.01	0.15±0.01	0.84±0.01	0.82±0.01	0.85±0.01
Cosine KNN	82.8	0.82	0.17±0.01	0.83	0.83	0.83±0.01
Cubic KNN	80.9±0.01	0.79	0.17±0.01	0.82±0.01	0.80±0.01	0.83±0.01
Weighted KNN	80.4±0.01	0.82±0.01	0.21±0.03	0.80±0.02	0.81±0.02	0.79±0.03
Bagged Tree	82.5±0.01	0.79±0.02	0.14±0.01	0.85±0.01	0.82±0.02	0.86±0.01
Subspace Discriminant	83.3±0.01	0.83±0.02	0.17±0.01	0.83±0.01	0.83±0.01	0.83±0.01
Subspace KNN	80.7±0.01	0.79±0.01	0.18±0.02	0.82±0.02	0.80±0.01	0.82±0.02

TABLE 7. Deep Learning Models classifying male and female patellae.

Model	Cross Validation			Test		
	Accuracy (%)	Sensitivity	Specificity	Accuracy (%)	Sensitivity	Specificity
VGG19	50.6±0.13	0.30±0.27	0.2±0.27	49.7	0	1
ResNet18	79.2±0.04	0.64±0.12	0.94±0.03	79.0	0.65	0.94
MobileV2	65.2±0.05	0.55±0.24	0.76±0.32	53.8	0.335	0.74

Gaussian SVM at 83.7% while the lowest was Weighted KNN at 80.4%. On the other hand, in WEKA the model with the highest validation accuracy is Jrip and OneR with 86.0% and the lowest accuracy is Logistic (81.2%). Hence, both WEKA and MATLAB provide models with similar ranging validation accuracy.

In addition, from the three models that were trained and tested in classification of images generated the best results with the accuracy of 79% for the ResNet18. Whereas the lowest accuracy obtained was for VGG19 with an accuracy of 49.7%.

DISCUSSION

The aim of this investigation was to conduct a comparative analysis of several sex estimation modalites and evaluate the efficacy of deep learning model to machine learning and logistic regression for sex estimation using the dry patellae. By combining morphometric and image-based analysis, this paper extends traditional statistical models and discriminant functions. The results obtained help reinforce that patella showcases apparent sexual dimorphism and can provide accuracy in the absence of the primary skeleton choice (pelvis and skull).

Furthermore, the advantage of using patella can be attributed to both the external factors and internal factors like hormones, genes and biomechanics that influence the osseous robustivity within each sex.[32] Therefore, other bones such as the patella can be of an alternative. The anatomical position, dense[33] composition and the articular cartilage of the patellae is composed of copious amounts of collagen and glycosaminoglycan[34], making it resilient to taphonomic changes. As a result it increases the probability of recovery during excavation.

Sexual dimorphism assessment and descriptive statistics

In this study, the morphometric examination confirmed that there is a significant difference between all eight parameters between males and females, which is consistent with several previous studies on the sexual dimorphism of this bone. Furthermore, the dimensions of the patella were larger in males than female, attributing to the hormonal factor like oestrogen which has a stimulatory effect to the increase in the periosteal growth in males but contrary in females, acts as inhibitory effect on the periosteal growth.[35] These results are in accordance with the findings of several previous researches conducted on different populations such as South African Whites[19], South African of European ancestry[18], Central Thai population[10], Northern Thai population[15], Spanish population[13], Southern Part of Andhra Pradesh, India[17] and Southern Italian population.[36]

Comparing the average between sexes within regional boundaries, the patella in this study exhibited a larger value for the HMRAF and BMAF in comparison to the Northern Thai region[15] and Central Thai region.[10] The hypothesis for this could be variation within a biology of a subgroup of a population but also it could be the indefinite landmark on the posterior aspect of the patella. Furthermore, comparing the mean values of this study to the findings from Southern Chinese population[37] and Korean population (total knee arthroplasty)[38] can infer that Asian population may have smaller patellar dimensions compared to Westerners, although there has also been a study on the Asian population with healthy knees that revealed no difference between Asians and Westerners.[39] The discrepancy in the measurement can be attributed to to the different genetic makeup, lifestyle and environment.[15]

In the forensic context, the differences in these datahighlight the importance of using population specific reference. Furthermore, with the presence of apparent sexual dimorphism, the patella could be used for sex estimation in the field of forensic anthropology. In the area where remains are found, if the primary bones are not available, patella can be the supplementary element to help determine the sex.

Difference in patellae side assessment

Furthermore, most of the findings in our study were consistent to both sides of the patella, indicating that any

side can be used in sex estimation. However, minor side specific difference was observed in BMAF for females and MAXH and BLAF in males. While these finding contradicts the information from previous research on Portuguese population[16] and Spanish population[13] reporting no significant difference, one plausible reason could be the existence of bilateral variation. Nevertheless, the differences in skeletal measurements may be limited or inconsistent across different parameters, suggesting a minimal impact. In addition, the variation of these significant differences between left and right side lacks comparative studies on the bilateral factor of patella, suggesting a need for further research.

Reliability analysis by interobserver and intraobserver

The ICC values for the interobserver and intraobserver measurement ranged from 0.502 to 0.995 and 0.842 to 0.997 respectively. The parameters that showcased the highest reliability for both male and female include MAXH, MAXB, and MAXT because these parameters have definite landmarks with methodology being repeated by many authors. In addition, research on the Thai population by Srinak and Sukvitchai[10] reported similar results. This reinforces the repeatability of the methodology being used. However, the lowest interobserver measurement can be attributed to the idea that even though the observers were using the same description, the inconsistency and subjectivity of the landmarks can cause the discrepancies. The VDAF parameter had the lowest value and this could be due to the posterior side of the patella having a smooth surface. Therefore, the margins are not as defined compared to the other parameters. Moreover, this also further suggest that parameters such as VDAF should probably be omitted when estimation sex of an unidentified individual.

Logistic regression

In the field of forensic science, two of the most common methods that are used in formulating a sex estimation equation are logistic regression and discriminant function analysis (DFA).[41] Thus, unlike most of the previous studies on sexual dimorphism on the patella that uses DFA, this study performed a logistic regression analysis to evaluate the parameters that estimate sex. The reason for this is that the data obtained in this study have parameters with measurements that are not normally distributed. Hence, logistic regression is more suitable because it can provide robust estimations even when assumptions are violated.[41] Additionally, logistic regression offers greater flexibility compared to DFA because the generated equation can be directly utilized to compute the probabilities of categorizing the sexes, thereby quantifying uncertainties. In contrast, DFA solely provides the equation and cut-off point, resulting in the absence of prior probabilities. This absence diminishes the confidence in classifying the sexes, potentially leading to misclassification.[42]

Many previous studies have reported that the measurement that exhibited the highest sexual dimorphism varied among populations as it included MAXH (South African White, Portuguese, Spanish)[13,16,19] MAXB (Portuguese)[16], MAXT (Italian and Spanish)[13,36], BLAF (Central Thai and Northern Thai).[10,15] However, in this study, TDAF was most sexually dimorphic with 12.15% on the left side and 11.53% on the right side. This is further iterated by the logistic regression analysis, producing the highest AUC but the lowest AIC value. TDAF is not a common parameter measured in previous studies but rather divided into the breadth of medial facet (BMAF) and breadth of lateral facet (BLAF). Furthermore, a stepwise regression analysis revealed the parameters identified as the most notable for sex estimation were MAXH.R, BLAF.L, and TDAF.L. The performance metrics for this LR model revealed a strong predictability model. With high sensitivity and specificity value as well as precision and negative predictive value, the model is very effective in accurately distinguishing between the sexes. A blind test of 100 samples revealed an accuracy 80.0%. Thus, these results are comparable with other populations such as the 73% - 89%.[10,13,15,19,21] The value obtained in this study is in the similar range of previous studies, but the methodologies used to obtain them were different, as all of these studies generated discriminant function equations. Nonetheless, as each population has different genetic composition, physical activities and biomechanics, it is important to obtain equations for each population. However, this study also possess some limitations like the lack of external validation. Having another collection for testing would help reinforce the accuracy. This would be a possible next step for future projects.

Studies conducted for sex estimation using the most ideal equation in each population were applied to different populations by different authors[13,19], revealed a lower accuracy in classifying the sexes.[16] Therefore, from the LR analysis, it can be suggested that patella provides an important significance in determining sex. It also signifies the importance of population specific determinator. This also provides the forensic anthropologist an alternative approach when profiling an unknown remain.

While, the logistic regression yielded a strong predictive value of 80%, the performance is lower compared

to pelvis and skull at approximately 92% to 95%. Thus, this suggest that patella while it can fulfill the role of being a supplementary bone for sex estimation, it cannot replace the primary models. Locally and internationally, the findings from studies conducted on patella has expressed it to be a viable bone in aiding the sex estimation procedure but only when the primary alternatives are missing, fragmented or unsuccessfully preserved.

Machine learning and deep learning

Logistic regression and discriminant function analysis are the prevailing methods in sex estimation but as times have been changing and there is an integration of artificial intelligence into every aspect of our daily lives. Hence, there has also been a growing tendency to use artificial learning models such as ML and DL in the field of forensic science.[12]

Previously, there have been two studies that utilized a ML approach predicting the best model for sex estimation: Knecht and colleagues[12] for an Italian population and Bidmos and colleagues[18] for a South African population with mixed European and African descent.

Comparing this current research with the work of Bidmos and colleagues[7], they achieved an overall accuracy of 89.61% using Random Forest, 89.23% using AdaBoost classifier and 90.77% using the Stacking model. The approach to the training model was different as a 5-fold cross validation training was used whereas this study used 10-fold cross validation. Using 5-fold cross validation reduced the overall accuracy for each classifier. Comparing this with the Italian population[12], their overall accuracy for their ML classifiers ranged from 85% to 91% for training data and 71% to 95% for their validation sample. The best classifier for their study was the Sequential minimal optimization (SMO) which is a type of support vector machine (SVM). Furthermore, their research only utilizes three parameters in classifying sex whereas, this one uses 8 parameters, broadening the horizon of computing differences. Nonetheless, the overall accuracies are in similar ranges showcasing the agreement of using ML as a means for classifying sex. Not only this, it further strengthens the applicability of using patella in sex estimation.

Furthermore, depending on the population, various classifiers would provide a different overall accuracy. This suggests that to study the Thai population, an ML model such as Coarse tree or SVM can be used.

Extending a step further, utilizing the images of the patella for sex estimation provides a flexible alternative to the primary choice of pelvis or the skull. ResNet 18 likely gained the highest accuracy because it is composed of residual network that help in training at deeper layer and this allows for information to skip layers allowing the model to learn more reliably.[43] However, VGG19 obtained the lowest accuracy because this is model is dense with large parameters and layers that could highly lead to overfitting on small dataset like 1,000 images.[44] Finally, the MobileNetV2 is a lightweight model which could reduce flexibility in trying to read complex images.[45]

However, looking at the results from a different perspective, the lower accuracy value can be ascribed to the type of image used in training and testing. The anterior and posterior images used in this study as seen in Fig 3 can be described to have a cluttered or noisy background. Prior studies have reported that training and testing a convolutional network with images with cluttered background can slower the training procedure and cause models to underperform.[43] Thus, this can act as a limitation not only for this study but a cautious point for future researches. The lower accuracy for our study can possibly be explained by a small dataset of 1000 images because in recent research using convolutional neural network to estimate sex from MRI images of patella, they utilized approximately 6,000 images.[20] Hence, a small dataset increases the risk of overfitting and weakens the generalizability factor. An MRI image is also more homogenous with a reduced background cluttering. Thus, these differences could also explain discrepancies and lower result yeilded.

The selection of models like the VGG19, ResNet18 and MobileNetV2 was to study a range of model. Not specifically for forensic anthropology but these models have been used to study age and sex estimation from dental radiograph in young adults, achieving an accuracy of approximately 80%.[44]

Computational forensics is a current developmental content that utilizes AI to identify human remains. Deep learning analysis usually employs radiographic datasets. Studies that have utilized CT scans[37] and MRI[45] in measuring the patella size to determine sexual dimorphism have been able to achieve high accuracy. Furthermore, another relevant study on using DL to classify sex in relation to the patella is the radiographs of tibiofemoral joints, that obtained 90.3%.[46] However, this study demonstrated a more flexible approach by using photographic imagery which is an easier, practical and much more cost-effective manner. However, with the limited sample size of 1,000 images the training and model can be developed to achieve higher accuracy. It is very important to always note the sample size, population and the methodologies utilized in each study before performing any comparison.

CONCLUSION

In conclusion, this study of sexual dimorphism of the patella within the Thai population using 250 paired patellae revealed sexual dimorphism exists in the patella as all measured parameters showed significant differences. Overall, comparing the traditional approach of logistic regression to advanced computational models such as ML and DL, it is evident that all these methods achieve an accuracy of 80% and above. These findings suggest that while logistic regression can be the basis of generating a robust sex estimation equation, ML and DL models can work to enhance the classification process. This suggests that even in the absence of the primary skeletal choice for sex determination (pelvis and skull), the patella can be used.

Data Availability Statement

The access of this data was approved through Siriraj Anatomical and Anthropological Bone Research Center (Si-AABRC) procedure. Data is available on reasonable request from Si-AABRC.

ACKNOWLEDGEMENTS

The authors of this research are thankful for Siriraj Anatomical and Anthropological Bone Research Center (Si-AABRC), Department of Anatomy, Faculty of Medicine Siriraj Hospital for allowing the authors to utilize the patellae from the bone collection. The authors are also grateful for the body donation program of Department of Anatomy, Faculty of Medicine Siriraj Hospital. This research would like to acknowledge Ms. Areeya Marie B. Wongla for her help.

DECLARATIONS

Grants and Funding Information

None.

Conflict of Interest

None.

Registration Number of Clinical Trial

Not Applicable.

Author Contributions

Conceptualization and methodology. S.H.; Investigation, S.H.; Formal analysis, S.H. P.R., P.J.,N.C.; Visualization and writing – original draft, S.H.; Writing – review and editing, P.R., N.C., P.J., S.R., R.M., N.S.; Supervision, N.S. All authors have read and agreed to the final version of the manuscript.

Use of Artificial Intelligence

This paper utilized ChatGPT-5 (Open AI) to generate initial code drafts for R-Studio and MATLAB.

ETHICAL STATEMENT

This research was approved by the Siriraj Institutional Review Board (SiRB), Faculty of Medicine Siriraj Hospital [SIRB Protocol No.331/2568 (IRB4)].

REFERENCES

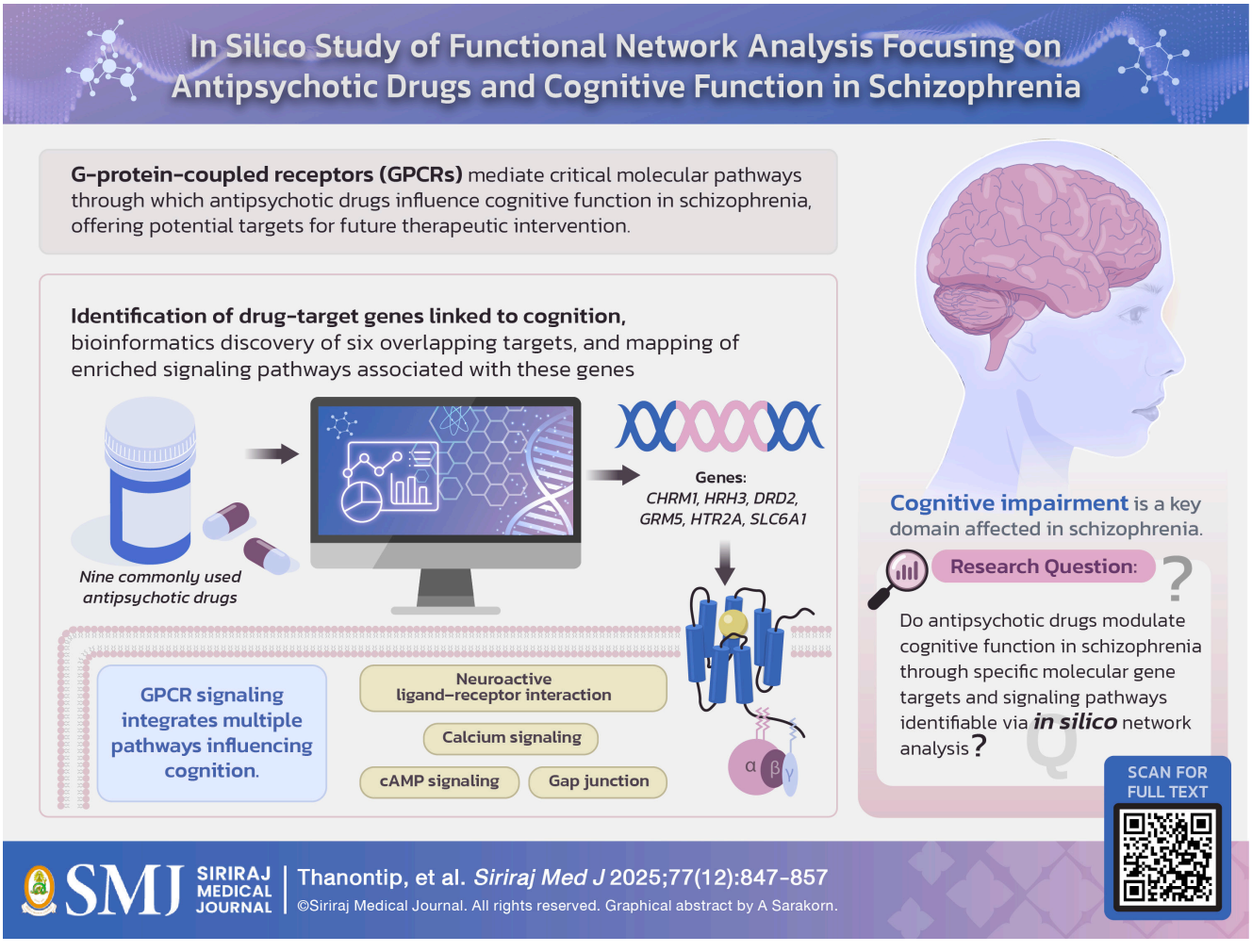
1. Kahana T, Hiss J. Identification of human remains: forensic radiology. J Clin Forensic Med. 1997;4(1):7–15.
2. de Boer HH, Blau S, Delabarde T, Hackman L. The role of forensic anthropology in disaster victim identification (DVI): recent developments and future prospects. Forensic Sci Res. 2018;4(4): 303-15.
3. Austin D, King RE. The Biological Profile of Unidentified Human Remains in a Forensic Context. Acad Forensic Pathol. 2016;6(3):370–90.
4. Krishan K, Chatterjee PM, Kanchan T, Kaur S, Baryah N, Singh RK. A review of sex estimation techniques during examination of skeletal remains in forensic anthropology casework. Forensic Sci Int. 2016;261:165.e1–8.
5. Ahmed AA. Estimation of sex from the lower limb measurements of Sudanese adults. Forensic Sci Int. 2013;229(1-3):169.e1–7.
6. Phenice TW. A Newly Developed Visual Method of Sexing the Os Pubis. Am J Phys Anthropol. 1969;30(2):297–301.
7. Bidmos MA, Olateju OI, Latiff S, Rahman T, Chowdhury MEH. Machine learning and discriminant function analysis in the formulation of generic models for sex prediction using patella measurements. Int J Legal Med. 2023;137(2):471–85.
8. Memarian A, Aghakhani K, Mehrpisheh S, Fares F. Gender determination from diagnostic factors on anteroposterior pelvic radiographs. J Chin Med Assoc. 2017;80(3):161–8.
9. Wangdee A, Thipdet W, Prasitwattanaseree S, Singsuwan P, Mahakkanukrauh P. Efficiency of sex determination by using external morphology of the pelvis in Thai population. BSCM. 2014;53(4):175–9.
10. Srinak N, Sukvitchai P. Sex estimation from patella using discriminant analysis in Central Thai population. Journal/ Journal - Canadian Society of Forensic Science. 2023;56(4):231–47.
11. Fox A, Wanivenhaus F, Rodeo S. The Basic Science of the Patella: Structure, Composition, and Function. J Knee Surg. 2012;25(2): 127–41.
12. Knecht S, Morandini P, Biehler-Gomez L, Nogueira L, Adalian P, Cattaneo C. Sex estimation from patellar measurements in a contemporary Italian population: a machine learning approach. Int J Legal Med. 2024;139(3):1371–80.
13. Peckmann TR, Meek S, Dilkie N, Rozendaal A. Determination of sex from the patella in a contemporary Spanish population. J Forensic Leg Med. 2016;44:84–91.
14. Silva RF, Franco A, Santos D, Picoli FF, Marinho DE de A. Human identification through the patella—Report of two cases. Forensic Sci Int. 2014;238:e11–4.
15. Chompoophuen H, Tipmala J, Duangto P, Mahakknukrauh P. Sex Determination from the Patella in a Thai Population. Int J Morphol. 2024;42(4):891–7.

16. Maio C, Cunha E, Navega D. Metric analysis of the patella for sex estimation in a Portuguese sample. Forensic Sci Res. 2024; 9(2):owae015.
17. Nagarjuna K, Mamatha K, Venkateswarlu B. Patellar Anthropometry in Sex Differentiation - A Study in the Southern Part of Andhra Pradesh, India. Indian Journal of Forensic Medicine and Toxicology. 2021;15(2):3113–8.
18. Olateju OI, Philander I, Bidmos MA. Morphometric analysis of the patella and patellar ligament of South Africans of European ancestry. South African Journal of Science. 2013;109(9/10):1–6.
19. Bidmos MA, Steinberg N, Kuykendall KL. Patella measurements of South African whites as sex assessors. Homo. 2005;56(1):69–74.
20. Cavlak N, Çınarer G, Erkoç MF, Kılıç K. Sex estimation with convolutional neural networks using the patella magnetic resonance image slices. Forensic Sci Med Pathol. 2025;21(2):628–39.
21. Dayal MR, Bidmos MA. Discriminating Sex in South African Blacks Using Patella Dimensions. J Forensic Sci. 2005;50(6): 1294-7.
22. The jamovi project (2024). jamovi (Version 2.5) [Computer Software]. Available from: https://www.jamovi.org
23. Govindaram D, Bharanidharan R, Ramya R, Rameshkumar A, Priyadharsini N, Rajkumar K. Root Length: As a determinant tool of sexual dimorphism in an ethnic Tamil population. J Forensic Dent Sci. 2018;10(2):96-100.
24. Bobak CA, Barr PJ, O'Malley AJ. Estimation of an inter-rater intra-class correlation coefficient that overcomes common assumption violations in the assessment of health measurement scales. BMC Med Res Methodol. 2018;18(1):93.
25. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. J Chiropr Med. 2016;15(2):155–63.
26. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull. 1979;86(2):420–8.
27. RStudio Team (2020). RStudio: Integrated Development for R. RStudio, PBC, Boston, MA, USA. Available from: http://www.rstudio.com
28. Verma R, Krishan K, Rani D, Kumar A, Sharma V, Shrestha R, et al. Estimation of sex in forensic examinations using logistic regression and likelihood ratios. Forensic Science International: Reports. 2020;2:100118.
29. Dietrichkeit Pereira JG, Fróes Lima K, Alves da Silva RH. Mandibular Measurements for Sex and Age Estimation in Brazilian Sampling. Acta Stomatologica Croatica. 2020;54(3):294–301.
30. The MathWorks Inc. MATLAB version: 9.14 (R2023a), Natick, Massachusetts: The MathWorks Inc., 2023. Available from: https://www.mathworks.com
31. Frank E, Hall MA, Witten IH. The WEKA Workbench. Online Appendix for "Data Mining: Practical Machine Learning Tools and Techniques", Morgan Kaufmann, Fourth Edition, 2016.
32. Sangchay N, Dzetkuličová V, Zuppello M, Chetsawang J. Consideration of Accuracy and Observational Error Analysis in Pelvic Sex Assessment: A Study in a Thai Cadaveric Human Population. Siriraj Med J. 2022;74(5):330–9.
33. Abdelaziz S, Khattab HM, AbdelHaq NA. Prediction of Sex from Patellar Parameters Obtained by Magnetic Resonance Imaging in a Sample of Adult Egyptians. Zagazig Journal of Forensic Medicine and Toxicology. 2024;22(2):1–16.
34. Kim Y-M, Joo Y-B. Patellofemoral Osteoarthritis. Knee Surg Relat Res. 2012;24(4):193–200.
35. Nieves JW, Formica C, Ruffing J, Zion M, Garrett P, Lindsay R, et al. Males Have Larger Skeletal Size and Bone Mass Than Females, Despite Comparable Body Size. J Bone Miner Res. 2004;20(3): 529–35.
36. Introna F, Di Vella G, Campobasso CP. Sex determination by discriminant analysis of patella measurements. Forensic Sci Int. 1998;95(1):39–45.
37. Shang P, Zhang L, Hou Z, Bai X, Ye X, Xu Z, et al. Morphometric measurement of the patella on 3D model reconstructed from CT scan images for the southern Chinese population. Chin Med J (Engl). 2014;127(1):96–101.
38. Kim MH, Yoo MJ, Seo JB, Park HG, Shim SH. Statistical Analysis of the Patellar Thickness in Adults by MRI. J Korean Orthop Assoc. 2005;40(6):646-51.
39. Pontoh LA, Dilogo IH, Rahyussalim AJ, Widodo W, Edwin R, Fiolin J. Evaluation of patellar dimension and Bristol Index in Asian population: An MRI study. Ann Med Surg (Lond). 2021;72:103072–2.
40. Kanchan T, Menezes RG, Moudgil R, Kaur R, Kotian MS, Garg RK. Stature estimation from foot dimensions. Forensic Sci Int. 2008;179(2-3):241.e1–5.
41. Bartholdy BP, Sandoval E, Hoogland MLP, Schrader SA. Getting Rid of Dichotomous Sex Estimations: Why Logistic Regression Should be Preferred Over Discriminant Function Analysis. J Forensic Sci. 2020;65(5):1685–91.
42. Antonogeorgos G, Panagiotakos DB, Priftis KN, Tzonou A. Logistic Regression and Linear Discriminant Analyses in Evaluating Factors Associated with Asthma Prevalence among 10- to 12-Years-Old Children: Divergence and Similarity of the Two Statistical Methods. Int J Pediatr. 2009;2009:952042.
43. Kc K, Yin Z, Li D, Wu Z. Impacts of Background Removal on Convolutional Neural Networks for Plant Disease Classification In-Situ. Agriculture. 2021;11(9):827–7.
44. Adedigba AP, Adeshina SA, Aina OE, Aibinu AM. Optimal hyperparameter selection of deep learning models for COVID-19 chest X-ray classification. Intell Based Med. 2021;5:100034.
45. Kalaiselvi K, Kasthuri M. Tuning VGG19 hyperparameters for improved pneumonia classification. The Scientific Temper. 2024;15(02):2231–7.
46. Ukwandu O, Hindy H, Ukwandu E. An evaluation of lightweight deep learning techniques in medical imaging for high precision COVID-19 diagnostics. Healthc Anal (N Y). 2022;2:100096.
47. Kim HE, Cosa-Linan A, Santhanam N, Jannesari M, Maros ME, Ganslandt T. Transfer learning for medical image classification: a literature review. BMC Med Imaging. 2022;22(1):69.
48. Yoo JH, Yi SR, Kim JH. The geometry of patella and patellar tendon measured on Knee MRI. Surg Radiol Anat. 2007;29(8):623–8.
49. Oura P, Junno JA, Hunt DR, Lehenkari P, Tuukkanen J, Maijanen H. Deep learning in sex estimation from knee radiographs – A proof-of-concept study utilizing the Terry Anatomical Collection. Leg Med (Tokyo). 2023;61:102211.

In Silico Study of Functional Network Analysis Focusing on Antipsychotic Drugs and Cognitive Function in Schizophrenia

Kamonwan Thanontip, M.Sc.[1], Banthit Chetsawang, Ph.D.[1], Vorasith Siripornpanich, M.D., Ph.D.[1], Vasunun Chumchua, D.Sc.[2], Jirapa Chetsawang, M.D., Ph.D.[3,*]

[1]Research Center for Neuroscience, Institute of Molecular Biosciences, Mahidol University, Salaya, Nakhon Pathom, Thailand, [2]National Institute for Child and Family Development, Mahidol University, Salaya, Nakhon Pathom, Thailand, [3]Department of Anatomy, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand.



*Corresponding author: Jirapa Chetsawang
E-mail: jirapa.cht@mahidol.ac.th
Received 16 August 2025 Revised 25 October 2025 Accepted 25 October 2025
ORCID ID:http://orcid.org/0000-0002-7776-6456
https://doi.org/10.33192/smj.v77i12.277103

ABSTRACT

Objective: Recent evidence has emphasized cognitive dysfunction in psychiatric disorders. This study explores bioinformatics analysis aimed at investigating gene targets and pathways related to antipsychotic drug effects on cognitive function in schizophrenia. The main purpose of conventional antipsychotics is to treat the positive symptoms, but they have limited evidence for the involvement of these drugs in cognitive function. Therefore, this study aimed to explore the pathway for the antipsychotic signaling in the cognitive function of schizophrenia.

Materials and Methods: Nine frequently used antipsychotics were included from the literature review. The GeneCards database was utilized to identify the drug target genes associated with cognitive function in schizophrenia. The gene lists were analyzed with two different tools, WebGestalt and STRING. The overlapping genes that result from the two previous steps may play an essential role in cognitive function. These genes were analyzed to find out the related pathway using the KEGG database.

Results: The potential gene list consists of CHRM1, HRH3, DRD2, GRM5, HTR2A, and SLC6A1. The functional enrichment in KEGG reveals six target proteins involved in several pathways, which play an essential role in cognitive function, including the neuroactive ligand-receptor interaction, the gap junction, the calcium signaling pathway, and the cAMP signaling pathway.

Conclusion: The results reveal that G-protein-coupled receptors (GPCRs) are the main target site of potential mechanisms or processes related to antipsychotics and cognitive function. Therefore, GPCRs might be a promising candidate for future research on potential therapeutic targets for intervention in neurological dysfunction-associated cognitive deficits in schizophrenia.

Keywords: In silico study; functional network analysis; antipsychotic drugs; cognitive function; schizophrenia (Siriraj Med J 2025; 77: 847-857)

INTRODUCTION

Schizophrenia is a psychiatric illness characterized by psychotic symptoms, including positive symptoms and negative symptoms.[1,2] Cognitive deficit is a neurological dysfunction that occurs in schizophrenia.[3-6] The cognitive domains that show differential impairment in schizophrenia include attention, executive function, language skills, working memory, and cognitive flexibility.[7,8] These determine the functional impairment characteristic of schizophrenia, which disturbs their daily life activity and quality of life.[9] Therefore, the most concerning clinical target in the treatment of schizophrenia is not only to reduce the psychotic symptoms but also to improve cognitive function.[10]

Cognitive impairment may develop from the pathology of the disease itself in some schizophrenia patients.[11] It has been implicated that conventional antipsychotics are effective in treating psychological symptoms; however, they cannot improve cognitive function.[12] The effects of antipsychotics on cognition are complex, which might depend on the stage of illness and the doses of antipsychotics. It has been reported that a dose reduction in a patient who is clinically in an advanced phase of remission, causing cognitive decline associated with antipsychotic drugs. On the other hand, dose reduction or discontinuation after reaching stable remission might improve cognitive function.[13] Several lines of evidence revealed that the usage of antipsychotics in high doses impairs cognitive function in schizophrenia patients.[14,15] On the contrary, a low dose of conventional medications such as haloperidol improves cognitive functions.[16]

According to the heterogeneity of antipsychotics in their pharmacologic properties, they have differential profiles of cognitive efficacy in patients with schizophrenia. Therefore, understanding the molecular mechanisms and mode of action of each antipsychotic related to cognitive processes is needed to address the gap in knowledge. Computational network analysis is a potential tool that may answer these knowledge gaps by providing a biological explanation for antipsychotic targets and insights for the putative molecular mechanism of antipsychotic agents, focusing on cognitive function. Therefore, this study aimed to explore the interaction of the target protein, which is related to the action of antipsychotic drugs and cognitive functions related to schizophrenia. The results of this study could be explored to discover more effective novel target treatment of neurological dysfunction involving neuropsychological-associated cognitive deficit.

MATERIALS AND METHODS

The conceptual framework was designed to gain

insight into the pathway of the potential protein involved in antipsychotics and cognitive function in schizophrenia. The framework starts to identify the commonly used antipsychotics for schizophrenia using a literature review and then using bioinformatics tools to identify gene targets for selected antipsychotics and genes associated with cognitive function in schizophrenia. Next step, the gene associated with antipsychotics and cognitive function in schizophrenia was identified. Then, protein-protein interaction network analysis and network clustering were performed, followed by the identification of a common gene that is associated with antipsychotics and cognitive function in schizophrenia from two different methods. Finally, the pathway that contains potential protein involves antipsychotics and cognitive function in schizophrenia was identified (Fig 1).

Identification of commonly used antipsychotics

In silico study of network analysis focusing on antipsychotic drugs and cognitive function in people with schizophrenia was started to identify commonly used antipsychotics for schizophrenia (step 1 of Fig 1) by reviewing published articles and textbooks.[17-20]

Identification of gene targets for selected antipsychotics

GeneCards (https://genecards.org) and a database for searching genes of interest (step 2 of Fig 1) were used to identify gene targets for selected antipsychotics. Target genes of selected antipsychotics were identified using the names of 9 antipsychotics as a keyword. The keyword for searching in this step is "[compounds] (Aripiprazole) OR [compounds] (Chlorpromazine) OR [compounds] (Fluphenazine) OR [compounds] (Haloperidol) OR [compounds] (Olanzapine) OR [compounds] (Quetiapine) OR [compounds] (Risperidone) OR [compounds] (Thiothixene) OR [compounds] (Ziprasidone)".

Identification of genes associated with cognitive function in schizophrenia by GeneCards Database

The GeneCards database[21] was used to identify genes associated with cognitive function in schizophrenia (step 3 of Fig 1). [publications] (schizophrenia) AND [disorders] (Schizophrenia AND [publications] (cognitive) AND [disorders] (cognitive) OR [functions] (cognitive) OR [pathways] (cognitive) is the keyword for searching in this step.

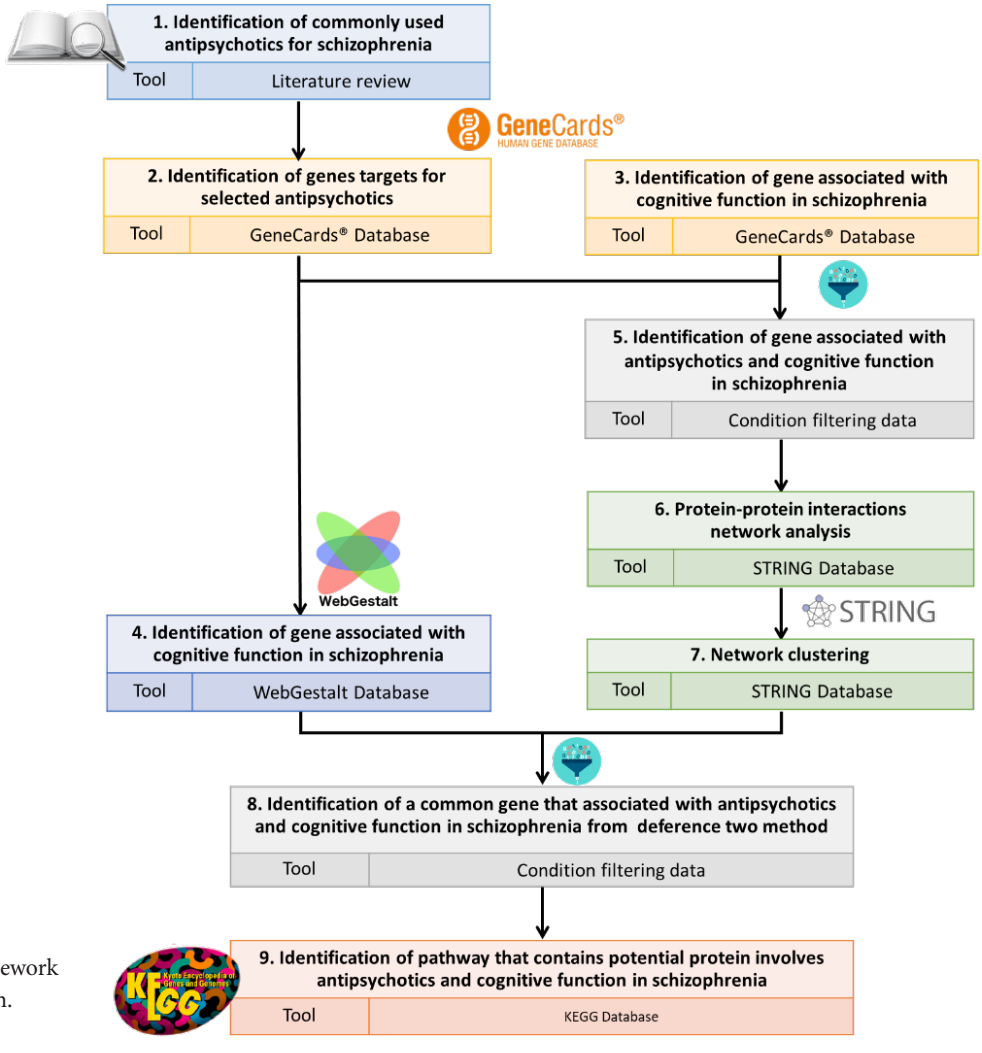


Fig 1. Conceptual framework of experimental design.

Identification of genes associated with cognitive function in schizophrenia by WebGestalt Database

The gene associated with cognitive function in schizophrenia was identified by the WebGestalt Database (http://www.webgestalt.org/) (step 4 of Fig 1), which stands for WEB-based GEne SeT AnaLysis Toolkit.[22] Over-representation analysis (ORA) is a method used in this study. The upload gene list is collected from the previous step (step 2 of Fig 1), which is the identification of gene targets for antipsychotics.

Identification of genes associated with antipsychotics and cognitive function in schizophrenia

This step is a consequence of step 2 (identification of gene targets for selected antipsychotics) and step 3 (identification of genes associated with cognitive function in schizophrenia by GeneCards Database). Gene findings from the previous step were condition formatting, then filtered by matching in both terms of antipsychotics and cognitive function in schizophrenia. The step-by-step data condition formatting and filtering data in Microsoft Excel is used to find common genes between two gene lists. The gene identification starts by selecting the data column > Conditional formatting in the Home ribbon > Highlight cell rule > Duplicate Values, and then pressing OK. Then, select the data column > Sort & Filter in the home ribbon > Filter to apply the filter dropdown icon to selected columns.

Protein-protein interactions network analysis

The gene list from the previous step (step 5 of Fig 1) was further analyzed to identify protein-protein interactions and construct a protein-protein interactions network using the STRING database. The step to identify protein-protein interactions of the gene of interest starts with inputting the name of the gene of interest into the multiple proteins search box and then selecting auto-detect organism and search. The input of the Gene list might appear in several organisms. The Homo sapiens organism was selected from the list. Many proteins in Homo sapiens might similar to the input data. The matching protein was selected. A protein-protein interactions network was generated and visualized only connected nodes in interaction evidence mode, using three interaction sources: experiment, database and co-expression. The default setting was applied for the minimum required interaction score and the maximum number of interactions to show in the network display mode.

Network clustering

The network was clustered using the Markov Clustering method (MCL), with three inflation parameters to establish the most interconnected cluster of the protein-protein interaction (PPI) network.[23]

Identification of a common gene associated with antipsychotics and cognitive function in schizophrenia from deference two methods

Gene findings from the WebGestalt and network clustering were condition formatting, then filtered by matching genes from the two lists. The condition formatting and filtering data settings in Microsoft Excel were used to find common genes between two gene lists.

Identification pathway of potential proteins involved in antipsychotics and cognitive function in schizophrenia

The results from the previous step of deference two methods, which involved the identification of a common gene associated with antipsychotics and cognitive function in schizophrenia, were used to identify the related pathway by the STRING database. The results of the related pathway were shown in functional enrichment in the analysis section.[24]

Comprehensive graphical user interface snapshots, input datasets, parameter configurations, filtering criteria, threshold values, and detailed procedural steps for each analysis are provided in the Supplemental Data.

RESULTS

The nine commonly used antipsychotics were chosen, including Aripiprazole, Chlorpromazine, Fluphenazine, Haloperidol, Olanzapine, Quetiapine, Risperidone, Thiothixene, and Ziprasidone.[17-20,24]

The results of gene targets for antipsychotics showed a total of 223 target genes (Fig 2). The results obtained from the GeneCards Database showed that 122 genes were associated with cognitive function in schizophrenia (Fig 2). The results from WebGestalt database analysis showed that 223 gene targets for nine selected antipsychotics were characterized into three main categories, including biological process, cellular component, and molecular function. Apart from 223 genes, the results showed that 31 genes were found to be associated with cognitive function in schizophrenia (Fig 2).

The 28 candidate genes associated with antipsychotics and cognitive function in schizophrenia were constructed into protein-protein interaction networks analysis using the STRING database. The results showed 28 Nodes of protein and 33 Edges, representing interactions between pairs of proteins. The protein-protein interaction was identified based on experimental, database, or coexpression evidence. Moreover, nine proteins, including CHAT,

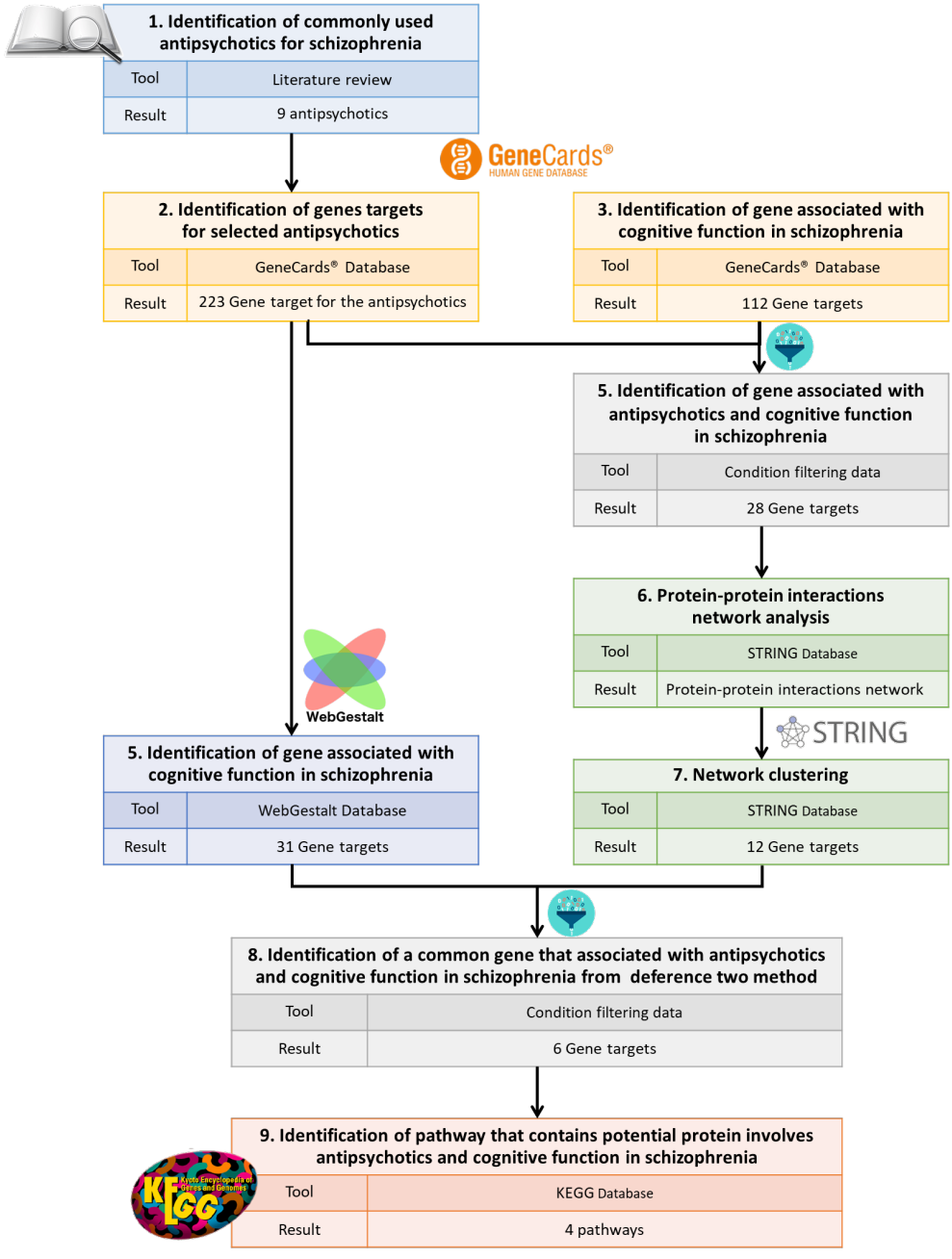


Fig 2. Flow diagram illustrating the summary of results.

APOE, GAD1, PLA2G6, SLC18A2, IL10, MTHFR, COMT, and TH, have no interaction with any protein (Fig 3A).

The network clustering of 28 candidate genes associated with antipsychotics and cognitive function in schizophrenia was constructed using the STRING database. The results exhibited four network clusters, and the largest cluster comprises 12 proteins. The 7 more proteins were clustered in one cluster of 3 proteins and two clusters of 2 proteins, respectively. The nine disconnected nodes were hidden from the network clustering (Fig 3B). The 12 proteins of the large clustering were selected for analysis in the next experiment.

The target gene associated with antipsychotics and cognitive function in schizophrenia, collected from 31 genes of the WebGestalt database analysis and 12 genes of network clustering, was taken into consideration using deference two methods. The six common genes associated with antipsychotics and cognitive function in schizophrenia were obtained from deference two methods, which include CHRM1, HRH3, DRD2, GRM5, HTR2A and SLC6A1. The gene symbol, description and antipsychotics interact with the 6 candidate genes are shown in Table 1.

The six common genes associated with antipsychotics and cognitive function in schizophrenia were further analyzed using KEGG (Kyoto Encyclopedia of Genes and Genomes) analysis. The KEGG analysis results in 4 pathways, including neuroactive ligand-receptor

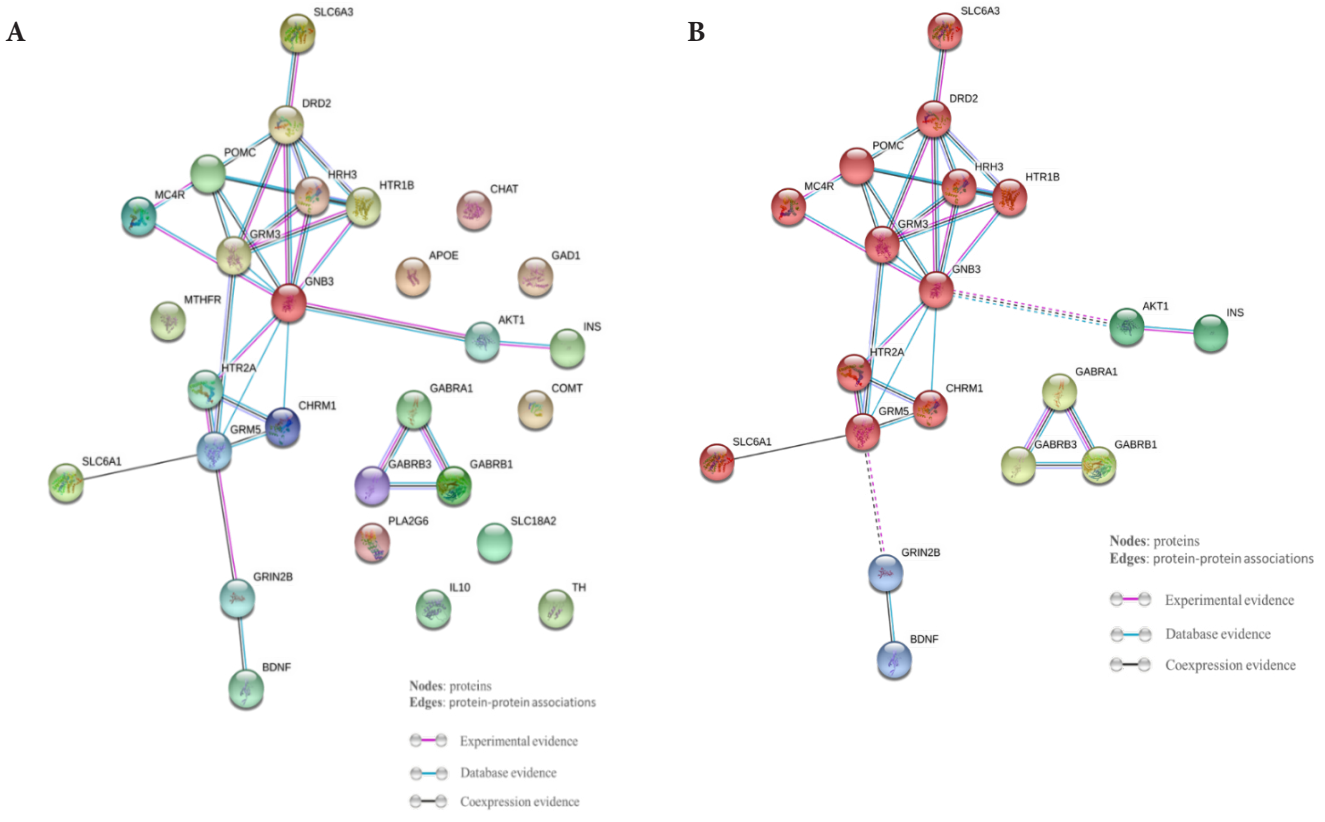


Fig 3. (A) Protein-protein interactions (PPI) network of proteins associated with antipsychotics and cognitive function in schizophrenia. The protein-protein interaction is characterized by three different types of evidence, including experimental, database, or coexpression. Each evidence of protein-protein interaction is shown in a different color of the connecting line. The experimental, database or coexpression evidence is indicated in purple, blue and black lines, respectively. (B) Protein-protein interactions network of proteins associated with antipsychotics and cognitive function in schizophrenia after network clustering and hiding disconnected nodes. The network clustering is indicated in color.
(A larger version of this figure is available in the Supplemental Data file.)

TABLE 1. Gene symbol, description, and list of antipsychotics interact with the common gene associated with antipsychotics and cognitive function in schizophrenia from deference two methods.

No.	Gene Symbol	Description	Antipsychotics interact with the gene
1	CHRM1	cholinergic receptor muscarinic 1	Aripiprazole, Chlorpromazine, Olanzapine, Quetiapine, Ziprasidone
2	HRH3	histamine receptor H3	Aripiprazole, Risperidone
3	DRD2	dopamine receptor D2	Aripiprazole, Chlorpromazine, Fluphenazine, Haloperidol, Olanzapine, Quetiapine, Risperidone, Thiothixene, Ziprasidone
4	GRM5	Glutamate metabotropic receptor 5	Haloperidol
5	HTR2A	5-hydroxytryptamine receptor 2A	Aripiprazole, Chlorpromazine, Fluphenazine, Haloperidol, Olanzapine, Quetiapine, Risperidone, Thiothixene, Ziprasidone
6	SLC6A1	solute carrier family 6 member 1	Haloperidol

interaction, gap junction, calcium signaling pathway, and cAMP signaling pathway (Table 2). The observed gene count, background gene count, false discovery rate and matching proteins in the network of each pathway are shown in Table 2.

DISCUSSION

Recently, several lines of evidence emphasized the functional role of antipsychotic drugs in the cognitive function of psychiatric disorders, including schizophrenia.[25-30,61-63] However, the mode of action and molecular mechanisms of antipsychotic drug-related cognitive function are not clearly understood. Bioinformatics is one of the tools that reduces the costs and time to study drugs and target-specific drugs.[28,29] Moreover, network analysis is a bioinformatics approach that is employed to gain a broad and in-depth picture of protein-protein interaction (PPI) networks and signaling pathways.[30]

In the present study, nine frequently used antipsychotics were selected from the literature review. The GeneCards database was used to identify the drug target gene and gene of interest involved in cognitive function in schizophrenia. The next step was to analyze the gene lists with two different tools: WebGestalt Database, a functional enrichment analysis web tool, and STRING database, a web tool for studying protein-protein interactions. The overlapping six genes from the two previous steps may play an essential role in cognitive function. The potential gene list consists of 6 genes including CHRM1 (cholinergic receptor muscarinic 1), HRH3 (histamine receptor H3), DRD2 (dopamine receptor D2), GRM5 (glutamate metabotropic receptor 5), HTR2A (5-hydroxytryptamine receptor 2A) and SLC6A1 (solute carrier family 6 member 1). These genes were analyzed to find the related pathway using the KEGG database. Functional enrichment in KEGG reveals 6 target proteins involved in several pathways, which play an important role in cognitive function, including neuroactive ligand-receptor interaction[31-36], gap junction[37], calcium signaling pathway[38-40], and the cAMP signaling pathway.[41]

CHRM1 (Cholinergic Receptor Muscarinic 1) is a G protein-coupled receptor involved in acetylcholine binding and is implicated in cognitive function. Reduced cortical CHRM1 levels have been linked to cognitive decline and executive dysfunction in schizophrenia.[42-45] HRH3 (Histamine Receptor H3) modulates histamine signaling in the CNS and regulates neurotransmitter release; H3 receptor antagonists show potential in treating cognitive impairment in schizophrenia without extrapyramidal side effects.[42,46,47] DRD2 (Dopamine Receptor D2) participates in locomotion, reward, memory, and learning; mutations and polymorphisms in DRD2 are associated with schizophrenia risk and cognitive deficits.[42,48,49] GRM5 (Glutamate Metabotropic Receptor 5), a glutamate receptor modulating synaptic plasticity, is linked to cognitive impairments in schizophrenia and may be a promising target to reverse these deficits.[42,50,51] HTR2A (5-Hydroxytryptamine Receptor 2A) regulates neural activity and cognition; serotonergic modulation via 5-HT receptors contributes to cognitive and motor function improvements in schizophrenia treatment.[42,52,53] SLC6A1 (Solute Carrier Family 6 Member 1) encodes a GABA transporter affecting synaptic inhibition, with impairments contributing to working memory deficits in schizophrenia.[42,54]

Nine antipsychotics bind to their protein target with different action mechanisms, and most of them are receptor antagonists. It has been noted that the large group of proteins involves antipsychotics and cognitive function in schizophrenia are receptors of the

TABLE 2. List of pathways that contain potential proteins involved in antipsychotics and cognitive function in schizophrenia

No.	#Term ID	Term description	Observed gene count	Background gene count	False discovery rate	Matching proteins in the pathway
1	hsa04080	Neuroactive ligand-receptor interaction	5	272	0.000000065	CHRM1, DRD2, GRM5, HRH3, HTR2A
2	hsa04540	Gap junction	3	87	0.0000186	DRD2, GRM5, HTR2A
3	hsa04020	Calcium signaling pathway	3	179	0.0001	CHRM1, GRM5, HTR2A
4	hsa04024	cAMP signaling pathway	2	195	0.0074	CHRM1, DRD2

neuroactive ligand-receptor interaction pathway, including CHRM1 (Cholinergic Receptor Muscarinic 1), DRD2 (Dopamine Receptor D2), GRM5 (Glutamate Metabotropic Receptor 5), HRH3 (Histamine Receptor H3), and HTR2A (5-Hydroxytryptamine Receptor 2A). These receptors are G-protein-coupled receptors (GPCRs).[55,56] (Fig 4) The downstream cascade of DRD2, HRH3, and CHRM1 is adenylate cyclase (AD), which activates Cyclic AMP (cAMP) and protein kinase A (PKA) in a cAMP signaling pathway.[57,58] PKA-mediated molecule that connects the cAMP signaling pathway and the gap junction. PKA might act on a gap junction as an inhibitor.[59] Gap junctions hold intercellular channels that allow direct communication and direct transfer of small molecules between the cytosolic of neighboring cells. The communication of these channels is essential for many biological processes, such as apoptosis and differentiation. Taken together, a cAMP-PKA signaling pathway is a crucial mediator of long-term memory formation.[41] The finding of new targets of antipsychotics in the present study is consistent with the recent evidence that glutamate metabotropic receptor modulators and serotonin agonists have the potential to improve cognitive impairments in schizophrenia.[60]

The involvement of GPCRs and their associated signaling pathways, such as neuroactive ligand-receptor interaction, calcium signaling, and cAMP signaling, in cognitive processes and antipsychotic drug action is well-established in schizophrenia research.[64] This study confirms these known mechanisms; however, the scientific contribution lies in the integrative identification of six overlapping target genes (CHRM1, HRH3, DRD2, GRM5, HTR2A, and SLC6A1) through dual bioinformatics approaches, refining specific drug-gene interaction candidates relevant to cognitive dysfunction.

Distinctly, the analysis highlights potential unique drug-gene interactions by mapping commonly prescribed antipsychotics to specific gene targets, some of which have been underexplored in population-specific contexts. For example, the inclusion of HRH3 and SLC6A1 as candidate genes suggests avenues beyond the classical dopaminergic and serotonergic targets, aligning with emerging therapeutic efforts to modulate histaminergic and GABAergic systems in schizophrenia cognition.[65]

Regarding the broad nature of the enriched pathways identified, this study advances understanding by contextualizing the role of these pathways specifically in cognitive impairment among schizophrenia patients. For instance, the cAMP-PKA signaling cascade identified here is crucial for synaptic plasticity and long-term memory formation, directly linking molecular targets to known cognitive deficits in schizophrenia.[66] Furthermore, the neuroactive ligand-receptor pathway's multiple GPCR targets illustrate how antipsychotics modulate neurotransmission collectively affecting cognitive domains.

Comparatively, prior bioinformatics and network pharmacology studies in schizophrenia have predominantly

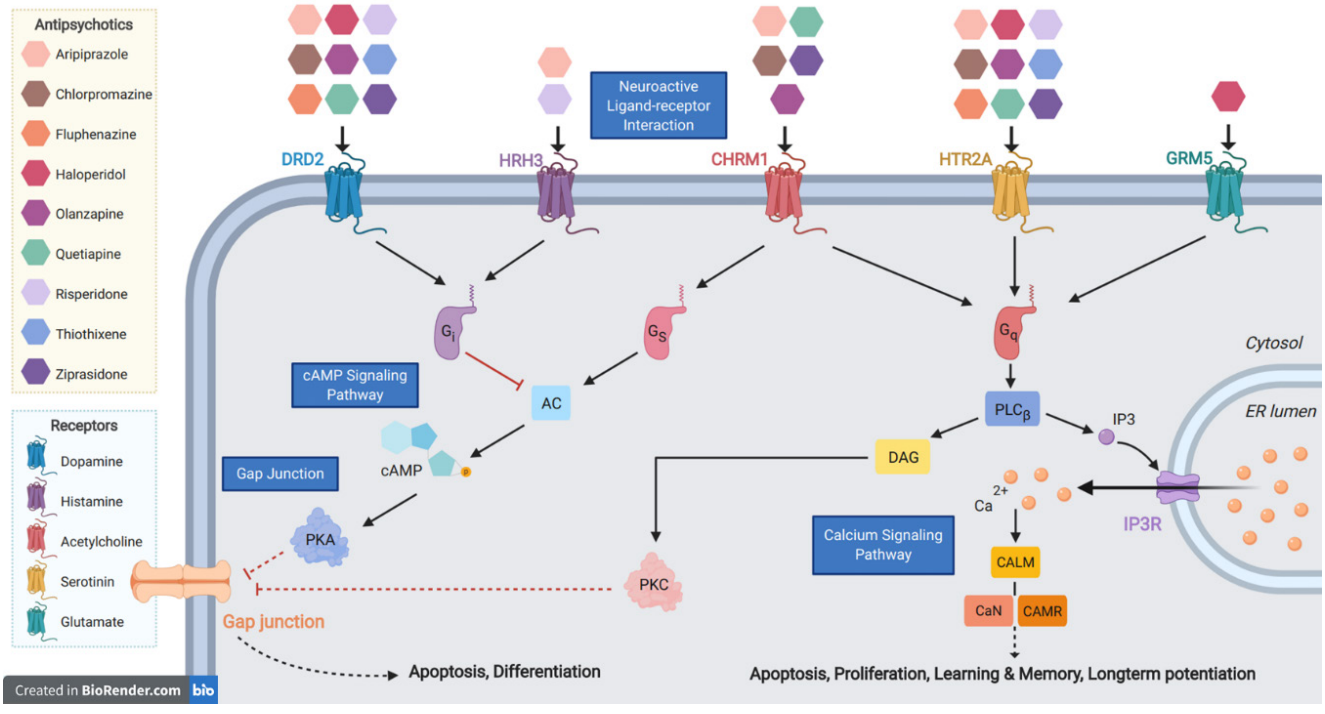


Fig 4. Schematic representation of antipsychotics' action on receptors and signal transduction pathways.

focused on genetic susceptibility or broad molecular networks without integration of specific antipsychotic drug targets.[67] This study contributes by bridging drug pharmacology and cognitive genomics, utilizing a combinatorial approach of gene target overlap, protein interaction networks, and pathway enrichment. Therefore, it provides a more targeted framework for identifying therapeutic candidates that could be further investigated in population-specific and pharmacogenetic contexts.[64]

Current evidence mostly reveals the role of antipsychotic action on psychotic symptoms and cognitive functions in schizophrenia patients.[63,68,69] However, the mechanism of action of antipsychotic drugs on GPCRs related to the neurological processes of cognitive function is limited. The results of this study highlight the interaction of the molecular mechanisms of antipsychotic drugs on the neurological processes associated with psychosis and cognitive function. Therefore, this crucial evidence might suggest the conceptual framework for further study on pharmacological and clinical applications of novel antipsychotic drugs with dual effects on psychotic symptoms and cognitive functions in schizophrenia.

CONCLUSION

The findings of this study support existing evidence that G-protein-coupled receptors (GPCRs) are critical targets through which antipsychotics exert their effects on cognitive function in schizophrenia. While the identification of enriched pathways such as neuroactive ligand-receptor interaction, calcium signaling, and cAMP signaling aligns with well-established biological mechanisms, this study refines these pathways by linking specific gene targets implicated in cognitive dysfunction in schizophrenia. Such integrative bioinformatics analyses enhance understanding of molecular interactions underlying cognitive deficits and provide a more precise molecular framework that connects antipsychotic pharmacodynamics with cognition-related genetic factors. This may facilitate the identification of candidate targets for future research aiming at developing therapies to address cognitive impairments in schizophrenia.

Data Availability Statement

Data will be made available on request.

DECLARATIONS

Grant and Funding Information

This research project is supported by National Research Council of Thailand (NRCT) and Mahidol University (grant number N42A650340 to B.C.).

Conflict of Interest

The authors declare no conflicts of interest related to this article.

Registration Number of Clinical Trial

None.

Author Contributions

Conceptualization and methodology, K.T., B.C., V.C., J.C.; Investigation, K.T.; Formal analysis, K.T., B.C., V.C.; Visualization and writing-original draft, B.C., J.C.; Writing-review and editing, B.C., J.C.; Supervision, B.C., V.C., V.S., J.C. All authors have read and agreed to the final version of the manuscript.

Use of Artificial Intelligence

None.

REFERENCES

1. Harvey PD, Lombardi J, Leibman M, White L, Parrella M, Powchik P, et al. Cognitive impairment and negative symptoms in geriatric chronic schizophrenic patients: a follow-up study. Schizophr Res. 1996;22(3):223-31.
2. Hughes C, Kumari V, Soni W, Das M, Binneman B, Drozd S, et al. Longitudinal study of symptoms and cognitive function in chronic schizophrenia. Schizophr Res. 2003;59(2-3):137-46.
3. Wu JQ, Chen DC, Tan YL, Xiu MH, De Yang F, Soares JC, et al. Cognitive impairments in first-episode drug-naive and chronic medicated schizophrenia: MATRICS consensus cognitive battery in a Chinese Han population. Psychiatry Research. 2016;238:196-202.
4. Seidman LJ, Buka SL, Goldstein JM, Tsuang MT. Intellectual decline in schizophrenia: evidence from a prospective birth cohort 28-year follow-up study. J Clin Exp Neuropsychol. 2006;28(2):225-42.
5. Friedman JI, Harvey PD, Coleman T, Moriarty PJ, Bowie C, Parrella M, et al. Six-year follow-up study of cognitive and functional status across the lifespan in schizophrenia: a comparison with Alzheimer's disease and normal aging. Am J Psychiatry. 2001;158(9):1441-8.
6. Goldberg TE, Ragland JD, Torrey EF, Gold JM, Bigelow LB, Weinberger DR. Neuropsychological assessment of monozygotic twins discordant for schizophrenia. Arch Gen Psychiatry. 1990;47(11):1066-72.
7. Riley EM, McGovern D, Mockler D, Doku VC, S OC, Fannon DG, et al. Neuropsychological functioning in first-episode psychosis--evidence of specific deficits. Schizophr Res. 2000;43(1):47-55.
8. Bilder RM, Lipschutz-Broch L, Reiter G, Geisler SH, Mayerhoff DI, Lieberman JA. Intellectual deficits in first-episode schizophrenia: evidence for progressive deterioration. Schizophr Bull. 1992;18(3):437-48.
9. Green MF. What are the functional consequences of neurocognitive deficits in schizophrenia? Am J Psychiatry. 1996;153(3):321-30.
10. Sharma T. Quetiapine-efficacy in different domains. Eur Neuropsychopharmacol. 2001;11 Suppl 4:S385-90.
11. Harvey PD, Lombardi J, Kincaid MM, Parrella M, White L,

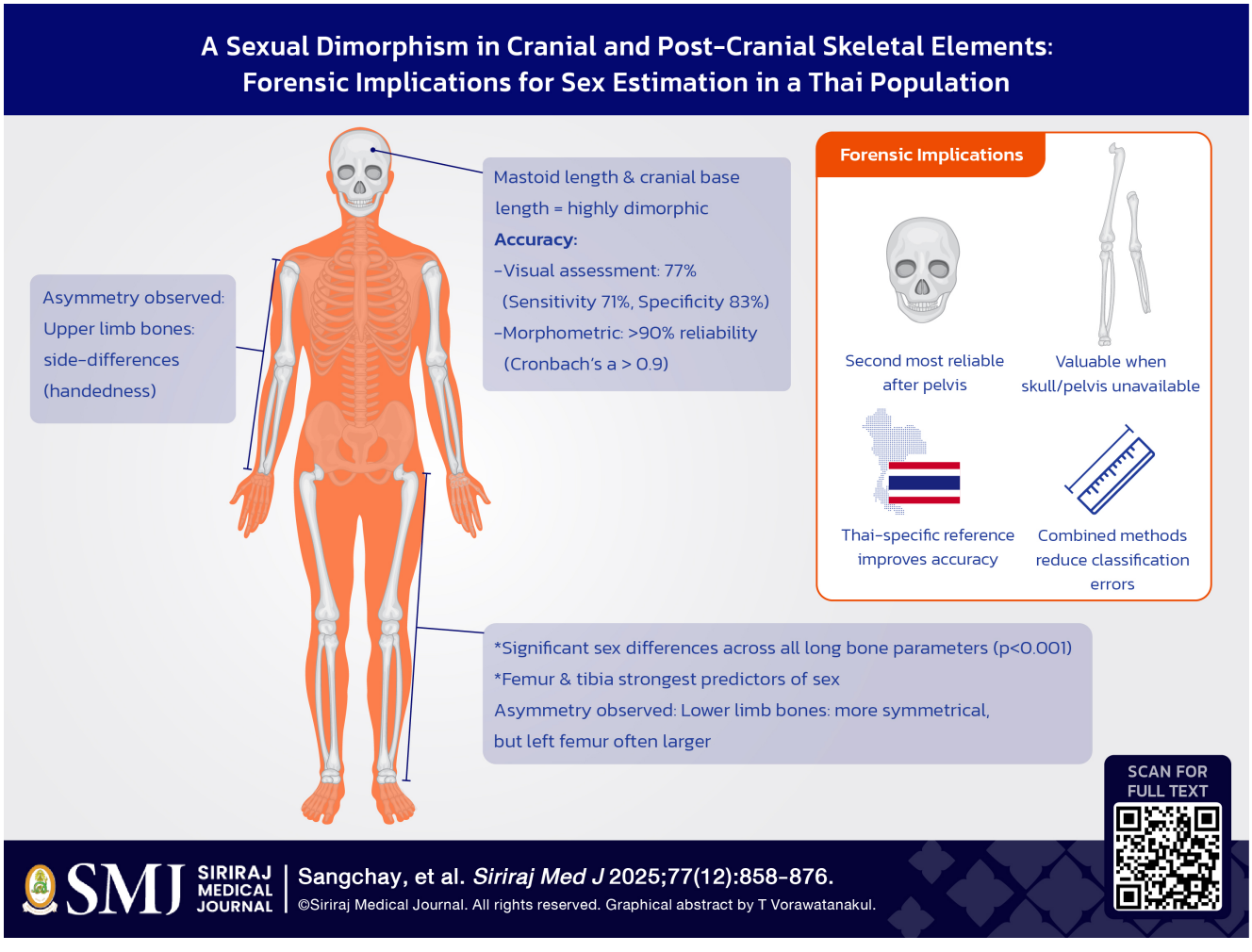
Powchik P, et al. Cognitive functioning in chronically hospitalized schizophrenic patients: age-related changes and age disorientation as a predictor of impairment. Schizophr Res. 1995;17(1):15-24.
12. Sharma T, Antonova L. Cognitive function in schizophrenia: Deficits, functional consequences, and future treatment. Psychiatr Clin North Am. 2003;26(1):25-40.
13. Faber G, Smid HG, Van Gool AR, Wiersma D, Van Den Bosch RJ. The effects of guided discontinuation of antipsychotics on neurocognition in first onset psychosis. Eur Psychiatry. 2012; 27(4):275-80.
14. Knowles EE, David AS, Reichenberg A. Processing speed deficits in schizophrenia: reexamining the evidence. Am J Psychiatry. 2010;167(7):828-35.
15. Elie D, Poirier M, Chianetta J, Durand M, Gregoire C, Grignon S. Cognitive effects of antipsychotic dosage and polypharmacy: a study with the BACS in patients with schizophrenia and schizoaffective disorder. J Psychopharmacol. 2010;24(7):1037-44.
16. Green MF, Marder SR, Glynn SM, McGurk SR, Wirshing WC, Wirshing DA, et al. The neurocognitive effects of low-dose haloperidol: a two-year comparison with risperidone. Biol Psychiatry. 2002;51(12):972-8.
17. Janicak PG, Marder SR, Tandon R, Goldman M. Schizophrenia recent advances in diagnosis and treatment. New York, NY: Springer; 2014.
18. John MECfCD, Communications S. AHRQ Comparative effectiveness reviews antipsychotic medicines for treating schizophrenia and bipolar disorder: A review of the research for adults and caregivers. Comparative Effectiveness Review Summary Guides for Consumers. Rockville (MD): Agency for Healthcare Research and Quality (US); 2005.
19. Chokhawala K, Stevens L. Antipsychotic Medications. StatPearls. Treasure Island (FL): StatPearls Publishing LLC.; 2019.
20. Andreasen NC, Pressler M, Nopoulos P, Miller D, Ho B-C. Antipsychotic Dose Equivalents and Dose-Years: A Standardized Method for Comparing Exposure to Different Drugs. Biol Psychiatry. 2010;67(3):255-62.
21. GeneCards - Human Genes | Gene Database | Gene Search [cited 2020 8 Feb]. Available from: https://www.genecards.org/.
22. Webgestalt(Web-based gene set analysis toolkit) [cited 2020 Feb 08]. Available from: http://www.webgestalt.org/.
23. STRING: functional protein association networks [cited 2020 Feb 8]. Available from: https://string-db.org/cgi/input.pl.
24. Jenraumjit R, Tethanyawarakool V, Kongkhamboot H, Piyatrakul N. Prescribing Pattern of Antipsychotic Long Acting Injections for the Treatment of Schizophrenia Patients in Suan Prung Psychiatric Hospital. J Psychiatr Assoc Thailand. 2018;63(4):335-48.
25. Sharma T. Cognitive effects of conventional and atypical antipsychotics in schizophrenia. Br J Psychiatry Suppl. 1999; 174(38):44-51.
26. Sakurai H, Bies RR, Stroup ST, Keefe RSE, Rajji TK, Suzuki T, et al. Dopamine D2 receptor occupancy and cognition in schizophrenia: analysis of the CATIE data. Schizophr Bull. 2013;39(3):564-74.
27. Knafo S, Venero C. Cognitive Enhancement: Pharmacologic, Environmental and Genetic Factors: Academic Press; 2014.
28. Dibyajyoti S, Bin ET, Swati P. Bioinformatics: The effects on the cost of drug discovery. Galle Medical Journal. 2013;18(1): 44-50.
29. Buchan NS, Rajpal DK, Webster Y, Alatorre C, Gudivada RC, Zheng C, et al. The role of translational bioinformatics in drug discovery. Drug Discov Today. 2011;16(9-10):426-34.
30. Wu J, Vallenius T, Ovaska K, Westermarck J, Mäkelä TP, Hautaniemi S. Integrated network analysis platform for protein-protein interactions. Nat Methods. 2009;6(1):75-7.
31. Bolton JL, Marioni RE, Deary IJ, Harris SE, Stewart MC, Murray GD, et al. Association between polymorphisms of the dopamine receptor D2 and catechol-o-methyl transferase genes and cognitive function. Behav Genet. 2010;40(5):630-8.
32. Wess J. Muscarinic acetylcholine receptor knockout mice: novel phenotypes and clinical implications. Annu Rev Pharmacol Toxicol. 2004;44:423-50.
33. Barco A, Bailey CH, Kandel ER. Common molecular mechanisms in explicit and implicit memory. J Neurochem. 2006;97(6):1520-33.
34. Ménard C, Quirion R. Group 1 metabotropic glutamate receptor function and its regulation of learning and memory in the aging brain. Front Pharmacol. 2012;3:182.
35. Sadek B, Saad A, Sadeq A, Jalal F, Stark H. Histamine H3 receptor as a potential target for cognitive symptoms in neuropsychiatric diseases. Behav Brain Res. 2016;312:415-30.
36. Zhang G, Stackman Jr RW. The role of serotonin 5-HT2A receptors in memory and cognition. Front Pharmacol. 2015;6:225.
37. Mitterauer BJ. Possible role of glia in cognitive impairment in schizophrenia. CNS Neurosci Ther. 2011;17(5):333-44.
38. Kawamoto EM, Vivar C, Camandola S. Physiology and pathology of calcium signaling in the brain. Front Pharmacol. 2012;3:61.
39. Suwanjang W, Phansuwan-Pujito P, Govitrapong P, Chetsawang B. Calpastatin reduces calpain and caspase activation in methamphetamine-induced toxicity in human neuroblastoma SH-SY5Y cultured cells. Neurosci Lett. 2012;526(1):49-53.
40. Joshi AU, Kornfeld OS, Mochly-Rosen D. The entangled ER-mitochondrial axis as a potential therapeutic strategy in neurodegeneration: a tangled duo unchained. Cell Calcium. 2016;60(3):218-34.
41. Titus D, Furones C, Kang Y, Atkins CM. Age-dependent alterations in cAMP signaling contribute to synaptic plasticity deficits following traumatic brain injury. Neuroscience. 2013; 231:182-94.
42. Stelzer G, Rosen N, Plaschkes I, Zimmerman S, Twik M, Fishilevich S, et al. The GeneCards suite: from gene data mining to disease genome sequence analyses. Curr Protoc Bioinformatics. 2016;54:1.30.1-1.30.33.
43. Liao D-L, Hong C-J, Chen H-M, Chen Y-E, Lee S-M, Chang C-Y, et al. Association of muscarinic m1 receptor genetic polymorphisms with psychiatric symptoms and cognitive function in schizophrenic patients. Neuropsychobiology. 2003;48(2):72-6.
44. Dean B, Soulby A, Evin GM, Scarr E. Levels of [3H] pirenzepine binding in Brodmann's area 6 from subjects with schizophrenia is not associated with changes in the transcription factor SP1 or BACE1. Schizophr Res. 2008;106(2-3):229-36.
45. Scarr E, Sundram S, Deljo A, Cowie TF, Gibbons AS, Juzva S, et al. Muscarinic M1 receptor sequence: preliminary studies on its effects on cognition and expression. Schizophr Res. 2012; 138(1):94-8.
46. Zhang M, Ballard ME, Pan L, Roberts S, Faghih R, Cowart M, et al. Lack of cataleptogenic potentiation with non-imidazole H3 receptor antagonists reveals potential drug–drug interactions

between imidazole-based H3 receptor antagonists and antipsychotic drugs. Brain Res. 2005;1045(1-2):142-9.
47. Jin C-Y, Anichtchik O, Panula P. Altered histamine H3 receptor radioligand binding in post-mortem brain samples from subjects with psychiatric diseases. Br J Pharmacol. 2009;157(1):118-29.
48. Cohen O, Weickert T, Hess J, Paish L, McCoy S, Rothmond D, et al. A splicing-regulatory polymorphism in DRD2 disrupts ZRANB2 binding, impairs cognitive functioning and increases risk for schizophrenia in six Han Chinese samples. Mol Psychiatry. 2016;21(7):975-82.
49. Nkam I, Ramoz N, Breton F, Mallet J, Gorwood P, Dubertret C. Impact of DRD2/ANKK1 and COMT polymorphisms on attention and cognitive functions in schizophrenia. PloS One. 2017;12(1):e0170147.
50. Matosin N, Newell KA, Quidé Y, Andrews JL, Teroganova N, Green MJ, et al. Effects of common GRM5 genetic variants on cognition, hippocampal volume and mGluR5 protein levels in schizophrenia. Brain Imaging Behav. 2018;12(2):509-17.
51. Moghaddam B. Targeting metabotropic glutamate receptors for treatment of the cognitive symptoms of schizophrenia. Psychopharmacology (Berl). 2004;174(1):39-44.
52. Shimizu S, Mizuguchi Y, Ohno Y. Improving the treatment of schizophrenia: role of 5-HT receptors in modulating cognitive and extrapyramidal motor functions. CNS Neurol Disord Drug Targets. 2013;12(6):861-9.
53. Blasi G, De Virgilio C, Papazacharias A, Taurisano P, Gelao B, Fazio L, et al. Converging evidence for the association of functional genetic variation in the serotonin receptor 2a gene with prefrontal function and olanzapine treatment. JAMA Psychiatry. 2013;70(9):921-30.
54. Lewis DA, Moghaddam B. Cognitive dysfunction in schizophrenia: convergence of γ-aminobutyric acid and glutamate alterations. Arch Neurol. 2006;63(10):1372-6.
55. Janušonis S. Functional associations among G protein-coupled neurotransmitter receptors in the human brain. BMC Neurosci. 2014;15:16.
56. Regard JB, Sato IT, Coughlin SR. Anatomical profiling of G protein-coupled receptor expression. Cell. 2008;135(3):561-71.
57. Sassone-Corsi P. The cyclic AMP pathway. Cold Spring Harb Perspect Biol. 2012;4(12):a011148.
58. Yan K, Gao LN, Cui YL, Zhang Y, Zhou X. The cyclic AMP signaling pathway: Exploring targets for successful drug discovery. Mol Med Rep. 2016;13(5):3715-23.
59. Pidoux G, Taskén K. Anchored PKA as a gatekeeper for gap junctions. Commun Integr Biol. 2015;8(4):e1057361.
60. Miyamoto A, Mikoshiba K. Probes for manipulating and monitoring IP3. Cell Calcium. 2017;64:57-64.
61. Feber L, Peter NL, Chiocchia V, Schneider-Thoma J, Siafis S, Bighelli I, et al. Antipsychotic Drugs and Cognitive Function: A Systematic Review and Network Meta-Analysis. JAMA Psychiatry. 2025;82(1):47-56.
62. Lee M, Cernvall M, Borg J, Plaven-Sigray P, Larsson C, Erhardt S, et al. Cognitive Function and Variability in Antipsychotic Drug–Naive First-Episode Psychosis. JAMA Psychiatry. 2024;81(5):468-76.
63. Kelebie M, Kibralew G, Tadesse G, Rtbey G, Aderaw M, Endeshaw W, et al. Effectiveness of antipsychotic medication in patients with schizophrenia in a real world retrospective observational study in Ethiopia. Sci Rep. 2025 Feb 7.
64. Nnadi CU, Malhotra AK. Individualizing Antipsychotic Drug Therapy in Schizophrenia: The Promise of Pharmacogenetics. Curr Psychiatry Rep. 2005;7(1):1-7.
65. Begni V, Marchesin A, Riva MA. IUPHAR review-Novel therapeutic targets for schizophrenia treatment: A Translational perspective. Pharmacol Res. 2025;214:107690.
66. Titus DJ, Furones C, Kang Y, Atkins CM. Age-dependent alterations in cAMP signaling contribute to synaptic plasticity deficits following traumatic brain injury. Neuroscience. 2013;231:182-94.
67. Faustmann TJ, Corvace F, Faustmann PM, Ismail FS. Influence of antipsychotic drugs on microglia-mediated inflammatory responses in schizophrenia. Front Psychiatry. 2025;16:1522128.
68. Pitanupong J, Karakate A, Tepsuan L, Sritrangnant G. Attitudes toward long-acting injectable antipsychotics among schizophrenia patients in southern Thailand: A multihospital-based cross-sectional survey. Siriraj Med J. 2022;74(3):193-201.
69. Ueapanjasin P, Thavornwattanayong W, Lertsirimunkong J, Chaiyakittisopon K.Cost-effectiveness analysis of long-acting injectable once-monthly of aripiprazole compared with long-acting injectable once-monthly paliperidone palmitate for the treatment of stable schizophrenia patients in Thailand. Siriraj Med J. 2023;75(10):725-35.

Sexual Dimorphism in Cranial and Post-Cranial Skeletal Elements: Forensic Implications for Sex Estimation in a Contemporary Thai Population

Napakorn Sangchay, M.D., Ph.D.[1], Kavin Tangmanpakdeepong, B.Sc.[1,3], Saidontree Boonyarud, B.Sc.[1,3], Sujinthara Wansopha, B.Sc.[1,3], Natipong Chatthai, M.Sc.[1,2], Patara Rattanachet, Ph.D.[1], Jirapa Chetsawang, M.D., Ph.D.[1,*]

[1]*Department of Anatomy, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand,* [2]*College of Medical Science, Western University, Bangkok, Thailand,* [3]*Mahidol University International College, Mahidol University, Nakhon Pathom, Thailand.*



*Corresponding Author: Jirapa Chetsawang
E-mail: napakorn.sac@mahidol.ac.th
Received 25 August 2025 Revised 17 September 2025 Accepted 20 September 2025
ORCID ID:http://orcid.org/0000-0002-7776-6456
https://doi.org/10.33192/smj.v77i12.277229

ABSTRACT

Objective: The analysis of human skeletal remains is instrumental in forensic and anthropological contexts, particularly for establishing biological profiles of unidentified individuals. Determining sex via skeletal examination is a fundamental component of this process and traditionally involves morphological assessment and metric analysis of pelvic and cranial bones. Nevertheless, the precision and reliability of these methodologies—whether through morphological evaluation or morphometric analysis—remain subjects of ongoing debate and scrutiny. This research set forth to investigate the efficacy of morphological and morphometric analysis in sex estimation by focusing on cranial and post-cranial long bones within contemporary Thai population.

Materials and Methods: The study sample comprised 204 skulls (105 from males, and 99 from females) and 200 sets of long bones of upper (humerus, radius, and ulna) and lower extremities (femur and tibia). Multiple measurements were systematically obtained from various anatomical regions of each bone, and measurements of extremity long bones were compared between the left and right sides.

Results: The analysis revealed statistically significant differences in these metrics between sexes, indicating the potential utility of this approach for sex classification. However, despite achieving high levels of accuracy, the studied methodology yielded some classification errors, which suggests some potential limitations.

Conclusion: The findings suggest that both inherent skeletal variability in cranial and post-cranial morphology within contemporary Thai population and the specific analytical techniques employed can markedly influence the accuracy of sex determination. These factors harbor and confer important implications for forensic and anthropological applications.

Keywords: Sexual dimorphism; cranial and post-cranial skeletal elements; forensic implications for sex estimation; contemporary Thai population (Siriraj Med J 2025; 77: 858-876)

INTRODUCTION

In forensic anthropology, constructing a biological profile for unidentified human remains is crucial for identification, including sex, age, stature, and ancestry estimations. These parameters are essential for matching ante-mortem records and other data. Notably, a swift and organized skeletal analysis for sex determination is vital for refining subsequent biological profile estimations.[1]

Skeletal sex estimation is a critical step that requires both expertise and precise decision-making due to its foundational role in the identification process. Sexual dimorphism in skeletal elements results from genetic and hormonal influences, which contribute to distinct traits differentiating male and female skeletons. To enhance accuracy and minimize subjectivity, sex assessments should be conducted and documented meticulously, thereby strengthening the identification process. The distinct skeletal traits influenced by genetic and hormonal factors differentiate males and females, underscoring the need for reliable and less subjective assessment techniques.[2]

In forensic investigations, accurate sex assessment from skeletal remains is essential for human identification. The pelvis and cranium are preferred for analysis due to their pronounced sexual dimorphism, enabling precise sex estimation. The os coxae is particularly favored for its marked dimorphism and high accuracy. However, when these elements are absent or severely fragmented, alternative bones must be considered. Human long bones, or postcranial elements, serve as a secondary option due to their sexual dimorphism, though morphometric methods using long bones are limited, especially in databases focusing on Asian genetic and ethnic backgrounds.[3,4]

The pelvis and skull are traditionally considered the most reliable skeletal components for sex determination due to well-defined dimorphism.[5] Accurate assessment requires complete or intact structures, often compromised in forensic settings due to taphonomic degradation or recovery limitations. In such cases, long bones are valuable substitutes due to their sexually dimorphic features influenced by musculoskeletal development and hormonal regulation. However, they lack specific morphological traits indicative of sex, limiting their reliability for sex estimation.[6]

Numerous studies have examined the use of the human skull and long bones of the extremities for sex estimation. However, a critical limitation of many such investigations is the lack of intra- and inter-observer error assessments, which impedes a thorough evaluation of the reliability of the methodology being studied or the methodologies being compared.[7] Cranial visual assessment of or for sexually dimorphic traits offers a rapid and practical method for sex estimation, but this

technique is highly subjective and less reproducible. This approach, particularly when evaluating the skull, involves analyzing varying morphological characteristics to distinguish between sexes, focusing on features such as the mastoid process, nuchal crest, and orbital traits.[8]

In addition to pelvic morphology, features in long bones, such as femoral head diameter, can indicate biological sex and are suitable for morphometric analysis. This method is effective with complete or nearly complete skeletal elements, but its reliability diminishes with fragmented remains due to taphonomic damage. Thus, exploring alternative methodologies, like digital imaging techniques to analyze long bone characteristics, becomes necessary.[9,10]

While gross morphological analysis of the skull offers acceptable accuracy for sex estimation, it also has some associated limitations as follows. This technique is susceptible to intra- and inter-observer variability, leading to inconsistent assessments. The reliance on subjective judgment complicates its utility, particularly in skeletal specimens with ambiguous features. To enhance precision and reliability, direct pelvic morphometric measurements are proposed as a more refined and reliable approach.[11,12]

Metric analysis of the skull provides greater precision in sex determination, often surpassing 90% accuracy, compared to visual morphological assessments. However, this method requires an intact or nearly complete skull to accurately assess landmarks. The mastoid process, in particular, demonstrates high reliability, with a sex determination success rate of 96.5%, supported by studies reporting approximately 90% accuracy from measurements in this region.[13]

Both non-metric and metric cranial analyses depend on identifiable skeletal features, which necessitate well-preserved bones. Current methodologies primarily derive from skeletal collections of specific populations, such as Black and White Americans, with limited data for Asian and other groups.[14] Researchers underscore the critical need for population-specific databases to improve sex estimation accuracy in diverse populations.[15]

Asymmetry between the bones on the left and right sides is attributed to handedness. Approximately 90% of the population is right-handed. Handedness influences bone morphology, with the dominant limb generally exhibiting larger bones. Forensic studies also show that the right side, including hand bones, tends to be larger, particularly in right-handed individuals.[16] Larger bones in the right hand are more common in right-handed individuals, with this difference being more pronounced in males. However, similar correlations are less consistent for left-handed individuals.[17]

This study evaluates sexual dimorphism in cranial and post-cranial elements for forensic sex estimation, assessing inter-observer reliability and error rates using gross morphological and metric approaches. Conducted in contemporary Thai population with dry skulls and extremity long bones, this study explores the method's applicability in forensic contexts. Additionally, morphometric analysis of the bilateral long bones examines the impact of asymmetry on sex estimation, considering the influence of sexual dimorphism.

MATERIALS AND METHODS

The study utilized human bone samples from the Siriraj Anatomical and Anthropological Bone Centre (Si-AABRC), of the Department of Anatomy, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand. The protocol for this study was exempted from Siriraj Institutional Review Board (SIRB) approval.

Cranial sexual dimorphism and morphometric measurements

Gross morphological assessment was conducted for each bone, focusing on traits such as the mastoid process, nuchal crest, and cranial contour, to estimate sex. Each cranium underwent detailed visual evaluation of specific dimensions and landmarks, including the supraorbital ridges, opisthocranion, and mastoid process. A sex classification was made if at least three traits indicated a particular sex. The supraorbital ridges were assessed for size and prominence, the opisthocranion for its most posterior point, and the mastoid process for size and robustness. All skulls were sequentially numbered and segregated to ensure unbiased subsequent craniometric analysis.

Following morphological assessment, cranial morphometric measurements were performed, as detailed in Table 1. An independent re-evaluation by two blinded investigators was conducted to enhance accuracy and reduce inter-observer variability. The craniometric parameters assessed are illustrated in Fig 1, with both observers unaware of prior measurements and the biological sex of each evaluated skull.

Post-cranial morphometric measurements

The upper extremity long bones included the humerus, radius, and ulna, on both sides, and the lower extremity bones included the femur and tibia, on both sides. Bones with deformities, fractures, osteophytes, or prosthetic modifications were excluded. Measurements were made using an osteometric board and digital vernier calliper. Post-cranial measurements are summarized in Table 2.

TABLE 1. Description of cranial landmarks and morphometric measurements.

Measurement of Cranium	Description
Cranial Length (Front to Back)	Measurement of distance from glabella to opisthocranion
Cranial Width	Measurement of distance from left eurion to right eurion
Cranial Index	(Cranial Width/Cranial Length) x 100
Cranial Base Length	Measurement of distance from nasion to opisthocranion
Mastoid Height	Measurement of porion to end of mastoid process
Inter-orbital Foramen Length	Measurement of supraorbital foramen or notch to infraorbital foramen
Inter-supraorbital Foramen Distance	Measurement of left supraorbital foramen or notch to right
Inter-infraorbital Foramen Distance	Measurement of left infraorbital foramen to right

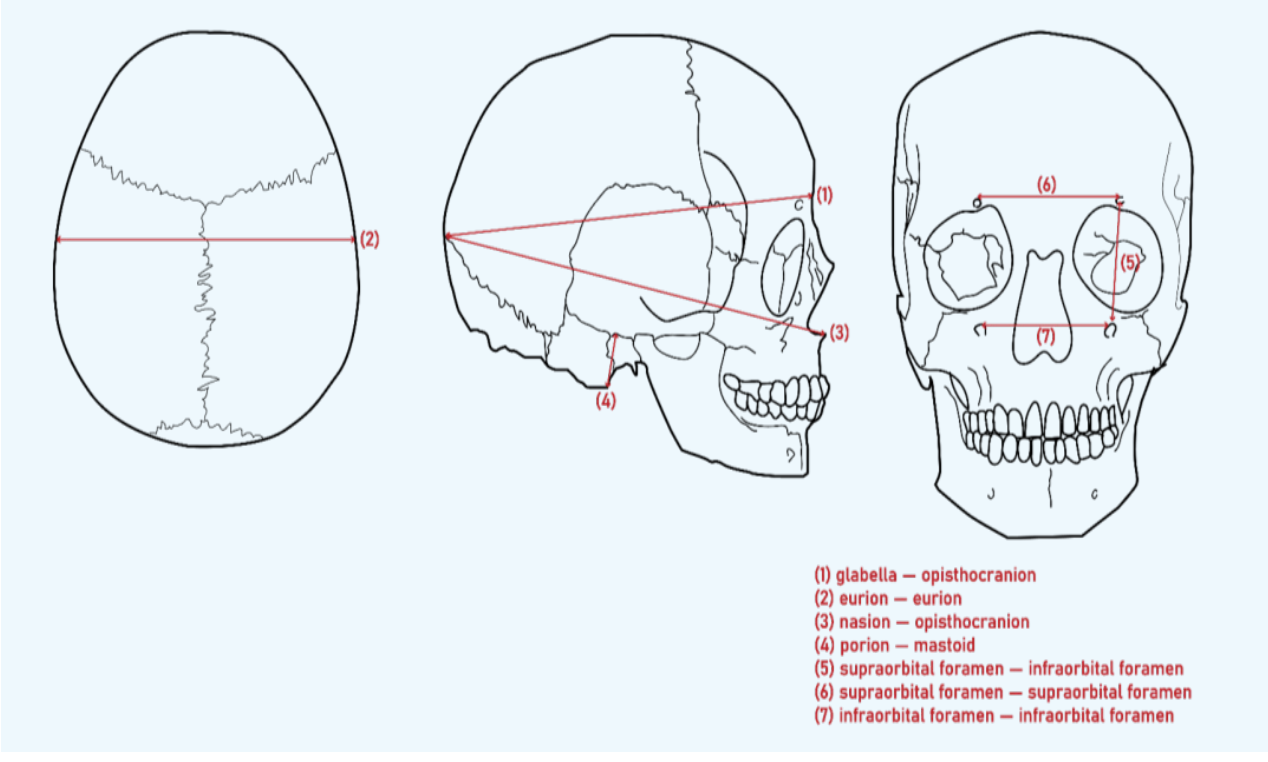


Fig 1. Cranial parameters used in cranial measurement method. All measurements were performed using the same digital vernier calliper and the same spreading calliper.

Statistical analyses

Cranial skeletal elements:

Prior to parameter measurements and analysis, two independent observers measured over 200 dry crania, yielding 204 viable samples. The study compared cranial measurements between sexes and evaluated classification accuracy via visual assessments measurement reliability. Statistical comparisons used either the independent t-test or Mann-Whitney U test, depending on data normality, to identify significant sexual dimorphism. The primary goal was to identify cranial features with pronounced sexual dimorphism.

A contingency table assessed classification accuracy, comparing predicted sex with actual sex. Metrics such as sensitivity, specificity, precision, and overall accuracy evaluated the sex determination method. Measurement reliability was assessed with Cronbach's alpha, and intraclass correlation coefficients (ICC) were calculated for normally

TABLE 2. Description of post-cranial long bone landmarks and morphometric measurements.

Measurement of Humerus	Abbreviation	Description
Maximum Anatomical Length	H-MAL	Distance between the proximal end of humerus head to the distal end of trochlea
Humeral Head diameter	HHD	Distance between superior part of humerus head to the inferior humerus head
Anatomical Neck Diameter	AND	Maximum diameter of anatomical neck
Surgical Neck Diameter	SND	Maximum diameter of surgical neck
Humerus Midshaft Diameter	H-MD	Maximum diameter of midshaft
Epicondylar Breadth	EB	Distance between medial epicondyle to lateral epicondyle
Capitular Diameter	CD	Distance between superior to inferior of capitulum
Olecranon Fossa Diameter	OFD	Distance of the maximum diameter of olecranon fossa
Measurement of Radius	**Abbreviation**	**Description**
Maximum Anatomical Length	R-MAL	Distance between the proximal end of the head to the distal end of radius
Radial Head Diameter	RHD	Maximum diameter of head
Radial Neck Diameter	RND	Maximum diameter of surgical neck
Radial Midshaft Diameter	R-MD	Maximum diameter of midshaft
Distal Radial Breadth	DRB	Distance between styloid process to ulnar notch
Distal Radial Width	DRW	Distance between superior of distal breadth to inferior of distal breath
Measurement of Ulna	**Abbreviation**	**Description**
Maximum Anatomical Length	U-MAL	Distance between the proximal end of olecranon process to distal end of styloid process
Ulna Midshaft Diameter	U-MD	Maximum diameter of midshaft
Maximum Olecranon Process Width	MOPW	Distance between maximum of olecranon process from medial to lateral side
Maximum Coronoid Process Width	MCPW	Distance between maximum of coronoid process from medial to lateral
Measurement of Femur	**Abbreviation**	**Description**
Maximum bone length	Fe-MBL	Maximum length of femur measured from the most superior part of the head of femur to the most inferior part of the furthest condyle
Maximum anatomical length	Fe-MAL	Maximum length of femur measured from the most superior part of the head of femur to the bicondylar contact point
Femoral midshaft diameter	Fe-MD	Maximum transverse diameter of femoral shaft measured at 50% length of Fe-MAL
Femoral head diameter	Fe-HD	Maximum diameter of the head of femur measured from the most superior to the most inferior points
Femoral neck diameter	Fe-ND	Maximum diameter of the neck of femur measured at the narrowest part from the most superior to the most inferior points
Bicondylar breadth	Fe-BB	Maximum length that crosses both femoral condyles, measured from the medialmost to the lateralmost points
Medial condylar width	Fe-MCW	Maximum width of medial condyle measured in vertical line from the most superior to the most inferior part of articular surface
Lateral condylar width	Fe-LCW	Maximum width of lateral condyle measured in vertical line from the most superior to the most inferior part of articular surface

TABLE 2. Description of post-cranial long bone landmarks and morphometric measurements. (Continue)

Measurement of Tibia	Abbreviation	Description
Maximum bone length	T-MBL	Maximum length of tibia measured from the most superior point of the furthest intercondylar tubercle on the tibial eminence to the most inferior point of the medial malleolus
Maximum anatomical length	T-MAL	Maximum length of tibia measured from the tibial bicondylar contact point to the most inferior point of the medial malleolus
Tibial midshaft diameter	T-MD	Maximum transverse diameter of tibial shaft measured at 50% length of T-MAL
Tibial plateau breadth	TPB	Maximum width of tibial plateau measured from the medialmost part of medial condyle to the lateralmost of lateral condyle
Tibial plateau medial width	TPMW	Maximum width of medial condyle measured from the most posterior to the most anterior part
Tibial plateau lateral width	TPLW	Maximum width of lateral condyle measured from the most posterior to the most anterior part
Distal tibial breadth	DTB	Maximum distance measured from the medialmost point of the medial malleolus to the medialmost point of the fibular notch
Distal tibial width	DTW	Maximum distance between the most anterior part to the most posterior part of inferior articular facet of distal tibia, measured at the middle of the inferior articular surface

distributed variables. ICC estimates the proportion of variance due to true differences, with reliability ranges, ensuring measurement consistency for valid comparisons.

Post-cranial skeletal elements:

The independent t-test compared two groups, with statistical significance set at $p<0.001$ ($\alpha=0.05$). Paired t-tests evaluated intra-individual differences between the left and right sides. Discriminant function analysis assessed sex prediction accuracy, selecting functions with the highest classification performance and significant variables.

Observation error and reliability analysis

Inter-observer measurement error was evaluated. Two observers independently measured the skeletal samples: the primary investigator and a faculty member of the Department of Anatomy with 10 years of experience. Each observer assessed all parameters blinded to the demographic data. The goal was to quantify measurement error and repeatability, as analyzed by SPSS Statistics software (SPSS, Inc., Chicago, IL. USA). Repeatability was evaluated using technical error of measurement (TEM), relative TEM (rTEM), and coefficient of reliability (R). An R-value approaching 1 indicates high agreement.[18]

RESULTS

Descriptive statistics for all measurements are summarized in Tables 3-5, including means and standard deviations. All measurements showed significant differences between males and females (all $p<0.05$).

Cranial skeletal elements

Table 3 provides descriptive statistics for cranial variables by sex. The independent t-test revealed that males have significantly larger cranial length, width, cranial index (Fig 2-A&B), cranial base length (Fig 2C), mastoid length (Fig 2D), inter-orbital foramen distance (Fig 2E), and inter-infraorbital foramen distance (Fig 2G). The Mann-Whitney U test additionally showed that the inter-supraorbital foramen distance (Fig 2F) was significantly larger in males ($p=0.005$), indicating strong sexual dimorphism suitable for sex estimation.

Diagnostic accuracy of visual assessment

Table 4 summarizes classification results: 75 male skulls correctly identified, 17 female skulls misclassified as male, 30 male skulls misclassified as female, and 82 female skulls correctly classified. The method's sensitivity was 71.43% (95% confidence interval [95% CI]: 61.79–79.82), and specificity was 82.83% (95% CI: 73.94–89.67), with a

TABLE 3. Descriptive statistics of cranial landmarks and morphometric variables ([a] Independent *t*-test, [b] Mann-Whitney U test).

Descriptive statistics: Measurements of cranial landmarks					
	Male (n=105)		Female (n=99)		
Morphometric parameters	**Mean**	**Std Dev**	**Mean**	**Std Dev**	***p*-value**
Cranial Length	17.15	0.73	16.24	0.61	< 0.001[a]
Cranial Width	14.02	0.61	13.59	0.65	< 0.001[a]
Cranial Index	81.83	4.34	83.74	4.36	< 0.01[a]
Cranial Base Length	17.02	0.77	16.18	0.75	< 0.001[a]
Mastoid Height	34.63	3.19	31.31	3.32	< 0.001[a]
Inter-orbital Foramen Length	44.26	2.75	43.53	2.72	0.056[a]
Inter-supraorbital Foramen Distance	113.8	4.56	90.52	5.12	0.005[b]
Inter-infraorbital Foramen Distance	59.77	4.23	57.03	4.07	< 0.001[a]

[a] Independent *t*-test, [b] Mann-Whitney U test

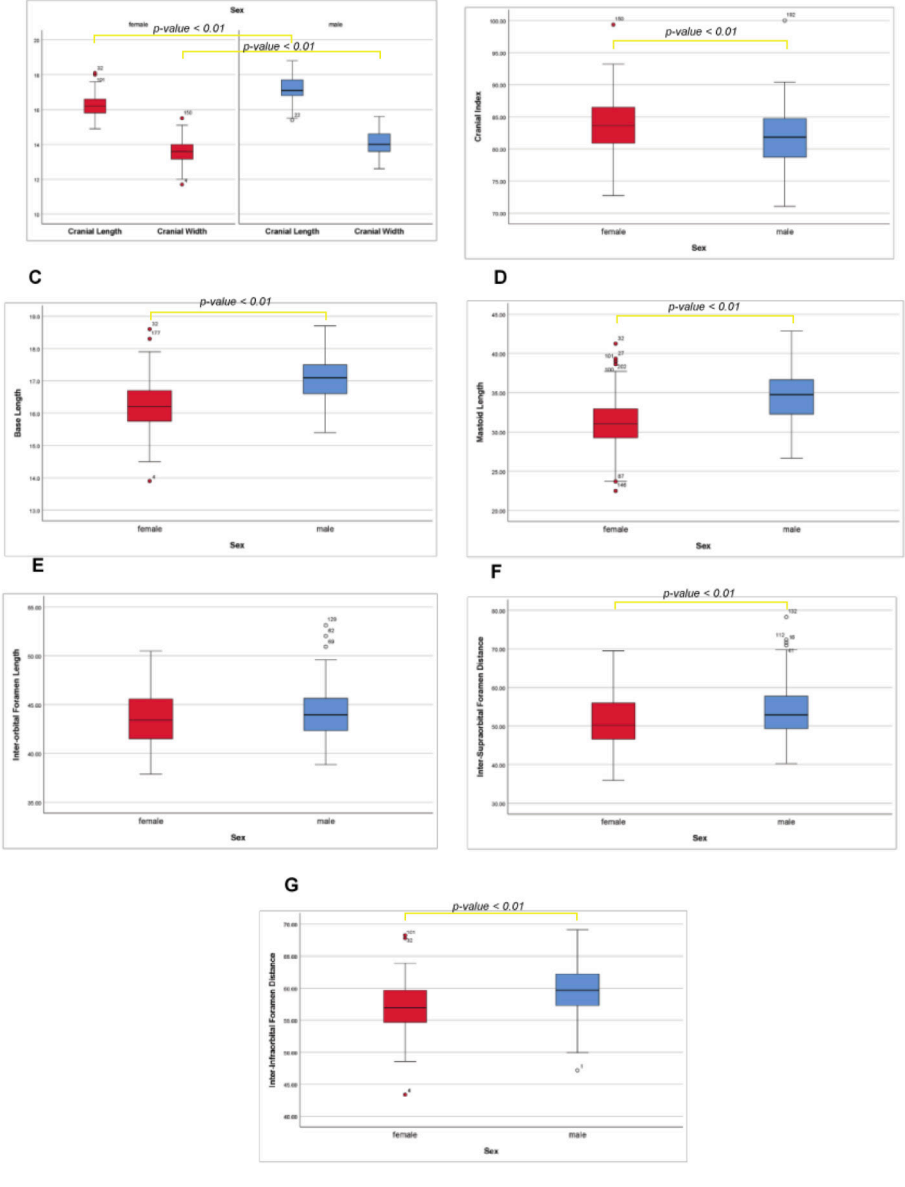


Fig 2. Comparative analysis of craniometric parameters. The figure displays a box and whisker plot comparing the means of craniometric parameters compared between male (n=105) and female (n=99) crania, namely: (A) Cranial length and width. (B) Cranial Index. (C) Cranial base length. (D) Mastoid length. (E) Inter-orbital foramen length. (F) Inter-supraorbital foramen distance. (G) Inter-infraorbital foramen distance. Significance levels ($p<0.01$) are indicated for each parameter.

TABLE 4. Contingency table of sex classification and diagnostic accuracy by visual assessment and reliability of cranial morphometric assessment.

Contingency table of sex classification by visual assessment		
	Actual Male	**Actual Female**
Predicted Male	75	17
Predicted Female	30	82

Diagnostic accuracy test of sex classification by visual assessment		
Statistic	**Value**	**95% C.I.**
Sensitivity	71.43%	61.79%–79.82%
Specificity	82.83%	73.94%–89.67%
Positive Likelihood Ratio	4.16	2.65–6.52
Negative Likelihood Ratio	0.34	0.25–0.47
Positive Predictive Value	81.52%	44.39%–58.51%
Negative Predictive Value	73.21%	66.60%–78.93%
Accuracy	76.96%	70.57%–82.55%

Reliability of craniometric parameters with Cronbach's alpha		
Parameter	**Cronbach's alpha**	**Interpretation**
Cranial Length (cm)	0.97	Excellent
Cranial Width (cm)	0.87	Good
Cranial Index	0.86	Good
Cranial Base Length (cm)	0.89	Good
Mastoid Length (mm)	0.83	Good
Inter-orbital Foramen Distance (mm)	0.92	Excellent
Inter-supraorbital Foramen Distance (mm)	0.87	Good
Inter-infraorbital Foramen Distance (mm)	0.91	Excellent

TABLE 5. Descriptive statistics of post-cranial landmarks and morphometric variables.

	Humerus Right					Left				
	Male (n=106)		Female (n=94)			Male (n=106)		Female (n=94)		
Morphometric parameters	Mean	Std Dev	Mean	Std Dev	*p*-value	Mean	Std Dev	Mean	Std Dev	*p*-value
HMAL	30.66	1.38	28.52	1.91	< 0.001	30.66	1.38	28.52	1.91	< 0.001
HHD	45.08	3.22	40.16	2.37	< 0.001	44.79	3.31	39.59	4.11	< 0.001
AND	45.68	2.70	40.54	2.44	< 0.001	45.37	2.76	40.23	2.73	< 0.001
SND	33.73	2.36	29.96	2.37	< 0.001	32.74	2.45	29.22	2.49	< 0.001
HMD	21.79	2.03	19.35	1.97	< 0.001	21.51	2.04	18.84	1.83	< 0.001
EB	60.59	5.46	54.53	3.49	< 0.001	59.94	5.14	53.81	3.32	< 0.001
CD	21.15	4.29	18.41	1.24	< 0.001	20.69	1.54	18.43	1.41	< 0.001
OFD	21.71	2.36	19.94	2.03	< 0.001	22.15	2.04	20.56	1.72	< 0.001
	Radius Right					Left				
	Male (n=106)		Female (n=94)			Male (n=106)		Female (n=94)		
Morphometric parameters	Mean	Std Dev	Mean	Std Dev	*p*-value	Mean	Std Dev	Mean	Std Dev	*p*-value
RMAL	24.35	1.27	22.49	1.65	< 0.001	24.25	1.26	22.26	1.30	< 0.001
RHD	22.31	1.94	19.59	1.39	< 0.001	22.1	1.45	19.48	2.77	< 0.001
RND	15	1.47	13.13	2.47	< 0.001	15.42	2.11	13.32	1.59	< 0.001
RMD	15.02	1.38	13.83	2.39	< 0.001	14.5	1.29	13.23	1.69	< 0.001
DRB	27.3	2.14	24.82	2.68	< 0.001	26.81	2.17	24.18	1.63	< 0.001
DRW	20.19	1.97	17.41	1.55	< 0.001	18.87	2.08	16.69	1.71	< 0.001
	Ulna Right					Left				
	Male (n=106)		Female (n=94)			Male (n=106)		Female (n=94)		
Morphometric parameters	Mean	Std Dev	Mean	Std Dev	*p*-value	Mean	Std Dev	Mean	Std Dev	*p*-value
MOPW	24.68	2.40	20.86	1.79	< 0.001	24.25	2.19	20.55	1.87	< 0.001
MCPW	23.32	2.18	20.32	1.75	< 0.001	23.05	2.16	20.09	1.65	< 0.001
UMD	16.08	1.44	14.31	1.67	< 0.001	15.77	1.40	13.87	1.69	< 0.001
	Femur Right					Left				
	Male (n=106)		Female (n=94)			Male (n=106)		Female (n=94)		
Morphometric parameters	Mean	Std Dev	Mean	Std Dev	*p*-value	Mean	Std Dev	Mean	Std Dev	*p*-value
Fe-MAL (cm)	42.9	2	40.1	1.8	< 0.001	43.1	2.1	40.3	1.8	< 0.001
Fe-MD (mm)	26.91	2.08	24.62	1.85	< 0.001	26.94	2.15	24.6	1.84	< 0.001
Fe-HD (mm)	45.41	2.52	40.09	2.01	< 0.001	45.15	2.45	39.89	1.96	< 0.001
Fe-ND (mm)	31.25	2.22	27.09	2.04	< 0.001	30.74	2.43	26.79	2.09	< 0.001
Fe-BB (mm)	78.91	4.98	70.6	3.78	< 0.001	78.45	4.89	69.92	3.64	< 0.001
Fe-MCW (mm)	35.39	3.41	31.54	2.47	< 0.001	36.27	3.19	32.85	2.31	< 0.001
Fe-LCW (mm)	32.87	2.86	29.09	1.94	< 0.001	34.57	2.84	30.82	2.27	< 0.001
	Tibia Right					Left				
	Male (n=106)		Female (n=94)			Male (n=106)		Female (n=94)		
Morphometric parameters	Mean	Std Dev	Mean	Std Dev	*p*-value	Mean	Std Dev	Mean	Std Dev	*p*-value
T-MAL (cm)	35.8	1.9	33.6	1.7	< 0.001	36	1.9	33.6	1.7	< 0.001
T-MD (mm)	25.21	2.65	22.05	2.18	< 0.001	23.93	2.49	21.36	1.97	< 0.001
TPB (mm)	74.2	4.5	66.99	3.95	< 0.001	74.19	4.36	67.05	3.88	< 0.001
TPMW (mm)	48.9	3.49	44.15	2.72	< 0.001	48.35	3.37	43.68	2.95	< 0.001
TPLW (mm)	43.47	3.07	39.05	2.72	< 0.001	42.92	2.88	38.61	2.69	< 0.001
DTB (mm)	44.24	2.85	39.65	2.41	< 0.001	44.06	2.75	39.59	2.32	< 0.001
DTW (mm)	36.51	2.54	32.65	1.85	< 0.001	36.49	2.39	32.58	1.94	< 0.001

positive likelihood ratio of 4.16 and a negative likelihood ratio of 0.34. The positive predictive value (PPV) was 81.52%, negative predictive value (NPV) was 73.21%, and overall accuracy was 76.96%.

Extremity long bones

Descriptive statistics (Table 5) and boxplots (Fig 3) show significant sex differences ($p<0.001$) across all humeral, radial, and ulnar parameters. Table 5 displays descriptive statistics for femoral and tibial measurements. Independent *t*-tests (Fig 5) confirmed significant sex differences for all parameters ($p<0.001$).

Bilateral comparisons

Table 5 presents bilateral measurements. Paired *t*-tests (Fig 4) showed significant symmetry, except for HuSND, HuHMD, HuEB, HuOFD, RaDRW and UlMAL in females; and, HuHMAL, HuSND, RaRMD, RaDRB, RaDRW and UlUMAL in males.

Paired *t*-tests (Fig 6) showed significant asymmetry, except for Fe-MD ($p=0.847$ in females, and $p=0.817$ in males). In tibiae, males demonstrated greater variability, with females showing differences in T-MD ($p<0.001$), T-PMW ($p=0.008$), T-PLW ($p=0.022$), while males differed significantly in five parameters, all $p<0.01$, as illustrated in Fig 6.

Reliability analysis: Technical error of measurement

The coefficient of reliability (R) for both cranial and post-cranial measurements is shown in Tables 4 and 6, respectively.

Reliability of craniometric measurements

Table 4 reports inter-observer reliability. Cronbach's alpha indicated high internal consistency, with parameters such as cranial length, inter-orbital distance, and inter-infraorbital distances all exceeding 0.90. Other metrics, including cranial width, cranial index, cranial base length, mastoid length, and supraorbital distances, showed good reliability (α 0.75-0.90), confirming measurement consistency.

Reliability of osteometric measurements of the extremity long bones

Table 6 illustrates the R-values for this interobserver reliability analysis. Most upper and lower extremity long bone morphometrics show a reliability above the cut-off point, (high to very high). However, two tibial morphometrics exhibit a reliability below the cut-off point, including TPLW ($r=0.64$) and DTW ($r=0.64$).

DISCUSSION

Sex estimation from skeletal remains can be achieved

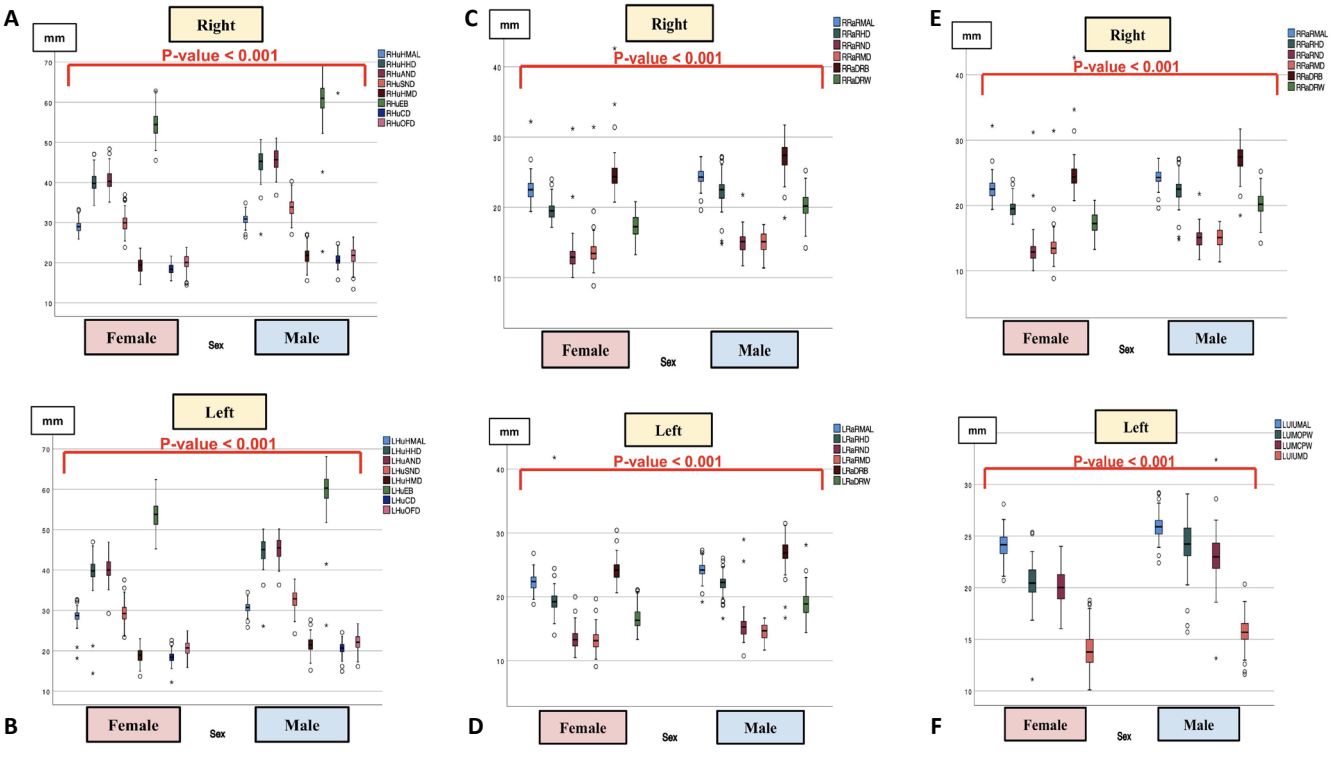


Fig 3. Boxplot comparing the results of the independent *t*-test between males and females. (A) Right side of humerus. (B) Left side of humerus. (C) Right side of radius. (D) Left side of radius. (E) Right side of ulna. (F) Left side of ulna.

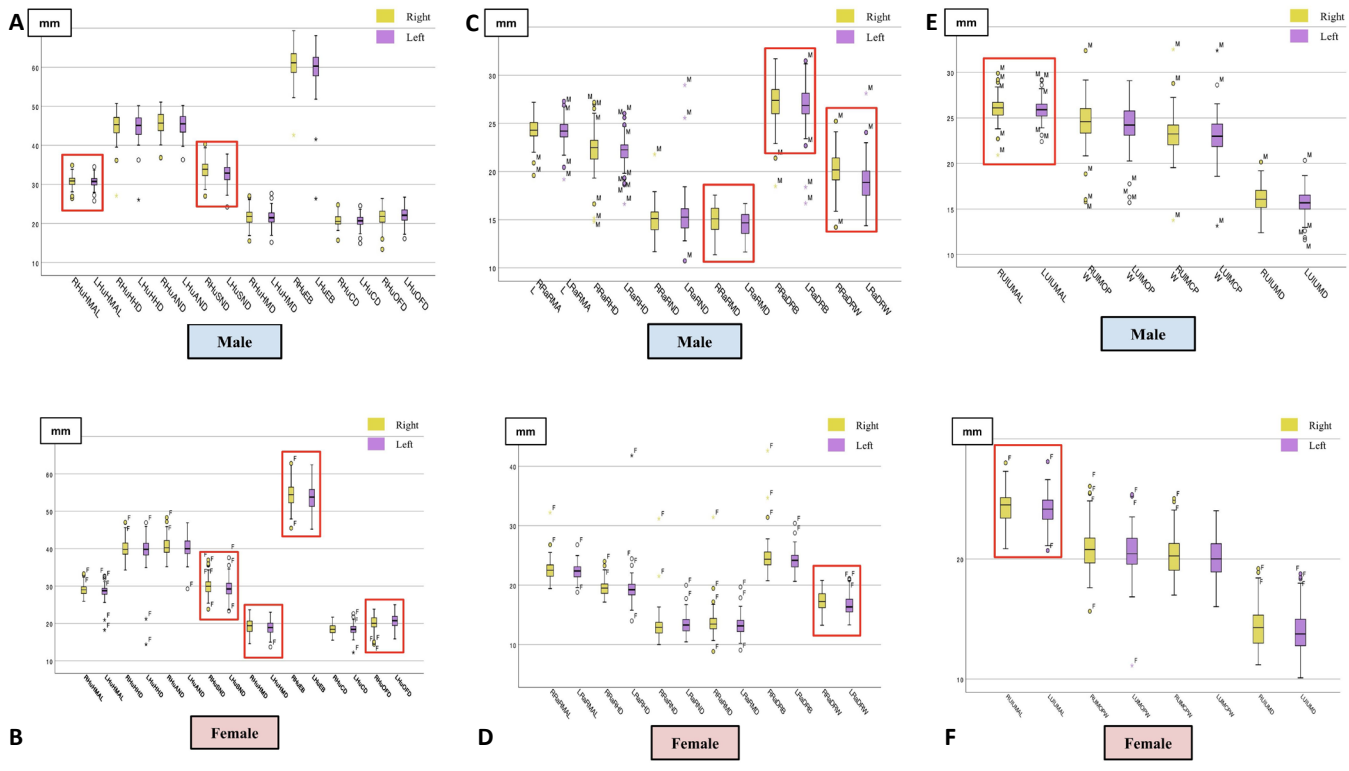


Fig 4. Boxplot comparing the results of the paired samples t-test between the left and right sides. (A) Humerus in males. (B) Humerus in females. (C) Radius in males. (D) Radius in females. (E) Ulna in males. (F) Ulna in females. (Red box indicates $p<0.001$).

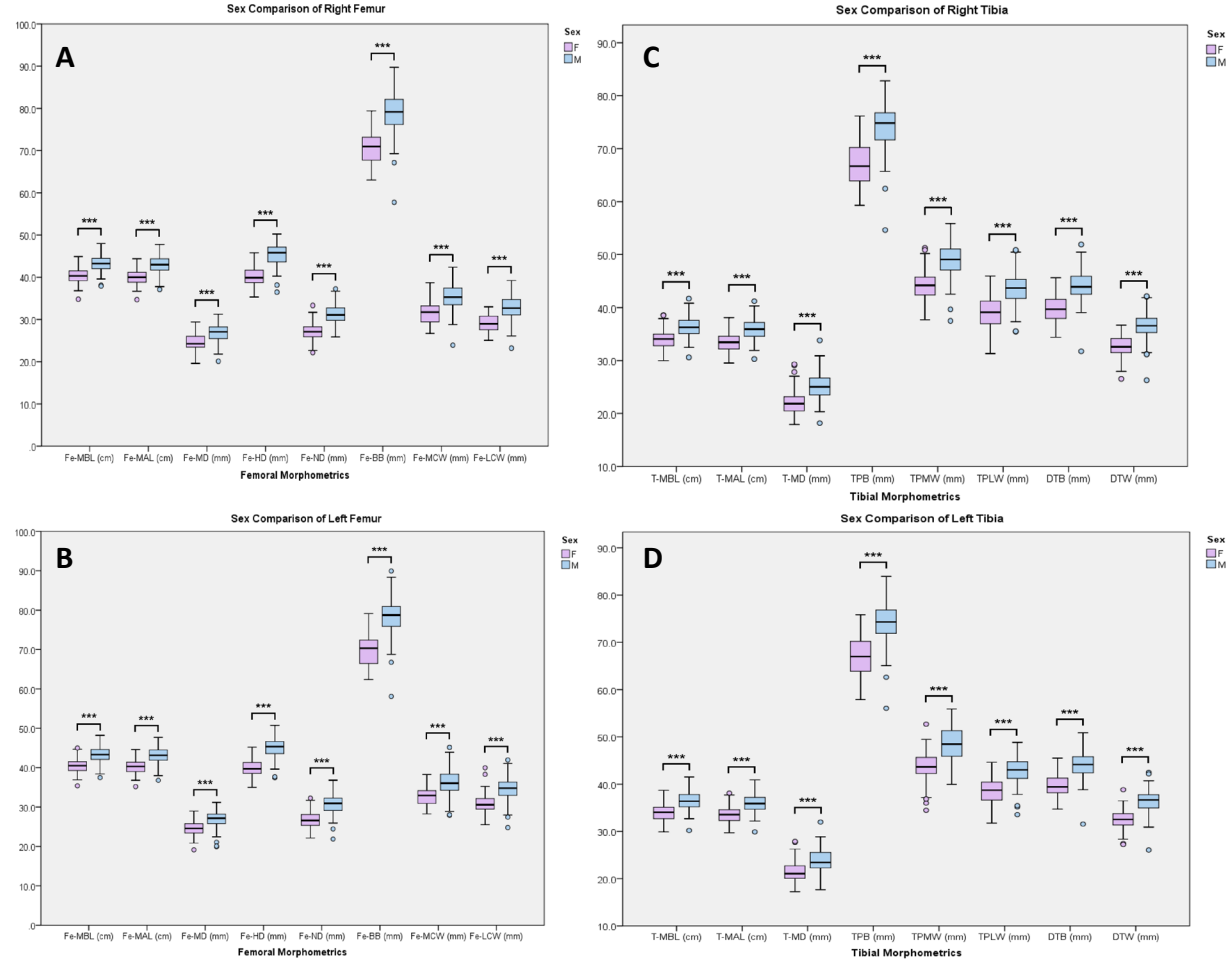


Fig 5. Boxplot comparing the results of the independent t-test between males and females. (A) Right side of femur. (B) Left side of femur. (C) Right side of tibia. (D) Left side of tibia. (***$p<0.001$).

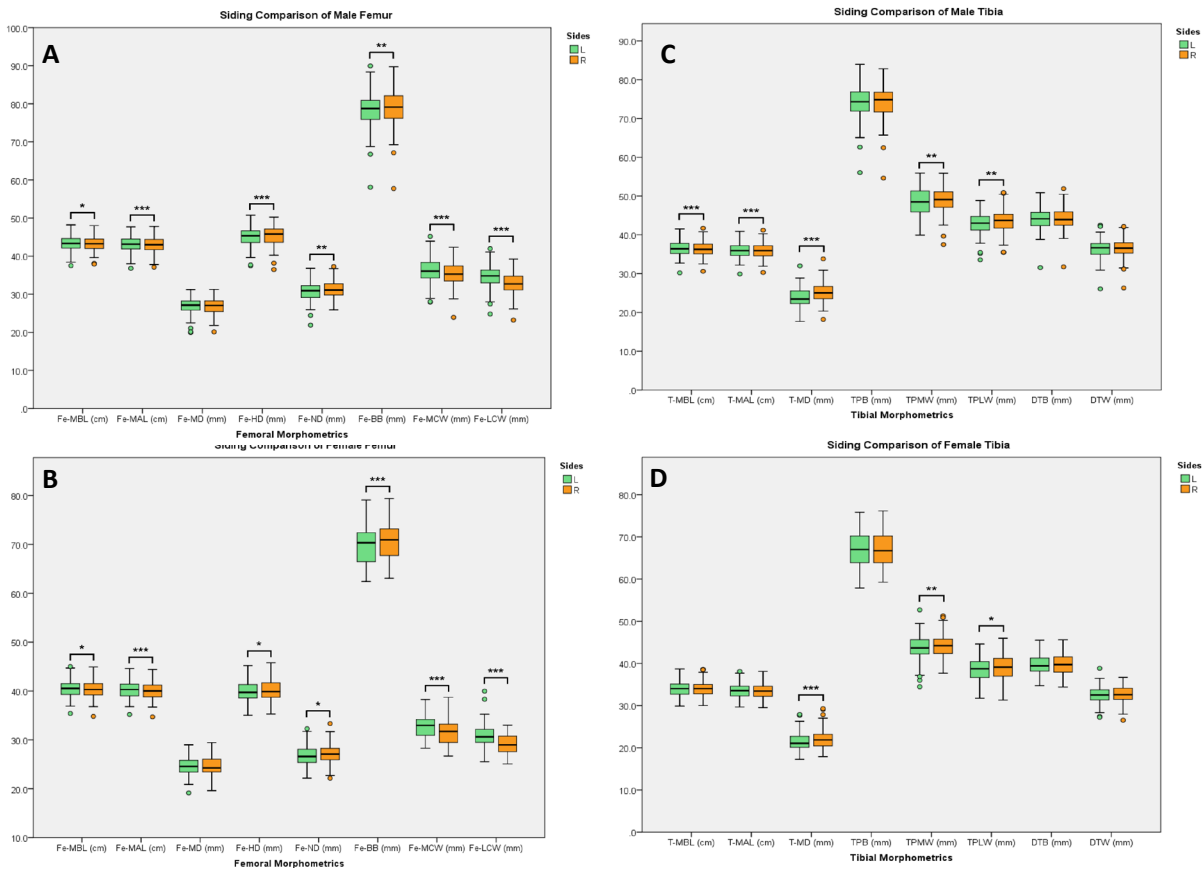


Fig 6. Boxplot comparing the results of the paired samples t-test between the left and right sides. (A) Femur in males. (B) Femur in females. (C) Tibia in males. (D) Tibia in females. (*p<0.05, **p<0.01, and ***p<0.001).

through the visual assessment of morphological variations stemming from sexual dimorphism. Key indicators include skeletal size and robustness, which are influenced by extrinsic factors such as biomechanical loads, as well as intrinsic factors like genetic makeup and hormonal influences. Both cranial and postcranial skeletal elements exhibit distinguishable sexual dimorphism in adult individuals. Among the available methodologies, combined non-metric and metric analysis of skull morphology is considered the second most reliable approach for sex determination in adult skeletal remains, following pelvic morphology assessment.[19] Within the cranium, features of the mastoid region are regarded as particularly informative. Morphological characteristics of the nuchal crest and mental eminence form the basis of cranial morphological sex estimation techniques, which are recognized for their high accuracy. Subsequent validation studies have reinforced and refined the reliability of skull morphological traits, leading to enhanced predictive accuracy and decreased classification errors.[20,21]

The findings of this study demonstrate that cranial measurements obtained by multiple observers, following anthropometric standards, exhibit high levels of reliability. Cronbach's alpha coefficients ranged from 0.832 to 0.971, indicating strong to excellent internal consistency between/among observers' measurements. These results are consistent with prior research, which recommended a reliability threshold of R>0.95 to ensure measurement dependability and accuracy. Notably, cranial length, inter-orbital foramen distance, and inter-infraorbital foramen distance showed the highest reliability coefficients at 0.971, 0.920, and 0.908, respectively. The elevated reliability scores of these parameters suggest that they are among the most sexually dimorphic features in the cranium. Furthermore, these results support the premise that craniometric measurements, when conducted following standardized protocols and independent assessments, can be both consistent and reproducible across different observers, thereby serving as robust tools for forensic sex determination. Overall, the results endorse the reliability and reproducibility of craniometric methods in forensic contexts.

The statistical analyses, employing independent *t*-tests and Mann-Whitney U tests, revealed statistically significant differences in various cranial measurements between males and females. Consistent with existing literature on the Thai population, male crania demonstrated larger mean measurements than females.[22,23] The high

TABLE 6. Assessment of the coefficients of reliability for post-cranial extremity long bones measurements.

Bone	Morphometric variables	Coefficient of reliability (r-value)	Interpretation
Humerus	HMAL	0.98	Very High
	HHD	0.96	Very High
	AND	0.94	Very High
	SND	0.85	High
	HMD	0.91	Very High
	EB	0.92	Very High
	CD	0.9	Very High
	OFD	0.85	High
Radius	RMAL	0.98	Very High
	RHD	0.91	Very High
	RND	0.82	High
	RMD	0.87	High
	DRB	0.86	High
	DRW	0.99	Very High
Ulna	UMAL	0.88	High
	MOPW	0.89	High
	MCPW	0.82	High
	UMD	0.88	High
Femur	Fe-MAL	0.91	Very High
	Fe-MD	0.92	Very High
	Fe-HD	0.96	Very High
	Fe-ND	0.94	Very High
	Fe-BB	0.94	Very High
	Fe-MCW	0.73	High
	Fe-LCW	0.79	High
Tibia	T-MBL	0.99	Very High
	T-MAL	0.92	Very High
	T-MD	0.68	Moderate
	TPB	0.84	High
	TPMW	0.78	High
	TPLW	0.64	Moderate
	DTB	0.92	Very High
	DTW	0.64	Moderate

reliability scores further reinforce the statistical validity of these findings. This concordance with prior studies underscores the efficacy of craniometric techniques in accurately sexing skulls and supports their application in forensic investigations. Although visual assessment yielded an accuracy of approximately 77%, with sensitivity and specificity of 71.43% and 82.83%, respectively, it remains inherently subjective and susceptible to inter-observer variability. This aligns with earlier research highlighting the potential for inaccuracies and examiner dependence associated with visual methods. In contrast, craniometric techniques offer objective, quantifiable data, reducing the likelihood of subjective bias. While direct comparison of the accuracy between these methods was limited by differing evaluation criteria, the superior reliability of craniometric measurement underscores its potential as an efficient and objective approach for forensic sex estimation.[24,25]

The morphologies of the mastoid process, orbital region, and nuchal crest are key indicators of sexual dimorphism. Additionally, other cranial features have been identified as potential sex determinants. These anatomical areas are notably resilient, often withstanding destructive forces. Employing morphometric measurements combined with stepwise discriminant functions generally achieves more accurate sex classification than visual assessments alone. However, this study found that morphometric analysis can be less accurate in sex differentiation compared to direct observational methods.

Sexual dimorphism and morphometric analysis of extremity long bone

Sexual dimorphism in the femur and tibia is strongly influenced by both sex hormones and biomechanical factors, leading to generally larger skeletal dimensions in males compared to females. This study revealed statistically significant differences between sexes across all measured femoral and tibial parameters ($p<0.001$), with males exhibiting greater mean values. The observed dimorphism is largely attributed to the influence of androgens and estrogens, particularly testosterone, which promote bone growth and increased bone mass in males, especially during puberty. This hormonal effect contributes to the development of more pronounced skeletal features, including key landmarks on the femur and tibia. Specifically, the femoral measurements in this study support findings from previous research in Thai population, showing similar values for midshaft diameter (Fe-MD), head diameter (Fe-HD), and neck diameter (Fe-ND).[26,27] However, the current study found that the maximum bone length of the femur (Fe-MBL) is notably greater, suggesting potential population-specific growth trends or methodological differences.

When comparing femoral bone lengths, including the maximum length (Fe-MBL) and anatomical length (Fe-MAL), the values observed in this study are smaller than those reported in Sri Lankan and Indian populations, yet larger than those found in Japanese population.[28] In comparison with an Ancient Anatolian population, the measurements for females are closely aligned, whereas male measurements are noticeably smaller. For midshaft diameter (Fe-MD), this study's results are slightly greater than those reported in Indian population and generally consistent with Sri Lankan data, although the Sri Lankan female values are slightly smaller.[29] Notably, in Sri Lankan population, male and female midshaft diameters are nearly identical, highlighting minimal dimorphism in that metric. Compared to ancient Anatolian population, midshaft diameters in this study are again closely aligned in females, but marginally smaller in males, suggesting possible population-specific variation or environmental influences on skeletal development.

On the distal femur (Fe-MCW and Fe-LCW), the measurements of this study were found to be slightly smaller in size compared to Korean population.[30,31] Importantly, it must be noted there is no comparative analysis of Fe-BB due to a difference in measurement methods since other studies used epicondylar breadth – the maximum length between medialmost to lateral most points on the epicondyle of the femur – instead of the definition used in this study.

Similarly, the tibial measurements in this study demonstrate marked sexual dimorphism, aligning with findings from previous research. When compared to earlier data from Thai population, the maximum tibial length (T-MBL) and distal tibia breadth (DTB) in this study are slightly smaller, while the proximal tibial measurements—specifically the medial and lateral condylar widths (TPMW and TPLW)—are slightly larger. The tibial plateau breadth (TPB), however, remains comparable between studies. In comparison with Sri Lankan population, this study reports a notably smaller anatomical length (T-MAL), while TPB measurements are larger.[32] Additionally, TPMW and TPLW values are slightly greater than those observed in Japanese and Brazilian populations, suggesting region-specific variation in proximal tibial morphology.[33] Compared to ancient Anatolian population, T-MBL values are marginally larger in females and slightly smaller in males, further highlighting population and sex-related differences in tibial dimensions.[34]

The midshaft diameter (T-MD) of the tibia in the

present study was found to be notably larger than that reported in the Indian population. When compared with ancient Anatolian population, female measurements were distinctly larger, while overall values were generally comparable. It is important to acknowledge, however, that methodological variations exist across studies, particularly in the anatomical landmarks used for measurement—many researchers take the diameter at the level of the nutrient foramen, which may influence reported values. Regarding the distal tibia, the breadth of the inferior articular surface (TPB) closely aligns with data from Sri Lankan population but is significantly larger than values recorded in Japanese population.[35-37] Similarly, the width of the inferior articular surface (DTW) in this study was found to be markedly greater than that of Kenyan population, highlighting inter-population variability in distal tibial morphology.[38]

From all of the aforementioned points, the femur and tibia can both be regarded as effective and reliable indicators for sex determination, and the measurement methods used for both the femur and tibia can be used in the process of sex determination.

TEM and reliability

It is well-recognized that observer error arising from visual or metric variables can result in inconsistencies in evaluating sexual dimorphism and determining sex. Recent studies indicate that visual assessment is prone to substantial inter-observer error due to vague variable definitions, heavy reliance on the observer's prior experience, and the seriation process used to categorize individuals. This error analysis demonstrates that geometric morphometrics achieves high levels of intra- and inter-observer agreement.

The results of the technical error of measurement (TEM) analysis, which is used to assess inter-assessor reliability, indicate variable levels of agreement across different skeletal parameters. For the humerus, six out of eight parameters—namely HMAL, HHD, AND, HMD, EB, and CD—demonstrated very high reliability, whereas SND and OFD exhibited only high reliability scores. In the case of the radius, half of the parameters, including RMAL, RHD, and DRW, showed very high interpretative consistency, whereas RND, RMD, and DRB were characterized as having only high reliability. For the ulna, all parameters—UMAL, MOPW, MCPW, and UMD—exhibited high reliability. This variation can be attributed to differences in the anatomical aspects of each bone, such as the specific landmarks used for measurement, including SND and OFD in the humerus, RND, RMD, and RBD in the radius, and UMAL, MOPW, MCPW, and UMD in the ulna. The instances where parameters yielded only high rather than very high reliability may stem from variability in landmark identification, often influenced by factors such as the absence of precise measurement landmarks—particularly in regions like the surgical neck of the humerus, where measurement relies heavily on the assessor's judgment.

Values reflect measurement consistency for both femoral and tibial parameters, except for the maximum bone length measurements of the femur (Fe-MBL) and tibia (T-MBL), due to the absence of comparable data. Using an R-value threshold of ≥0.90, most femoral measurements exceeded this cutoff, indicating high interobserver reproducibility, except for Fe-MCW and Fe-LCW. In contrast, most tibial measurements did not reach this threshold with only two parameters—T-MAL and DTB—demonstrating high reliability. These results suggest that femoral morphometric measurements generally exhibit greater interobserver consistency, whereas tibial measurements show variable reliability, with only selected parameters achieving robust repeatability.

Diagrams were used to clarify the definitions of each parameter during measurements; however, despite standardization, measurement errors remained possible due to limitations in observer interpretation. A key source of error is the variation in observer experience. In this study, one observer had no prior experience in skeletal measurement, whereas the other was an expert in forensic anthropology, which likely contributed to discrepancies, particularly in the identification of anatomical landmarks. Variations in understanding and interpreting the diagrams may have further contributed to minor differences in landmark annotation. Additionally, the morphology of the landmarks themselves influences measurement reliability. Many tibial measurements, along with two femoral parameters (Fe-MCW and Fe-LCW), involve landmarks with curved or rounded contours, which can be ambiguous and challenging to delineate precisely, thereby reducing measurement accuracy.

Cranial morphologies can often be consistently distinguished even by observers without prior experience. However, morphologies that are not clearly defined present significant challenges for consistent determination. Our findings indicate that even experienced observers analyzing skulls for sexual dimorphism may inconsistently interpret coordinate landmarks for sex assessment. While the robustness of the mastoid process and the pronounced nuchal crest are strong indicators of male sex, morphological trait evaluation still leads to varying levels of prediction accuracy between/among observers. This suggests that anatomical landmarks on cranial

bones may be misinterpreted. The description of these landmarks varies across the literature. It is crucial to describe quantitative methodologies that utilize cranial features exhibiting sexual dimorphism since they are pertinent to forensic practitioners. Additionally, understanding the impact of cranial sexual dimorphism on classification accuracy in sex estimation is essential.[39-41]

The results of the reliability analysis of craniometric measurements taken by two inter-observers (the intraclass correlation coefficient [ICC], specifically Cronbach's alpha), was used to assess the internal consistency and reliability of the measurements across various cranial parameters. That analysis revealed that the craniometric measurements exhibit high internal consistency, with all parameters showing good to excellent reliability. When interpreted according to the ranges established by Koo and Li, the Cronbach's alpha values for cranial length, inter-orbital foramen distance, and inter-infraorbital foramen distance indicate excellent reliability ($\alpha>0.90$).[42] Other parameters, including the cranial width, cranial index, cranial base length, mastoid length, and inter-supraorbital foramen distance showed good reliability ($0.75 \leq \alpha \leq 0.90$). These high reliability scores suggest that the measurements are consistent and repeatable across different observers.[43,44]

Furthermore, the definition of specific landmarks measured on the cranium and long bones of the extremities, such as the tibia, presents a recurring challenge in certain regions. To ensure measurement accuracy, all assessments must be both reproducible and independent. Consequently, future research should include evaluations of interobserver error in these measurements to minimize the risk of sex misclassification.

Effects of skeletal asymmetry and siding comparison

The paired samples t-test results of the upper extremity long bones indicate that two of the eight parameters—HMAL and SND—exhibit statistically significant differences between the left and right sides in males, with *p*-values less than 0.001. Additionally, analysis of the radius bones revealed three parameters—RMD, DRB, and DRW—that demonstrate significant bilateral differences at *p*-values below 0.001. Conversely, examination of the ulna shows that only UMAL exhibits significant asymmetry with a *p*-value under 0.001. The variation in the degree of bilateral differences, ranging from high to very high significance, is likely influenced by factors, such as handedness, which predominantly affects bone size disparity between sides. Supporting this, Walters *et al.* (1998) reported that right-handed individuals tend to have larger bones in their dominant hand, highlighting the functional role of dominance in skeletal asymmetry.[45-47]

Humans, as bipeds, typically distribute body weight evenly across the lower limbs during upright stance. Minimal bilateral differences in femoral and tibial morphometrics are, therefore, generally expected, so many previous studies have therefore analyzed only one limb for sex estimation. However, our findings revealed significant asymmetry between sides for both femoral and tibial measurements, which can be attributed to factors such as differential biomechanical loading, environmental influences, habitual limb preferences, and localized mechanical stress - all of which can induce side-specific bone remodeling. Degenerative joint conditions, notably osteoarthritis affecting the hip and knee, may further contribute to asymmetries, with parameters, such as femoral head diameter, bicondylar breadth, and proximal tibial measurements, showing consistent side differences. In the femur, almost all parameters demonstrated significant asymmetry with *p*-values below 0.05, which is consistent with prior studies that reported greater measurements on the left side, particularly in the lower limbs.[48,49]

Similarly, most femoral measurements in our study were larger on the left side in both sexes. Carvallo and Retamal (2020) studied a Chilean population and found bilateral symmetry predominantly in proximal femoral parameters, but they did not explore the underlying causes.[50] In the tibia, asymmetry was also present, but less pronounced, with males showing greater asymmetry than females. These findings align with van der Gaast, *et al.* (2022), who identified asymmetry in the tibial plateau potentially related to previous injuries, while the distal tibia exhibited high symmetry, consistent with the findings of Verbakel, *et al.* (2024). Nonetheless, caution in interpretation is warranted due to the relatively limited sample size.[51,52]

Given the observed inherent asymmetry, measurements should be taken bilaterally to ensure accuracy - especially for the femur. Despite the tibia generally demonstrating greater symmetry than the femur, bilateral assessment of the tibia remains advisable. Due to the possibility of unpredictable laterality in fragmentary remains at scene investigations, it is prudent to prioritize measurements of parameters known for bilateral symmetry when only partial femoral or tibial fragments are available.

Although most adult skeletons display characteristic sexual dimorphism, the accuracy of sex estimation is influenced by several factors, including population variability, age, and pathological or taphonomic changes. The degree and expression of sexual dimorphism can vary significantly both within and across different populations. Therefore,

it is crucial to employ population-specific data when applying these techniques in forensic contexts. The extent of sexual dimorphism and the differences between sexes differ across populations, and both morphological and metric methods are used for sex estimation. Morphological attributes—such as shape, specific traits, and relative size differences—are key indicators. Techniques focusing on pelvic shape, measurements, and the presence or absence of distinctive pelvic features are generally preferred since these tend to demonstrate greater sexual dimorphism and higher accuracy. Other morphological traits may also suggest sex, but they are often less reliable than those with pronounced dimorphic features. Employing appropriate instrumentation, standardized protocols, advanced analytical software, and integrating multiple measurements through multivariate approaches can improve the reliability of sex assessments, although individual measurements can still yield reasonably accurate results.

Sample size is a crucial consideration since the validation of classification functions necessitates cross-population studies to assess their robustness and generalizability. The use of approximately 400 individuals in this study may affect prediction accuracy and elevate the risk of technical errors. Expanding the sample size could mitigate these issues and enhance the overall effectiveness of the method.

Lastly, understanding the factors that influence prediction accuracy and classification errors—particularly in cranial and post-cranial sex estimation—is vital, given that the cranium is primarily governed by hormonal rather than mechanical factors - unlike the extremity long bones. This makes it especially useful in cases where the population of origin from which an unidentified skeleton originates from is unknown. Future research should focus on collecting and analyzing cooperative datasets from diverse populations, both for pelvic parameters and other skeletal features that display sexual dimorphism, to refine and validate forensic sex estimation methods across different population groups.

Limitations and suggestions

Future research should include measurements of additional osteometric landmarks, such as cranial landmark distances of cranium and fibular morphometrics of post-cranial bones and assess both intra- and interobserver reliability across all measurement methods.

Larger regionally diverse, and ethnically varied samples are recommended to evaluate the reliability, repeatability, and accuracy of sex estimation methods. The resulting data may facilitate the development of discriminant functions and help to estimate other biological profiles, such as stature.

The fact that our sample size of 204 crania is slightly lower than Yamane's suggested 214 likely exerted minimal impact on the results of our study. However, variation in observer expertise significantly influenced measurement accuracy in our study. The author's limited experience may have affected the accuracy of visual assessment, and despite efforts to reduce variability, differences in examiner skill and physical and mental condition contribute to inter-observer variability. Standardized training and calibration are, therefore, necessary to enhance consistency.

Outliers in descriptive statistics, particularly for the inter-supraorbital foramen distance (p=0.056), can be attributed to population-specific variations, such as the presence of multiple foramina or notches. These variations may have affected measurement accuracy, highlighting the need for further research to account for such differences in diverse populations.

CONCLUSION

Determining sex from skeletal remains is a fundamental component of human identification in forensic investigations. Although the skull is regarded as the second most preferable skeletal element for sex estimation, it is essential to account for measurement errors, which can impact the accuracy of sex determination. The adoption of alternative morphometric approaches, such as direct measurement of long bones of the extremities, may offer a viable alternative with superior accuracy—particularly when the pelvis or skull are not available. It is crucial to identify and understand the instrumental and methodological factors that influence accurate sex prediction. The selection of an appropriate analytical approach is paramount since it significantly influences the overall process of forensic human identification and helps reduce the likelihood of misclassification.

Data Availability Statement

The data used in this study was obtained from the Siriraj Anatomical and Anthropological Bone Research Centre (Si-AABRC) database, Department of Anatomy, Faculty of Medicine Siriraj Hospital, Mahidol university, Bangkok, Thailand. Access to this data was approved through Si-AABRC data access procedures. The data is not publicly available but can be requested from the Si-AABRC with appropriate permissions.

ACKNOWLEDGEMENT

None.

DECLARATIONS

Grants and Funding Information

None.

Conflicts of Interest

None.

Registration Number of Clinical Trial

Not Applicable.

Author Contributions

Conceptualization and methodology, N.S. and J.C.; Investigation, K.T., S.B., S.W., N.C. and J.C.; Formal analysis, K.T., S.B., S.W., N.C. and P.R.; Visualization and writing – original draft, K.T., S.B., S.W., N.C. and P.R.; Writing – review and editing, N.S.; Supervision, N.S. All authors have read and agreed to the final version of the manuscript.

Use of Artificial Intelligence

No artificial intelligence was used in this study.

Ethics/IRB Approval

SIRB Protocol No. 195/2564 (Date of Proof: 22 March 2021) and 680/2564 (Date of Proof: 25 August 2021).

REFERENCES

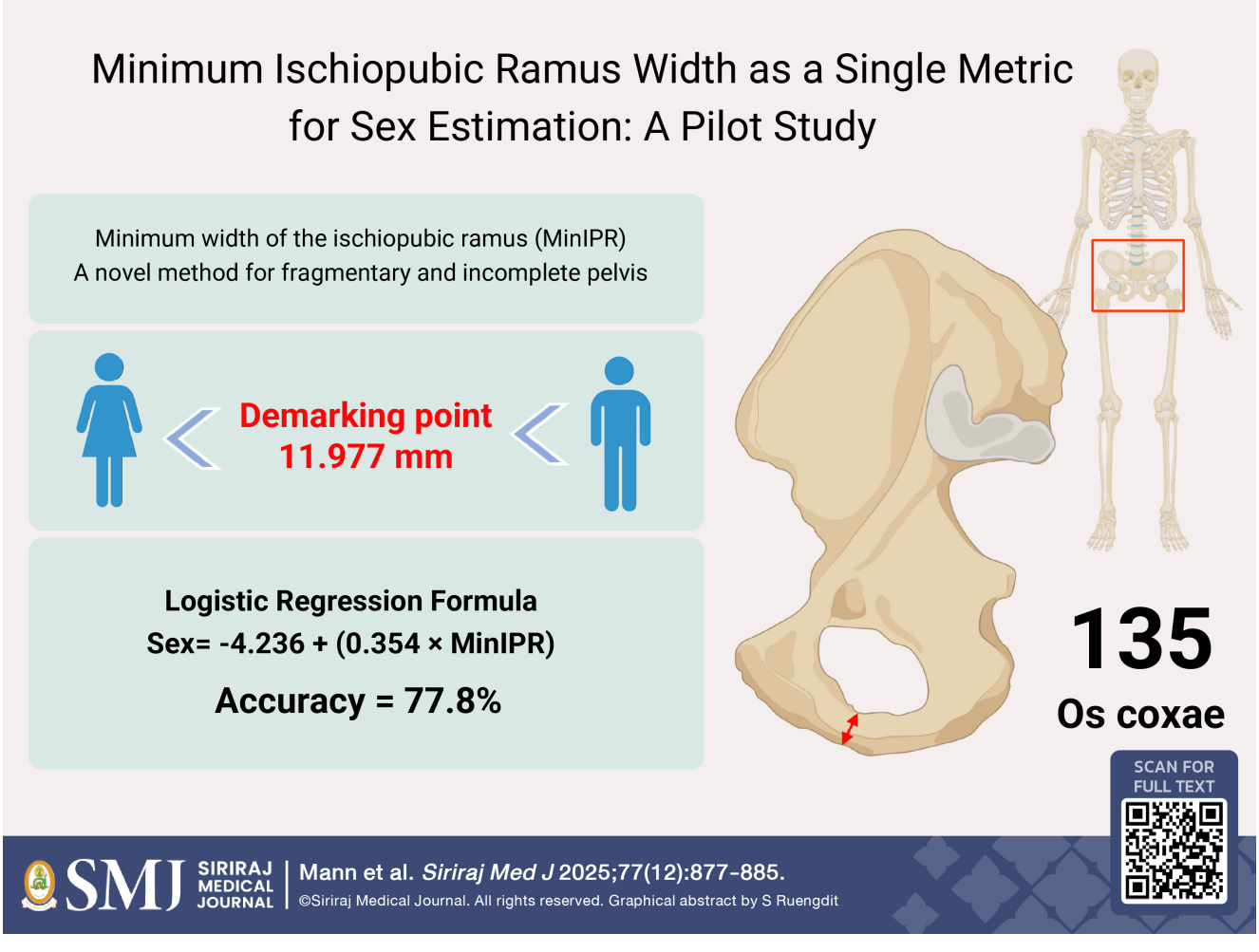
1. Wang X, Liu G, Wu Q, Zheng Y, Song F, Li Y. Sex estimation techniques based on skulls in forensic anthropology: a scoping review. Egypt J Forensic Sci. 2024;19(12):e0311762.
2. Lye R, Min H, Dowling J, Obertová Z, Estai M, Bachtiar NA, et al. Deep learning versus human assessors: forensic sex estimation from three-dimensional computed tomography scans. Sci Rep. 2024;14:30136.
3. Duangto P, Mahakkanukrauh P. Sex estimation from upper limb bones in a Thai population. Anat Cell Biol. 2020;53(1):36–43.
4. Krishan K, Chatterjee PM, Kanchan T, Kaur S, Baryah N, Singh RK. A review of sex estimation techniques during examination of skeletal remains in forensic anthropology casework. Forensic Sci Int. 2016;261:165.e1–8.
5. Selliah P, Martino F, Cummaudo M, Indra L, Biehler-Gomez L, Campobasso CP, et al. Sex estimation in middle and late adulthood: reliability of pelvic morphological traits and long bone metrics on an Italian skeletal collection. Forensic Sci Int. 2020; 134(5):1683-690
6. Blanc M, Knecht S, Nguyen K, Poulain C, Quatrehomme G, Alunni V, et al. Sexual dimorphism of the humerus bones in a French sample: comparison of several statistical models including machine learning models. Int J Legal Med. 2025;139(3):1395–408.
7. Houston SK, Brits D, Myburgh J, Liebenberg L. The impact of age-related changes in the skull on sex estimation using morphoscopic traits. Int J Legal Med. 2025.
8. Lewis CJ, Garvin HM. Reliability of the Walker cranial nonmetric method and implications for sex estimation. J Forensic Sci. 2016;61(3):743-51.
9. Wysocka J, Cieslik A, Danel D. Sex estimation using measurements of the proximal femur in a historical population from Poland. Anthropological Review. 2023;86(1):37-49.
10. Fliss B, Lüthi M, Fürnstahl P, Christensen AM, Sibold K, Thali M, et al. CT-based sex estimation on human femora using statistical shape modeling. Am J Phys Anthropol. 2019;169(2): 279–86.
11. Setiawati R, Rahardjo P, Ruriana I, Guglielmi G. Anthropometric study using three-dimension pelvic CT scan in sex determination among adult Indonesian population. Forensic Sci Med Pathol. 2022;19(1):24-33.
12. Torimitsu S, Makino Y, Saitoh H, SakumaA, Ishii N, Yajima D, et al. Morphometric analysis of sex differences in contemporary Japanese pelves using multidetector computed tomography. Forensic Sci Int. 2015;257:530.e.1-7.
13. Houston SK, Brits D, Myburgh J, Liebenberg L. The impact of age-related changes in the skull on sex estimation using morphoscopic traits. Int J Legal Med. 2025.
14. Tallman SD. Cranial nonmetric sexual dimorphism and sex estimation in East and Southeast Asian individuals. Forensic Anthropology. 2019;2(4):204–21.
15. Lye R, Obertová Z, Bachtiar NA, Franklin D. Validating the use of clinical MSCT scans for cranial nonmetric sex estimation in a contemporary Indonesian population. Int J Legal Med. 2024;138(4):1559–71.
16. Pérez-Criado L, Rosas A, Bastir M, Pastor JF. Humeral laterality in modern humans and Neanderthals: a 3D geometric morphometric analysis. Anthropol Sci. 2017;125(3):117–28.
17. Kumar S, Voracek M, Singh M. The effects of hand preference and sex on right–left asymmetry in dorsal digit lengths among adults and children. Early Hum Dev. 2021;153:105293.
18. Langley NR, Jantz LM, McNulty S, Maijanen H, Ousley SD, Jantz RL. Error quantification of osteometric data in forensic anthropology. Forensic Sci Int. 2018;287:183-9.
19. Sangchay, N, Dzetkuličová, V, Zuppello M, Chetsawang J. Consideration of Accuracy and Observational Error Analysis in Pelvic Sex Assessment: A Study in a Thai Cadaveric Human Population. Siriraj Med J. 2022;74(5):330–9.
20. Rogers TL. Sex determination of human remains through cranial morphology. J Forensic Sci. 2005;50(3):493-500.
21. Kruger GC, L'Abbe' EN, Stull KE, Kenyhercz MW. Sexual dimorphism in cranial morphology among modern South Africans. Int J Legal Med. 2015;129(4):869–75.
22. Mahakkanukrauh P, Sinthubua A, Prasitwattanaseree S, Ruengdit S, Singsuwan P, Praneatpolgrang S, et al. Craniometric study for sex determination in a Thai population. Anat Cell Biol. 2015;48(4): 275–83.
23. Sinthubua A, Ruengdit S, Das S, Mahakkanukrauh P. A new method for sex estimation from maxillary suture length in a Thai population. Anat Cell Biol. 2017;50(4):261–4.
24. Saini V, Srivastava R, Rai RK, Shamal SN, Singh TB, Tripathi SK. Sex estimation from the mastoid process among North Indians. J Forensic Sci. 2012;57(2):434-9.
25. Ekizoglu O, Hocaoglu E, Inci E, Can IO, Solmaz D, Aksoy S, et al. Assessment of sex in a modern Turkish population using cranial anthropometric parameters. Leg Med (Tokyo). 2016;21:45-52.
26. Monum T, Prasitwattanaseree S, Das S, Siriphimolwat P, Mahakkanukrauh P. Sex estimation by femur in modern Thai population. Clin Ter. 2017;168(3):e203–e207.

27. Duangto P, Mahakkanukrauh P. Sex estimation from upper limb bones in a Thai population. Anat Cell Biol. 2020;53(1):36–43.
28. Ozer I, Katayama K. Sex determination using the femur in an ancient Japanese population. Coll Antropol. 2008;32(1):67-72.
29. Ranaweera L, Cabral E, Dissanayake DMPV, Lakshan WSV. Estimation of sex from the osteometric measurements of the femur in a contemporary Sri Lankan population. Int J Morphol. 2022;40(4):1009–17.
30. Cho H-J, Kwak D-S, Kim I-B. Morphometric evaluation of Korean femurs by geometric computation: comparisons of the sex and the population. BioMed Res Int. 2015;2015:730538.
31. Kim J-B, Lyu S-J, Kang H-W. Are Western knee designs dimensionally correct for Korean women? A morphometric study of resected femoral surfaces during primary total knee arthroplasty. Clin Orthop Surg. 2016;8(3):254–61.
32. Nanayakkara D, Vadysinghe AN, Nawarathna LS, Sampath H. Determination of sex from the tibia in a contemporary Sri Lankan population. J Forensic Sci Med. 2019;5(1):24–28.
33. Uehara K, Kadoya Y, Kobayashi A, Yamano Y. Anthropometry of the proximal tibia to design a total knee prosthesis for the Japanese population. J Arthroplasty. 2002;17(8):1028–32.
34. Özer BK, Özer İ, Sağır M, Güleç E. Sex determination using the tibia in an ancient Anatolian population. Mediterr Archaeol Archaeom. 2014;14(2):329–36.
35. Gupta C, Nayak N, Kalthur SG, D'Souza AS. A morphometric study of tibia and its nutrient foramen in South Indian population with its clinical implications. Saudi J Sports Med. 2015;15(3):244–8.
36. Nanayakkara D, Vadysinghe AN, Nawarathna LS, Sampath H. Determination of sex from the tibia in a contemporary Sri Lankan population. J Forensic Sci Med. 2019;5(1):24–28.
37. Tiwari A, Mahendru A, Priya A. An anatomical study of the tibia in the North Indian population. Int J Hum Anat. 2019;2(1):1–7.
38. Misiani MK, Amuti T, Darbar S, Mandela P, Maranga E, Obimbo M. Sex determination from dimensions of distal tibiae in adult Kenyans: A discriminant function analysis. Translational Research in Anatomy. 2020;20:100075.
39. Torres HR, Morais P, Fritze A, Burkhardt W, Kaufmann M, Oliveira B, et al. Anthropometric landmarking for diagnosis of cranial deformities: Validation of an automatic approach and comparison with intra- and interobserver variability. Ann Biomed Eng. 2022;50(9):1022–37.
40. Kotěrová AP, Santos F, Bejdová S, Rmoutilová R, Attia MH, Habiba A, et al. Prioritizing a high posterior probability threshold leading to low error rate over high classification accuracy: the validity of MorphoPASSE software for cranial morphological sex estimation in a contemporary population. Int J Legal Med. 2024;138(2):1759–68.
41. Arigbabu OA, Liao IY, Abdullah N, Mohamad Noor MH. Computer vision methods for cranial sex estimation. IPSJ Trans Comput Vis Appl. 2017;9:19.
42. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med. 2016;15(2):155–63.
43. Richard AH, Parks CL, Monson KL. Accuracy of standard craniometric measurements using multiple data formats. Forensic Sci Int. 2014;242:177–85.
44. Kizilgoz V, Aydin S, Aydemir H, Keles P, Kantarci M. Interobserver and intraobserver reliability of skull base angles measured on magnetic resonance images. World J Clin Cases. 2024;12(34):6687–95.
45. Walters J, Koo WWK, Bush A, Hammami M. Effect of hand dominance on bone mass measurement in sedentary individuals. J Clin Densitom. 1998;1(4):359–67.
46. Nandi ME, Olabiyi O, Okubike EA, Cyprain IE. A study of bilateral asymmetry of upper extremity and its effects on stature reconstruction amongst Nigerians. Aus J Forensic Sci. 2018;1(8):978–88.
47. Auerbach BM, Ruff CB. Limb bone bilateral asymmetry: variability and commonality among modern humans. J Hum Evol. 2006;50(2):203–18.
48. Marques S, Pinto C, Ferreira MT, Garcia S, Curate F. Sex Estimation from the Fibula and Tibia: A Study in Three Portuguese Reference Collections. Forensic Sci. 2023;5(1):2.
49. Mittino G, Langstaff H, García-Donas JG. Sex and Stature Estimation on the Tibia: A Virtual Pilot Study on a Contemporary Hispanic Population. J R Anthropol Inst. 2024;30(3):1–15.
50. Carvallo D, Retamal R. Sex estimation using the proximal end of the femur on a modern Chilean sample. Forensic Sci Int Rep. 2020;2:100077.
51. van der Gaast N, Dunning H, Huitema JM, Waters A, Jaarsma RL, Doornberg JN, et al. The symmetry of the left and right tibial plateau: a comparison of 200 tibial plateaus. Eur J Trauma Emerg Surg. 2023;49(1):69–74.
52. Verbakel J, Boot MR, van der Gaast N, Dunning H, Bakker M, Jaarsma RL, et al. Symmetry of the left and right tibial plafond; a comparison of 75 distal tibia pairs. Eur J Trauma Emerg Surg. 2024;50(6):2877–82.

Minimum Ischiopubic Ramus Width as a Single Metric for Sex Estimation: A Pilot Study

Robert W. Mann, Ph.D.[1], Sittiporn Ruengdit, Ph.D.[2,*], Patara Rattanachet, Ph.D.[3,4], Napakorn Sanchay, M.D., Ph.D.[3,4]

[1]Department of Anatomy, Biochemistry, and Physiology, John A. Burns School of Medicine, The University of Hawai`i at Mānoa, Honolulu, Hawaii, USA, [2]Department of Forensic Medicine, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand, [3]Department of Anatomy, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand, [4]Siriraj Anatomical and Anthropological Bone Research Centre, Department of Anatomy, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand.



*Corresponding author: Sittiporn Ruengdit
E-mail: neu.sittiporn@gmail.com
Received 21 September 2025 Revised 2 November 2025 Accepted 9 November 2025
ORCID ID:http://orcid.org/0000-0002-3220-9146
https://doi.org/10.33192/smj.v77i12.277804

ABSTRACT

Objective: The present study assesses the effectiveness of the minimum width of the ischiopubic ramus as a single measurement for sex estimation.

Materials and Methods: Visual and metrical examination of 135 known-identity male and female os coxae was investigated by measuring the minimum mediolateral width of the ischiopubic ramus, a novel and simple measurement of the narrowest part of the ischiopubic ramus.

Results: The results revealed females typically have a narrow ("pinched") ischiopubic ramus in comparison to males, with a statistically significant difference in the minimum mediolateral width of the ischiopubic ramus between sexes. Logistic regression, which delivered the best sex estimation model among traditional statistical analysis and machine learning approaches, provided an accuracy of 77.8%, a sensitivity of 83.33% and a specificity of 73.33% with a demarking point of 11.977 mm.

Conclusion: Although the method yields moderate accuracy, the minimum width of the ischiopubic ramus provides a straightforward and practical approach that may assist in sex estimation when only a fragment of the ischiopubic ramus is available. Further validation using larger and more diverse samples is recommended to confirm its applicability.

Keywords: Ischiopubic ramus; sex estimation; fragmented bones; machine learning (Siriraj Med J 2025; 77: 877-885)

INTRODUCTION

Estimating the biological sex of an individual through visual and metrical examination of the human skeleton is one of the four components of the biological profile, consisting of age at death, sex, ancestry, and stature. Establishing the biological profile of skeletal remains is important to medicolegal authorities as it provides them with information that may help narrow the possibilities or lead to the identification of a decedent. Reported accuracies in estimating sex of a complete skeleton range from 64 to 100%[1], with the most reliable visual features being present in the pelvis, followed by the cranium.[2] The pelvis, due to its sexually dimorphic size and shape features that are related to child birthing, can be used to reliably estimate sex in about 99% of individuals compared to 91% using the cranium. Constant re-examination of existing metrical and visual methods and technologies, and the development of new methods for estimating sex through examination of the human skeleton, continue to provide medicolegal analysts with higher accuracy rates[3] for establishing sex in both complete and incomplete human skeletal remains.

In forensic investigations, complete skeletal remains are not always recovered, requiring the use of skeletal analysis methods tailored to fragmented bones. The pelvis, being one of the skeletal structures most susceptible to fractures[4], is often involved in these cases. Most sex estimation methods—whether traditional morphoscopic and morphometric techniques or shape analysis using geometric morphometrics—typically focus on the entire innominate bone or require a complete pubic or ischiopubic complex.[4] However, in archaeological contexts for example, the preservation of the pubic region rarely exceeds 30%.[5] Additionally, the upper half of the ischiopubic ramus is more commonly preserved compared to the lower half. Therefore, the present study aims to assess the effectiveness of the minimum width of the ischiopubic ramus as a single measurement for sex estimation.

MATERIALS AND METHODS

This research protocol has been certified as exempt by the Research Ethics Committee of the Faculty of Medicine, Chiang Mai University (Exemption 0525/2025). All procedures were conducted in accordance with the ethical standards of the committee and with the Declaration of Helsinki. A total of 135 adult male and female left os coxae (innominate or hip bones) in the Mann-Labrash Osteology Collection at the John A. Burns School of Medicine University of Hawaii, Siriraj Anatomical and Anthropological Bone Research Centre at the Department of Anatomy at Siriraj Hospital School of Medicine, Mahidol University, and Osteology Research and Training Center in the Department of Anatomy at Chiang Mai University School of Medicine, Thailand were randomly selected and examined. The sample consisted of 73 adult males and 62 adult females of known age, sex, and identity donors in the Willed Body Program at the three universities. The Mann-Labrash Osteology Collection consists of known-identity individuals representing the population diversity of Hawaii who died between 1974 and 2020, including individuals of Japanese, Chinese,

Korean, Vietnamese, Hawaiian, African American, and European ancestry.[6] The osteological collection curated at the Osteology Research and Training Center at Chiang Mai University School of Medicine consists of Thais, primarily from Northern Thailand, and who died between 2003 and 2014.[6] The Siriraj Anatomical and Anthropological Bone Research Centre is composed of modern Thai and Chinese-descent individuals who donated their bodies for medical research between 1933 and 2020.[7] Author RM measured the Hawaii sample, and SR measured the Northern Thai samples. PR and NS compiled data for the Siriraj samples. An additional holdout sample of 24 cases (12 males and 12 females) was selected to validate the efficiency of the obtained equation or model and its demarcation point.

The minimum mediolateral width of the ischiopubic ramus (minIPR) was obtained using a Mitutoyo SC-6"C sliding caliper by visually and metrically locating the narrowest mediolateral portion of the ramus, approximately 3 cm lateral to the most inferior part of the pubic symphysis. The narrowest mediolateral part of the IPR in many individuals was visible as a "pinched", concave or biconcave area formed by the inferior margin of the IPR and the inferior margin of the obturator foramen in the area where the ischium and pubic bones join. The measurement was obtained while avoiding any bony projections, osteophytes, or evidence of trauma such as a fracture or healing callus. The minIPR was, however, rarely affected or altered by localized or systemic bone disease such as cancer or osteoporosis. To obtain the minimum width of the IPR the blade-like portion of the arms of the caliper was held perpendicular to the long axis and ventral surface of the IPR (Fig 1) and moved slightly in all directions to obtain the narrowest mediolateral width. The measurement was rounded to the nearest millimeter. Each minIPR was measured without knowing the sex of an individual.

Data analysis

Statistical analyses were conducted using R software (version 4.5.1).[8] Inter-observer error analysis was performed to evaluate their reliability and repeatability. Twenty cases, 10 males and 10 females, were randomly selected and remeasured by the third author (PR). All measurements recorded by the second author (SR), acting as the first observer, and the third author (PR), acting as the second observer, were compared using the mean absolute difference between repeated measurements, the mean difference expressed as a percentage of average bone size, as well as the technical error of measurement (TEM) and the coefficient of reliability (R).[9-11] Descriptive statistics were computed to summarize the data. An independent samples t-test was performed to compare the minIPR variable between males and females, with statistical significance set at $p < 0.05$.

Numerous machine learning approaches have been applied in forensic anthropology and related disciplines. For this study, we selected the most commonly used

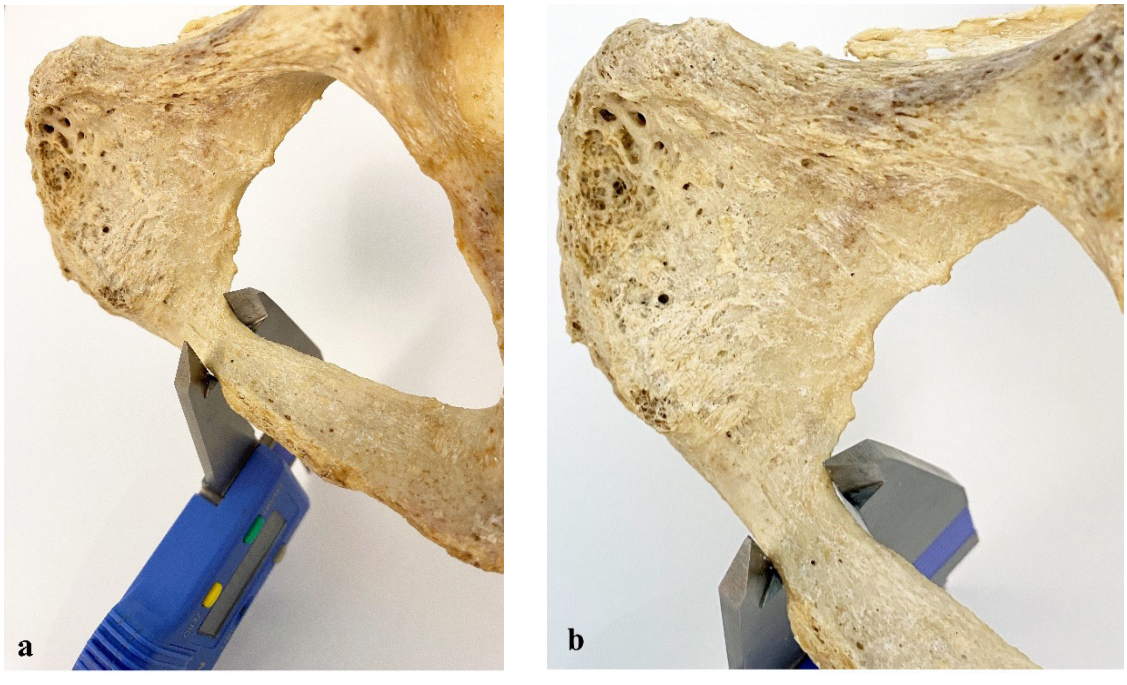


Fig 1a-b. Position of the sliding caliper when measuring the left ischiopubic ramus in an elderly female. (a) Note that the caliper is held perpendicular to the ventral surface of the ischiopubic ramus at its narrowest point to obtain the smallest measurement. It's easier to take this measurement when holding the calipers upside down as in this image. (b) Slightly different angle showing the position of the caliper arms against the ischiopubic ramus.

predictive models, including seven classifiers: XGBoost (XGB), Support Vector Machine (SVM), Random Forest (RF), Logistic Regression (LR), Artificial Neural Network (ANN), Linear Discriminant Analysis (LDA), and Decision Tree (DT). The performance of both classical statistical methods and machine learning classifiers was then evaluated.

The dataset was partitioned into a training set (80%) for model construction and a test set (20%) for evaluating classification accuracy.

Model evaluation

The models' performance was evaluated using accuracy, precision, recall, and F1-score metrics. Accuracy measures the overall agreement between the actual and predicted sex by dividing the number of correctly predicted instances by the total number of instances. In medical literature, precision—also known as predictive value—indicates the probability that a sex estimation is accurate based on the algorithm's prediction. Recall, a group-specific performance metric, also known as sensitivity, represents the probability that the algorithm correctly estimates sex. Lastly, F1-score is a metric used to evaluate the performance of a classification model, especially in scenarios where the data is imbalanced. It is the harmonic mean of precision and recall, offering a balance between the two.

RESULTS

Table 1 presents the median and mean absolute differences, as well as the mean differences expressed as a percentage of the mean bone size, for the interobserver error analysis. Technical Error of Measurement (TEM) and the correlation coefficient (R) were also calculated to assess measurement repeatability. The obtained R value of 0.90 indicates an acceptable level of reliability for the minIPR measurement.[11] Although the mean percentage difference was moderately high (9.59%), it was comparable to the TEM value (0.73 mm) when evaluated against the grand mean bone size, and notably lower than the mean difference in minIPR observed between males and females (2.7 mm).

Table 2 contains descriptive statistics including means and standard deviations for minIPR measurements of males and females. Independent t-test revealed the average measurements in males are statistically significantly wider than those in females ($p < 0.05$) with a mean difference of 2.7 mm, highlighting the presence of sexual dimorphism.

The classification accuracy of all models ranged from 63.0% to 77.8%, with the Random Forest (RF) model showing the lowest accuracy, while Logistic Regression (LR) and Linear Discriminant Analysis (LDA) achieved the highest accuracy. Notably, LR and LDA also demonstrated superior performance in precision (71.4%), recall (83.3%), and F1 score (76.9%). In contrast,

TABLE 1. Interobserver error statistics for the minimum mediolateral width of the ischiopubic ramus.

Measure	*n*	Median abs diff (mm)	Mean abs diff (mm)	TEM (mm)	R	%Mean diff
minIPR	40	0.81	1.04	0.73	0.90	9.59

Note: Median abs diff refers to Median absolute difference; Mean abs diff refers to Mean absolute difference; TEM refers to Technical error of measurement; %Mean diff refers to Mean difference as percentage of mean size.

TABLE 2. Descriptive statistics for the minimum mediolateral width of the ischiopubic ramus.

	n	Mean	Median	SD	Min	Max
Male	73	13.7	13.9	2.84	7.3	19.4
Female	62	11.0	10.5	2.51	4.9	18.3

RF recorded the lowest values across these metrics. Regarding recall, the k-Nearest Neighbors (KNN) and Decision Tree (DT) models matched the highest values achieved by LR and LDA (Fig 2).

The Shapiro-Wilk test was used to assess normality, and the p-values were not statistically significant ($p > 0.05$), indicating that the data followed a normal distribution. The total correct classification accuracy obtained from both sexes was 77.8%. Logistic regression model fitting is tabulated in Table 3. The model using minIPR as a predictor variable is described by the following equation (females classified with negative values, whereas males are classified with positive values):

$$Sex = -4.236 + (0.354 \times minIPR)$$

The model accurately classified the sex of 77.8% of individuals, achieving a sensitivity of 83.33% and a specificity of 73.33% with a demarking point of 11.977 mm. This demonstrates a strong discriminative ability, with an area under the curve (AUC) obtained from ROC curve of 0.85 and a Kappa value of 0.557. In the logistic regression model, a demarking point was established, indicating that a minIPR value of less than 11.977 mm suggests a female, while a value greater than this threshold suggests a male. Additionally, twenty-four holdout samples were randomly selected to evaluate the logistic regression model and its demarking point. The model demonstrated moderate performance, similar to that obtained from the LR model developed from the training set, achieving an accuracy of 75%.

DISCUSSION

The interobserver error statistics for the minIPR show relatively moderate to high error; however, in terms of millimeters, this error appears to be smaller than the sexual dimorphism observed in the minIPR

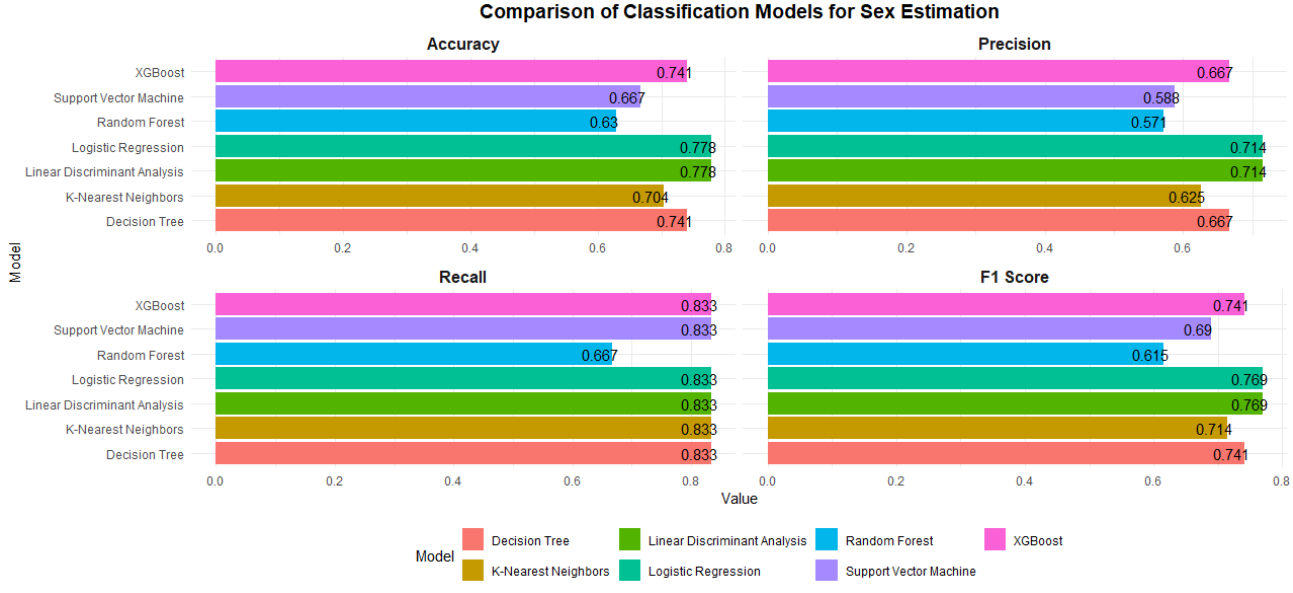


Fig 2. Comparison of classification model performance for sex estimation using the minimum ischiopubic ramus width. Seven statistical and machine learning classifiers—XGBoost, Support Vector Machine, Random Forest, Logistic Regression, Linear Discriminant Analysis, K-Nearest Neighbors, and Decision Tree—were evaluated based on Accuracy, Precision, Recall, and F1 Score.

TABLE 3. Coefficient obtained from logistic regression.

	Estimate std.	Std. error	Z value	p-value
Intercept	-4.23587	1.06625	-3.973	7.11e-05
minIPR	0.35367	0.08506	4.158	3.21e-05

Hosmer and Lemeshow goodness of fit (GOF) test
X-squared = 13.069, df = 8, p-value = 0.1095

measurements between males and females. Therefore, the potential measurement error is likely to have a minimal impact on sex estimation when using this measurement. Additionally, strictly adhering to the definition and carefully obtaining the absolute minimum IPR measurements by repeatedly measuring the areas visually assessed as the 'narrowest' could help reduce the potential for error.

The Phenice method has proven to be an accurate and reliable method for estimating sex in adult skeletons since it was developed. All three features of the Phenice method have been shown to yield high accuracy rates, particularly the ventral arc (96%)[12-14], making the anterior pelvis the most reliable area in the skeleton for estimating sex.

A few studies have investigated sexual dimorphism using the minimum mediolateral width of the ischiopubic ramus[15,16], which is in the anterior pelvic region. Dixit et al. reported high significance in Wilk's lambda and F-ratio tests, consistent with Luo's findings[15], which achieved higher accuracy (81.7%) than the present study. The terminology and definition of the ischiopubic ramus employed in these studies closely matched that proposed in the present study. Dixit et al.[16] refined Luo's vague definition of "the minimum thickness of the ischiopubic ramus" by providing clearer measurement criteria, though neither study addressed interobserver measurement reliability.[15] The definition described in Dixit's study was "the least straight distance between the ischiopubic ramus and nearest obturator foramen margin". Although both Luo and Dixit's studies proposed a relatively similar definition of ischiopubic ramus width as compared to the present study, they did not report the classification accuracy for this measurement alone. All the high classification accuracies derived from their studies still required a combination of measurements. For instance, Luo's study required four measurements, including "the minimum thickness of ischiopubic ramus" as they defined it, to obtain 81.7% accuracy. Similarly, Blake et al.[17] included the ischiopubic ramus measurement among pubic measurement combinations analyzed using linear discriminant function analysis, achieving a cross-validated classification accuracy of 87.1%. However, like Luo and Dixit's studies, the study did not specify the classification accuracy when using IPR alone. Ischiopubic ramus thickness (IPR), the term used in Blake et al.[17], is a pubic measurement similar to minIPR. As mentioned in Blake et al[17], the IPR was defined as "a distance between the inferior-most point of medial obturator foramen to the narrowest point inferior to the pubic symphysis" while the minIPR, used in the present study, is the minimum mediolateral width of the ischiopubic ramus locating the narrowest mediolateral portion of the ramus, approximately 3 cm lateral to the most inferior part of the pubic symphysis. Although Luo, Dixit, and Blake's studies reported higher classification accuracies (>80%) using similar definitions of ischiopubic ramus width, their methods relied on multiple measurements, whereas the present study demonstrates the utility of a single minimum ischiopubic ramus width (77.8%). Patriquin et al.[18] proposed an LDA equation based on pubic measurements, including pubic bone width, with an accuracy of 69–79% for White and Black populations. However, this equation required measuring pubic bone height, and its definition of pubic bone width differed from the minIPR used in the present study.

Apart from traditional morphometric measurements, geometric morphometric analysis of the ischiopubic region has been explored.[4,19,20] Black analyzed landmarks and semi-landmarks along the ischiopubic ramus, however, they did not clearly define or apply the narrowest area of the ischiopubic ramus measured mediolaterally as defined in the present study.[4] Although Gonzalez and colleagues investigated the sexual dimorphism efficiency of the ischiopubic complex and achieved a high prediction accuracy of 93.4%, they used landmarks along the borders of the ischiopubic region.[20] Bytheway and Ross[19] also examined various os coxal shapes, including the ischiopubic ramus, which they defined as the narrowest point inferior to the pubic symphysis, using a geometric morphometric approach. However, their study did not report the sex estimation accuracy derived from the ischiopubic ramus shape alone.

Numerous studies have proposed multivariate and univariate equations using pelvic measurements for sex estimation, achieving high classification accuracy (> 80%).[21,22] However, most of these methods still require an almost intact pelvis. Similar to the ischiopubic ramus index, a promising proportion requiring only two measurements has been proposed for sex estimation from fragmented pelves.[23,24] However, this index also necessitates an almost complete ischiopubic region. Additionally, the uncertainty in defining the acetabular point increases the potential for observer error with this index.[23] Therefore, a single-metric approach, as proposed in the present study, offers an alternative for sex estimation from fragmented or incomplete pelves, despite its moderate classification accuracy (77.8%).

Table 4 presents a comparison of sex classification accuracies using a single measurement of the ischiopubic region across different studies. The classification accuracy of 77.8% achieved in the present study falls within the medium to high range compared to other studies. This

TABLE 4. A comparison of sex classification accuracies using a single measurement of the ischiopubic region across different studies.

Measurement	References	Accuracy (%)
Minimum thickness of the ischiopubic ramus	Luo et al.[15]	81.7
Minimum ischiopubic ramus width	**Present study**	**77.8**
Acetabular diameter and width	Steyn and İşcan[25], Patriquin et al.[18], Mahakkanukrauh[22]	77.0-89.0
Ischial length	Patriquin et al.[18], Franklin et al.[26], Mahakkanukrauh[22]	79.4-81.2
Pubic length	Mahakkanukrauh[22]	70.0
Subpubic angle	Franklin et al.[26], Torimisu et al.[21]	93.2, 98.1
Angle of greater sciatic notch	Franklin et al.[26], Torimisu et al.[21]	85.2, 83.7
Maximum ischio-pubic length	Arun et al.[27]	68.0

accuracy is approaching 80%, which is an acceptable rate for sex estimation in forensic anthropology. Furthermore, it is nearly identical to the accuracy obtained from the maximum diameter of the femoral head, a widely used single measurement for sex estimation.[25]

In forensic investigations, complete skeletal remains are not always recovered, necessitating skeletal analysis methods for fragmented or incomplete bones. The pelvis is among the skeletal structures that are particularly prone to fractures.[4] Most sex estimation methods focus on the entire innominate or require a complete pubic or ischiopubic complex.[4,22,23,28] The pubic rami are among the most fractured structures in cases of moderate to severe trauma.[29,30] Isolated fractures of a single ramus, particularly the ischial ramus, are also observed, with these fractures being more frequent in the elderly.[31] Furthermore, the superior rami are more often fractured than the inferior rami. As a result, the minimum mediolateral width of the ischiopubic ramus, typically located in the upper part of the ischiopubic ramus about three centimeters below the inferior border of the pubic symphysis, shows promise as a useful measurement for sex estimation in fragmentary pelves. Future studies should empirically evaluate the reliability of the minIPR on experimentally fragmented or taphonomically altered os coxae. Such validation would provide critical evidence of the method's robustness and practical applicability in forensic contexts where pelvic completeness is often compromised.

While this method can be used in addition to the Phenice method, its strength as an indicator of sex is when dealing with highly fragmented or incomplete skeletal remains. Essentially, if a 2-to-3-centimeter portion of the ischiopubic ramus is present for examination, an accurate sex assessment using this feature alone can be achieved 77.8%. When dealing with burned remains, however, the morphology of the IPR may be reflective of sex, but the size of the IPR will almost certainly be reduced in size, rendering unusually small measurements that more likely fall within the female range. Using this method for badly burned pelves, therefore, is not recommended.

The shape of the IPR is a feature and morphology that is visualized during examination of the pelvis, is reflective of and part of the size and shape of the obturator foramen, and formation of the subpubic concavity. This part of the IPR is represented variously by a parallel, biconcave, or convex IPR and its narrowest width is often along the inferior ramus of the pubic bone but sometimes occurs at or near the junction (synchondrosis) of the ischiopubic bones. This study revealed that the narrowest width of the IPR varies from the pubic bone near the anterior portion of the obturator foramen to more inferior-posteriorly near the intersection of the two ischium and pubic bones. The authors found a few methods or mention of any use of the width of the IPR, visually or metrically, used to establish sex.[15,16,25,31]

While the width and depth of the IPR are incorporated into the size and shape morphology of the obturator foramen, this feature appears to have escaped mention as it relates to estimating sex of the pelvis. The morphology of the ischiopubic ramus is, however, an integral part

of the size and shape of the obturator foramen that is based, in part, on whether the IPR is convex, parallel, concave, or biconcave.

Several limitations should be noted. First, the interobserver technical error of measurement (9.6%) indicates moderate variability between observers, which may result from subtle differences in landmark placement and identification along the irregular contour of the ischiopubic ramus. Second, the validation set was relatively small (n = 24), which limits the generalizability of the present findings and warrants replication in larger and more diverse samples. Third, only left os coxae were analyzed; therefore, potential bilateral variation in the minimum ischiopubic ramus width remains untested. Future research incorporating both coxae and larger population samples would strengthen the reliability and applicability of this method.

Although machine learning models, including XGB, SVM, RF, ANN, and DT, were evaluated alongside traditional classifiers such as LDA and LR, the relatively small sample size of this study likely limited the ability of ML models to capture more complex patterns. The comparable accuracy observed among these methods indicates that classical statistical approaches remain reliable and efficient for sex estimation in small forensic datasets. Nevertheless, the inclusion of machine learning algorithms provides a comparative baseline and highlights their potential for future research using larger and more heterogeneous samples, where their capacity for non-linear modeling and feature interaction analysis could offer additional benefits.

CONCLUSIONS

This measurement provides an additional, straightforward tool that can be applied in addition to established sex estimation methods—such as the Phenice method, cranial and postcranial indicators—and may be particularly useful in cases involving fragmented or incomplete skeletal remains.

Data Availability Statement

The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.

ACKNOWLEDGEMENTS

The authors sincerely thank the curators, faculty, staff, and assistants for granting access to the collections and for their dedication to acquiring and preserving the materials referenced in this study. We also extend our appreciation to the two anonymous reviewers for their valuable comments and suggestions.

DECLARATIONS

Grants and Funding Information

None.

Conflict of Interest

The authors declare no conflicts of interest.

Registration number of clinical trial

None.

Author Contributions

Conceptualization and methodology, R.W.M.; Investigation, R.W.M., S.R., P.R., N.S.; Formal analysis, S.R.; Visualization and writing – original draft, R.W.M.; Writing – review and editing, R.W.M., S.R., P.R., N.S.; Supervision, R.W.M. All authors have read and agreed to the final version of the manuscript.

Use of Artificial Intelligence

During the preparation of this manuscript, the authors used ChatGPT to improve readability and grammar. All content was subsequently reviewed and verified by the authors, who take full responsibility for the final version.

REFERENCES

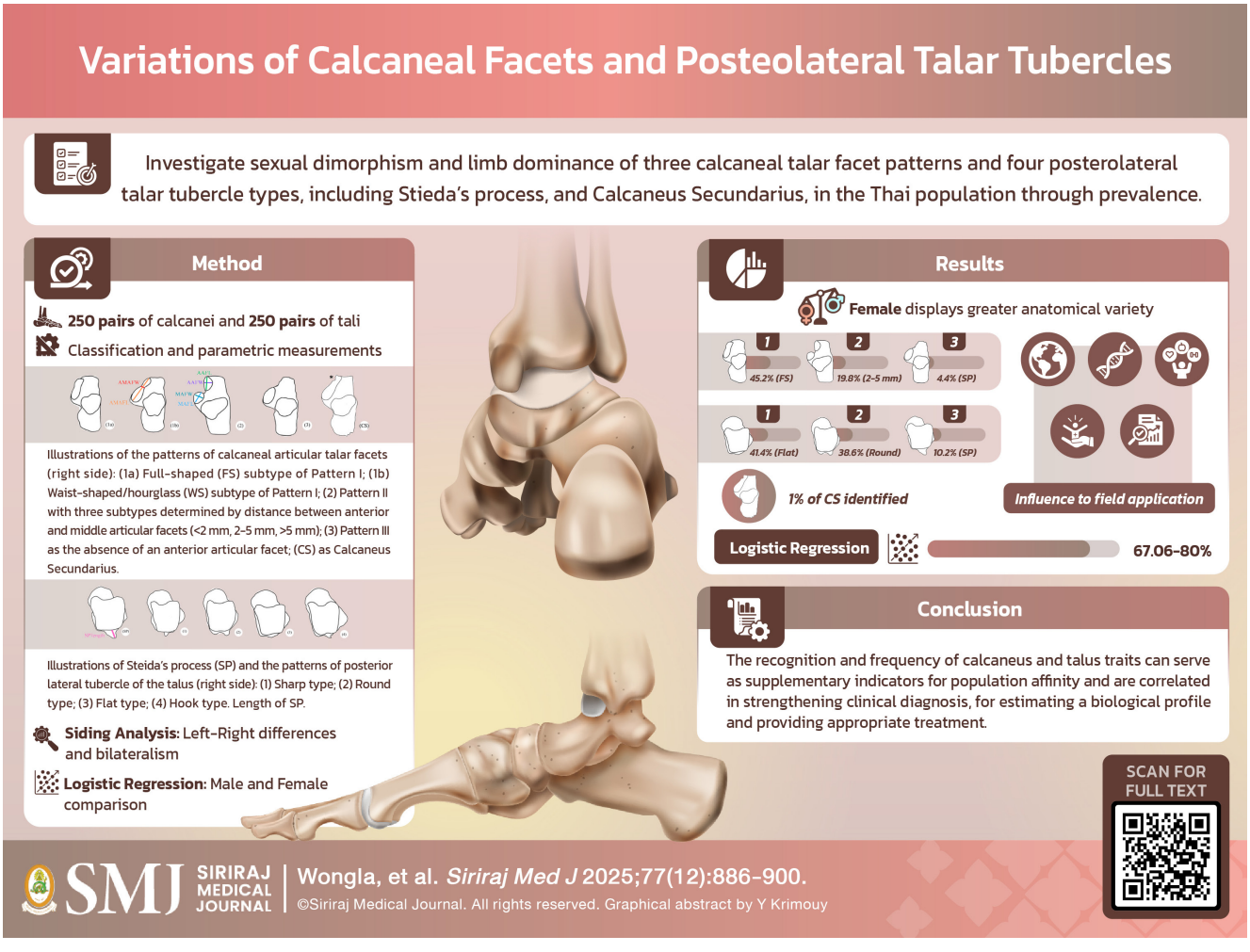
1. Klales AR. Sex estimation using pelvis morphology. In: Klales AR, ed. Sex estimation of the human skeleton: history, methods, and emerging techniques, Cambridge: Academic Press, 2020. p. 75-93.
2. Buikstra JE, Ubelaker DH. Standards for Data Collection from Human Skeletal Remains, Fayetteville: 1994.
3. Krishan K, Chatterjee PM, Kanchan T, Kaur S, Baryah N, Singh RK. A review of sex estimation techniques during examination of skeletal remains in forensic anthropology casework. Forensic Sci Int. 2016;261:165.e1-.e8.
4. Black V. Sex estimation using geometric morphometrics: evaluation of elements of the Pubis. Forensic Anthropology. 2021;4(3):47.
5. Waldron T. A report on the human bone from Merton Priory: English Heritage; 1985.
6. Mann RW, Koel-Abt K, Dhody A, Mahakkanukrauh P, Mann VJ, Techataweewan N, et al. The importance of human osteological collections: Our past, present, and future. Forensic Sci Int. 2021;325:110895.
7. Chomean S, Chatthai N, Sangchay N, Kaset C. Enhancing forensic sex identification through AI-based analysis of the foramen magnum. Forensic Science International: Reports. 2025;11:100411.
8. R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2025. Available from: https://www.R-project.org/
9. Harris SM, Case DT. Sexual dimorphism in the tarsal bones: implications for sex determination. J Forensic Sci. 2012;57(2):295-305.
10. Knapp TR. Technical error of measurement: a methodological critique. American Journal of Physical Anthropology. 1992;87(2):235-6.
11. Ward R, Jamison P. Measurement precision and reliability in

craniofacial anthropometry: implications and suggestions for clinical applications. J Craniofac Genet Dev Biol. 1991;11(3):156-64.
12. Lovell NC. Test of Phenice's technique for determining sex from the os pubis. Am J Phys Anthropol. 1989;79(1):117-20.
13. Suchey JM, Sutherland LD. Use of the ventral arc in pubic sex determination. J Forensic Sci. 1991;36(2):501-11.
14. Ubelaker DH, Volk CG. A test of the phenice method for the estimation of sex. J Forensic Sci. 2002;47(1):19-24.
15. Luo Y-C. Sex determination from the pubis by discriminant function analysis. Forensic Sci Int. 1995;74(1):89-98.
16. Dixit S, Kakar S, Agarwal S, Choudhry R. Sexing of human hip bones of Indian origin by discriminant function analysis. J Forensic Leg Med. 2007;14(7):429-35.
17. Blake KA, Hartnett-McCann K. Metric assessment of the pubic bone using known and novel data points for sex estimation. J Forensic Sci. 2018;63(5):1472-8.
18. Patriquin M, Steyn M, Loth S. Metric analysis of sex differences in South African black and white pelves. Forensic Sci Int. 2005;147(2-3):119-27.
19. Bytheway JA, Ross AH. A geometric morphometric approach to sex determination of the human adult os coxa. J Forensic Sci. 2010;55(4):859-64.
20. Gonzalez PN, Bernal V, Perez SI. Geometric morphometric approach to sex estimation of human pelvis. Forensic Sci Int. 2009;189(1-3):68-74.
21. Torimitsu S, Makino Y, Saitoh H, Sakuma A, Ishii N, Yajima D, et al. Morphometric analysis of sex differences in contemporary Japanese pelves using multidetector computed tomography. Forensic Sci Int. 2015;257:530.e1-e7.
22. Mahakkanukrauh P, Ruengdit S, Tun SM, Case DT, Sinthubua A. Osteometric sex estimation from the os coxa in a Thai population. Forensic Sci Int. 2017;271:127.e1-e7.
23. Vacca E, Di Vella G. Metric characterization of the human coxal bone on a recent Italian sample and multivariate discriminant analysis to determine sex. Forensic Sci Int. 2012;222(1–3):401.e1-e9.
24. Novotný V. Sex determination of the pelvic bone: a system approach. Anthropologie Paris. 1986;24(2-3):197-206.
25. Steyn M, İşcan M. Metric sex determination from the pelvis in modern Greeks. Forensic Sci Int. 2008;179(1):86.e1-e6.
26. Franklin D, Cardini A, Flavel A, Marks MK. Morphometric analysis of pelvic sexual dimorphism in a contemporary Western Australian population. Int J Legal Med. 2014;128(5):861-72.
27. Arun M, Nagesh K, Kumar GP. Estimation of sex from fragments of os coxa by metric analysis. Australian Journal of Forensic Sciences. 2012;44(2):145-53.
28. Sangchay N, Dzetkuličová V, Zuppello M, Chetsawang J. Consideration of accuracy and observational error analysis in pelvic sex assessment: A study in a Thai cadaveric human population. Siriraj Med J. 2022;74(5):330-339.
29. Lüthje P, Nurmi N, Kataja M, Heliövaara M, Santavirta S. Incidence of pelvic fractures in Finland in 1988. Acta Orthop Scand. 1995;66(3):245-8.
30. Ragnarsson B, Jacobsson B. Epidemiology of pelvic fractures in a Swedish county. Acta Orthop Scand. 1992;63(3):297-300.
31. Kane WJ. Fracture of the pelvis. In: Rockwood CA, Green DP, eds. Fractures in Adults, Philadelphia: J.B. Lippincott, 1984.p.1093-209.
31. Rogers T, Saunders S. Accuracy of sex determination using morphological traits of the human pelvis. J Forensic Sci. 1994;39(4):1047-56.

Calcaneal Articular Talar Facets, Stieda's Process, and Calcaneus Secundarius: Variations Found in Thai Population

Areeya Marie B. Wongla, B.Sc.[1], Patara Rattanachet, Ph.D.[1], Saranya Honghimaphan, B.Sc.[1], Natipong Chatthai, M.Sc.[1,2], Parawee Jitrabeab, M.Sc.[1], Napakorn Sangchay, M.D., Ph.D.[1,*]

[1]*Siriraj Anatomical and Anthropological Bone Research Centre (Virapan Davivongs), Department of Anatomy, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand,* [2]*Colleges of Medical Science, Western University Bangkok Campus, Bangkok, Thailand.*



*Corresponding author: Napakorn Sangchay
E-mail: napakorn.sac@mahidol.ac.th
Received 3 September 2025 Revised 4 November 2025 Accepted 9 November 2025
ORCID ID:http://orcid.org/0000-0001-9250-3702
https://doi.org/10.33192/smj.v77i12.277342

ABSTRACT

Objective: The calcaneus and talus have been studied in anatomy and pathology, but research is limited compared to its forensic use. This study examines the prevalence of three types of variants (calcaneal talar facet types, Stieda's process, and calcaneus secundarius) to find correlations with sex and siding of the calcaneus and talus in a Thai population.

Materials and Methods: A total of 250 specimens from the bone collection of Siriraj Anatomical and Anthropological Bone Research Centre (Si-AABRC), Thailand, were used. The calcaneus was classified into three types based on its articular talar facets, and the talus was categorized based on the presence of Stieda's process and four types of posterolateral tubercles. Presence of a calcaneus secundarius was checked for a crescent-shaped notch and an accessory ossicle.

Results: Pattern I, the most common facet type, accounted for 60.23% of calcanei, with a 45.2% average of "full-shaped" facets. Pattern III was the least common, averaging 4.4%. The flat posterior processes of the talus were the most frequent trait (41.4%), while the hook type was the least common (2.4%). Calcaneus secundarius (CS) and Stieda's process (SP) were the rarest traits.

Conclusion: The aforementioned calcaneal and talar traits can aid in investigating sides, as left or right, but not as a specific sex characteristic. The calcaneus secundarius has no correlation to either sex or siding due to limited samples. The outcomes of prevalence align with prior studies.

Keywords: Calcaneus; talus; facet; prevalence; anatomy; morphology (Siriraj Med J 2025; 77: 886-900)

INTRODUCTION

The calcaneus is irregularly cuboidal and the largest among the tarsals. It is located inferiorly to the talus, and therefore susceptible to fractures from high instant loads of force.[1,2] Three articular facets of the calcaneus are described in anatomical nomenclature as: anterior, medial, and posterior articular facets.[3] As part of the subtalar joint, these three facets aid in the maintenance of the foot's longitudinal arch, alongside the talus, to prevent flatfoot abnormality.[4-6] The talus is shaped similar to a saddle and is second to the largest of the tarsals. It articulates with the calcaneus, essential to joint movement.[7] Just as the calcaneus noticeably has variations of its articular surfaces to the talus, so does the talus to the calcaneus— but with only two facets, namely the medial and posterior.[8]

Population-specific and cultural behaviors influence the occurrence of different facet types of the calcaneus and talus, such as squatting in Asia and Africa.[9,10] Although the revelation is not new, the different shapes of facets could cause interference when studying parameters as the measurements of each calcaneal talar type do not have a comprehensive outline.[11,12] Yang et al. (2019)[12] quotes a study by Lawrence et al. (1989)[13] that details how facet classification will be advantageous, especially in the fields of surgery and diagnosis. Nonetheless, there are studies that focus on the morphological features of the calcaneal facets that can link to sexual dimorphism.[14-16]

Foot pathologies are studied to understand their association with facet diversity.[16-18] Syndromes of the foot can be linked to accessory bones, which may later fuse to bones as osteophytes. Skeletal foot variation supports trauma interpretation and population affinity estimations in anthropology[15,19], leading to anatomical and morphological study categories applicable to various research fields.

In the calcaneus, the articular facets are extensively studied to investigate incidence and functional implications.[5,10,12,20,21] Standard anatomy textbooks, like Gray's Anatomy and Netter Atlas of Human Anatomy, and morphometric research analyze the superior portion of the calcaneus into patterns. However, three variants are generally recognized: fusion of the anterior and middle facets, separation of the anterior and middle articular facets, and absence of the anterior articular facet, as per the talus and calcaneonavicular joint.[3,21-24] A trait called 'calcaneus secundarius' refers to the notch in the anterior articular facet filled by an accessory bone. In foot radiography, the trait is evident through diagnosis as the medial border of the anterior articular facet is "chipped off," presenting an ossicle. In dry bones, the porosity and shape of the notch are considered, with or without the presence of the accessory bone.[25,26] In metric studies, such as Ouamthong et al.[27] and Scott et al.[28], the measurements of length, breadth, and height of calcanei were conducted to determine applicability in

determining sex or population reference in forensic cases. This study pertains to only measure the articular facets.

In the talus, the diversities of the posterior process are subjected to more frequent research as opposed to the calcaneus secundarius.[26] An elongated ossification from bone spur formation of the posterior lateral tubercle, is known as 'Stieda's process'. Named by German anatomist Ludwig Stieda in 1869, several studies have been conducted to learn more of the process's influence on foot pathology and prevalence in sex and population. Through radiography, Stieda's process is defined as having a lateral tubercle at least 5 mm in length.[29] Other variants of the posterior process describe the shapes of the tubercle, including sharp, round, flat, and hook.[30,31] The length of the posterior lateral tubercle will be considered for Stedia's process and other variants, not solely based on visual observation.

This study aims to examine the prevalence of three calcaneal talar facet types, Stieda's process, and calcaneus secundarius in the Thai population. It also investigates the correlation between limb dominance and sex to understand calcaneal and talar sexual dimorphism.

MATERIALS AND METHODS

Materials

A total of 250 specimens were used for this study, consisting of 250 pairs of dry calcanei and 250 pairs dry tali, each for male and female. The age at death of the samples ranged from 16 to 95 years in males and 16 to 99 years in females. The mean age in males and females is 69 and 68, respectively. The samples were chosen to exclude those with structural damage that could hinder parameter measurement, such as osteoporosis, severe spur, and postmortem damage. An incomplete set of bones, where either side of the calcaneus or talus is missing, is also part of the exclusion criteria. The bones were obtained from the bone collection of the Siriraj Anatomical and Anthropological Bone Research Centre (Si-AABRC), Department of Anatomy, Faculty of Medicine Siriraj Hospital, Bangkok, Thailand.

Methodology

From the third edition of Human Osteology[22], three patterns of calcaneal talar facets are considered and used in this study for classification:

(1) Pattern I— anterior and middle articular facets are fused together without constriction as a single facet, but with a separated posterior facet, consisting of two subtypes; (Fig 1a-1b)

(2) Pattern II— anterior, middle, and posterior articular facets are separated and consist of three subtypes, determined by the distance between the anterior and middle articular facets; (Fig 1(2))

(3) Pattern III—anterior articular facet is absent or too minimal in size to be considered a facet, while the middle articular facet and posterior articular facet stand independently. (Fig 1(3))

In Pattern I, fusion of the anterior and middle facets meets the criterion of no gap between the two landmarks, and can sometimes be seen with a smooth "ivory-like, shiny patch" that can often be associated as eburnation.[22] There are two subtypes that can visibly categorize Pattern I, "full-shaped" (FS) to characterize the fusion as one union with a continuous surface, and "waist-shaped/hourglass" (WH) to characterize the fusion of both the anterior and middle articular facets but a slight boundary or faintly "raised rims" can be seen to indicate the formerly two facets and is directly named by its shape (Fig 2).[3,22,32]

To expand on Pattern II, the distances between the anterior and middle articular facets are categorized into subtypes derived from identical assessments done by Agarwal et al.[31], Boyan et al.[9] and Koshy, Vettivel, & Selvaraj[30]: Pattern IIa has a distance less than 2 mm; Pattern IIb has a distance between 2-5 mm; and Pattern IIc has a distance more than 5 mm— the distance is measured from the most posterior point of the anterior facet to the most anterior point of the middle facet (Fig 3).

In Pattern III, an anterior articular facet of the calcaneus would articulate at the head of the talus (subtalar joint),

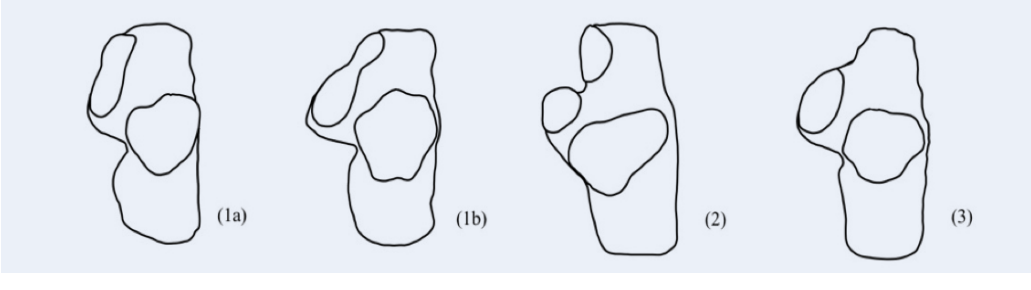


Fig 1. Illustration of the patterns of calcaneal articular talar facets (right side): **(1a)** Full-shaped (FS) subtype of Pattern I; **(1b)** Waist-shaped/hourglass (WS) subtype of Pattern I; (2) Pattern II with three subtypes determined by distance between anterior and middle articular facets; (3) Pattern III as the absence of an anterior articular facet.

Fig 2. Pattern I calcaneal talar facet subtypes of the right calcaneus. See (1a) and (1b) in Fig 1 for illustration.

Fig 3. Pattern II calcaneal talar facet subtypes of the right calcaneus depicting the least distance between anterior and middle articular facets to the most distant (left to right). See (2) in Fig 1 for illustration.

not its neck (talocalcaneal neck articulation). Therefore, the absence of an anterior articular facet (AbAAF) is the most prominent feature of Pattern III (Fig 4).

Stieda's process (SP) refers to the extended ossification of the posterior lateral tubercle of the talus.[33] When evaluating whether the SP is present, the talus is placed on a flat surface then viewed superiorly (Fig 5). If the posterolateral tubercle is enlarged or prominent in length, it is then marked present. For samples marked with the presence of SP, the lowest tip of the talar trochlea to the most posterior tip of the lateral talus process is measured to determine the average length (Fig 7). Absence of SP is preceded by noting down the anatomical variants derived from the criteria by Yang et al.[34] and Kalbouneh et al.[33] Four types are observed as follows:

(1) Sharp: lateral tubercle is a small, pointed projection;
(2) Round: lateral tubercle has a blunt, rounded end but short in size to not be mistaken as SP, with a groove for the flexor hallucis longus tendon;
(3) Flat: lateral tubercle and medial tubercle are faintly seen with the groove for flexor hallucis longus tendon shallow in appearance;
(4) Hook: lateral tubercle is wide and has a "neck" or the corner of the tubercle is chipped and curved towards or away from the groove for flexor hallucis longus tendon.

Fig 4. Pattern III calcaneal talar facet type of the right calcaneus. See (3) in Fig 1 for illustration.

Calcaneus secundarius (CS) is evaluated by the crescent-shaped notch at the medial border of the anterior articular facet alongside an accessory ossicle.[32] In the case that the ossicle is not found, the criteria to note down the presence of CS are the smooth roundedness/triangulation of the notch's shape and porosity at the notch.[26,32,35] (Fig 6).

Morphometric Measurement

To determine the trait by sex estimation, the calcaneus and talus are subjected to measurement using a vernier caliper (mm). In Pattern I, the anterior articular facet

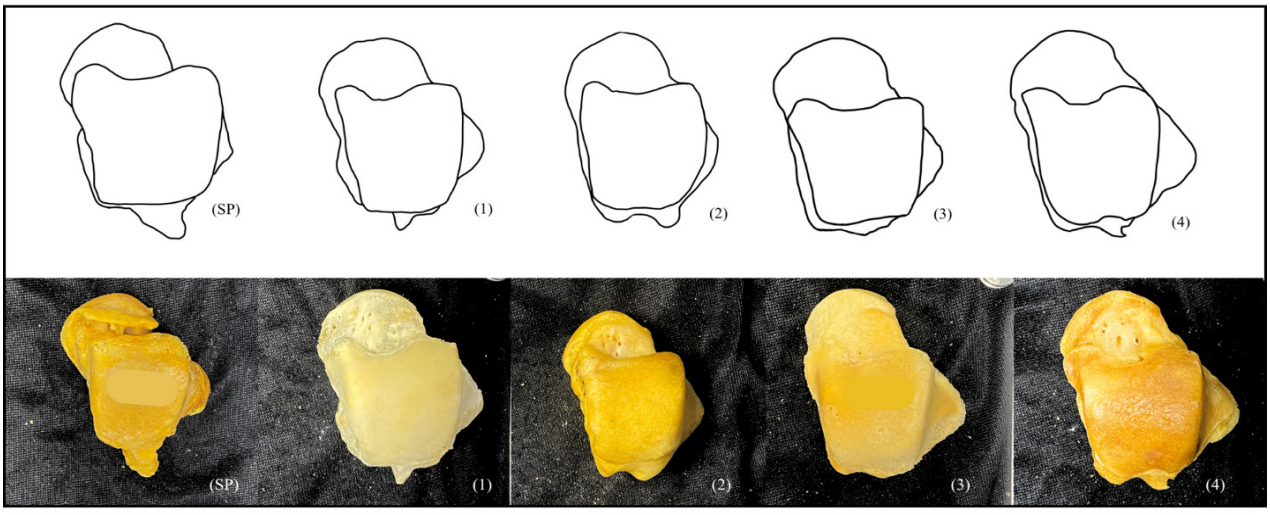


Fig 5. Illustration and photographs of Stieda process (SP) and the patterns of posterior lateral tubercle of the talus (right side): (1) Sharp type; (2) Round type; (3) Flat type; (4) Hook type.

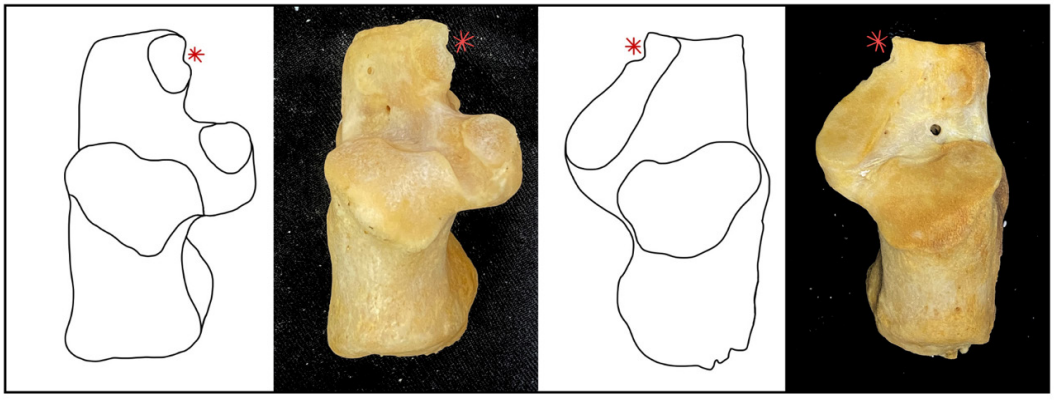


Fig 6. Illustration and photographs of the calcaneus secundarius (CS) of the left and right calcaneus. Red asterisk indicates presence of CS.

(AAF) and middle articular facet (MAF) are fused together, therefore the conjoined facets (AMAF) are then measured from the most anterior point to the most posterior point (AMAFL), and the most medial point to the most lateral point of the fused facets (AMAFW). In Pattern II, AAF is measured by the length of the anterior articular facet (AAFL) and the width of the anterior articular facet (AAFW), applicable to MAF by measuring the length (MAFL) and width (MAFW) from its anterior border to posterior border of the facet. Since Pattern III has an absent anterior articular facet (AbAAF), only MAFW and MAFL were measured. For samples marked with the presence of SP, the lowest tip of the talar trochlea to the most posterior tip of the lateral talus process is measured to determine the average length (Fig 7).

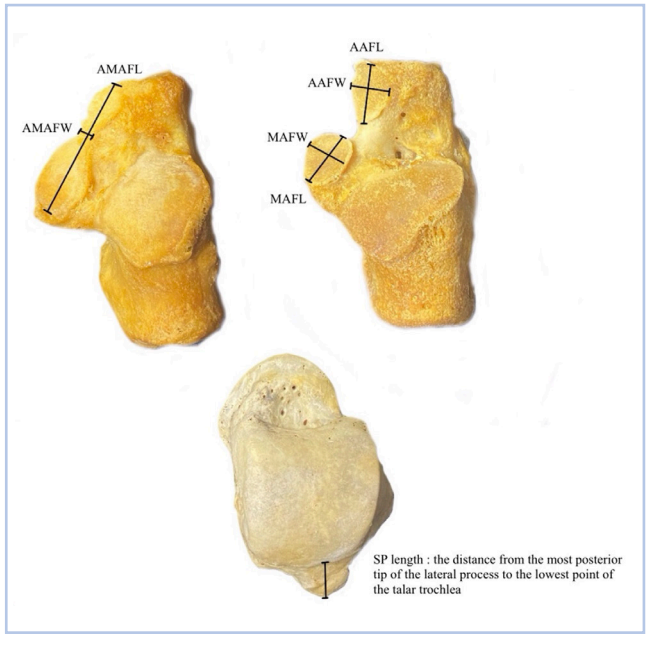


Fig 7. Parameters of calcaneal articular talar facet measurements and SP measurement.

Statistical analyses

Data is collected and submitted into Jamovi Version 2.5 (2024)[36], an open statistical software, to analyze

descriptive values of the calcaneus and talus including mean and standard deviation (SD). The data of the calcaneus and talus is assessed for normal distribution by the Shapiro-Wilk normality test. The formula to calculate the degree of sexual dimorphism (DoSD) was originated from Ricklan & Tobias (1986), which was used in former bone studies, namely the skull, scapula and clavicle.[37,38]

$$\left(\frac{X_{male} - X_{female}}{X_{male}}\right) \times 100 = \text{degree of sexual dimorphism (\%)}$$

To determine sex differences, the data is subjected to independent t-test, but if assumption of distribution is violated, then the Mann–Whitney U test is implemented. Siding (left and right side) of the calcaneus and talus to sex is assessed by paired t-test for normal distribution and Wilcoxon Signed Rank for violation of normal distribution. Chi-square test is applied for frequency data to determine the significance of prevalence. The value $p < 0.05$ determines statistical significance in this study.

Logistic regression analysis

A logistic regression equation was formulated using Posit Cloud[39], formerly known as R Studio (Version 2025.5.1.513.3). The equation is developed by generalized linear model function – glm(). This model is then evaluated for the log of odd ratio (logit) and odds ratio (OR), where a logit equation for each side of a parameter was developed. The ROC curve was created where the model that presented with the lowest AIC (Akaike Information Criterion) and highest Area Under the Curve (AUC) was selected, by using pROC package.[40] For each calcaneal pattern, the best parameters developed an equation where accuracy, AIC, and AUC were input into tables to determine the parameters with the most efficacy.

In the case of small sample number, the logistic regression models were computed to prevent overfitting by applying R package logistf by Heinze & Ploner (2024)[41] into Firth's bias-reduction method.[42]

RESULTS

The prevalence of calcaneal talar facet types is presented in Table 2. Pattern I was the most prevalent, accounting for 58.4% of samples. Within this pattern, 42.8% of specimens exhibited a full-shaped (FS) facet while 15.6% demonstrated an hourglass or waist-shaped (WS) facet, both subtypes with similar representation between males and females. Pattern II represented the second most common type, observed in 37.2% of cases. Among these, 7.2% displayed an average distance between facets at 1.66 mm, with an equal distribution between sexes; 19.6% showed a mean distance of 3.96 mm, occurring more frequently in females than males; and 10.4% exhibited a mean distance of 5.97 mm, equally comprised in males and females. Pattern III was the least frequent, observed in only 4.4% of samples characterized by the absence of an anterior articular facet, with higher occurrence in males than in females. In the odds ratio column, males are more likely to feature Patterns I and II than females. However, the frequency of types indicates no significance as p-value >0.05.

In Table 3, the order of prevalence is similar to Table 2. Pattern I exhibits the highest prevalence of calcaneal talar types, accounted in 62.4% of samples. FS is found in 47.6% of samples with the frequency of females higher than males. WS is identified at 14.8% of samples, the number higher in females than males. Pattern II is in 33.2% samples. The first subtype exhibits in 4.8% of samples, an average distance between articular facets at 1.52 mm with 3 found in males and 9 in females; 20% of the samples have the second subtype with an average distance of 4.20 mm, 30 in males and 20 in females; the third subtype represents 8.4% of samples with an average at 7.06 mm between articular facets, comprising 13 in males and 8 in females. Pattern III is deemed the least varied, with a prevalence of 4.4% in samples found with 6 in males and 5 in females. By gathering the mean values of male and female parameters, Pattern II has the highest degree of sexual dimorphism at 8.38%. The odds ratio and significance of type frequency imply similar results to Table 2.

Sexual dimorphism statistics are summarized in Table 4. With the left and right sides of the calcaneus, eight out of sixteen variables have significant differences with p-value <0.05, which are: AMAFL (L&R) and AMAFW (L) of Pattern I, AAFW (L), MAFL (L&R), and MAFW (L&R) of Pattern II, and MAFW (L) of Pattern III. Together with Table 2 and Table 3, the average of sexual dimorphism degrees of the calcaneus are: Pattern I with 2.32%, Pattern II with 7.41%, and Pattern III with 6.68%.

In Table 5, the siding to sex prevalence of posterolateral tubercles showcases the flat variant to be the highest in prevalence at 41.4% or 207 tali which is followed closely by the round variant at 38.6% or 193 tali. The hook variant, or neck-like variant, is the least prevalent of types at 2.4% (12 tali) with the sharp variant higher at 7.4% (37 tali). SP is identified to be present in 10.2% of the talus samples (51 tali), which is the third highest prevalence of traits identified. The highest degree of sexual dimorphism found is the sharp posterior process

TABLE 1. List of abbreviations and descriptions.

Abbreviations	Description
Anterior articular facet (AAF)	The anterior articular facet situated on the lateral surface of the calcaneus, positioned medially.
Middle articular facet (MAF)	The middle articular facet situated on the lateral surface of the calcaneus, positioned medially.
Anterior and middle articular facet (AMAF)	A grouped abbreviation for Pattern I, FS and WS.
Anterior articular facet length (AAFL)	The distance from the anterior border to the posterior border of the anterior articular facet
Anterior articular facet width (AAFW)	The distance from the medial border to the lateral border of the anterior articular facet
Middle articular facet length (MAFL)	The distance from the anterior border to the posterior border of the middle articular facet
Middle articular facet width (MAFW)	The distance from the medial border to the lateral border of the middle articular facet
Anterior and middle articular facet length (AMAFL)	The distance from the anterior border of the anterior articular facet to the posterior border of the posterior articular facet
Anterior and middle articular facet width (AMAFW)	The distance from the medial border to the lateral border of the fused anterior and middle articular facet, also known as the “width of constriction” according to Mcshane[43]
Full-shaped facet (FS)	Fusion of the AAF and MAF of the calcaneus into one continuous surface.
Waist-shaped facet (WS)	Fusion of the AAF and MAF of the calcaneus into one surface, either continuous or shows signs of individual articular facets. The gap between AAF and MAF is filled but is minimal in size, presenting the articular facet to be shaped like an hourglass.
Absent anterior articular facet (AbAAF)	The anterior articular facet is absent or too minimal in size to be considered an AAF based on articulation to the head of the talus.
Stieda’s process (SP)	An elongated projection of the posterior process of the talus, specifically the lateral tubercle
Calcaneus secundarius (CS)	An accessory bone (ossicle) located near the anterior process of the calcaneus with the ossicle fitting into the notch of AAF

TABLE 2. Contingency Tables for Calcaneal Articular Talar Facet Types and Prevalence of the Left Calcaneus.

Facet Type	Sex (L)		Total	Prevalence (%)	Significance of prevalence	Odds ratio	Average distance between articular facets (mm)
	Male	Female					
Pattern Ia (FS)	54	53	107	42.8	0.701	1.10	-
Pattern Ib (WS)	19	20	39	15.6	0.593	0.826	-
Pattern IIa	9	9	18	7.2	0.892	0.892	1.66
Pattern IIb	23	26	49	19.6	0.873	0.950	3.96
Pattern IIc	13	13	26	10.4	0.541	0.779	5.97
Pattern III (AbAAF)	7	4	11	4.4	0.355	0.557	-
Total	125	125	250	100	-	-	-

Abbreviations: FS = Full-shaped facet; WS = Waist-shaped facet; AbAAF = Absent anterior articular facet

TABLE 3. Contingency Tables for Calcaneal Articular Talar Facet Types and Prevalence of the Right Calcaneus.

Facet Type	Sex (R)		Total	Prevalence (%)	Significance of prevalence	Odds ratio	Average distance between articular facets (mm)
	Male	Female					
Pattern Ia (FS)	57	62	119	47.6	0.704	0.908	-
Pattern Ib (WS)	17	20	37	14.8	0.597	1.21	-
Pattern IIa	3	9	12	4.8	0.150	2.35	1.52
Pattern IIb	30	20	50	20	0.527	0.818	4.20
Pattern IIc	13	8	21	8.4	0.512	0.749	7.06
Pattern III (AbAAF)	6	5	11	4.4	0.758	0.826	-
Total	125	125	250	100	-	-	-

Abbreviations: FS = Full-shaped facet; WS = Waist-shaped facet; AbAAF = Absent anterior articular facet

TABLE 4. Descriptive statistics, Shapiro-Wilk normality test of each parameter for both sexes and both sides of the calcaneus.

Parameters		Side	Male (n = 250) Mean	SD	Female (n = 250) Mean	SD	Mean difference	P value	DoSD (%)	Cohen's D
Pattern I	AMAFL	L	32.48	2.72	32.81	3.01	2.780	**<.001**	1.01	-0.875
		R	29.58	3.79	30.51	3.38	2.455	**<.001**	3.14	-0.818
	AMAFW	L	9.60	2.17	9.12	2.28	0.734	**0.043**	5.00	0.359
		R*	8.83	2.12	8.82	2.27	0.897	0.082	0.11	0.281
Pattern II	AAFL	L	12.43	2.10	11.81	1.81	0.140	0.183	4.99	0.274
		R	12.40	2.23	11.92	3.15	0.665	0.170	3.87	0.153
	AAFW	L*	9.38	1.75	8.38	1.77	0.006	**0.006**	10.66	0.569
		R*	9.01	2.04	9.29	4.88	0.360	0.577	3.11	-0.079
	MAFL	L*	20.14	3.05	17.93	2.78	0.444	**<.001**	10.97	0.780
		R	20.56	2.85	18.02	2.27	2.538	**<.001**	12.35	0.997
	MAFW	L*	12.07	2.46	11.40	1.34	0.671	**0.037**	5.55	0.780
		R	12.09	1.46	11.15	1.18	0.937	**<.001**	7.78	0.698
Pattern III (AbAAF)	MAFL	L	21.5	0.79	21.3	1.66	0.220	0.841	0.93	0.645
		R	20.5	1.12	20.9	1.87	0.815	0.552	3.69	-0.193
	MAFW	L	11.9	0.82	10.4	0.79	2.150	**0.003**	16.8	1.442
		R	11.5	1.26	10.7	1.58	0.618	0.614	5.31	0.585

*n = Sample size (pair of calcaneus); * = Shapiro Wilk (Normality Test): violation of normality assumption proceeds with Wilcoxon Signed Rank test. Bold text indicates significant difference.*

TABLE 5. Prevalence and degree of sexual dimorphism of the posterolateral tubercle types to sex.

Posterolateral tubercle types		Left (n = 250)	Right (n = 250)	Total	Prevalence (%)	DoSD (%)	Significance of prevalence	Odds ratio
Sharp	M	8	9	37	7.4	15.0	0.599	0.835
	F	10	10				0.599	0.835
Round	M	50	43	193	38.6	7.53	0.389	0.854
	F	52	48				0.389	0.854
Flat	M	55	49	207	41.4	0.96	0.885	0.974
	F	51	52				0.885	0.974
Hook	M	0	4	12	2.4	1.0	0.240	2.040
	F	3	5				0.240	2.040
Stieda's process (SP)*	M	12	20	51	10.2	4.11	0.057	0.563
	F	9	10				0.057	0.563

**The length of Stieda's process in this study ranges from 5.11 - 6.92 mm*

of the talus at 15% and the least to be the hook posterior process. The presence is seen to be higher in males than females. However, all the four types of posterolateral tubercle types and SP do not have significant statistical difference nor significance in prevalence as all p-value exceeds 0.05.

Table 6 portrays the bilateralism and siding impact of the posterolateral tubercle types to the left and right sides of the talus. All traits have no significance to siding since the p-value exceeds 0.05.

The logistic regression analysis in Table 7 is conducted by standard logistic regression, and indicates the lowest AIC value and highest AUC value to be the most effective parameter to identify sex of the calcaneus by articular facet. In Pattern I, AMAFL has the strongest discrimination value, favoring the left side (AUC = 0.796). MAFL of Pattern II has the highest AUC values, with the left dominating over the right (AUC = 0.759).

Given the small sample size (n = 11), all analyses of Pattern III is considered exploratory until future research strengthens evidence of morphological differences between sexes in a bigger dataset. Firth's logistic regression is computed to reduce bias as Pattern III has an insufficient amount of data. In single logistic regression, MAFW.L demonstrated the strongest relationship with sex (AUC = 0.792), indicating a possible trend where greater width may increase the likelihood of being male while other variables showcased weaker associations (AUC ≤ 0.667).

DISCUSSION

There are considerable variations of anterior and middle articular facet patterns in different cohorts of people. Previous studies have also yielded results indicating a higher proportion of female samples to be more diverse compared to male samples, although the disparity may not be significant.[44] Pattern I resulted with the highest incidence found amongst the samples, which is consistent with numerous studies conducted within Asian populations.[4,6,11,14,31,45] In contrast, other populations exhibit Pattern II by the highest incidence, such as a Bulgarian population[5], Anatolian population[9], Turkish population[44], and South Africans of European descent.[15] European female samples have Pattern I as a majority whereas Indian and African female sample groups have a higher ratio of Pattern II.[44] Pattern III is the least prevalent type of calcaneal talar facet, in agreement with a study dating back to 1974[11,45] and McShane's (2021) comprehensive literature review in demographic studies, where the siding evaluation shows an average prevalence of 4%.[43]

Ancestral differences of the articular facets of the calcaneus have been acknowledged because of repeated impressions in different populations, but no consensus has been formally established.[17,46] A plausible explanation of how the lengths of each parameter have significance could be from how the fusion to a continuous facet is formed into two types, FS and WS. FS and WS are

TABLE 6. Bilateralism and siding significance of the posterolateral tubercle types of the talus.

Posterolateral tubercle types		Total	Side			P-value	Odds ratio
			Unilateral (L)	Unilateral (R)	Bilateral		
Sharp	L	18	1	2	16	0.864	0.943
	R	19				0.864	0.943
Round	L	102	17	4	92	0.856	1.030
	R	91				0.856	1.030
Flat	L	106	11	7	94	0.716	0.936
	R	101				0.716	0.936
Hook	L	3	3	10	1	0.080	3.070
	R	9				0.080	3.070
Stieda's process*	L	21	2	11	12	0.184	1.490
	R	30				0.184	1.490

Bold text indicates significant difference.

TABLE 7. Logistic Regression Model of calcaneus parameters for sex estimation.

Single Logistic Regression						
Parameters		**Side**	**Logit**	**OR (95% CI)**	**AIC**	**AUC**
Pattern I	AMAFL	L	-9.981 + 0.318 x AMAFL.L	1.375 (1.21-1.59)	173.90	0.769
		R	-8.437+0.266 x AMAFL.R	1.306 (1.16-1.49)	180.28	0.744
	AMAFW	L	-1.619+0.168 x AMAFW.L	1.183 (1.01-1.39)	197.49	0.598
		R	-0.911+0.095 x AMAFW.R	1.099 (0.95-1.28)	200.47	0.558
Pattern II	AAFL	L	-2.084+0.176 x AAFL.L	1.191 (0.96-1.50)	119.33	0.601
		R	-0.690+0.059 x AAFL.R	1.060 (0.90-1.26)	121.30	0.587
	AAFW	L	-2.979+0.341 x AAFW.L	1.407 (1.09-1.87)	114.65	0.674
		R	0.222+(-0.021) x AAFW.R	0.978 (0.85-1.10)	121.69	0.535
	MAFL	L	-5.507+0.293 x MAFL.L	1.341 (1.13-1.63)	109.48	0.711
		R	-7.768+0.405 x MAFL.R	1.500 (1.23-1.89)	102.38	0.759
	MAFW	L	-1.764+0.153 x MAFW.L	1.165 (0.94-1.50)	119.96	0.608
		R	-6.296+0.544 x MAFW.R	1.723 (1.22-2.56)	111.80	0.681
*Pattern III	MAFL	L	5.254+(-2.039) x MAFL.L	0.787 (0.36-1.38)	13.645	0.417
		R	1.839+(-0.073) x MAFL.R	0.930 (0.48-1.65)	14.299	0.500
	MAFW	L	-10.330+ 0.959 x MAFW.L	2.609 (0.89-17.4)	12.264	0.792
		R	-4.022+0.396 x MAFW.R	1.485 (0.59-4.94)	14.646	0.667

Multiple Logistic Regression					
Parameters		**Logit**	**Accuracy**	**AIC**	**AUC**
Pattern I	AMAFL.L + AMAFL.R	-14.993 + 0.267 x AMAFL.L + 0.210 x AMAFL.R	67.06	165.10	0.822
Pattern II	MAFL.L + MAFL.R	-13.377 + 0.298 x MAFL.L + 0.406 x MAFL.R	74.13	94.99	0.793
*Pattern III	MAFL.L + MAFW.L	-3.893 + (-0.471) x MAFL.L + 1.24 x MAFW.L	80	10.39	0.917

**Firth's penalized logistic regression*

distinct in shape, especially with the slimmer constriction width of WS. The ossification of the two articular facets is part of the talocalcaneal joint, suggested to stem from genetic determination from a study conducted on fetal calcanei[47,48], but lifestyle and posture can also influence the development of surfaces.[3,12,23,24] A study by Iamsaard et al.[49] states Pattern I as the most common in Northeastern Thai population, possibly linking to the region of abundant rural activity where most of society practices agriculture.[28] Further supported by Harper et al. (2022)[50], Pattern I is wider as a consequence to higher load on the calcaneus from non-sedentary lifestyle of the population. Thus, the attribution of similarities may possibly be from geographical proximity[51], but with current data conducted in a few Asian groups such as Indian, Thai, and Chinese, the correlation is yet to considered as definitive evidence.

Conversely, the contrast of occurrence in different nations supports a theory about subtalar joint stability. Shweta et al.[46] discusses the hypothesis of the Indian population to be more susceptible to subtalar arthritis, as their findings contribute to an earlier study by Drayer-Verhagen in 1993 stating Pattern I is the most prevalent type.[52,53] The smooth, continuous surface of the facet enhances the mobility of the calcaneus and talas articulation, with the exception of overexerting the surrounding ligaments and muscles and thus, initiates the early stages of osteoarthritis and trauma.[52] In relation to clinical relevance, the knowledge of variants to treatment development is essential to the dynamics of the joint.[12,17,31,51] In a Bangledeshi population[51], Pattern I is believed to be a mechanical influence of the talocalcaneal coalition as it causes instability of movement and painful flatfoot.[54] Flatfoot is not suited for weightbearing and disadvantageous as opposed to having an arched-foot, as an arch restricts overpronation.[55] For Pattern I, AMAF narrows or diminishes the presence of the sinus tarsi and tarsal canal, deforming to flatfoot.[54] Additionally, the "easy gliding" in Pattern I chronically stresses a chance for subtalar arthritis.[55] Calcanei and osteonomy are discussed in CT images by Koh et al. (2024)[56], in Southeast Asian population. Surgical correlation is pertinent as Pattern II forms a well-supported base, or an "architeturally structured tripod," for tali orientation, effectively preventing the progression of osteoarthritis.[18,57]

However, our study lacks age analysis, occupational records of the donated bodies, and further confirmation. Future studies could explore the possibility of calcaneal variants as diagnostic criteria, similar to the Indian study.[46,52,58] Future research should use a larger sample size to evaluate risk factors for subtalar joint stability, especially the development of subtalar arthritis, and subsistence studies in anthropology by regional affiliation and sex.[3,59]

Sexual dimorphism of bone variants has been a recurring topic of study, especially in the establishment of patterns in anatomical literature of the articular surfaces.[16,18,20] Logistic regression is conducted as it is one of the methods to estimate sex in the field of forensics. The AUC values from multivariate equations are higher than univariate equations, suggesting that individual calcaneal variables should be considered for sex estimation when taphonomic change limits the use of multivariate equations.[28] However, this research showcased all parameters of the calcaneal articular facets to have moderate to low discriminatory power for clinical relevance as AUC values are under 0.8[60], possibly due to varied amount of samples in each pattern. To derive forensic significance, variants can be incorporated as supplementary indicators rather than replace conventional biological profile estimation parameters. Ideally, a larger sample size for each parameter of facet type would enhance the accuracy of developing sex estimation based on morphometry.

A literature review covering a large database by Ogut in 2022[61] encompasses the incidence of SP found in the talus range from 12 - 36%. The development of SP stems from a secondary ossification center at varied ages. The length of SP increases the likelihood of fracture, pain, and hinders range of motion (plantar flexion) when strenuous activities are part of daily life, such as the lifestyle of dancers and athletes.[61,62] In a radiography case report, the SP is labelled in a patient when the length is measured to be approximately 5.0 mm but there is no standard to the level of elongation to be considered SP despite clinical relevance.[29,63] The insignificance of SP presence to sex and siding is closely consistent to preceding studies. However, in relation to posteolateral tubercles of the talus, most were comparative studies with patients with ankle impingement and CT scans. Closely in line by ratio is a study in the Chinese population[34], describing the flat variant to be second highest in incidence at 36.29% and the round variant at 38.39%, with the sharp and hook variants following respectively. Over 1,000 ankles, Kalbouneh et al. (2021)[33] discovered the flat variant to be most frequent (46.1%), followed by SP (26.1%) in CT scans of Jordanian population. In a Netherland sample group[64], about 1,200 ankles with "SP-like morphology" occurred in 34.9% of the collection. 16.7 % of SP prevalence was found in a Turkish population.[29] Within Asia, data from Yang et al. (2022)[34] proposes how the occurrence of prevailing types of the talus's posterolateral aspect has biomechanical differences and effect in surgery. The different types may lead to contrasting types of fracture and treatment; such as how ankle arthroscopy should be performed on a displaced hook variant. The hook variant is hypothesized to be highly vulnerable to fracture and

cracks based on a stress distribution program because irregularly shaped bone growth, to minimize risk of debris.[34]

CS, an uncommonly documented ossicle of the foot, is infrequent and unprioritized in foot diagnosis compared to other ossicles.[59] Due to its rarity, the incidence of calcaneus secundarius in this collection is deemed low. In this study, 5 confirmed CS out of 500 calcanei are found with the traits as based on the evaluation method from Mann (1990)[65], 3 found on the left calcaneus and 2 found on the right calcaneus. In a previous study conducted on the Northeastern population of Thailand[47], morphometric measurements and the presence of the CS were recorded. 2 specimens out of 300 samples were found to have a positive identification of CS, which is a lower incidence of 0.66% as opposed to this study's 1% of CS found. The low incidence may be due to the higher percentage of CS found in cadavers, as dry bone collections are prone to missing bones and mistaken discardment, and bigger sample size increasing chances of variation found in collection. Limited CS prevents determining significance and correlation to sex, corresponding to past papers.[26,32] Nonetheless, Candan et al.[66] reiterates CS for differential diagnosis as a handful of case reports mistake the ossicle as an anterior process fracture of the calcaneus, hindering suitable care for persisting foot pain. A case study reconstructing a patient's foot to confirm a case of CS indicates how accessory ossicles can be detrimental if not properly diagnosed. Given the patient's recurrent history of foot pain and unsuccessful attempts at physiotherapy, the excision of CS eliminated the tenderness of direct palpation of the anterior aspect of the sinus tarsi.[67]

CONCLUSION

Over the years, the anatomical variety of the calcaneus and talus has been a topic of study. The frequency of calcaneal and talar morphological variations were compared between sex (male vs. female) and sides (right vs. left) in a sample of Thai dry bones. Nine out of sixteen variables of the three calcaneal talar facet types show sexual dimorphism, which are: AMAFL (L&R) and AMAFW (L) of Pattern I, AAFW (L), MAFL (L&R), and MAFW (L&R) of Pattern II, and MAFW (L) of Pattern III. Pattern I calcaneal articular facet has the highest frequency of prevalence whereas Pattern III has the lowest. The flat posterior processes of the talus is the most frequent trait and the hook type to be the least frequent trait. The four variants of the posterolateral tubercle of the talus do not show sex dimorphism, neither do CS nor SP. The utilization of the calcaneal and talar traits can assist in determining siding, but not much as a characteristic of a specific sex. Logistic regression results are considered exploratory as sample size is limited. CS still holds to its reputation of rarity, and no relationship is found due to limited samples found in the study. However, the outcomes of this research are consistent with prior studies. The recognition of morphology and their frequency can influence in the study of various disciplines, such as diagnosis for specific population-appropriate treatment and anthropology for subsistence affiliation. Therefore, tailored techniques based on the anatomy of the population are essential for providing appropriate care and estimating biological profile.

Data Availability Statement

The data that support the findings of this study are available from Siriraj Anatomical and Anthropological Bone Research Centre (Si-AABRC), upon reasonable request.

ACKNOWLEDGEMENTS

The authors of this research are grateful for the bone collection of Siriraj Anatomical and Anthropological Bone Research Centre (Virapan Davivongs) and Siriraj's body donation program. Special thanks to Dr. Robert W. Mann and Asst. Prof. Dr. Sittiporn Ruengdit for the idea and consultation of the study.

DECLARATIONS

Grants and Funding Information

This research has not received any grants or funds.

Conflict of Interest

The authors of this study have no conflict of interest.

Registration Number of Clinical Trial

This research does not require a registration number for a clinical trial.

Author Contributions

Conceptualization and methodology, A.W., N.C. ; Investigation, A.W., P.J. ; Formal analysis, A.W., S.H. ; Visualization and writing – original draft, A.W. ; Writing – review and editing, A.W., S.H., P.R. ; Supervision, N.S. All authors have read and agreed to the final version of the manuscript.

Use of Artificial Intelligence

This paper applied the assistance of Consensus: AI Search Engine for Research to further explore for relevant papers and ChatGPT-5 (Open AI) for preliminary

R-Studio codes. The generated drafts were verified by A.W.

Ethical Approval

This study was approved by the institutional ethics committee, Faculty of Medicine Siriraj Hospital, Mahidol University [523/2568 (IRB4)].

REFERENCES

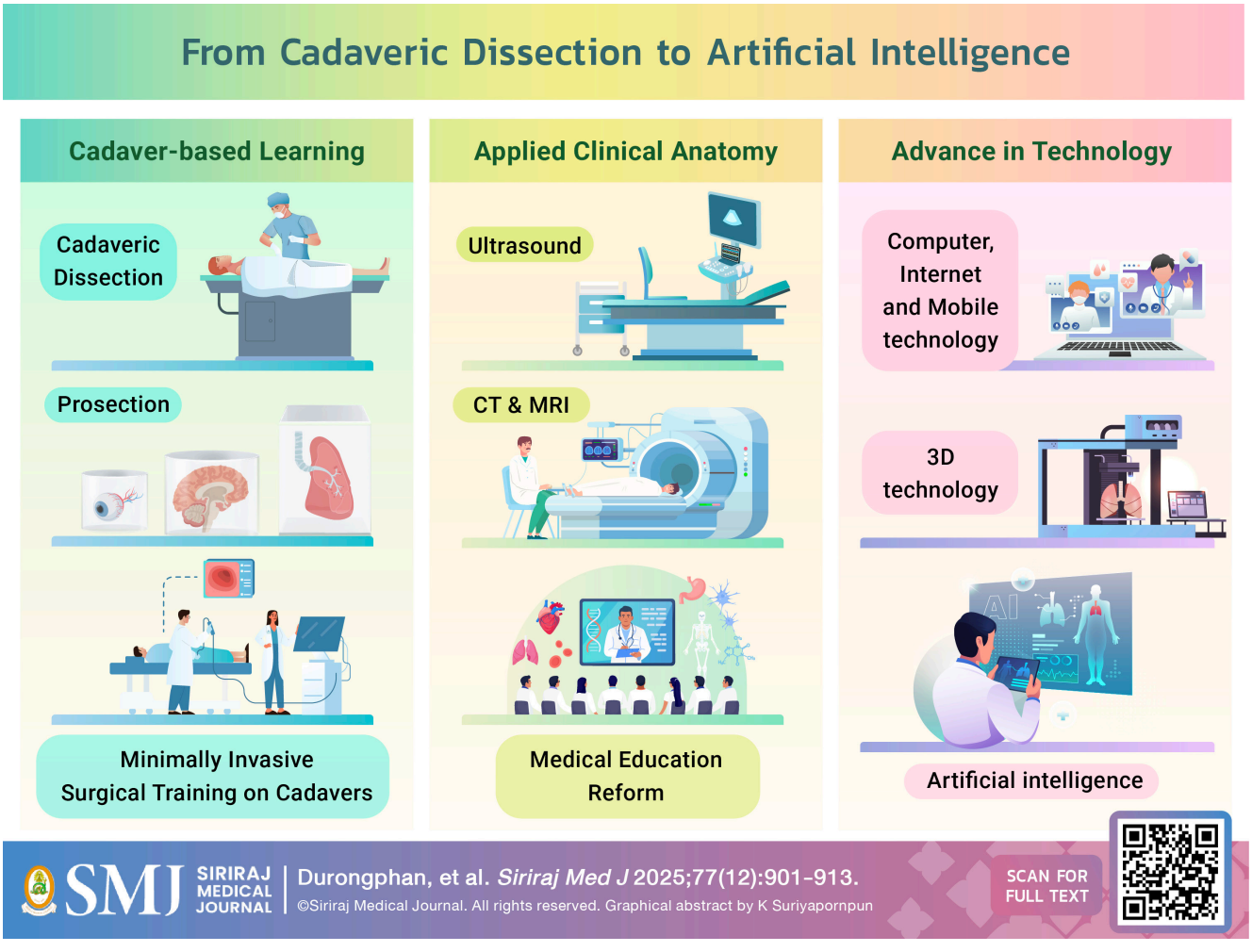
1. Hall RL, Shereff MJ. Anatomy of the calcaneus. Clinical Orthopaedics and Related Research (1976-2007). 1993;290:27–35.
2. Gupton M, Özdemir M, Terreberry RR. Anatomy, bony pelvis and lower limb: calcaneus. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2025 Jan.
3. Cockerill SJ, Arnay-de-la-Rosa M, González-Reimers E. An atlas of anatomical variants of the human calcaneus. J Morphol. 2024;285(5):e21706.
4. Artkengkla T, Piyawinijwong S. Variations of Calcaneal Articular Facets in Thais Related to Navicular Facet ความแตกต่างของชนิดข้อต่อบนกระดูกแคลคาเนียส และความเกี่ยวข้องกับข้อต่อบนกระดูกนาวิคูลาร์ในคนไทย. 34th National Graduate Research Conference. 2015.
5. Angelova M, Marinova D. Anatomical Variations of the Articular Surfaces of the Calcaneus among Bulgarian Population. Acta Morphologica et Anthropologica. 2024;31:1–2.
6. Prasad SA, Rajasekhar S. Morphometric analysis of talus and calcaneus. Surg Radiol Anat. 2019;41(1):9–24.
7. Hegazy AAM, Hegazy MA. Talus bone: Unique anatomy. International Journal of Cadaveric Studies and Anatomical Variations. 2022;3(2):52–5.
8. Khan IA, Varacallo MA. Anatomy, bony pelvis and lower limb, foot talus. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2025 Jan.
9. Boyan N, Ozsahin E, Kızılkanat E, Soames R, OĞUZ Ö. Morphometric measurement and types of articular facets on the talus and calcaneus in an Anatolian population. International Journal of Morphology. 2016;34(4).
10. Srivastava S, Khan AZ, Chigurupalli K, Singh KV, Arora NK, Haque M. Comparative Analysis of Squatting Facets on Femur, Tibia, and Talus: Insights from A Population in Northwest Uttar Pradesh and Cross–Population Comparisons. International Journal of Pharmaceutical and Clinical Research. 2023;15(6):87–96.
11. El-Eishi H. Variations in the talar articular facets in Egyptian calcanei. Cells Tissues Organs. 1974;89(1):134–8.
12. Yang Y, Cheng H-W, Xiong Z-R, Liu N, Liu Y, Wang Y, et al. Classification and morphological parameters of the calcaneal Talar facet: which type is more likely to cause osteoarthritis in Chinese population? Biomed Res Int. 2019;2019:6095315.
13. Lawrence RC, Hochberg MC, Kelsey JL, McDuffie FC, Medsger Jr TA, Felts WR, et al. Estimates of the prevalence of selected arthritic and musculoskeletal diseases in the United States. J Rheumatol. 1989;16(4):427–41.
14. Boonruangsri P, Woraputtaporn W, Namking M. The pattern of talar articular facets in Northeastern Thai calcanei. Srinagarind Hosp Med J. 1992;7(1):28–34.
15. Bidmos M. Metrical and non-metrical assessment of population affinity from the calcaneus. Forensic Sci Int. 2006;159(1):6–13.
16. Nozaki S, Watanabe K, Kamiya T, Katayose M, Ogihara N. Sex-and age-related morphological variations in the talar articular surfaces of the calcaneus. Ann Anat. 2020;229:151468.
17. Garg R, Dagal N, Kumar S, Shekhawat S. Study of patterns of talar articular facets of human calcanei and their clinical implications in population of Rajasthan. Indian J Basic Appl Med Res. 2013;7(2):643–50.
18. Vučinić N, Teofilovski-Parapid G, Erić M, Tubbs RS, Radošević D, Jovančević B. Morphometric analysis of the patterns of calcaneal facets for the talus in Serbian population. PLoS One. 2020; 15(10):e0240818.
19. Spradley MK. Metric Methods for the Biological Profile in Forensic Anthropology: Sex, Ancestry, and Stature. Acad Forensic Pathol. 2016;6(3):391–9.
20. Singh DV. Assessment of morphometric study of calcaneus and its articular facets. Int J Acad Med Pharm. 2023;5(1):801–3.
21. Kumar S, Singh AK, Fatima N, Akhtar J, Ratnesh R, Kumar V. A morphological study on patterns of human calcaneal articular facets for talus in population of bihar and its clinical implications. Journal of Evolution of Medical and Dental Sciences. 2017; 6(56):4193–7.
22. White TD, Black MT, Folkens PA. Human osteology: Academic press; 2011.
23. Mann RW, Hunt DR, Lozanoff S. Photographic regional atlas of non-metric traits and anatomical variants in the human skeleton: Charles C Thomas Publisher; 2016.
24. Mann RW. The bone book: A photographic lab manual for identifying and siding human bones: Charles C Thomas Publisher; 2017.
25. Mann R. Calcaneus secundarius. Variation of a common accessory ossicle. J Am Podiatr Med Assoc. 1989;79(8):363–6.
26. Silva AM, Curate F. Accessory foot bones in a Portuguese identified skeletal collection. Sci Rep. 2024;14(1):17169.
27. Ouamthong R, Mahacharoen T, Inthasan C, Srisinghasongkram J, Singsuwan P, Mahakkanukrauh P. A Comparison of Effectiveness of Sex Estimation from the Calcaneus and Talus in a Thai Population. Int J Morphol. 2023;41(1):268-77.
28. Scott S, Ruengdit S, Peckmann TR, Mahakkanukrauh P. Sex estimation from measurements of the calcaneus: Applications for personal identification in Thailand. Forensic Sci Int. 2017; 278:405.e1–e8.
29. Derin Cicek E, Bankaoglu M. Prevalence of Elongated Posterior Talar Process (Stieda Process) Detected by Radiography. International Journal of Morphology. 2020;38(4).
30. Koshy S, Vettivel S, Selvaraj K. Estimation of length of calcaneum and talus from their bony markers. Forensic Sci Int. 2002;129(3): 200–4.
31. Agarwal S, Garg S, Vasudeva N. Subtalar joint instability and calcaneal spurs associated with the configuration of the articular facets of adult human calcaneum in Indian population. J Clin Diagn Res. 2016;10(9):AC05-AC09.
32. Mann RW, Ubelaker DH. The forensic anthropologist. FBI L Enforcement Bull. 1990;59:20.
33. Kalbouneh H, Alsalem M, Hani MB, Alhusamiah H, Momani Y, Massad T, et al. A Comprehensive Study of the Anatomical Variations of the Posterolateral Tubercle of Talus. International Journal of Morphology. 2021;39(3).
34. Yang H, Liao L, Xue F, Li Y, Hu G. Anatomical observation, classification, fracture and finite element analysis of the posterior process of the Asian adult talus. J Orthop Surg Res. 2022;17(1):444.
35. Keles-Celik N, Kose O, Sekerci R, Aytac G, Turan A, Güler F.

Accessory ossicles of the foot and ankle: disorders and a review of the literature. Cureus. 2017;9(11).
36. Jamovi. The jamovi project (Version 2.5).
37. Koukiasa AE, Eliopoulos C, Manolis SK. Biometric sex estimation using the scapula and clavicle in a modern Greek population. Anthropol Anz. 2017;74(3):241-6.
38. Ricklan D, Tobias P. Unusually low sexual dimorphism of endocranial capacity in a Zulu cranial series. Am J Phys Anthropol. 1986;71(3):285–93.
39. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J-C, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics. 2011;12:77.
40. Heinze G, Ploner M, Dunkler D, Southworth H, Heinze MG. Package 'logistf'. 2025. Available from: https://cran.r-project.org/web/packages/logistf/logistf.pdf
41. Firth D. Bias reduction of maximum likelihood estimates. Biometrika. 1993;80(1):27–38.
42. McShane PA. The prevalence of absent anterior facet of the calcaneus and a suggested convention for the naming of the configuration of the superior facets of the calcaneus. Int J Anat Res. 2021;9(1.3):7928–34.
43. Uygur M, Atamaz F, Celik S, Pinar Y. The types of talar articular facets and morphometric measurements of the human calcaneus bone on Turkish race. Arch Orthop Trauma Surg. 2009;129(7):909–14.
44. Sharada R, Sneha K, Gupta C, Pai SR, Rairam G. Non-metrical study of the pattern of talar articular facets in south Indian dry calcanei. Surg Radiol Anat. 2012;34(6):487–91.
45. Shweta J, Rashvaita K, Krunal R, Meenakshi B. Patterns of talar articular facets on calcaneum and its clinical implication. Int J Anat Physiol. 2013;2(4):23–6.
46. Das S, Sinthubua A, Ruengdit S, Mahakkanukrauh P. Anomalous Calcaneus Secundarius: Anatomical and Clinical Considerations. International Medical Journal. 2018;25(5).
47. Rühli FJ, Solomon L, Henneberg M. High prevalence of tarsal coalitions and tarsal joint variants in a recent cadaver sample and its possible significance. Clin Anat. 2003;16(5):411–5.
48. Iamsaard S, Uabundit N, Boonruangsri P, Sawatpanich T, Hipkaeo W. Types of Facets on the Superior Articular Surface of Isan-Thai Dried Calcanei. Int J Morphol. 2015;33(4):1549–52.
49. Harper CM. Human calcaneal variation relative to subsistence strategy, activity level, and footwear. Front Earth Sci. 2023;11:1213374.
50. Mowtoshee NR, Naushaba H, Kishwara S, Kamal AM, Rahman M, Singha PS. Variations In Talar Articular Facets on Dry Adult Human Left Calcaneus. Bangladesh Journal of Anatomy. 2020;18(1):17–20.
51. Drayer-Verhagen F. Arthritis of the subtalar joint associated with sustentaculum tali facet configuration. J Anat. 1993;183(Pt 3):631-4.
52. Agarwal P, Agarwal V, Kumar A, Kumar S, Gupta R. A Study of Morphological Pattern of the Talar Articular Facets in Dry Human Calcanei and Its Clinical Implications. International Journal of Health Sciences. 2022;6(S4):12676–83.
53. Cheng KY, Smitaman E, Resnick DL. Developmental Talocalcaneal Coalitions and Associated Conditions. MRI Web Clinic; 2022. Available from: https://radsource.us/wp-content/uploads/2022/06/2206_Tarsal-Coalition-Resnick_FINAL.pdf
54. Bruckner J. Variations in the Human Subtalar Joint. J Orthop Sports Phys Ther. 1987;8(10):489–94.
55. Anbumani S, Sridharan R, Selvi A. An anatomical study of morphology and morphometric analysis of calcaneum and its talar articular surfaces. International Journal of Anatomy and Research. 2017;5(3.2):4223–9.
56. Harnroongroj T, Chuckpaiwong B, Angthong C, Nanakorn P, Sudjai N, Harnroongroj T. Displaced articular calcaneus fractures: classification and fracture scores: a preliminary study. J Med Assoc Thai. 2012;95(3):366-77.
57. Ohishi T, Fujita T, Nishida T, Asukai M, Suzuki D, Matsuyama Y. Tibial spastic varus foot caused by os calcaneus secundarius: A case report. Foot (Edinb). 2019;39:92–5.
58. Çorbacıoğlu ŞK, Aksel G. Receiver operating characteristic curve analysis in diagnostic accuracy studies: A guide to interpreting the area under the curve value. Turk J Emerg Med. 2023;23(4):195–8.
59. Ogut E. The Stieda process of the talus: the anatomical knowledge and clinical significance of an overlooked protrusion. Bulletin of the National Research Centre. 2022;46(1):280.
60. Brodsky AE, Khalil MA. Talar compression syndrome. Am J Sports Med. 1986;14(6):472–6.
61. Jilani LZ, Istiyak M, Siddiqui YS. Stieda Process as a Source of Posterior Ankle Pain: A Case Report with Its Structural and Clinical Implication. Journal of Foot and Ankle Surgery (Asia Pacific). 2024;11(3):147–51.
62. Zwiers R, Baltes TP, Opdam KT, Wiegerinck JI, van Dijk CN. Prevalence of os trigonum on CT imaging. Foot Ankle Int. 2018;39(3):338–42.
63. Mann RW. Calcaneus secundarius: description and frequency in six skeletal samples. Am J Phys Anthropol. 1990;81(1):17–25.
64. Candan B, Torun E, Dikici R. The prevalence of accessory ossicles, sesamoid bones, and biphalangism of the foot and ankle: a radiographic study. Foot Ankle Orthop. 2022;7(1):24730114211068792.
65. Ceroni D, De Coulon G, Spadola L, De Rosa V, Kaelin A. Calcaneus secundarius presenting as calcaneonavicular coalition: a case report. J Foot Ankle Surg. 2006;45(1):25–7.

From Cadaveric Dissection to Artificial Intelligence: A Chronological Review of Advances in Anatomy Education

Anuch Durongphan, M.D.

Department of Anatomy, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok 10700, Thailand.



*Corresponding author: Anuch Durongphan
E-mail: anuch.dur@mahidol.ac.th
Received 1 May 2025 Revised 28 July 2025 Accepted 28 July 2025
ORCID ID:http://orcid.org/0000-0002-7216-1052
https://doi.org/10.33192/smj.v77i12.275193

ABSTRACT

Cadaveric dissection has traditionally been the cornerstone of anatomy education. However, in recent decades, its role has come under increasing scrutiny. Comprehensive historical analyses explaining this shift remain limited. This chronologically structured review traces the evolution of anatomy teaching from its origins with Herophilus in the 3rd century BC to the advent of contemporary digital and artificial intelligence (AI)-enhanced approaches. Key milestones include the 1910 Flexner Report, which emphasized integrated biomedical sciences, widespread curricular reforms, increased reliance on prosection, the rise of minimally invasive surgery, and the incorporation of cadaveric simulation in postgraduate training. Advances in imaging technologies such as computed tomography (CT) and magnetic resonance imaging (MRI), alongside innovations in three-dimensional (3D) printing, digital dissection platforms, and virtual or augmented reality (VR/AR), have significantly reshaped anatomy instruction. Internet-based learning and mobile technologies have further transformed self-directed study through accessible, interactive resources. Most recently, AI has introduced capabilities such as personalized tutoring, performance prediction, automated assessment, and intraoperative anatomical guidance. This review highlights the value of a blended approach that integrates cadaveric experience with technological innovations. Future curricula should prioritize optimizing the sequence and combination of these modalities while ensuring equitable access and preparing students for real-world clinical challenges.

Keywords: Anatomy education; cadaveric dissection; clinical anatomy; medical curriculum reform; problem-based learning (PBL); Artificial intelligence; ChatGPT (Siriraj Med J 2025; 77: 901-913)

INTRODUCTION

Anatomical science has shaped medical education for over two millennia. In the 5th century BC, Hippocrates laid the foundation of medical practice through the Hippocratic Oath and formulated principles for disease treatment.[1,2] In the 3rd century BC, Herophilus emphasized the importance of learning through dissections.[3,4] However, for over 1,700 years, the practice remained largely absent until the Renaissance, when Andreas Vesalius revived it with the publication of De Humani Corporis Fabrica. This landmark work, based on cadaveric dissection, marked the recognition of dissection as a fundamental tool in medical education.[5] In contrast, over the past two decades, the necessity of cadaveric dissection and the time allocated to anatomy instruction have been increasingly debated.[6] This narrative review will investigate these changes by addressing the following guiding questions: 1. What are the key changes in cadaver-based learning, applied clinical anatomy, medical education, and technologies that intersect with anatomy education? 2. What theoretical and technological foundations drive these changes? 3. What is currently known about the pedagogical value and implementation challenges of these emerging modalities? This review adheres to narrative review conventions and follows the Scale for the Assessment of Narrative Review Articles (SANRA) criteria to ensure clarity, coherence, and critical engagement with the literature.[7]

Literature search strategy

Relevant literature was identified through systematic searches of PubMed and Scopus, supplemented by targeted searches in Google Scholar and Google. Citation tracking of included studies was also performed to identify additional sources. Searches were conducted from 1 April to 30 November 2023 and updated on 2 July 2025. A flow diagram of the search process is provided in the supplementary material. The search covered publications from 1 January 1990 to 2 July 2025. Keywords included "anatomy", "anatomy education", "medical student", "cadaveric dissection", "prosection", "ultrasound", "computed tomography", "magnetic resonance imaging," "3D-printing," "virtual reality", "augmented reality", "Problem-based learning", "ChatGPT", and "artificial intelligence". As a narrative review, inclusion was guided by thematic relevance to the evolution of anatomy education.[7,8] Eligible publications met at least one of the following criteria: early seminal works or highly cited studies, reports of notable educational outcomes, or credible grey literature offering conceptual or contextual value. Only English-language sources were considered. Studies were excluded if they duplicated content without providing new insights.

Changes in Cadaver-based Learning

Cadaveric Dissection

The shift towards hands-on study emerged during

the Renaissance period. Before this, students primarily observed dissections, as evidenced by a 13th-century painting.[9,10] Hands-on cadaveric dissection became standard practice following William Hunter's adoption of the Paris method in the 18th century, which involved providing each pupil with an entire body for dissection and periodically inspecting their work.[11] This paradigm began to change after the publication of the Flexner Report in 1910, which emphasized the integration of basic sciences into medical education.[12] Several factors have contributed to the gradual shift away from cadaveric dissection in modern pedagogy. These factors include the development of alternative teaching methods and supplementary learning tools, which are discussed later in this review, and the adoption of systems-based curricula with increased emphasis on clinical applications.[13,14] Additional challenges include a shortage of trained anatomy educators[15], the considerable costs associated with cadaver management[15,16], heightened ethical concerns regarding body donation, such as the requirement for informed consent from the donor and the recommended practice of obtaining co-signature from the next of kin[17,18], cadaver shortage in some countries[19,20], and health risks related to formaldehyde exposure.[21,22] A 2019 survey of anatomy education in the United Kingdom and the Republic of Ireland reported an average teaching time of 85 ± 6 hours dedicated for gross anatomy, with only one out of 39 medical schools using dissection as the primary teaching method.[23] In contrast, the use of cadavers in residency training has increased in parallel with the rise of minimally invasive surgery, a trend that began with the first laparoscopic cholecystectomy in the 1980s and continues to shape modern surgical practice.[24-26] Minimally invasive surgery presents a unique anatomical perspective compared to open surgery, employing modern surgical tools such as microscopes and endoscopes. Since the 1990s, cadaver-based procedural training has expanded beyond core techniques to advanced surgeries.[27-29] Training with cadavers has been shown to enhance competency and confidence among future surgeons.[30-32] Globally, institutions have integrated cadaver labs into their residency curricula.[33-35]

Prosection

Prosection, the use of pre-dissected cadaveric specimens, serves as an alternative to student-performed dissection.[36] A renowned prosection collection from the Enlightenment era was developed by William Hunter in the 18th century and has been preserved to the present day.[37,38] By the 19th century, the British General Medical Council had removed the requirement for whole-body dissection, prompting evaluations of prosection's efficacy as an alternative.[39] A search of MEDLINE, Embase, Scopus, and Google Scholar identified only two randomized controlled trials (RCT) comparing the effectiveness of prosection and dissection. The search strategy focused on the core concept of prosection, dissection, RCT, and comparative effectiveness. The first RCT comparing cadaveric dissection and prosection in teaching applied surgical anatomy assessed instruction in truncal anatomy through 6 clinically relevant procedures: inguinal hernia repair, cholecystectomy, laparotomy, right hemicolectomy, and Hartmann's procedure. The dissection group achieved significantly higher post-course test scores (69.8% ± 10.8%) than the prosection group (62.3% ± 10.6%; $p < 0.05$), as well as higher one-year retention scores (59.8% ± 10.4% vs. 49.8% ± 10.1%; $p < 0.05$). However, after adjusting for the time spent in sessions, these differences were no longer statistically significant.[40] Another study reported higher practical and multiple-choice question (MCQ) scores in the dissection group (169 ± 1.99) compared to the prosection group (142 ± 1.78), though it lacked a pre-intervention assessment and used only 2 cadavers for 50 students.[41] A meta-analysis comparing anatomy laboratory pedagogies found no significant difference in outcomes between traditional cadaveric dissection and prosection across 4 studies, with a pooled standardized mean difference (SMD) of –0.33 (95% confidence interval [CI], –0.76 to 0.11, $p = 0.15$). However, the included studies varied widely in focus, ranging from human limb anatomy to canine thoracic limb dissection, and included comparisons involving computer-assisted instruction versus traditional methods.[42] These findings suggest that it remains inconclusive whether prosection is more effective than dissection for learning human anatomy, underscoring the need for more rigorous research.

Changes in Applied Clinical Anatomy: Surface Anatomy, Physical Examination, Ultrasound, and Imaging Modalities

Surface anatomy

The observation of living humans is one of the simplest methods for learning clinical anatomy. Leonardo da Vinci documented surface anatomy through detailed sketches during the 15th century.[43,44] Evidence of anatomy instruction using live subjects is also depicted in 18th-century paintings.[45] These methods have been reintroduced in the 21st century through methods such as body painting and peer physical examination in clinically integrated anatomy courses.[46,47]

Ultrasound

Ultrasound enhances anatomy education by illustrating the relationship between surface anatomy and underlying structures and by demonstrating real-time physiology and organ function. In 1990, Hannover Medical School became the first institution to implement ultrasound into anatomy education, with 48.7% of first-year medical students reporting improved topographical knowledge of abdominopelvic organs.[48] Subsequent studies have consistently reported positive student feedback on ultrasound-based learning.[49,50] A randomized study comparing prosection and ultrasound found both methods to be equally effective in improving cardiac anatomy test scores.[51] At Mayo Medical School, an outcome-based ultrasound course trained students in echocardiography using handheld devices. Over the course of 3 weeks, students' accuracy in labeling ultrasound anatomy improved from 3.7% to 91.0%, while satisfactory image quality increased from 57.1% to 78.6%, p=0.04.[52] A 2023 non-randomized study reported that students who participated in additional point-of-care ultrasound (POCUS) sessions scored significantly higher on cardiovascular anatomy exams compared to those in the standard curriculum (92.1 vs. 88.5, $p < 0.05$). However, no significant differences were observed for gastrointestinal anatomy exams or in USMLE Step 1 and clinical examination scores.[53] Recent publications have explored the use of ultrasound in teaching head and neck, musculoskeletal, and limb anatomy.[54] Despite its expanding applications, the role of ultrasound as a supplement to or substitute for cadaver-based learning remains unclear in much of the existing literature. One RCT addressed this question by comparing cardiac anatomy instruction using ultrasound versus cadaveric prosection in first-year medical students who had not received prior training in cardiac anatomy. Post-test scores showed no significant difference between the groups (cadaveric 85% ± 15.7, ultrasound 85.1% ± 13.5, $p = 0.95$).[51] Despite its demonstrated benefits, gaps remain in understanding the optimal integration of ultrasound into anatomy curricula and its long-term impact on knowledge retention and clinical performance.

Imaging modalities in anatomy teaching

Wilhelm Röntgen discovered X-rays in 1895 and published the first radiograph, an image of his wife's hand, in 1896.[55,56] During World War I, Marie Curie developed the Radiological Ambulance to offer surgical guidance on the battlefield, laying the groundwork for portable X-ray machines.[57] In 1971, inspired by the idea of visualizing the contents of a closed box, Sir Godfrey Hounsfield invented the computed tomography (CT) scanner, which enabled detailed cross-sectional imaging of internal body structures.[58] Around the same time, magnetic resonance imaging (MRI) was being developed.[59]

The use of imaging to enhance anatomical understanding was first introduced at McMaster University in 1969 as part of its Problem-based learning (PBL) curriculum.[60] A 1990 study evaluating the integration of diagnostic imaging into dissection laboratories reported an improvement in student performance, with correct responses increasing from 16.7% (pre-test) to 88.1% (post-test), and a retention rate of 73.5% retention after 14–17 weeks.[61,62] Subsequent studies investigating the incorporation of imaging modalities into anatomy education support these findings. In a five-year study, Chew et al. reported higher anatomy MCQ scores during the two academic years in which radiology was integrated into small-group sessions. These sessions included brief radiologic anatomy lectures and rotating quiz stations focused on clinically relevant pathology aligned with weekly anatomical content. Student performance improved considerably, with mean anatomy scores rising to 68.97% in 2017 and 73.77% in 2018, compared to 65.85% in 2015 and 62.23% in 2019, when no radiology instruction was provided.[63] Similarly, Larsen and Engle examined the impact of integrating radiologic content into anatomy education through lectures, laboratory sessions, and assessments. Students were evaluated using a 12-question anatomy test incorporating imaging modalities such as X-ray, CT, and MRI. Across the 2016 and 2017 cohorts, mean pre-test scores increased from 12.05% to 41.3% following the intervention.[64] Integrating imaging-based anatomy early in medical education, particularly alongside cadaveric dissection, has shown promise in enhancing the clinical applicability of anatomical knowledge. However, further research is needed to evaluate student-perceived confidence in diagnosis and the long-term educational outcomes associated with this approach

Changes in Medical Education

In 1870, Christopher Columbus Langdell, Dean of Harvard Law School, pioneered a new learning approach by encouraging students to analyze and discuss cases to understand underlying principles.[65,66] This initiative laid the foundation for the Harvard Business School's Case Method, introduced in 1921, which emphasized real-life scenario analysis to develop problem-solving skills beyond textbook knowledge and conventional ideas.[65,67,68] A similar shift occurred in medical education when, in 1900, a Harvard Medical School student criticized reliance on didactic lectures and advocated for a Case System involving

real patient data and analytical discussions.[69] By 1906, the Case System had been implemented in the United States medical education.[70] A major transformation came in 1969 when Howard Barrows at McMaster University introduced PBL, which encouraged students to analyze clinical scenarios and develop solutions independently, with instructors serving as facilitators to support self-directed learning.[60,71] A 1990 publication described an educational reform at Harvard Medical School that replaced formal lectures with problem-oriented, case-based tutorials and small group activities, aiming to integrate basic and clinical sciences by eliminating their temporal separation.[72]

Despite its widespread adoption, studies evaluating its effectiveness have reported mixed outcomes. A 2005 meta-analysis reported no significant benefit for acquiring basic factual knowledge, with a weighted effect size (ES) of –0.04. In contrast, PBL showed a strong positive effect on understanding underlying principles (ES= +0.75), and a moderate, though not statistically significant, benefit in applying knowledge to real-world problems (ES = +0.34).[73] A 2010 systematic review of 30 studies, conducted primarily in North America and Western Europe over 22 years, found that 73.3% of included studies reported no significant difference in examination performance between PBL and traditional teaching methods.[74] In contrast, a 2022 scoping review that included studies from a broader international context reported that 69% of publications found PBL to be superior to Lecture-based learning in terms of learning outcomes and knowledge acquisition.[75]

In the context of anatomy education, findings have been similarly mixed. One study compared anatomy test scores between two institutions with different teaching approaches. Institution A employed a traditional model, including lectures, guided dissections, and weekly small-group discussions. Institution B implemented a PBL approach with unguided dissections and a single weekly integrated preclinical-clinical session without lectures. Students in the traditional curriculum scored significantly higher (mean 37.5, range 25–46) than those in the PBL curriculum (mean 32.35, range 18–45, $p < 0.001$).[76] A 2014 systematic review focusing on anatomy education included studies from Western countries and the Middle East and evaluated active learning strategies such as PBL and Team-based learning (TBL). Among these, 40% of the studies reported improved examination performance, while 50% found no significant difference when compared with traditional teaching.[77] Taken together, these findings indicate that conclusions about the effectiveness of PBL remain inconclusive, which may be due in part to differences in the geographical representation of the studies included. In the context of anatomy education, there remains a need for further analysis focusing specifically on the isolated impact of PBL. At present, it is difficult to conclude that PBL is superior to conventional methods for teaching anatomy. However, existing evidence indicates that active learning approaches are more effective than traditional Lecture-based instruction.[77]

Beyond method-level evaluations, broader curricular trends raise additional concerns for the future of anatomy education. An analysis of Association for Medical Education in Europe (AMEE) guides published between 1992 and 2025 revealed that only one guide directly addresses anatomy education.[78] Additionally, AMEE's movement toward Lecture-free curricula and the growing shift toward online learning may further deprioritize hands-on anatomical training,[79,80] which itself constitutes a form of active learning. These changes raise critical questions. Is the current trajectory of anatomy education optimal? Can PBL or online methods fully replace hands-on skills gained through dissection? How should effectiveness be measured, and what defines sufficient anatomical knowledge for medical students? Most importantly, how can educators ensure that students are well-prepared to meet real-world clinical demands?

Changes in technology: digital anatomy, 3D-printing, the internet era to mobile technology, and artificial intelligence

In the 1930s, Alan Turing laid the foundation for modern computing by developing the theoretical principles of digital computation.[81,82] He later proposed the concept of AI, envisioning machines capable of autonomous thought.[83,84] By the 1970s, technological advances had produced smaller and more affordable computers, making them accessible to household user[85,86] and driving rapid changes across multiple fields.

Digital anatomy

A pivotal development in digital anatomy was the launch of the Visible Human Project (VHP) in 1988. The initiative aimed to address the limitations of traditional dissection, which restricts learners to a single perspective and does not allow for review of the dissection process.[87] Utilizing CT, MRI, and 3D rendering technologies, the VHP produced a detailed digital cadaver dataset that can be viewed from multiple angles, virtually dissected, and used to simulate clinical procedures such as colonoscopy.[88,89] Modern 3D models now incorporate virtual reality (VR), augmented reality (AR), and mixed reality (MR), further enhancing immersive learning. A 2025 systematic review

reported that VR improved anatomical knowledge in 50% of studies assessing learning outcomes and outperformed traditional methods, including lectures, prosections, two-dimensional models (2D), 3D-printing, cadaver dissection, and AR, in 75% of comparative studies.[90] A 2024 meta-analysis found that, compared with traditional methods overall, VR and AR showed a significant benefit (SMD = 0.40, $p < 0.001$), with the greatest effect when used as complementary tools (SMD = 0.52, $p < 0.001$). Cybersickness, resulting from sensory conflict between visual and vestibular systems, was reported in 25% of studies but was generally mild and transient.[91] However, a 2021 meta-analysis reported no significant difference in anatomical test scores between AR and control groups. Student motivation and engagement outcomes also remained inconclusive due to variability in the assessment tool.[92] Based on current evidence, VR and AR appear most effective when used to complement rather than replace traditional teaching methods. Importantly, none of the reviewed studies were conducted in low-income countries, suggesting limited access to these technologies in resource-constrained educational settings.[93]

3D-printed

The origins of 3D-printing date back to 1980, when Hideo Kodama proposed using ultraviolet light to form 3D objects from photosensitive resin, laying the foundation for stereolithography (SLA). In 1984, Charles Hull developed the first functional SLA printer.[94] Shortly thereafter, Scott Crump invented fused deposition modeling,[95] a technique in which heated filament is extruded layer by layer. Since the 1990s, 3D-printed models have been used in surgical planning.[96] In anatomical education, a 2016 RCT comparing painted cadaveric skulls, 3D-printed models, and 2D atlases found that students using 3D-printed models achieved the highest median post-test scores (31.5 vs. 29.5 and 27.75, $p = 0.04$).[97] Another RCT at Monash University reported significantly higher scores in the 3D-printed cardiac model group compared to those using prosected specimens or combined materials (60.83% vs. 44.81% and 44.62%, adjusted $p = 0.01$).[98] In functional simulation, a study in Singapore found that students using a 3D-printed knee joint simulator achieved higher spatial understanding scores compared to those using text and skeletal models (85.03% vs. 70.71%, $p < 0.05$).[99] A 2020 systematic review found that 3D-printed models improved learning outcomes in neuroanatomy, cardiac, and abdominal domains, with SMD of 0.69 compared to cadavers and 1.05 compared to 2D resources ($p < 0.05$).[100] With ongoing advances and falling production costs, 3D-printing continues to show promise in anatomy education. Reported learning gains fall within the medium to large effect size range, supporting its educational value. These developments also parallel broader clinical applications, including surgical simulation models, patient-specific prostheses, and emerging work in bioprinting.[101-103]

Internet and mobile technology

The internet's foundation was the Advanced Research Projects Agency Network (ARPANET), developed in the 1960s to enable connectivity among universities.[104] Following expanded public access in the 1980s, internet service providers began offering individual users connectivity via modems in the 1990s. In 1997, the University of Washington developed a web-based atlas featuring 2D images, 3D animations, and interactive quizzes, which became a globally recognized resource.[105]

The 2000s marked the advancement of smartphones, which integrated mobile phone functionality with computing capabilities such as user-friendly web browsing, email, and multimedia applications.[106] The launch of the iPhone, the introduction of Android smartphones, and the debut of mobile application platforms, including the App Store and Android Market in the late 2000s, along with the release of the iPad in 2010, accelerated the integration of mobile computing into daily life and educational practice.[106-108] Evidence from the early 2010s indicates that pre-clinical students most often preferred self-directed learning tools, with study notes as the most frequently used resource, followed by textbooks, atlases, websites, videos, and educational software.[109] A decade later, a shift in preference was reported toward anatomy websites, followed by interactive sessions, lecture notes, textbooks, and computer-assisted tools.[110] Social media has also contributed to interactive learning environments, with platforms such as Facebook, YouTube, Twitter, and Instagram used to share anatomy-related content and encourage student engagement through content creation.[111-114]

The COVID-19 pandemic highlighted the critical role of online learning, as lockdowns necessitated a rapid shift from traditional classrooms to digital platforms globally.[115-117] By 2024, improvements in infrastructure and affordability had enabled 66% of the global population to access the internet.[118] In the post-pandemic era, a 2024 study found that 51% of students identified the Visible Body application, an interactive 3D anatomy and physiology platform available on the web and tablet, as the most beneficial self-study tool. Other frequently used resources included Kenhub, Noted Anatomist, and AnatomyZone.[119]

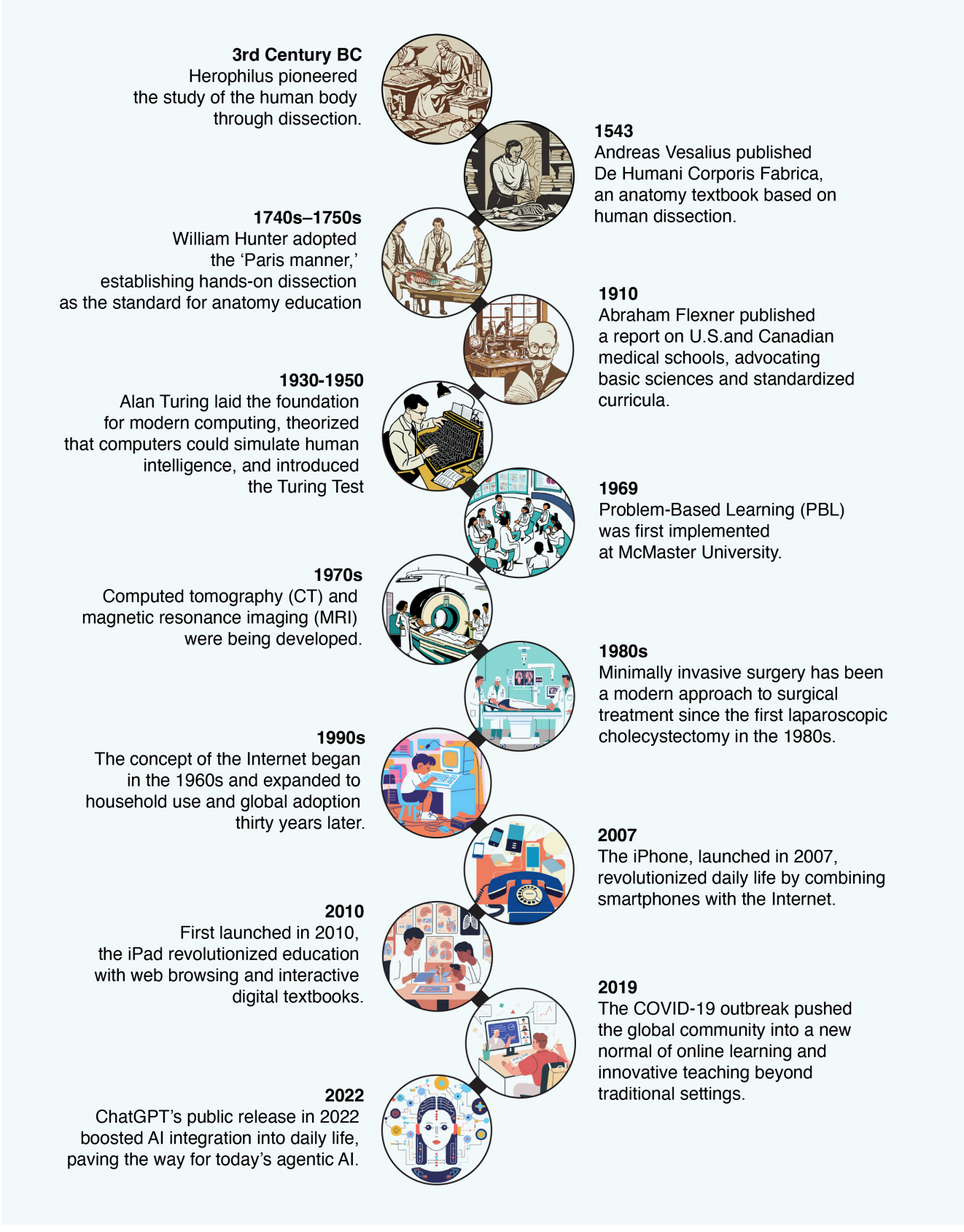


Fig 1. Timeline of the Evolution of Anatomy Education: From Cadaveric Dissection to Artificial Intelligence. The illustrations featured in this timeline were created using Canva's AI-powered image generator, Magic Media. (https://www.canva.com/apps/generate_image/magic-media).

Artificial intelligence

In the 1940s, Turing proposed the idea that machines could learn like children and gradually develop independent thinking[83,84], which was later termed AI.[120] To assess whether a machine possessed such intelligence, Turing introduced the "imitation game", suggesting that a machine could be considered intelligent if its responses were indistinguishable from those of a human.[83]

Since the 1990s, AI has shown transformative potential in medicine, initially through tasks such as interpreting electrocardiograms and detecting breast cancer via imaging.[121] By the 2010s, AI matched or exceeded physician-level performance in diagnosing conditions including skin cancer, diabetic retinopathy, colonic polyps, and various radiologic pathologies.[122] It has also been assessed for surgical decision-making and intraoperative anatomical guidance.[123,124] By the 2020s, AI had been integrated across the continuum of care, from triage to outpatient follow-up.[125-127]

AI applications in anatomy and medical education have grown rapidly, with momentum accelerating after ChatGPT's public release in November 2022, driving widespread engagement with large language model (LLMs) technologies.[128,129] In 2021, Li et al. developed the Artificial Intelligence Support System (AISS), a chatbot trained on syllabi from the UK Anatomical Society and the Chinese University of Hong Kong. Designed to enhance anatomy learning and student–teacher communication, AISS could answer questions, generate quizzes, and provide instant feedback. Although accuracy metrics were not reported, students who used the system during a one-hour session on upper limb anatomy reported increased confidence, with self-assessment scores rising from 2.10 to 3.84 on a 5-point Likert scale.[130] In 2023, Fries et al. demonstrated AI's ability to monitor learning and predict academic performance by analyzing student interactions with a virtual microscope. Their model accurately predicted practical exam outcomes in first-year medical and dental students by identifying key learning behaviors, including time spent on specific regions, study sequence, and session duration.[131] At McMaster University, researchers explored AI's potential to automate grading in anatomy practical examinations, achieving an average accuracy of 94.49% across 54 questions.[132] Another study applied AI to analyze students' reflective writing about themselves and body donors, using sentiment and emotional tone analysis to better understand student attitudes toward human anatomy.[133] These findings highlight the growing role of AI in medical education as a study companion, a tool for performance monitoring, a mechanism for early identification of students needing support, and a means of personalized learning. AI also offers the potential for reducing educator workload and generating new insights from large-scale data.

Recent advancements in 2025 LLMs, such as Google's Gemini, ChatGPT, and Perplexity, allow them to generate outputs ranging from simple explanations to full research papers.[134-136] However, several concerns about using AI have been reported since its wide adoption. A study conducted between 2022 and 2023 evaluating medical references generated by various LLMs identified fabricated or erroneous citations across multiple providers.[137] A 2023 case report showed that LLMs can fabricate healthcare-related content, highlighting hallucination risks with potential implications for scholarly integrity and clinical decision-making.[138] Accordingly, AI ethical guidelines emphasize the importance of maintaining a "human-in-the-loop" approach, whereby the user is responsible for monitoring and evaluating the results generated by AI.[139,140] Bias in AI-generated outputs has also been pointed out, reflecting the nature of its training data and algorithms.[139,140] Moreover, the World Health Organization (WHO) has raised concerns about AI's unpredictable behavior and the unclear processes behind its outputs, highlighting the importance of robust data protection and security.[139,141] Despite these concerns, AI has also been recognized as a valuable tool for idea generation, content refinement, literature synthesis, data analysis, and manuscript preparation.[142] Transparent disclosure of AI use in academic work is recommended to promote accountability.[142] In conclusion, the rapid adoption of AI in academia is unavoidable, making it essential to be aware of both its transformative potential and inherent risks.

CONCLUSION

Cadaveric dissection has been the cornerstone of anatomy teaching for nearly seven centuries, yet the three decades since the early 1990s have brought unprecedented and rapid transformation. Advances in medical imaging technologies, the development of cadaveric 3D models and AR/VR/MR, and affordable 3D printing have coincided with the rise of minimally invasive surgery, PBL curricula, financial constraints in maintaining cadaver-based systems, and evolving ethical considerations around body donation as well as safety standards in anatomy laboratories. Together, these factors have shifted many programs away from a dissection-only paradigm. Current evidence suggests that cadavers remain the gold standard for surgical simulation, spatial understanding, whereas prosection allows more time-efficient coverage but has not shown

equivalence across all body regions. Early integration of diagnostic imaging with gross anatomy enhances short-term clinical reasoning, though its long-term impact on practice remains unclear. Similarly, PBL and TBL foster self-directed inquiry, but whether these approaches can fully replicate the psychomotor and affective gains of dissection remains unresolved, particularly in an era where AI delivers information rapidly yet cannot replace hands-on skills. Immersive VR and AR, together with increasingly affordable 3D prints, show promise as complementary tools within blended learning environments. AI tutors offer potential for personalized and scalable feedback, supporting instructors by automating repetitive, non-instructive tasks, though risks of hallucinations, bias, and data governance challenges require careful oversight. To conclude, 5 key gaps warrant attention in the coming years. 1. Determining which anatomical regions and competencies require cadaveric dissection versus those that can be taught effectively through alternative pedagogies. 2. Identifying clinically applied imaging modalities that, when integrated early, reinforce anatomical knowledge, support long-term retention, and enhance diagnostic reasoning and procedural skills. 3. Developing cost-effective, globally accessible solutions to democratize advanced anatomy education. 4. Exploring, monitoring, and keeping pace with rapid AI advancements 5. Defining learning outcomes aligned with healthcare system needs for future undergraduate medical graduates and identifying pedagogical strategies to achieve them in an era where AI rapidly generates and delivers knowledge. Addressing these priorities will help shape anatomy curricula that preserve the enduring value of cadaveric heritage while embracing 21st-century innovation, ensuring graduates are clinically proficient, ethically grounded, and globally prepared.

Data Availability Statement

This study is a narrative review and did not generate any primary datasets. All sources cited in this review are publicly available in the published literature and can be accessed through the referenced journals and databases.

Limitations

While this review selectively incorporates key RCTs, systematic reviews, and grey literature, it is structured as a narrative synthesis to provide a broad, historically informed perspective that complements the focused scope of systematic reviews and meta-analyses. Future studies should assess the comparative educational efficacy of traditional and emerging modalities to build on this historical synthesis.

ACKNOWLEDGEMENT

The author is grateful to Assistant Professor Dr. Benjaporn Pamornpol for her thoughtful editorial guidance on this manuscript.

DECLARATIONS

Grants and Funding Information

No funding was received to conduct this review.

Conflict of Interest

The author declares no relevant financial or non-financial conflicts of interest.

Author Contributions

The sole author was responsible for all aspects of the work, including conceptualization, data curation, formal analysis, reviewing, project administration, resources, software, supervision, validation, visualization, writing the original draft, review, and editing.

Use of Artificial Intelligence

The author used Canva's AI-powered image generator, Magic Media (https://www.canva.com/apps/generate_image/magic-media), to create the illustration in Fig 1. Timeline of the Evolution of Anatomy Education: From Cadaveric Dissection to Artificial Intelligence.

REFERENCES

1. Garrison FH. An introduction to the history of medicine : with medical chronology, suggestions for study and bibliographic data. 4th ed. Philadelphia: W.B. Saunders; 1929. 996 p.
2. Elizondo-Omana RE, Guzman-Lopez S, Garcia-Rodriguez Mde L. Dissection as a teaching tool: past, present, and future. Anat Rec B New Anat. 2005;285(1):11-5.
3. Bay NS, Bay BH. Greek anatomist herophilus: the father of anatomy. Anat Cell Biol. 2010;43(4):280-3.
4. Standring S. A brief history of topographical anatomy. J Anat. 2016;229(1):32-62.
5. Toledo-Pereyra LH. De Humani Corporis Fabrica surgical revolution. J Invest Surg. 2008;21(5):232-6.
6. Ghosh SK. Cadaveric dissection as an educational tool for anatomical sciences in the 21st century. Anat Sci Educ. 2017;10(3):286-99.
7. Baethge C, Goldbeck-Wood S, Mertens S. SANRA-a scale for the quality assessment of narrative review articles. Res Integr Peer Rev. 2019;4:5.
8. Gregory AT, Denniss AR. An Introduction to Writing Narrative and Systematic Reviews - Tasks, Tips and Traps for Aspiring Authors. Heart Lung Circ. 2018;27(7):893-8.
9. Toledo-Pereyra LH. Medical Renaissance. J Invest Surg. 2015;28(3):127-30.
10. Bonnel F, Lavabre-Bertrand T, Bonnel C. The teaching of anatomy in Montpellier University during VIII centuries (1220-2020). Surg Radiol Anat. 2019;41(10):1119-28.
11. Gelfand T. The "Paris manner" of dissection: student anatomical

dissection in early eighteenth-century Paris. Bull Hist Med. 1972;46(2):99-130.
12. Flexner A. Medical education in the United States and Canada. From the Carnegie Foundation for the Advancement of Teaching, Bulletin Number Four, 1910. Bull World Health Organ. 2002;80(7):594-602.
13. Ratanayotha A, Oo EM. Chronicle of Anatomical Education in Thailand: Experiences at Siriraj Medical School. Siriraj Med J. 2022;74(7):463-71.
14. Sinha SR, Entezampour M, Davis K, Vyas RM. Clinically Applied Anatomy: An Impactful Initiative in the Evolution of Medical Education. Ann Plast Surg. 2020;84(5):476-80.
15. Kim IB, Joo KM, Song CH, Rhyu IJ. A Brief Review of Anatomy Education in Korea, Encompassing Its Past, Present, and Future Direction. J Korean Med Sci. 2024;39(20):e159.
16. Gabriel AG, Chia SL. High Price of Perfection that Is Anatomy: Why Studying and Teaching the Human Body Is a Financial Muscle. Journal of Innovations in Medical Research. 2025;4(2):84-8.
17. Jayakumar N, Athar S, Ashwood N. Where do these cadavers come from? Clin Anat. 2020;33(6):872-5.
18. Anatomists IFoAo. Recommendations of good practice for the donation and study of human bodies and tissues for anatomical examination 2012 [cited 2025 July 1]. Available from: https://ifaa.net/wp-content/uploads/2017/09/IFAA-guidelines-220811.pdf.
19. Chen D, Zhang Q, Deng J, Cai Y, Huang J, Li F, et al. A shortage of cadavers: The predicament of regional anatomy education in mainland China. Anat Sci Educ. 2018;11(4):397-402.
20. Zdilla MJ, Balta JY. Human body donation and surgical training: a narrative review with global perspectives. Anat Sci Int. 2023;98(1):1-11.
21. Durongphan A, Amornmettajit N, Rungruang J, Nitimanee E, Panichareon B. One academic year laboratory and student breathing zone formaldehyde level, measured by gas-piston hand pump at gross anatomy laboratory, Siriraj Hospital, Thailand. Environ Sci Pollut Res Int. 2020;27(14):16521-7.
22. Durongphan A, Chongkolwatana W, Ngamskulrungroj P, Pochnasomboon T, Pinkaew J, Pamornpol B, et al. A Pilot Comparative Study of Submerge vs. Non-Submerg Saturated Salt Solution Human Cadavers Embalming Method by Gross, Histological, and Microbiological Evaluation. Siriraj Med J. 2022;74(7):431-9.
23. Smith CF, Freeman SK, Heylings D, Finn GM, Davies DC. Anatomy education for medical students in the United Kingdom and Republic of Ireland in 2019: A 20-year follow-up. Anat Sci Educ. 2022;15(6):993-1006.
24. Peter SDS, Holcomb III GW. History of minimally invasive surgery: Philadelphia: Elsevier Health Sciences; 2008.
25. Jani K, Rajan PS, Sendhilkumar K, Palanivelu C. Twenty years after Erich Muhe: Persisting controversies with the gold standard of laparoscopic cholecystectomy. J Minim Access Surg. 2006;2(2):49-58.
26. Memon I. Cadaver Dissection Is Obsolete in Medical Training! A Misinterpreted Notion. Med Princ Pract. 2018;27(3):201-10.
27. Anastakis DJ, Regehr G, Reznick RK, Cusimano M, Murnaghan J, Brown M, et al. Assessment of technical skills transfer from he bench training model to the human model. Am J Surg. 1999;177(2):167-70.
28. Lewis CE, Peacock WJ, Tillou A, Hines OJ, Hiatt JR. A novel cadaver-based educational program in general surgery training. J Surg Educ. 2012;69(6):693-8.
29. Gasco J, Holbrook TJ, Patel A, Smith A, Paulson D, Muns A, et al. Neurosurgery simulation in residency training: feasibility, cost, and educational benefit. Neurosurgery. 2013;73 Suppl 1:39-45.
30. Kim SC, Fisher JG, Delman KA, Hinman JM, Srinivasan JK. Cadaver-Based Simulation Increases Resident Confidence, Initial Exposure to Fundamental Techniques, and May Augment Operative Autonomy. J Surg Educ. 2016;73(6):e33-e41.
31. Camp CL, Krych AJ, Stuart MJ, Regnier TD, Mills KM, Turner NS. Improving Resident Performance in Knee Arthroscopy: A Prospective Value Assessment of Simulators and Cadaveric Skills Laboratories. J Bone Joint Surg Am. 2016;98(3):220-5.
32. Farah GJ, Rogers JL, Lopez AM, Brown NJ, Pennington Z, Kuo C, et al. Resident Training in Spine Surgery: A Systematic Review of Simulation-Based Educational Models. World Neurosurg. 2023;174:81-115.
33. London IC. London School of Surgery Training Programme 2023 [cited 2025 Jun. 23]. Available from: https://www.imperial.ac.uk/medicine/research-and-impact/facilities/london-school-of-surgery-training-programme/training-and-assessment/ear-nose--throat-surgical-training/.
34. University HMS. Didactic Curriculum in Otolaryngology–Head and Neck Surgery 2023 [cited 2025 Jun. 23]. Available from: https://oto.hms.harvard.edu/didactic-curriculum-otolaryngology.
35. University GSM. Curriculum 2023 [cited 2025 Jun. 23]. Available from: https://graduate.mahidol.ac.th/inter/prospective-students/?p=curriculum&id=2035N03G.
36. Estai M, Bunt S. Best teaching practices in anatomy education: A critical review. Ann Anat. 2016;208:151-7.
37. Mitchell PD, Boston C, Chamberlain AT, Chaplin S, Chauhan V, Evans J, et al. The study of anatomy in England from 1700 to the early 20th century. J Anat. 2011;219(2):91-9.
38. Roth H, Smith RA, Mackay S. Modern day relevance of William Hunter's approach to teaching "The organ of hearing". Clin Anat. 2013;26(5):551-5.
39. Sinclair DC. An Experiment in the Teaching of Anatomy. 1965 May. Report No.: 0022-2577 (Print) 0022-2577 (Linking).
40. Koh ZJ, Yeow M, Srinivasan DK, Ng YK, Ponnamperuma GG, Chong CS. A randomized trial comparing cadaveric dissection and examination of prosections as applied surgical anatomy teaching pedagogies. Anat Sci Educ. 2023;16(1):57-70.
41. Abdellatif H. Time Spent in Practicing Dissection Correlated with Improvement in Anatomical Knowledge of Students: Experimental Study in an Integrated Learning Program. Cureus. 2020;12(4):e7558.
42. Wilson AB, Miller CH, Klein BA, Taylor MA, Goodwin M, Boyle EK, et al. A meta-analysis of anatomy laboratory pedagogies. Clin Anat. 2018;31(1):122-33.
43. Clayton M. Leonardo da Vinci as an anatomist 2001. Available from: https://media.rct.uk/sites/default/files/file-downloads/Leonardo%20da%20Vinci%20Anatomist.pdf?_ga=2.35543076.461267984.1751959990-560325555.1751959990.
44. Martin C, Philo R. Leonardo da Vinci: the mechanics of man. London Los Angeles: Royal Collection Enterprises; J. Paul Getty Museum; 2010. 160 p.
45. McLachlan JC, Patten D. Anatomy teaching: ghosts of the past, present and future. Med Educ. 2006;40(3):243-53.
46. Op Den Akker JW, Bohnen A, Oudegeest WJ, Hillen B. Giving color to a new curriculum: bodypaint as a tool in medical

education. Clin Anat. 2002;15(5):356-62.
47. Bergman EM, Sieben JM, Smailbegovic I, de Bruin AB, Scherpbier AJ, van der Vleuten CP. Constructive, collaborative, contextual, and self-directed learning in surface anatomy education. Anat Sci Educ. 2013;6(2):114-24.
48. Teichgraber UK, Meyer JM, Poulsen Nautrup C, von Rautenfeld DB. Ultrasound anatomy: a practical teaching system in human gross anatomy. Med Educ. 1996;30(4):296-8.
49. Swamy M, Searle RF. Anatomy teaching with portable ultrasound to medical students. BMC Med Educ. 2012;12:99.
50. So S, Patel RM, Orebaugh SL. Ultrasound imaging in medical student education: Impact on learning anatomy and physical diagnosis. Anat Sci Educ. 2017;10(2):176-89.
51. Griksaitis MJ, Sawdon MA, Finn GM. Ultrasound and cadaveric prosections as methods for teaching cardiac anatomy: a comparative study. Anat Sci Educ. 2012;5(1):20-6.
52. Wittich CM, Montgomery SC, Neben MA, Palmer BA, Callahan MJ, Seward JB, et al. Teaching cardiovascular anatomy to medical students by using a handheld ultrasound device. JAMA. 2002;288(9):1062-3.
53. Haidar DA, Kessler R, Khanna NK, Cover MT, Burkhardt JC, Theyyunni N, et al. Association of a longitudinal, preclinical ultrasound curriculum with medical student performance. BMC Med Educ. 2022;22(1):50.
54. Johnson CD, Davison L, Graham EC, Sweeney EM. Ultrasound technology as a tool to teach basic concepts of physiology and anatomy in undergraduate and graduate courses: a systematic review. Adv Physiol Educ. 2025;49(1):11-26.
55. Bradley WG. History of medical imaging. Proc Am Philos Soc. 2008;152(3):349-61.
56. Rontgen WC. On a New Kind of Rays. Science. 1896;3(59):227-31.
57. Mould RF. The discovery of radium in 1898 by Maria Sklodowska-Curie (1867-1934) and Pierre Curie (1859-1906) with commentary on their life and times. Br J Radiol. 1998;71(852):1229-54.
58. Schulz RA, Stein JA, Pelc NJ. How CT happened: the early development of medical computed tomography. J Med Imaging (Bellingham). 2021;8(5):052110.
59. Geva T. Magnetic resonance imaging: historical perspective. J Cardiovasc Magn Reson. 2006;8(4):573-80.
60. Hamilton JD. The McMaster curriculum: a critique. Br Med J. 1976;1(6019):1191-6.
61. Erkonen WE, Albanese MA, Smith WL, Pantazis NJ. Gross anatomy instruction with diagnostic images. Invest Radiol. 1990;25(3):292-4.
62. Erkonen WE, Albanese MA, Smith WL, Pantazis NJ. Effectiveness of teaching radiologic image interpretation in gross anatomy. A long-term follow-up. Invest Radiol. 1992;27(3):264-6.
63. Chew C, O'Dwyer P, Young D, Gracie J. Radiology teaching improves anatomy scores for medical students. BR J Radiol. 2020;93(1114):20200463.
64. Larsen RJ, Engle DL. The impact of a radiological anatomy-based intervention in a gross anatomy course for undergraduate medical students. Education in the Health Professions. 2020;3(2):47-53.
65. Merseth KK. The Early History of Case-Based Instruction: Insights for Teacher Education Today. Journal of Teacher Education. 1991;42(4):4.
66. School HL. The Case Study Teaching Method [cited 2025 Jun. 23]. Available from: https://casestudies.law.harvard.edu/the-case-study-teaching-method/#:~:text=The%20case%20method%20in%20legal,that%20furthered%20principles%20or%20doctrines.
67. Normand Y. THE HISTORY OF THE CASE STUDY AT HARVARD BUSINESS SCHOOL: Harvard Business School Online; [cited 2025 Jun. 23]. Available from: https://online.hbs.edu/blog/post/the-history-of-the-case-study-at-harvard-business-school.
68. School HB. CASE METHOD 100 YEARS: Harvard Business School; [cited 2025 Jun. 23]. Available from: https://www.hbs.edu/case-method-100.
69. W.B. Cannon. The Case Method of Teaching Systematic Medicine. Boston Medical and Surgical Journal. 1900(142):4.
70. The Case System in Medical Teaching Boston Medical and Surgical Journal. 1906;155:2.
71. Neufeld VR, Barrows HS. The "McMaster Philosophy": an approach to medical education. J Med Educ. 1974;49(11):1040-50.
72. Tosteson DC. New pathways in general medical education. N Engl J Med. 1990;322(4):234-8.
73. Gijbels D, Dochy F, Van den Bossche P, Segers M. Effects of problem-based learning: A meta-analysis from the angle of assessment. Review of Educational Research. 2005;75(1):27-61.
74. Hartling L, Spooner C, Tjosvold L, Oswald A. Problem-based learning in pre-clinical medical education: 22 years of outcome research. Med Teach. 2010;32(1):28-35.
75. Trullas JC, Blay C, Sarri E, Pujol R. Effectiveness of problem-based learning methodology in undergraduate medical education: a scoping review. BMC Med Educ. 2022;22(1):104.
76. Hinduja K, Samuel R, Mitchell S. Problem-based learning: is anatomy a casualty? Surgeon. 2005;3(2):84-7.
77. Williams JM. Is Student Knowledge of Anatomy Affected by a Problem-Based Learning Approach? A Review. Journal of Education and Training Studies. 2014;2(4):108-12.
78. Louw G, Eizenberg N, Carmichael SW. The place of anatomy in medical education: AMEE Guide no 41. Med Teach. 2009;31(5):373-86.
79. Parmelee D, Roman B, Overman I, Alizadeh M. The lecture-free curriculum: Setting the stage for life-long learning: AMEE Guide No. 135. Med Teach. 2020;42(9):962-9.
80. MacNeill H, Masters K, Nemethy K, Correia R. Online learning in Health Professions Education. Part 1: Teaching and learning in online environments: AMEE Guide No. 161. Med Teach. 2024;46(1):4-17.
81. JP B. Alan Turing: Founder of Computer Science. Engineering Trustworthy Software Systems. 2017.
82. Copeland BJ. The modern history of computing: Stanford University; 2006 [Available from: https://plato.stanford.edu/entries/computing-history/#ACE
83. Turing AM. Computing machinery and intelligence. Mind. 1950; LIX(236):433–60.
84. Carpenter BE, Doran R. AM Turing's ACE report of 1946 and other papers. Massachusetts Institute of Technology; 1986. Report No.: 0262031140.
85. Aspray W. The Intel 4004 microprocessor: What constituted invention? IEEE Annals of the History of Computing. 1997;19(3):4-15.
86. Batson E. Apple IIe and IBM PC. Comparison of America's most popular personal computers. Postgrad Med. 1984;75(5):279-82.

87. Ackerman MJ. The Visible Human Project. Inf Serv Use. 2022;42(1):129-36.
88. Ackerman MJ. The Visible Human Project: a resource for education. Acad Med. 1999;74(6):667-70.
89. Rowe PM. Visible Human Project pays back investment. Lancet. 1999;353(9146):46.
90. Adnan S, Benson AC, Xiao J. How virtual reality is being adopted in anatomy education in health sciences and allied health: A systematic review. Anat Sci Educ. 2025;18(5):496-525.
91. García-Robles P, Cortés-Pérez I, Nieto-Escámez FA, García-López H, Obrero-Gaitán E, Osuna-Pérez MC. Immersive virtual reality and augmented reality in anatomy education: a systematic review and Meta-Analysis. Anatomical Sciences Education. 2024;17(3):514-28.
92. Bolek KA, De Jong G, Henssen D. The effectiveness of the use of augmented reality in anatomy education: a systematic review and meta-analysis. Sci Rep. 2021;11(1):15292.
93. Bank W. World Bank Country and Lending Groups Washington, DC: World Bank; 2025 [cited 2025 July 3]. Available from: https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups.
94. Al'Aref SJ. 3D printing applications in cardiovascular medicine.
95. Crump SS. Apparatus and method for creating three-dimensional objects. United States Patent. 1992;6(9). Available from: https://patents.google.com/patent/US5121329A/en11
96. McGurk M, Amis AA, Potamianos P, Goodger NM. Rapid prototyping techniques for anatomical modelling in medicine. Ann R Coll Surg Engl. 1997;79(3):169-74.
97. Chen S, Pan Z, Wu Y, Gu Z, Li M, Liang Z, et al. The role of three-dimensional printed models of skull in anatomy education: a randomized controlled trail. Sci Rep. 2017;7(1):575.
98. Lim KH, Loo ZY, Goldie SJ, Adams JW, McMenamin PG. Use of 3D printed models in medical education: A randomized control trial comparing 3D prints versus cadaveric materials for learning external cardiac anatomy. Anat Sci Educ. 2016;9(3):213-21.
99. Cai B, Rajendran K, Bay BH, Lee J, Yen CC. The Effects of a Functional Three-dimensional (3D) Printed Knee Joint Simulator in Improving Anatomical Spatial Knowledge. Anat Sci Educ. 2019;12(6):610-8.
100. Ye Z, Dun A, Jiang H, Nie C, Zhao S, Wang T, et al. The role of 3D printed models in the teaching of human anatomy: a systematic review and meta-analysis. BMC Med Educ. 2020;20(1):335.
101. Blohm JE, Salinas PA, Avila MJ, Barber SR, Weinand ME, Dumont TM. Three-Dimensional Printing in Neurosurgery Residency Training: A Systematic Review of the Literature. World Neurosurg. 2022;161:111-22.
102. VanKoevering KK, Hollister SJ, Green GE. Advances in 3-Dimensional Printing in Otolaryngology: A Review. JAMA Otolaryngol Head Neck Surg. 2017;143(2):178-83.
103. Raees S, Ullah F, Javed F, Akil HM, Jadoon Khan M, Safdar M, et al. Classification, processing, and applications of bioink and 3D bioprinting: A detailed review. Int J Biol Macromol. 2023;232:123476.
104. Michael H. Behind the net: the untold history of the ARPANET and computer science 2006 [cited 2025 Jun. 23]. Available from: https://www.columbia.edu/~hauben/book-pdf/CHAPTER%207.pdf.
105. Brinkley JF, Bradley SW, Sundsten JW, Rosse C. The digital anatomist information system and its use in the generation and delivery of Web-based anatomy atlases. Comput Biomed Res. 1997;30(6):472-503.
106. Goggin G. Google phone rising: The Android and the politics of open source. Continuum. 2012;26(5):741-52.
107. West J, Mace M. Browsing as the killer app: Explaining the rapid success of Apple's iPhone. Telecommunications Policy. 2010;34(5-6):270-86.
108. George P, Dumenco L, Dollase R, Taylor JS, Wald HS, Reis SP. Introducing technology into medical education: two pilot studies. Patient Educ Couns. 2013;93(3):522-4.
109. Choi-Lundberg DL, Low TF, Patman P, Turner P, Sinha SN. Medical student preferences for self-directed study resources in gross anatomy. Anat Sci Educ. 2016;9(2):150-60.
110. Leung BC, Williams M, Horton C, Cosker TD. Modernising Anatomy Teaching: Which Resources Do Students Rely On? J Med Educ Curric Dev. 2020;7:2382120520955156.
111. Jaffar AA. YouTube: An emerging tool in anatomy education. Anat Sci Educ. 2012;5(3):158-64.
112. Lasker R, Vicneswararajah N. Using Twitter to teach problem-based learning. Med Educ. 2015;49(5):531.
113. Chytas D. Use of social media in anatomy education: A narrative review of the literature. Ann Anat. 2019;221:165-72.
114. Joseph MA, Natarajan J, Al Zaabi O, Sirasanagandla SR. Instagram Reels improve students' knowledge, motivation, perception, and satisfaction with anatomy and physiology. Anat Sci Educ. 2025;18(6):533-43.
115. Pather N, Blyth P, Chapman JA, Dayal MR, Flack N, Fogg QA, et al. Forced Disruption of Anatomy Education in Australia and New Zealand: An Acute Response to the Covid-19 Pandemic. Anat Sci Educ. 2020;13(3):284-300.
116. Harmon DJ, Attardi SM, Barremkala M, Bentley DC, Brown KM, Dennis JF, et al. An Analysis of Anatomy Education Before and During Covid-19: May-August 2020. Anat Sci Educ. 2021;14(2):132-47.
117. Natsis K, Lazaridis N, Kostares M, Anastasopoulos N, Chytas D, Totlis T, et al. "Dissection Educational Videos" (DEVs) and their contribution in anatomy education: a students' perspective. Surg Radiol Anat. 2022;44(1):33-40.
118. Pelchen L. Internet usage statistics in 2024 Forbes Magazine; 2024 [updated Mar 1, 2024; cited 2025 Jun. 23]. Available from: https://www.forbes.com/home-improvement/internet/internet-statistics.
119. Pettersson A, Karlgren K, Hjelmqvist H, Meister B, Silen C. An exploration of students' use of digital resources for self-study in anatomy: a survey study. BMC Med Educ. 2024;24(1):45.
120. Ertel W. Introduction to artificial intelligence: Springer; 2018.
121. Ramesh AN, Kambhampati C, Monson JR, Drew PJ. Artificial intelligence in medicine. Ann R Coll Surg Engl. 2004;86(5):334-8.
122. Nagendran M, Chen Y, Lovejoy CA, Gordon AC, Komorowski M, Harvey H, et al. Artificial intelligence versus clinicians: systematic review of design, reporting standards, and claims of deep learning studies. BMJ. 2020;368:m689.
123. Loftus TJ, Tighe PJ, Filiberto AC, Efron PA, Brakenridge SC, Mohr AM, et al. Artificial Intelligence and Surgical Decision-making. JAMA Surg. 2020;155(2):148-58.
124. Madani A, Namazi B, Altieri MS, Hashimoto DA, Rivera AM, Pucher PH, et al. Artificial Intelligence for Intraoperative Guidance: Using Semantic Segmentation to Identify Surgical Anatomy During Laparoscopic Cholecystectomy. Ann Surg.

2022;276(2):363-9.
125. Sahni NR, Carrus B. Artificial Intelligence in U.S. Health Care Delivery. N Engl J Med. 2023;389(4):348-58.
126. Solaphat H, Kittipoch S, Pasit J, Chanyanart K-O, Kantapat C, Arisara A, et al. Development and evaluation of the DMIND questionnaire: Preparing for AI integration into an effective depression screening tool. Siriraj Med J. 2024;76(9):620-9.
127. Sa-ngiamwibool P, Laohawetwanit T. Comparative Analysis of ChatGPT and Human Expertise in Diagnosing Primary Liver Carcinoma: A Focus on Gross Morphology. Siriraj Med J. 2025; 77(2):119-29.
128. OpenAI. ChatGPT San Francisco, CA: OpenAI; 2022 [cited 2025 Jun. 23]. Available from: https://openai.com/index/chatgpt/.
129. Khan N, Khan Z, Koubaa A, Khan MK, Salleh RB. Global insights and the impact of generative AI-ChatGPT on multidisciplinary: a systematic review and bibliometric analysis. Connection Science. 2024;36(1):2353630.
130. Li YS, Lam CSN, See C. Using a Machine Learning Architecture to Create an AI-Powered Chatbot for Anatomy Education. Med Sci Educ. 2021;31(6):1729-30.
131. Fries A, Pirotte M, Vanhee L, Bonnet P, Quatresooz P, Debruyne C, et al. Validating instructional design and predicting student performance in histology education: Using machine learning via virtual microscopy. Anat Sci Educ. 2024;17(5):984-97.
132. Bernard J, Sonnadara R, Saraco AN, Mitchell JP, Bak AB, Bayer I, et al. Automated grading of anatomical objective structured practical examinations using decision trees: An artificial intelligence approach. Anat Sci Educ. 2024;17(5):967-78.
133. Rechowicz KJ, Elzie CA. The use of artificial intelligence to detect students' sentiments and emotions in gross anatomy reflections. Anat Sci Educ. 2024;17(5):954-66.
134. Google. Gemini Deep Research: Google; 2025 [cited 2025 July 3]. Available from: https://gemini.google/overview/deep-research/?hl=en.
135. OpenAI. Introducing GPT-4.1 in the API: OpenAI; 2025 [updated 2025 April 14; cited 2025 July 3]. Available from: https://openai.com/index/gpt-4-1/?utm_source=chatgpt.com.
136. Team P. Introducing Perplexity Deep Research 2025 [cited 2025 July 3]. Available from: https://www.perplexity.ai/hub/blog/introducing-perplexity-deep-research.
137. Aljamaan F, Temsah MH, Altamimi I, Al-Eyadhy A, Jamal A, Alhasan K, et al. Reference Hallucination Score for Medical Artificial Intelligence Chatbots: Development and Usability Study. JMIR Med Inform. 2024;12:e54345.
138. Colasacco CJ, Born HL. A Case of Artificial Intelligence Chatbot Hallucination. JAMA Otolaryngol Head Neck Surg. 2024;150(6): 457-8.
139. Yu KH, Healey E, Leong TY, Kohane IS, Manrai AK. Medical Artificial Intelligence and Human Values. N Engl J Med. 2024; 390(20):1895-904.
140. Howell MD, Corrado GS, DeSalvo KB. Three Epochs of Artificial Intelligence in Health Care. JAMA. 2024;331(3):242-4.
141. Van Booven D, Cheng-Bang C, Meenakshy M. Limitations of artificial intelligence in healthcare. Artificial Intelligence in Urologic Malignancies: Elsevier; 2025. p. 231-46.
142. Khalifa M, Albadawy M. Using artificial intelligence in academic writing and research: An essential productivity tool. Computer Methods and Programs in Biomedicine Update. 2024:100145.