Received: 27 March 2024 | Revised: 4 June 2024 | Accepted: 3 July 2024

DOI: 10.1111/1556-4029.15587

JOURNAL OF FORENSIC SCIENCES AAFS

ORIGINAL PAPER

Anthropology

Sexual dimorphism of pelvic scarring: A new method of adult biological sex estimation

Georgina Ives MSc | Sarah E. Johns PhD | Chris Deter PhD

School of Anthropology and Conservation at the University of Kent, Canterbury, UK

Correspondence
Georgina Ives, School of Anthropology and Conservation at the University of Kent, Canterbury CT2 7NZ, UK.
Email: g.ives@kent.ac.uk

Abstract

Estimating biological sex is a crucial aspect of forensic anthropology, and is pivotal in forensic investigations. Presently, the most frequently adopted osteological sex estimation methods focus on the anterior pelvis, which is easily susceptible to postmortem damage, revealing a need for additional accurate methods. This study introduces a novel method for estimating adult sex through metric pelvic scar analysis, using a known skeletal sample (169 females; 51 males). Relationships between sex and scar dimensions were subjected to Kendall's tau-B testing, and the strongest associated measurements were further analyzed using binary logistic regression to determine their predictive capacity. The final estimation method was tested on an additional known-sex sample of 43 males and 43 females from the Spitalfields skeletal collection. All associations between biological sex and scar measurements were significant, with the preauricular sulcus and newly defined inferior interosseous cavity presenting the strongest relationships (τb 0.223–0.504). Individual regression models using the approximate volume of each feature predicted sex with over 80% accuracy, but when combined in a single regression model, the accuracy increased to an impressive 97.1%. When then applied to the validation sample, the final estimation model achieved an accuracy of 90.7%. These results highlight the high estimation accuracy achieved by simultaneously utilizing the approximate volume of the sulcus and the inferior cavity. This is not only highly accurate but also utilizes the sturdier posterior pelvis, making it a promising tool for forensic investigations and the wider field of osteology.

KEYWORDS
biological profile, forensic anthropology, interosseous cavity, pelvic scarring, preauricular sulcus, sex estimation

Highlights

- Pelvic scar severity is associated with biological sex.
- Females present with more severe scarring at all sites except for the pubic tubercle.
- Posterior scarring is the most dimorphic, indicating higher sacroiliac tension in females.
- Analysis of scarring around the auricular surface can estimate sex with over 90% accuracy.

1 | INTRODUCTION

Biological sex estimation is a central process in identifying unknown individuals in forensic cases, and thus, the ability to confidently estimate sex is vital. The most widely adopted estimation methods focus on the pelvis, which is known to provide highly accurate results [1–3]. Such methods typically involve the analysis of the shape of pelvic features known to respond to the pubertal changes associated with the reproductive role of mature females. In Buikstra & Ubelaker's [4] publication on the standards for osteological data collection, well-known pelvic techniques for estimating sex are highlighted, which include analysis of pubic morphology, greater sciatic notch morphology, and preauricular sulcus formation [4–6]. However, Klales [7] reports that methods involving the categorical and morphological analysis of dry-bone pubic features are the most frequently used today. The most popular examples of such methods include that from Phenice [5] and the more recent adaptation from Klales [7, 8]. However, as Walker [6] and Dar & Hershkovitz [9] explain in focusing on the anterior pelvis, which is more likely to suffer postmortem damage, such methods are not always appropriate.

Moving away from methods that use key structural features of the pelvis for biological sex assessment, some researchers have investigated the dimorphic significance of more specific pelvic scar features. This allows for a focus on smaller pelvic details, including potential scarring around the auricular surface that may be more likely to escape postmortem damage.

1.1 | Pelvic scarring and biological sex

Pelvic scarring occurs at localized sites of soft tissue attachment, including ligamentous attachments at the anterior and posterior pelvis, and tendinous attachments on the pubis. Historically, scarring is believed to be associated with hormonal changes or physical stress caused by obstetric events [10]—attributed to the widespread skeletal laxity, increased pressure upon the musculoskeletal pelvis and abdominal muscles, and significant expansion of the pelvic structures associated with gravidity (pregnancy) and parity (birth) [11–15]. However, many researchers suggest reservations about linking scarring with obstetric events [16–22].

As a result, many studies have considered the theory that pelvic scarring at varying sites is evidence of osteological dimorphism [19, 20, 22–31]. The modern female pelvis is believed to have evolved to compensate for the difficulty in birthing introduced by bipedal adaptations [15, 32], resulting in the development of a significantly broader, more gynecoid pelvis [32]—potentially increasing connective tissue tension and resulting in bony alterations. This theory is partially supported by studies that found an association between broader pelvis and preauricular sulcus development [31, 33], with one also noting a similar relationship with dorsal pubic pitting [33]. Maxwell [34] also discovered that relaxation of pelvic connective tissues and subsequent articular diastasis occurs during cyclical periods of heightened estrogen, while Micussi et al. [35] found that pelvic floor muscle tone increases during the follicular and luteal phases of the menstrual cycle. Thus, it is possible that scarring could also arise, or at least be exacerbated by, regular periods of simultaneous muscle tension and pelvic flexibility.

Considering the presence of preauricular scarring on the ilium, most studies agree that the mere binary assessment of sulcus presentation cannot confidently determine biological sex [19, 20, 23, 25, 26, 30, 31]. However, upon consideration of the variable nature of sulcus presentation, dimorphism in degree of expression has been widely reported in both morphological studies [23, 25, 30, 31] and metric studies [19, 22, 27]—associating female sex with increased preauricular sulcus severity. In a rare dispute, Spring et al. [20] recorded significant preauricular cavitation in few female cases, suggesting a reduction in the dimorphic significance.

A similar general trend can be observed in the dimorphic presentation of pitting on the dorsal pubis [19, 22, 27, 28]. In 1986, Andersen [19] thoroughly investigated dorsal pitting, highlighting that feature absence is likely in biological males while the presence is common in over half of all females. Further studies have since supported Andersen's findings [27, 28] while Praxmarer et al. [22] observed a higher female incidence of pitting. Furthermore, Andersen [19] and Praxmarer [22] also metrically assessed scar severity at the dorsal pubis, both noting that pitting was consistently larger in biological females.

This interosseous groove develops at the site of the posterior sacroiliac ligament insertion, starting at the dorsal aspect of the superior demi face of the auricular surface and extending along the inferior dorsal border—but always terminating before the posterior inferior iliac spine [24]. Işcan & Derrick [24] were some of the first to note that the interosseous groove was observed in almost all females but only a small percentage of males, suggesting that the presence of such is highly indicative of being biologically female and vice versa—a theory supported by later studies [19, 29]. Regarding scar severity at this site, research suggests that females are also more likely to present with increased interosseous scarring [27, 29]. In contrast, Gohil et al. [26] found only a weak association between biological sex and interosseous groove presence. However, as discussed by Andersen [19] and Houghton [23], discrepancies in conclusions may be influenced by the challenge of identifying this feature, given the rugged nature of the retroauricular surface.

Further studies have also considered scarring in the form of iliac tuberosity or pubic tubercle extension [19, 22, 24, 26, 27]. Işcan & Derrick [24] first highlighted that sexual dimorphism was less pronounced when analyzing the iliac tuberosity compared to other scar features. Similarly, Gohil et al. [26] found a significant but weak association between this feature and biological sex, while others found no significant relationship [19, 27]. Praxmarer et al. [22] and Mass & Friedling [27] also metrically investigated the possibility of an association between the extension of the pubic tubercle and sex. Interestingly, opposing the trend observed through the analysis of other significant scar features, these studies found a positive association between tubercle extension and being biologically male.

1.2 | Summary and study aims

A review of relevant literature highlights the potential for further investigation into the sexual dimorphic presentation of common scar sites—particularly the preauricular sulcus, interosseous area, dorsal pubis, and pubic tubercle, which previous studies largely agree have some association with biological sex. Despite some promising results, few studies have applied a uniform metric analysis method across multiple sites. Therefore, exploration into the predictive power of pelvic scarring has been limited, and no frequently adopted sex estimation method using pelvic scar analysis has been developed to date.

This study aims to metrically assess these scar sites to investigate pelvic scar development from a biological sex perspective, and develop the most accurate possible sex estimation method based on those scars identified as having the highest predictive potential. This process involves the examination of a skeletal collection of known sex while enhancing the solely categorical grading methodology of many previous studies. In doing so, scar features were assessed individually and in combination where appropriate, while maintaining awareness of the issues associated with existing multi-feature methods.

2 | MATERIALS AND METHODS

2.1 | Skeletal samples

A sample of 169 females and 51 males from The Texas State Donated Skeletal Collection (TXSTDSC) were examined, facilitating the analysis of a large modern American population with known personal data, including confirmed biological sex. Only adult donors were selected for this study to ensure the most explicit osteological expression of sexually dimorphic features, with samples excluded in instances where bilateral pubic symphyseal or sacroiliac fusion or skeletal trauma made it impossible to observe all pelvic scar sites. As a result, the age range of the full sample was 21–103 years old, with a mean age of 68.2. The age range for the female sample only was also 21–103, but with a mean of 67 years. For males, the range was reduced to 29–97, with a mean age of 69.2. However, the exact age at death of six males was unknown, but they had been confirmed to be adults at the time of donation. Unfortunately, 90% of the individuals in the sample were white, representative of the ancestry bias of the skeletal collection.

To later validate the sex estimation method developed as part of this research, an additional skeletal sample from the Christ Church Spitalfields Collection, comprised of individuals interred across the 18th and 19th centuries, was used. This sample included 43 males and 43 females, aged between 21 and 89 at the time of death, as established based on the coffin plate associated with each individual. Among the male validation sample, ages spanned from 21 to 81, with a mean age of 53.1. In contrast, the female sample age ranged from 23 to 89, with an average age of 56.3. Those selected were well preserved and conformed to the exclusion criteria established with the initial sample. The ancestry information for this sample was not known.

2.2 | Data collection process

During the original data collection process, and the later validation study, measurements were taken from all scar sites on the left side of the pelvis where possible (see Figure 1). Where scar features on the left side were too damaged to analyze or the associated pelvic elements were missing entirely, scarring on the right side was assessed. Measurements were conducted blindly to avoid unintentional bias as a result of knowing the sex of each individual during the data collection process, thereby maximizing result validity. In collecting the data, existing metric methods from Snodgrass & Galloway [21], Maass [36], and Waltenberger et al. [10, 37] were used as a guide, with some adaptations. However, the technique used in this study to analyze pelvic scarring resulting from dorsal sacroiliac ligament interaction is novel. Here, we focused solely on cavitation confined to the dorsal borders of the auricular surface, rather than considering changes across the retroauricular surface.

Intra-observer testing of the first 20 cases in the original sample was carried out before continuing data collection for the remaining 200 to ensure no significant differences among repeated measures. There was a 1-week delay between initial and secondary measurements, and the latter results were recorded blind to those obtained during the first assessment. Where there was a difference between the two measurements, the mean value was calculated and used as the final variable result in each case.

2.3 | Dorsal pubic pitting

Maximum pitting measurements on the dorsal surface of the pubis were recorded following a similar process to that outlined by Waltenberger et al. [10, 37]. Pit width was recorded as the measurement perpendicular to the symphyseal surface, and depth as the maximum distance from the dorsal pubic surface to the base of the deepest pit using the caliper depth rod. Pitting length was measured as the full measurement of the pit parallel to the surface of the symphysis (see Figure 2) (note that "length" was used instead of "height," as seen in some studies, to maintain terminology consistency throughout this research). Where there were multiple independent pits, length measurements were taken across all of them, with an additional length measurement taken across the longest pit separately.

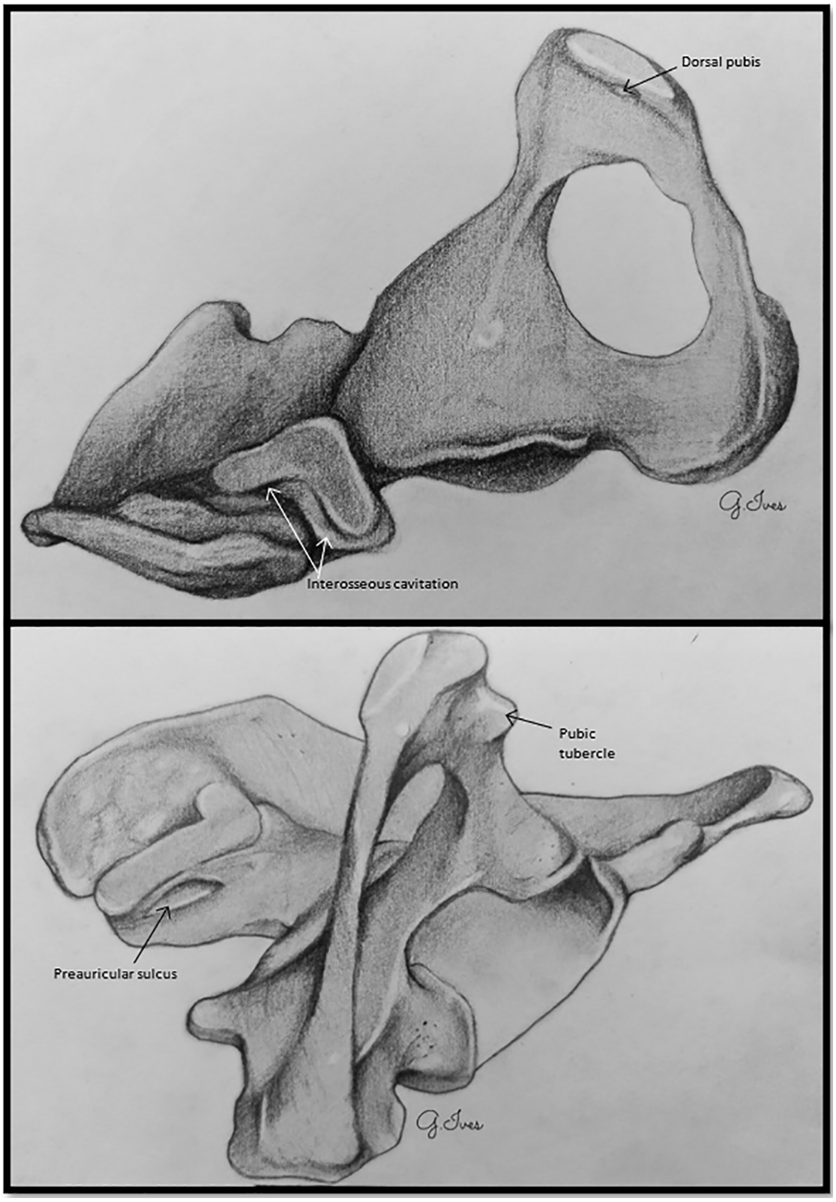

FIGURE 1 Diagrams of the left os coxa indicating all four scar sites investigated in this study.

2.4 | Preauricular sulcus

Sulcus width and depth measurements were taken for all individuals following the methodology outlined by Maass [36]. The maximum width was measured at the widest point perpendicular to the anterior edge of the auricular surface, and maximum depth was recorded as the deepest point of the sulcus floor from the unaffected surface level. In this research, an additional measurement of maximum length was also taken, which was observed as the measurement of the depression as it runs parallel to the auricular surface (see Figure 3). Note that in most cases, this depression should be clearly visible—but where it is exceptionally shallow, the edges of the sulcus should still be palpable and traceable to facilitate measurement. This is the first known study to utilize all three measurements from the preauricular sulcus.

2.5 | Pubic tubercle extension

When measuring pubic tubercle extension, the methodology developed by Snodgrass & Galloway [21] was followed with minimal deviations. Measurement was taken from the line of the natural pubis curve to the center of the tubercle extension apex (see Figure 4). However, unlike previous studies, this research preserves metric readings and does not place results into general descriptive categories.

2.6 | Interosseous cavitation

Traditionally, the study of interosseous scarring involves width and depth assessment across the full attachment site of the

JOURNAL OF FORENSIC SCIENCES AAFS

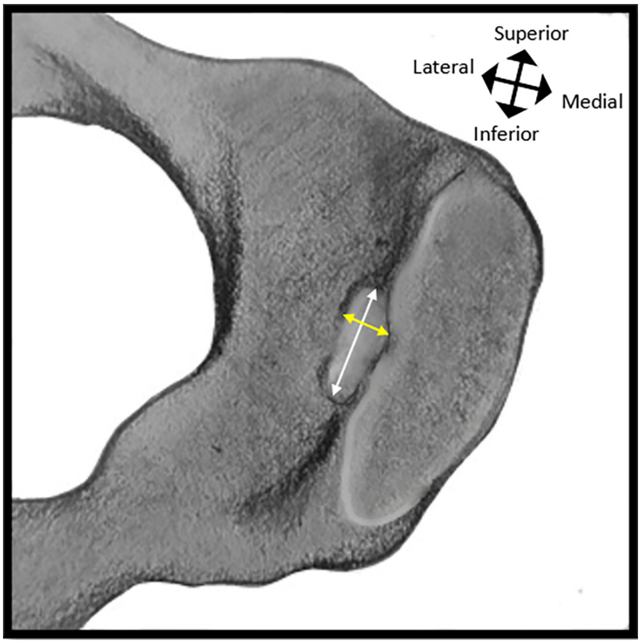

FIGURE 2 Diagram of the left dorsal pubic surface featuring a large single pit. The white line represents the maximum length measurement, and the yellow line indicates the maximum width.

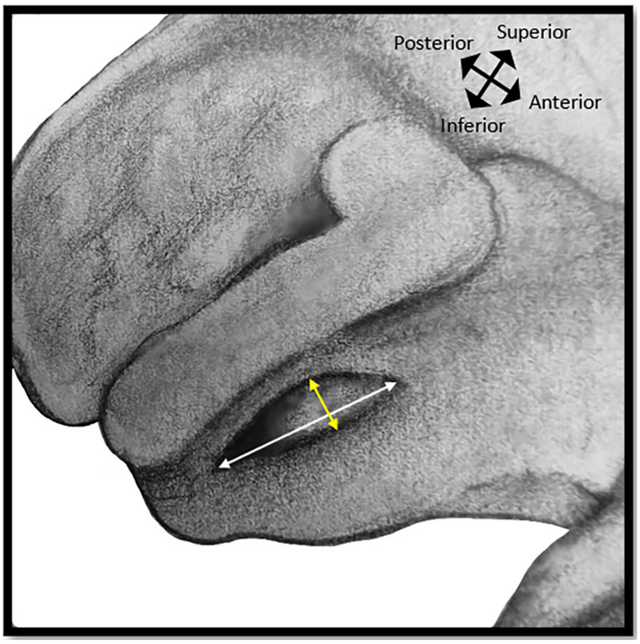

FIGURE 3 Diagram of the left auricular area with preauricular sulcus. Once again, the white line represents the maximum length measurement, and the yellow line indicates the maximum width.

interosseous ligament [10, 23, 24, 36, 37]. Given that scarring on the retroauricular surface can be challenging to identify and analyze, this research focuses solely on the cavity formed by auricular surface lipping at the dorsal edge(s) first discussed by Houghton [23] (hereon referred to as interosseous cavitation) (see Figure 5). When analyzing interosseous cavitation, maximum length measurement was comparable to that of Maass' [36] interosseous

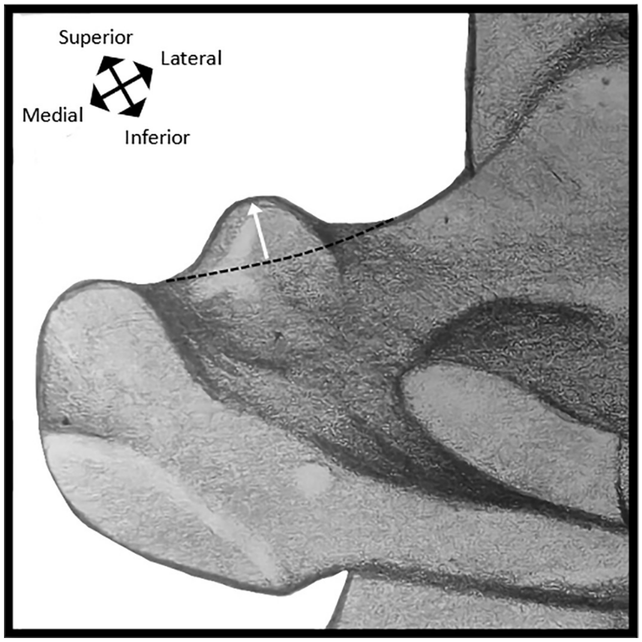

FIGURE 4 Diagram of the anterior view of the left pubis with a large pubic tubercle. The black dashed line indicates the natural pubis curve across the base of the tubercle, from which the white line presents the measurement of the maximum extension to the center of the extension apex.

analysis method in accounting for the entire length of the inferior ramus cavity. Specific to this research, a second measurement was also taken along the superior dorsal edge in cases where a cavity was present. This was not the case in all donors, but where it was, it acted as a continuation of the inferior cavity following a small, although variable, hiatus. Auricular lipping determined depth and width measurements for each cavity. Even in minor cases, lipping facilitates a maximum width measurement between it and the retroauricular floor over which it extends. Meanwhile, the maximum depth was defined as the measurement from the edge of the auricular lip to the deepest point of the cavity, where measurement is not impeded by iliac tuberosity thickening. In cases where cavitation is especially shallow or narrow, there may be small pauses in cavitation along the separate inferior or superior lengths, but for the purpose of simplifying maximum length measurements, these were disregarded.

2.7 | Statistical analysis

All statistical analyses were conducted using IBM SPSS Statistics version 28, with results considered significant at the standard threshold of $p < 0.05$. Following data collection for the first 20 cases, measurements were retaken, and initial intraobserver checks were carried out using Cronbach's Alpha internal consistency test [38]—with Phi interpreted based on the criteria established by Rea and Parker [39]. The categories are as follows: negligible (<0.099), weak (0.1–0.199), moderate (0.2–0.399), moderately strong (0.4–0.599),

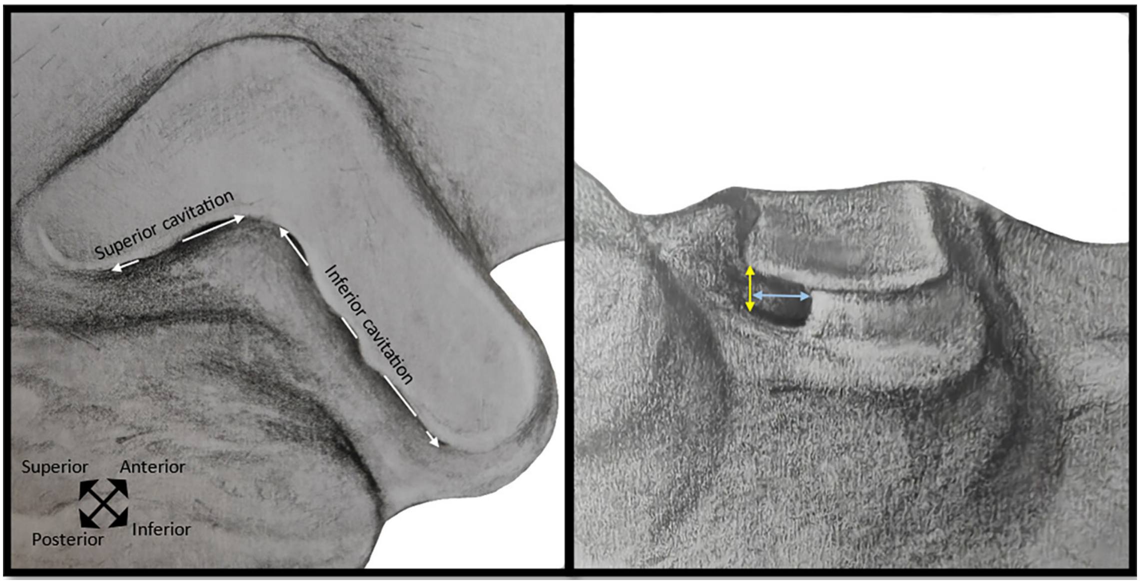

FIGURE 5 Diagrams of the medial view (left illustration) and inferior view (right illustration) of the left auricular surface indicating cavitation. The illustration on the left shows maximum length measurements for both the superior and inferior cavities, represented by white lines. The adjacent illustration presents the inferoposterior view of the inferior cavity, from the base of the auricular demiface apex. This image indicates the maximum cavity width (yellow line) and depth (blue line) measurements for the inferior cavity specifically.

strong (0.6–0.799), and very strong (>0.8). Meanwhile, Kendall's Tau-b (τb) testing was used to gain insight into relationships between sex and scar measurement variables, where significant correlations are classified as strong at 0.3 or higher [40]. The data met all required assumptions for each analysis.

In preparation for further analyses, measurements indicating a strong association with sex were checked for potential multicollinearity. This was done by assessing inter-variable correlations and consultation with collinearity diagnostics as guided by Field [41]. As appropriate, variables were then utilized in the production of logistic regression models to determine their predictive potential—adhering to a standardized residual limit of $\pm$2.5 recommended by Filler & DiGabriele [42], to reduce outlier influence but not overly restrict the effect of natural biological variation. Independent variable-logit linearity was also confirmed for final regression model predictor variables using the Box-Tidwell [43] procedure with Bonferroni correction [44].

3 | RESULTS

3.1 | Intra-observer error

Table S1 shows the intra-observer results as mean values for each scar measurement variable across the first 20 samples, with a maximum absolute difference of 0.58 mm for sulcus width and a maximum relative difference of 12.7% for pit width. Repeat measure testing for all scar variables across all 20 cases produced a high Cronbach's Alpha of 0.999, reflecting excellent consistency between test pairs and high measurement precision [45].

3.2 | Descriptive statistics and metric correlations

Table 1 presents the descriptive statistics for all individual scar measurements for both male and female groups separately, including maximum, mean, and standard deviation values. The percentage of males and females for which each feature was absent has been presented in place of the minimum value, as this was zero in all cases. Table 1 shows that there is almost no dimorphic variation in pubic tubercle extension absence. All remaining features were absent in a higher percentage of males than females. However, across both sexes, the absence count was high for both dorsal pubic pitting and the superior interosseous cavity, and low for the inferior interosseous cavity. The preauricular sulcus absence was most dimorphic, with a percentage-point difference of 38.6. Meanwhile, the maximum, mean, and standard deviation values were consistently higher across all measurement variables for females compared to males. The only exception was pubic tubercle mean, which was higher in the biological male group.

All tau-b correlations between variable measurements and biological sex were significant, (see Table S2). The most prominent feature was the preauricular sulcus, which presented a strong correlation (τb >0.3) with biological sex across all three measurements. This was followed by the inferior interosseous cavity, where width and depth displayed strong correlations (τb = 0.448 and 0.444), while the length was moderately correlated (τb = 0.223). Consequently, these scar features were selected as the focus of further analysis. Figure 6 provides reference photographs of these scar features to further assist in their analysis.

Ahead of logistic regression analysis, the key measurements were checked for multicollinearity. Upon consultation with standard

TABLE 1 Scar measurement statistics for all 51 males (M) and 169 females (F) separately.

Measurement variable	Scar feature absent count		Maximum value (mm)		Mean value (mm)		Std. deviation	
	M	F	M	F	M	F	M	F
Pit length – Single pit	48 (94.1%)	128 (75.7%)	9.45	18.78	0.44	1.53	1.878	3.725
Pit length – Multiple pits			9.45	26.36	0.55	2	2.211	4.794
Pit width			2.43	6.47	0.11	0.76	0.462	1.574
Pit depth			1.34	4.54	0.07	0.4	0.267	0.834
Sulcus length	23 (45.1%)	11 (6.5%)	26.23	54.06	8.23	24.04	8.462	10.409
Sulcus width			5.26	18.75	1.52	6.38	1.594	3.074
Sulcus depth			3.54	4.76	0.55	1.54	0.695	0.992
Superior cavity length	43 (84.3%)	113 66.9%)	11.27	29.24	0.99	4.02	2.605	6.583
Superior cavity width			1.90	4.61	0.2	0.59	0.505	0.992
Superior cavity depth			2.41	4.37	0.18	0.49	0.477	0.926
Inferior cavity length	10 (19.6%)	4 (2.4%)	32.18	36.62	14.18	20.66	9.959	7.817
Inferior cavity width			2.40	7.51	1.16	2.68	0.7152	1.247
Inferior cavity depth			2.86	8.94	0.84	2.87	0.773	1.69
Pubic tubercle extension	0 (0%)	3 (1.8%)	7.52	9.18	4	3.16	1.392	1.739
Sulcus approximate volume	**23 (45.1%)**	**11 (6.5%)**	**218.79**	**3074.6**	**23.04**	**318.63**	**36.983**	**378.5**
Inferior cavity approximate volume	**10 (19.6%)**	**4 (2.4%)**	**138.95**	**968.05**	**25.1**	**206.44**	**30.722**	**207.294**

FIGURE 6 Photos of the preauricular sulcus (left) and inferior interosseous cavity (right) as evident on samples from the Spitalfields skeletal collection courtesy of the Trustees of the Natural History Museum, London [March 06, 2024].

regression collinearity diagnostics, no significant findings were noted, but the strongest variable pairs for both the sulcus and inferior cavity were strongly intercorrelated at $r=0.781$ and $r=0.721$, respectively. Therefore, to avoid potential cases of discrete multicollinearity while allowing for the use of all measurements for both key variables, their dimension measurements were cubed to create an approximate volume variable for use in regression analysis. Correlation statistics for these two new variables can also be found in Table S2, for which further descriptive statistics have been presented in Table 1 (in bold).

3.3 | Single-variable logistic regression models

Table 1 reveals that the preauricular sulcus is absent in over 45% of males but only 6.5% of females, meanwhile the inferior interosseous cavity is absent in a reduced 19.6% of males and 2.4% of females. Furthermore, it shows that biological females had significantly higher maximum, mean, and standard deviation values for these two key approximate volume variables.

Data regarding the single-variable logistic regression models generated using these variables can be found in Tables 2 and 3 (data not in bold). Both single-variable models were statistically significant ($p \leq 0.001$). The approximate volume of the preauricular sulcus alone can explain a substantial proportion of the variation in biological sex (67.9%). In identifying an increase of 1 mm in sulcus volume associated with a 4.7% increase in the odds of an individual being biologically female, this model accurately classified 86.5% of cases, with a higher sensitivity than specificity. A lower odds ratio was observed in the inferior interosseous cavity model, with a 3.9% increase in the odds of being biologically female per 1 mm measurement increment. Separately, inferior interosseous cavity volume explained a slightly lower 52.2% biological sex variation than the sulcus volume model. In this case, specificity was reduced to just over 56%—and as a result, the overall accuracy of the inferior cavity model was decreased to 80.2%.

3.4 | Multi-variable logistic regression model

A final regression model was developed using 210 individuals from the original sample following the removal of residuals ±2.5. This involved the assessment of both key features in combination, conforming to the same biological sex threshold as the single-variable models. The information in bold in Tables 2 and 3 provides the final regression model statistics. Figure 7 offers the prediction graph corresponding with the multivariable equation statistics in bold in Table 3, to be used in the estimation of unknown samples by plotting the intersecting measurement point.

Cases where one of the two regression variables was absent were included in the analysis. Just three males did not present with either of the final key scar features, equating to only 5.9% of the male sample. Meanwhile, just one female had neither scar feature (0.6% of the female sample)—although this female was one of 10 samples excluded as residuals in the development of the final model.

Combining these variables produced the most robust prediction model, again highly statistically significant ($p \leq 0.001$). When both approximate volume measurements were utilized simultaneously, the analysis identified that for every 1 mm increase in sulcus volume, there was a 9.5% increase in the odds of the individual being biologically female, while the same measurement increase in inferior cavity volume was linked to a 10% increase in the odds of being biologically female. This model produced a final accuracy of 97.1%, significantly reducing the disparity between sensitivity and specificity percentages—both of which were over 91%.

3.5 | Method validation—Application of the final predictive model

Table 4 presents the results of the final sex estimation method as applied to the validation sample. These results have been categorized based on the likelihood of each individual being female given their approximate preauricular sulcus and inferior interosseous cavity

TABLE 2 Summary and classifications for single-predictor and combined logistic regression models (sulcus approximate volume and inferior cavity approximate volume).

Predictor variable(s) for the model	Sample (n)	Max ZResid. (1.d.p)	Chi-squared ($\chi 2$)	Sig. (p)	Nagelkerke (R^2)	% males (0) correctly predicted (specificity)	% females (1) correctly predicted (sensitivity)	Total % correctly predicted	Positive predictive value (%)	Negative predictive value (%)
Sulcus approx. volume	215	None	99.221	<0.001	0.679	73.9% (34 of 46)	89.9% (152 of 169)	86.5%	92.7%	66.7%
Inferior cavity approx. volume	217	None	90.437	<0.001	0.522	56.3% (27 of 48)	87% (147 of 169)	80.2%	87.5%	55.1%
Sulcus approx. volume and inferior cavity approx. volume	**210**	**None**	**200.795**	**<0.001**	**0.935**	**91.7% (44 of 48)**	**98.8% (160 of 162)**	**97.1%**	**97.6%**	**95.7%**

JOURNAL OF FORENSIC SCIENCES

TABLE 3 Variables in the equation for single-predictor and combined logistic regression models (sulcus approximate volume and inferior cavity approximate volume).

Predictor variable for model	Equation variable	B	S.E.	Wald	Sig. (*p*)	Exp (*B*)	95% C.I. For Exp (*B*) – Lower and upper boundaries	
Sulcus approx. volume	Predictor	0.046	0.009	27.247	<0.001	1.047	1.029	1.065
	Constant	−1.249	0.333	14.068	<0.001	0.287		
Inferior cavity approx. volume	Predictor	0.039	0.008	24.873	<0.001	1.039	1.024	1.055
	Constant	−0.741	0.294	6.376	0.012	0.476		
Sulcus approx. volume and inferior cavity approx. volume	**Sulcus vol. (predictor 1)**	**0.091**	**0.025**	**13.311**	**<0.001**	**1.095**	**1.043**	**1.15**
	Inferior cavity vol. (predictor 2)	**0.095**	**0.028**	**12.035**	**<0.001**	**1.1**	**1.042**	**1.161**
	Constant	**−8.739**	**2.499**	**12.231**	**<0.001**	**0.000**		

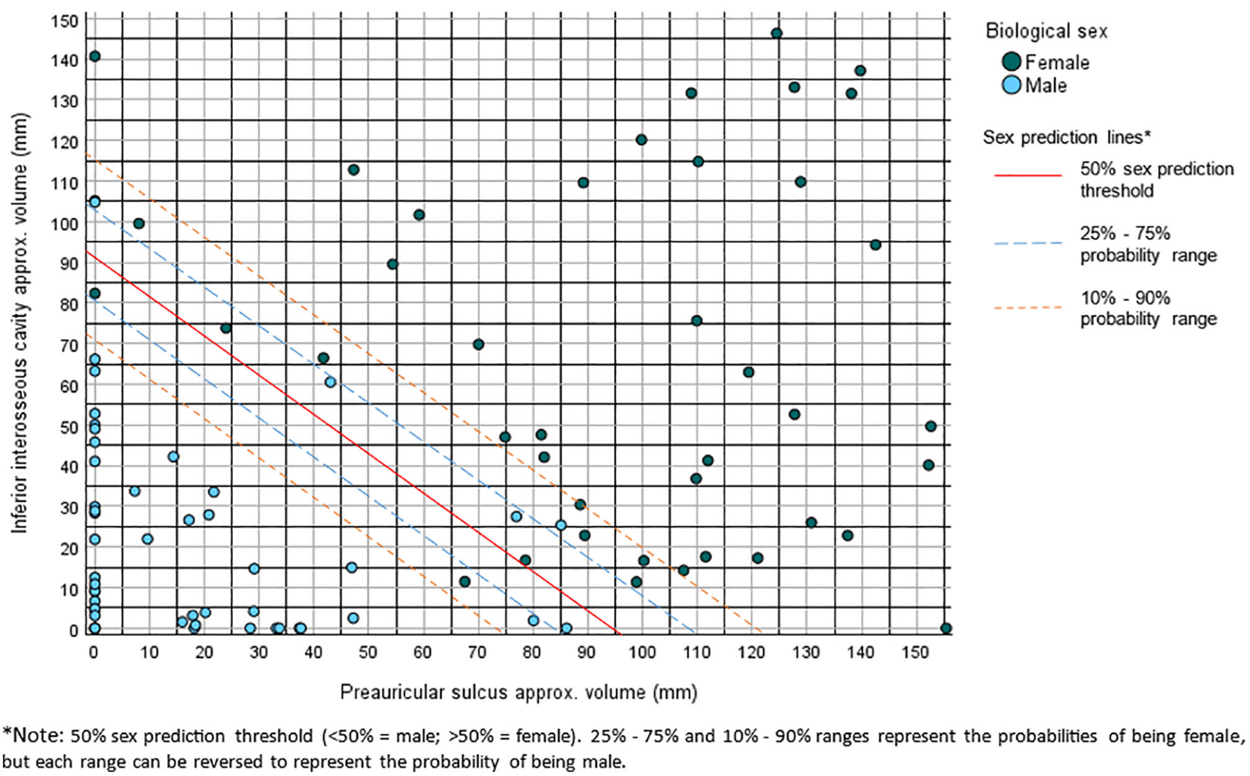

FIGURE 7 Graphical presentation of biological sex prediction based on interactions between the approximate volume of the inferior interosseous cavity and the approximate volume of the preauricular sulcus with axis measurements limited to 150 mm to aid prediction boundary interpretation.

measurements. Individuals identified as "very likely" male fell within the less than 10% probability range for being female, while "likely" males were classified as just 10–25% likely to be female. Conversely, "likely" females were identified based on a 75–90% predicted likelihood of being female, while an over 90% likelihood indicated that an individual is "very likely" female. Around the 50% separation line, individuals over 25% but less than 50% likely to be female were categorized as "most likely" male, versus those categorized as "most likely" female at between 50% and 75% likely to be female.

Table 4 reveals a final estimation accuracy of 90.7% across the sample—with a higher accuracy rate for males than females, at 97.67% and 83.73%, respectively. This equates to the incorrect biological sex estimation of just one male and seven females around the 50% predicted probability line assigned by the method. Focusing on the male sample only, over 86% of all biological males, and 88.1% of those correctly estimated, were categorized as very likely male. Meanwhile, 76.74% of the total number of biological females, or 91.7% of all of those correctly estimated, fell into the very likely female category.

4 | DISCUSSION

4.1 | Data analysis and method development

Initial analyses revealed that caution should be taken in associating scar occurrence with biological sex, with similar absence percentages in both males and females across almost all scar feature variables. Therefore, scar presence or absence cannot confidently

TABLE 4 Spitalfields sample validation classifications for males ($n=43$) and females ($n=43$), based on the probability of being female according to the final estimation method.

Sample	Classification ranges: Very likely male (<10%)	Likely male (10–25%)	Most likely male (25–50%)	Most likely female (50%–75%)	Likely female (75%–90%)	Very likely female (>90%)	
Biological males ($n=43$)	37	4	1	0	0	1	
Accumulative male accuracy (left to right)	86.05%	95.35%	97.67%				
Biological females ($n=43$)	4	2	1	3	0	33	
Accumulative female accuracy (right to left)				83.73%	76.74%	76.74%	
Total ($n=86$)	41	6	2	3	0	34	Total accuracy = 90.7%

indicate sex either way, as indicated by previous studies [19, 20, 22, 23, 25–28, 30, 31]. However, all individual scar measurements were significantly correlated with biological sex. Most of these were positive, enabling us to deduce that larger values are associated with being biologically female. The only exception was pubic tubercle extension, where a larger measurement was more likely to be observed in biological males. These findings concur with those of previous studies [19, 22, 23, 25, 27, 29–31]. The preauricular sulcus and inferior interosseous cavity presented the strongest relationship with biological sex, indicating that pelvic scar presentation is most sexually dimorphic at the sacroiliac joint of the pelvic girdle in agreement with Houghton [23].

The approximate volumes of the preauricular sulcus and the inferior interosseous cavity each proved effective in predicting biological sex, with overall accuracies of 86.5% and 80.2%, respectively. These results are comparable to the Walker [6] method of sex estimation using the sciatic notch, which has been found to produce an overall accuracy of between 80% and 89% [6, 46, 47]. When comparing the models, the decrease in accuracy of the inferior cavity model can be primarily attributed to the prediction accuracy for males (model specificity). While the sex of biological females was predicted with 85–87.7% accuracy, male group accuracy reduced from 73.9% using sulcus volume, to just 56.3% when using inferior cavity volume. Therefore, if analyzing only one of these scar features to estimate biological sex, it would be advisable to concentrate on the sulcus volume, which, with an accuracy rate of 86.5%, surpasses the average accuracy reported for existing sulcus-based methods of 81.4% [25, 30, 48].

Estimation accuracy was increased further when both approximate volume variables were integrated into the ultimate regression model. This model significantly improved prediction percentages, with a final accuracy of 97.1%—correctly identifying 98.8% of biological females and 91.7% of biological males. This accuracy exceeds that of the Phenice [5] method, which initially achieved a 96% accuracy in development, but later method validations have produced accuracies ranging from 59% to 88.4% [1, 9, 49].

The final model incorrectly classified just six individuals around the 50% sex prediction threshold, as evident in Figure 7, which were responsible for the 2.9% error in the original sample estimation. This highlights four males wrongly identified as females, and two females as males. Four cases fell close to the 25–75% prediction range near the central prediction line, identifying each as around 75% likely to be the incorrect biological sex. Meanwhile, one male and one female approached 90% likely to be the opposite sex and were therefore identified as such. No causal link was discovered upon analysis of associated personal data—suggesting standard idiosyncratic variation as a cause of outliers. The significant increase in specificity noted in this model also indicates that the majority of incorrectly predicted male cases in the earlier models were not shared across the two. This demonstrates that many of the volume measurement values responsible for incorrectly predicting sex in either of the single-variable models appear to have been offset by the other volume measurement in the final model. It is also crucial to note that no additional male outliers were excluded in comparison to the earlier models, thereby supporting the conclusion of cross-variable inaccuracy offset as opposed to the potential for accuracy increase via case removal. Furthermore, with only 5.9% of the male sample absent of both final method variables, we can be confident that there was a minimal effect of male scar feature absence on method accuracy.

4.2 | Method validation

The application of the final estimation method to the validation sample saw a reduction in overall estimation accuracy to 90.7%, although this remains slightly higher than many validation studies for frequently used morphological sex estimation methods, proving the method to be robust. Interestingly, our validation study found a higher degree of accuracy for males (97.67%) compared to females (83.73%)—which contrasts with the findings of the initial method development sample. However, there was a higher degree of confidence within the correctly estimated female sample.

It was anticipated that the accuracy of the method might decrease when tested on an archaeological population, having been developed using a modern one, with changes in population lifestyle and health over time, which may be evident throughout the skeleton.

However, one female in the sample displayed particularly strong male features according to the Phenice [5] and Walker [6] methods, as well as the method proposed in this study. Furthermore, five of seven incorrectly estimated females presented with no preauricular sulcus at all, despite this feature being highly likely in females to some degree [19, 23, 30]. Thus, we could consider the possibility that interment or exhumation errors could have occurred—with biological sex determined based on the coffin plate associated with each individual, potentially resulting in false negatives during osteological analysis. However, this may simply be a result of idiosyncratic variation outside of the standard phenotypic female presentation.

4.3 | Study limitations

In terms of the original skeletal samples used, there is a clear sex bias, with the inclusion of 169 females and 51 males. This was a result of data collection required for supplementary research within a broader study framework, as well as the time constraints related to the international travel required. This was partially compensated for by ensuring that the validation sample was comprised of an equal number of males and females. In doing so, it was concluded that the sex bias of the original sample did not have a significant impact on the sex estimation method developed, as the method was able to correctly estimate the biological sex of the validation sample while maintaining a high degree of accuracy.

As previously stated, the validation sample was temporally older than that used to develop the method, which may present issues in terms of a potential variation in sexually dimorphic skeletal expression over time. However, the Christ Church Spitalfields Collection was selected as it is one of very few large skeletal collections that are both well preserved and have associated known biological data for researcher reference. Therefore, given these qualities and the availability at the time when methodological validation was possible, it was deemed the most suitable option despite temporal incongruity.

Regrettably, due to the time and logistical restraints associated with the primary data collection, an inter-observer study was not conducted as part of this research, although intra-observer testing was diligently carried out. While it has been acknowledged that a lack of inter-observer testing can promote validity and replicability concerns, the outcome of intra-observer study reveals a high level of consistency. The methodology has also been written in a way that is hoped to facilitate further validation testing.

4.4 | Conclusion

This research has identified that the preauricular sulcus and the inferior interosseous cavity are particularly significant in biological sex estimation, and results are impressive when using the approximate volume of both scar features in combination. In the method development stage, the proposed method was shown to be the most accurate to date, and the validation process presented a slightly reduced, although still high, degree of accuracy. These findings therefore support the theory that pelvic scarring—especially that around the auricular surface—is likely a simple example of dimorphic presentation on the skeleton, with more extensive presentation indicating that an individual is biologically female. This would suggest a relationship between a broader female pelvis, the increased natural flexibility linked to reproductive processes, and subsequent increased musculoskeletal strain associated with these scar sites.

This method focuses on a specific and often well-preserved area of the pelvis, thus avoiding issues associated with multi-site methods and fragmented remains, making it applicable in more cases than traditional macroscopic methods. We have also presented an estimation graph to facilitate easy simultaneous interpretation of the approximate cubic values without the need for calculations. However, the logistic regression equation would be recommended where final estimations lie close to the central prediction line, which requires reference to Table 3, and the use of the following equation:

$$\text{Probability} = P = \frac{1}{1 + e^{-(\beta 0 + \beta 1 X1 + \beta 2 X2)}}$$

In this equation, the constant coefficient (β0) is −8.739 and the approximate volume variable coefficients (β1 and β2) are 0.091 for the preauricular sulcus and 0.095 for the inferior interosseous cavity. X1 and X2 refer to the corresponding input measurements for each case.

Going forward, further testing of this method using additional modern known-sex samples is suggested to further validate the findings of this research, and supplementary biological factors should be considered to identify any that might be influential. It would be of particular interest to investigate whether the more extensive scarring in females is further enhanced by obstetric events or simply serves as evidence of gravid potential due to a naturally broader and more flexible pelvic architecture. Similarly, it is important to consider additional biological factors that may result in a more typically female scar expression in biological males.

ACKNOWLEDGMENTS

The authors would like to thank FACTS and the Natural History Museum, London, for permitting the use of their skeletal collections for this research. They would also like to thank Dr. Daniel Wescott and Dr. Deborah Cunningham for their support during the initial data collection period and Dr. Rachel Ives for their invaluable support and encouragement when conducting the validation study.

CONFLICT OF INTEREST STATEMENT

The authors have no conflicts of interest to report.

ORCID

Georgina Ives https://orcid.org/0000-0001-6783-936X
Sarah E. Johns https://orcid.org/0000-0002-7715-7351

REFERENCES

1. Lovell NC. Test of Phenice's technique for determining sex from the os pubis. Am J Phys Anthropol. 1989;79:117–20. https://doi.org/10.1002/ajpa.1330790112
2. 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. https://doi.org/10.1111/j.1556-4029.2010.01374.x
3. White TD, Black MT, Folkens PA. Human osteology. 3rd ed. Burlington, MA: Elsevier Academic Press; 2012. p. 408–15.
4. Buikstra JE, Ubelaker DH. Standards for data collection from human skeletal remains. Arkansas Archeological Survey Research Series No. 44. Fayetteville, AR: Arkansas Archeological Survey; 1994. p. 15–20.
5. Phenice TW. A newly developed visual method of sexing the os pubis. Am J Phys Anthropol. 1969;30:297–301.
6. Walker PL. Greater sciatic notch morphology: sex, age, and population differences. Am J Phys Anthropol. 2005;127(4):385–91. https://doi.org/10.1002/ajpa.10422
7. Klales AR. Sex estimation using pelvis morphology. In: Klales AR, editor. Sex estimation of the human skeleton: history, methods, and emerging techniques. London, UK: Elsevier; 2020. p. 75–93.
8. Klales AR, Ousley SD, Vollner JM. A revised method of sexing the human innominate using Phenice's nonmetric traits and statistical methods. Am J Phys Anthropol. 2012;149(1):104–14. https://doi.org/10.1002/ajpa.22102
9. Dar G, Hershkovitz I. Sacroiliac joint bridging: simple and reliable criteria for sexing the skeleton. J Forensic Sci. 2006;51(3):480–3. https://doi.org/10.1111/j.1556-4029.2006.00119.x
10. Waltenberger L, Rebay-Salisbury K, Mitteroecker P. Are parturition scars truly signs of birth? The estimation of parity in a well-documented modern sample. Int J Osteoarchaeol. 2022;32(3):619–29. https://doi.org/10.1002/OA.3090
11. Gilleard WL, Brown JMM. Structure and function of the abdominal muscles in primigravid subjects during pregnancy and the immediate postbirth period. Phys Ther. 1996;76(7):750–62. https://doi.org/10.1093/PTJ/76.7.750
12. Dehghan F, Haerian BS, Muniandy S, Yusof A, Dragoo JL, Salleh N. The effect of relaxin on the musculoskeletal system. Scand J Med Sci Sports. 2014;24(4):220–9. https://doi.org/10.1111/SMS.12149
13. Talbot L, Maclennan K. Physiology of pregnancy. Anaest Intens Care Med. 2016;17(7):341–5. https://doi.org/10.1016/J.MPAIC.2016.04.010
14. Theodorsen NM, Strand LI, Bø K. Effect of pelvic floor and transversus abdominis muscle contraction on inter-rectus distance in postpartum women: a cross-sectional experimental study. Physiotherapy. 2019;105(3):315–20. https://doi.org/10.1016/J.PHYSIO.2018.08.009
15. Mitteroecker P, Fischer B. Evolution of the human birth canal. Am J Obstet Gynecol. 2022;230(3):841–55. https://doi.org/10.1016/j.ajog.2022.09.010
16. Stewart TD. Personal identification in mass disasters. Washington, DC: National Museum of Natural History, Smithsonian Institution; 1970. https://doi.org/10.5479/sil.30678.39088001440254
17. Holt CA. A re-examination of parturition scars on the human female pelvis. Am J Phys Anthropol. 1978;49(1):91–4. https://doi.org/10.1002/AJPA.1330490114
18. Suchey JM, Wiseley DV, Green RF, Noguchi TT. Analysis of dorsal pitting in the os pubis in an extensive sample of modern American females. Am J Phys Anthropol. 1979;51(4):517–39. https://doi.org/10.1002/AJPA.1330510404
19. Andersen BC. Parturition scarring as a consequence of rlexible pelvis architecture. [Doctoral dissertation]. Burnaby, Canada: Simon Fraser University 1986.
20. Spring DB, Lovejoy CO, Bender GN, Duerr M. The radiographic preauricular groove: its non-relationship to past parity. Am J Phys Anthropol. 1989;79(2):247–52. https://doi.org/10.1002/AJPA.1330790212
21. Snodgrass JJ, Galloway A. Utility of dorsal pits and pubic tubercle height in parity assessment. J Forensic Sci. 2003;48(6):1226–30. https://doi.org/10.1520/JFS2003027
22. Praxmarer EM, Tutkuviene J, Kirchengast S. Metric and morphological analysis of pelvic scars in a historical sample from Lithuania: associations with sex, age, body size and pelvic dimensions. Int J Osteoarchaeol. 2020;30(5):629–41. https://doi.org/10.1002/oa.2887
23. Houghton P. The relationship of the pre-auricular groove of the ilium to pregnancy. Am J Phys Anthropol. 1974;41:381–90. https://doi.org/10.1002/ajpa.1330410305
24. Işcan YM, Derrick K. Determination of sex from the sacroiliac joint: a visual assessment technique. Fla Sci. 1984;47(2):94–8.
25. Novak L, Schultz JJ, McIntyre M. Determining sex of the posterior ilium from the Robert J. Terry and William M. Bass collections. J Forensic Sci. 2012;57(5):1155–60. https://doi.org/10.1111/j.1556-4029.2012.02122.x
26. Gohil D, Dangar K, Sp R, Jethwa K, Singal G. A study of morphological features of ilium for sex determination in Gujarat state. J Res Med Dent Sci. 2014;2(4):75–8. https://doi.org/10.5455/jrmds.20142415
27. Maass P, Friedling LJ. Scars of parturition? Influences beyond parity. Int J Osteoarchaeol. 2016;26(1):121–31. https://doi.org/10.1002/OA.2402
28. McArthur TA, Meyer I, Jackson B, Pitt MJ, Larrison MC. Parturition pit: the bony imprint of vaginal birth. Skeletal Radiol. 2016;45(9):1263–7. https://doi.org/10.1007/S00256-016-2418-3
29. Mahadevappa RG, Shivalingaiah N. Sex determination by post auricular sulcus in South Karnataka. Ind J Forensic Commun Med. 2017;4(3):176–80. https://doi.org/10.18231/2394-6776.2017.0038
30. Karsten JK. A test of the preauricular sulcus as an indicator of sex. Am J Phys Anthropol. 2018;165(3):604–8. https://doi.org/10.1002/ajpa.23372
31. Canty SE. The preauricular sulcus in relation to sexual dimorphism, pregnancy and parturition in humans. [Doctoral dissertation]. Liverpool, England: Liverpool John Moores University 2020.
32. Fischer B, Mitteroecker P. Allometry and sexual dimorphism in the human pelvis. Anat Rec. 2017;300(4):698–705. https://doi.org/10.1002/ar.23549
33. Cox MJ. An evaluation of the significance of "scars of parturition" in the Christ Church Spitalfields sample. [Doctoral dissertation]. London, England: University College London 1989.
34. Maxwell AF. Abstract of discussion (joints during pregnancy, Thorpe and Fray, 1938). JAMA. 1938;111:1165–6.
35. Micussi TM, FreiTas PR, Angelo HP, Soares ME, LeMos MT, Maranhão MT. Is there a difference in the electromyographic activity of the pelvic floor muscles across the phases of the menstrual cycle? J Phys Ther Sci. 2015;27:2233–7. https://doi.org/10.1589/jpts.27.2233
36. Maass P. The bony pelvis: scars of parturition and factors influencing their manifestation. [Master's dissertation]. Capetown, South Africa: University of Cape Town 2012.
37. Waltenberger L, Pany-Kucera D, Rebay-Salisbury K, Mitteroecker P. The association of parturition scars and pelvic shape: a geometric morphometric study. Am J Phys Anthropol. 2021;174(3):519–31. https://doi.org/10.1002/AJPA.24196
38. Cronbach LJ. Coefficient alpha and the internal structure of tests. Psychometrika. 1951;16(3):297–334.
39. Rea LM, Parker RA. Designing and conducting survey research – a comprehensive guide. 4th ed. San Fransisco, CA: Jossey-Bass; 2014. p. 203–34.

40. Sharma RD. Tests of significance for the ordinal level of variables. Elements of statistics: a hands-on primer. Newcastle upon Tyne. U.K: Cambridge Scholars Publishing; 2017. p. 144–57.
41. Field AP. Discovering statistics using SPSS: (and sex and drugs and rock "n" roll). 3rd ed. London, England: SAGE Publications; 2009. p. 233–50.
42. Filler MG, DiGabriele JA. A quantitative approach to commercial damages. Hoboken, NJ: Wiley & Sons; 2012. p. 51–9.
43. Box GEP, Tidwell PW. Transformation of the independent variables. Dent Tech. 1962;4(4):531–50. https://doi.org/10.2307/1266288
44. Tabachnick BG, Fidell LS. Using multivariate statistics. 6th ed. New York, NY: Pearson Education; 2014. p. 33–52.
45. Zeller RA. Measurement error, issues and solutions. In: Kempf-Leonard K, editor. Encyclopedia of social measurement. Amsterdam, Netherlands: Elsevier Science; 2005. p. 665–76.
46. Bonczarowska JH, Bonicelli A, Papadomanolakis A, Kranioti EF. The posterior portion of the ilium as a sex indicator: a validation study. Forensic Sci Int. 2019;294:216.e1–216.e6. https://doi.org/10.1016/j.forsciint.2018.10.031
47. DesMarais A, Obertova Z, Franklin D. The influence of age on greater sciatic notch morphology: testing the Walker method in an Australian population. Int J Leg Med. 2023;138(1):239–47. https://doi.org/10.1007/s00414-023-02988-1
48. Inskip S, Scheib CL, Wohns AW, Ge X, Kivisild T, Robb J. Evaluating macroscopic sex estimation methods using genetically sexed archaeological material: the medieval skeletal collection from St John's Divinity School, Cambridge. Am J Phys Anthropol. 2019;168(2):340–51. https://doi.org/10.1002/ajpa.23753
49. Maclaughlin SM, Bruce MF. The accuracy of sex identification in European skeletal remains using the Phenice characters. J Forensic Sci. 1990;35(6):1384–92. https://doi.org/10.1520/JFS12974J

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Ives G, Johns SE, Deter C. Sexual dimorphism of pelvic scarring: A new method of adult biological sex estimation. J Forensic Sci. 2024;69:1959–71. https://doi.org/10.1111/1556-4029.15587