J Forensic Sci, March 2008, Vol. 53, No. 2
doi: 10.1111/j.1556-4029.2008.00656.x
Available online at: www.blackwell-synergy.com

Laura A. Wilson,[1] B.Sc.; Norman MacLeod,[2] Ph.D.; and Louise T. Humphrey,[2] Ph.D.*

Morphometric Criteria for Sexing Juvenile Human Skeletons Using the Ilium

ABSTRACT: Previous attempts to sex juvenile skeletons have focused on the application of qualitative or semi-quantitative techniques. This study applies a variety of geometric morphometric methods, including eigenshape analysis, to this problem. Six metric criteria for the ilia were tested with the aim of investigating previous ideas concerning sexually diagnostic characters. This study uses 25 ilia from juveniles of known age and sex from Christ Church, Spitalfields, London. Ninety-six percent of juvenile ilia were correctly identified as male or female using the shape of the greater sciatic notch. Identification accuracy is shown to improve with age for several criteria. Males were identified to a higher accuracy than females. Application of geometric techniques improves the understanding of the relationship between age, sex, and shape and the clarity with which these relationships can be quantified. Archaeological and forensic relevance of the results are discussed with recommendations for future application.

KEYWORDS: forensic science, forensic anthropology, juvenile osteology, sex determination, ilium, geometric morphometrics

The adult human skeleton can be sexed using morphological and metrical traits found on almost every bone (1). However, forensic scientists and biological anthropologists have struggled to identify criteria that might allow sex determination of juvenile human skeletons at a level of accuracy and reproducibility comparable with those documented for adults (2). The secondary sexual characteristics that provide a basis for sex determination of adults develop mainly during puberty and do not clearly differentiate juvenile males and females (3). Nonetheless Humphrey (4) demonstrates there is variation in the relative contribution of sexual differences in growth rate and duration to adult sexual dimorphism in the postcranial skeleton, with some parts of the skeleton manifesting divergent growth patterns prior to puberty. This inability to reliably sex juveniles is a serious shortcoming as one of the primary tasks in the analysis and description of human skeletal remains, whether adult or juvenile, is the determination of sex. The ability to identify the sex of juvenile individuals is critical for forensic investigations. Anthropological applications include the reconstruction of palaeodemographic patterns of childhood survivorship and mortality, growth and development, and disease stress (3,5).

Early attempts at developing suitable criteria for sexing juveniles focused on the application of techniques known to be successful with adult skeletons. In particular, early studies described various indicators that displayed marked sexual dimorphism in the adult pelvis (6,7). The innominate has long been recognized to be the most diagnostic skeletal region for estimation of sex as differences are directly related to the functional adaptation of the female pelvis for childbirth (8–10). The male pelvis exhibits a high, narrow shape having evolved to ensure economical bipedality. The female pelvis is modified to ensure obstetric success and thus is transversely oval in shape with a relatively wider inlet and greater pelvic diameter and outlet (11,12).

Studies of sexually diagnostic features of the pelvis in juveniles and fetal material have focused on the unfused ilium. Boucher (13) considered the application of basic metric techniques to identify sex differences in the fetal pelvis. Weaver (8) subsequently applied these techniques to investigate six metric characters and one nonmetric trait, achieving an accuracy of 91% for male and 75% for female identification using elevation of the auricular surface, the nonmetric trait. Hunt (10) later questioned this result, claiming that auricular surface elevation was inherently associated with growth, not sex, after re-testing the technique using an unknown Arikara Indian sample from South Dakota. More recently, Schutkowski (14) identified four characteristics of the ilium that could be useful for sex estimation; the angle of the greater sciatic notch, depth of the greater sciatic notch, curvature of the iliac crest, and the "arch criterion"—the positioning of the arch formed by drawing a cranial extension from the vertical side of the greater sciatic notch. With the exception of iliac crest curvature, all the criteria had previously displayed high levels of sex discrimination in studies of the adult skeleton (e.g., 15). Schutkowski (14) tested these characteristics on 61 juveniles of known age and sex from Spitalfields, London and was able to identify male samples with 95% and female samples with 71.4% accuracy using the angle of the greater sciatic notch. The other criteria discriminated less well between males and females, achieving 62–76% accuracies for females and 73–81% accuracies for males.

In an early attempt to apply geometric morphometrics to sex determination of juvenile skeletons, Holcomb & Konigsberg (16) used a convex hull to locate semilandmarks in a study of sexual dimorphism in the sciatic notch of 133 fetal ilia from Washington. The use of semilandmark point data enabled Holcomb & Konigsberg (16) to adjust for, and control, size such that the coordinate points produced, represented Bookstein shape variables (17). Consequently, more complex statistical methods were employed to analyze the shape variation data; Holcomb & Konigsberg used skewness and kurtosis to illustrate that male values deviated from a bivariate normal distribution and ANCOVA to deduce that

[1]Department of Biology, Imperial College London, London SW7 2AZ, U.K.
[2]Department of Palaeontology, The Natural History Museum, Cromwell Road, London SW7 5BD, U.K.
*Present Address: Universität Zürich, Palaontologisches Institut und Museum, Karl Schmid-Strasse 4, CH-8006 Zürich, Switzerland.

Received 15 April 2007; and in revised form Aug. 18 2007; accepted Aug. 19 2007.

size-related variation in the anterior/posterior position of the maximum sciatic notch depth differed by sex (16). The results indicated sex could be identified correctly in 60% of cases. This relatively low success rate was a function of the large amount of overlap observed between the male and female data.

The advent of advanced imaging software has created more intuitive geometric morphometric methods for collecting outline data than those employed by Holcomb & Konigsberg (16). Moreover, data-collection software readily records measurements in a format suitable for analysis using statistical software packages and algorithms. One such application is eigenshape analysis (18–20), a highly effective method of testing explicit shape-based hypotheses that has not yet been applied to the problem of identifying shape variation between male and female juvenile pelvic bones. Eigenshape analysis summarizes variation in shapes of objects by the multivariate analysis of a matrix of landmark-delimited outline shape functions while retaining the information content of both outlines and landmarks (see 20). Using the pairwise matrix of covariances between shape functions as the basis for a singular value decomposition, the technique re-expresses the observed shape variation in a sample as a series of hierarchical vectors such that the greatest amount of shape variation can be represented on the fewest number of independent latent shape axes. Thus, the vector array produced displays the shape variation present within a sample in a succinct, uncomplicated manner. Subsequent projection of the original shape functions into the "shape space" formed by these axes (see 21) allows the distribution of shapes to be visualized as simple ordination plots on which shape similarity and difference relations can be portrayed. Furthermore, this covariance-optimized shape space can be used to support secondary and tertiary numerical–statistical analysis of the group ordinations.

This study utilizes ilia from a known sex and age sample of juvenile skeletons from Christ Church, Spitalfields, London. The sample is analyzed using a variety of geometric morphometric methods, including eigenshape analysis, with the aim of testing previous conclusions regarding sexually diagnostic characters (1,3,8,10,13,14,16). Further application of statistical models is employed to improve upon the statistical analysis presented by previous studies in order to produce a more accurate, relevant quantification of sexual dimorphism, independently, in terms of magnitude and shape.

Materials and Methods

This study was completed as a blind test using ilia from a known sex and age sample of juvenile skeletons from the Christ Church, Spitalfields Collection, held at the Natural History Museum, London. The collection represents burials in the vaults of Christ Church between 1729 and 1852 (22). Excavations of the crypt yielded a total of 968 skeletons of which 387 could be identified from coffin plates.

After exclusion of poorly preserved material and material exhibiting gross pathology, a sample of 25 juvenile individuals was available for study. The sample consisted of 17 males and eight females, ranging from birth to 7.88 years (Table 1). The left ilium was preferentially selected, but in 10 cases where the left ilium was absent or state of preservation was less than suitable, the right ilium was used in replacement. Each of the samples was oriented in a consistent manner and photographed in two dimensions using Image Pro Plus© (Media Cybernetics, Inc., Silver Spring, MD), a dedicated image analysis software package (Figs. 1 and 2). Image measurement tools within the Image Pro Plus© package were used to collect all geometric data for this study. Eigenshape analysis was applied to each set of data with the exception of the nonlinear greater sciatic notch angle data. Eigenshape 2.1© Software (Norman MacLeod, Natural History Museum, London, U.K.) was used to transform the Cartesian *x–y* coordinate data into the ϕ form of the Zahn and Roskies shape function (23), the data type usually used for eigenshape analysis. This step is intended to remove shape differences due to object position, orientation, and size (23). Eigenshape analysis was used to determine a set of eigenscores, covariances between the observed shapes and the Eigenshape functions (20). These were used to compute a discriminant function, dependent on *a priori* knowledge of individual sex, to find the orientation of an axis along which males and females are separated the most and inflated the least. Six separate areas of measurement were identified as follows.

TABLE 1—*Age and sex details for the study sample.*

Sample number	Sex	Age (years)
1	F	0.04
2	F	0.04
3	F	0.10
4	F	1.10
5	F	1.14
6	F	1.24
7	F	1.64
8	F	3.21
9	M	0.00
10	M	0.06
11	M	0.06
12	M	0.12
13	M	0.42
14	M	0.83
15	M	0.92
16	M	1.29
17	M	1.50
18	M	2.46
19	M	2.56
20	M	2.61
21	M	4.17
22	M	4.50
23	M	4.55
24	M	5.40
25	M	7.88

F, female; M, male.

FIG. 1—*Ilium viewed from the ventral aspect. This orientation of the bone was used to record measurements for the shape of the greater sciatic notch, auricular surface and angle of the greater sciatic notch criteria.*

FIG. 2—*Orientation of the ilium as photographed in two dimension. A trace tool was used to record the outline of the iliac crest in this orientation. Equal distance coordinate points were identified along the outline.*

Greater Sciatic Notch Angle

Three landmark points were identified to characterize the angle of the notch (Fig. 3); these were standardized to be (a) the point at which the notch shape is distinct from the curvature of the crest, (b) the point of maximum curvature—the vector of the angle, and (c) the point at which the notch transitions into the basal attachment area. For each landmark point, $x–y$ coordinate data were obtained and transformed into a corresponding angle measurement using the dot product. The angle measurements were analyzed using the methodology for circular scale data discussed by Zar (24); the data set was subject to Watson and Williams Parametric Two-Sample test and Watson's U^2 test to enable conclusions to be drawn with regard to the relationship between the two groups within the sample and the mean angles for each of the two groups. Two ilia were selected at random from the sample population, to address and quantify the intra-observer error consequent in the orientation and identification of the landmark points. For each bone the orientation, image capture, and landmark point identification was repeated 20 times. Rayleigh's test was used to assess the significance of the mean angle, calculated from the 20 measurements. This enabled intra-observer error to be quantified (24), so that confidence in the original data collection method could be gauged.

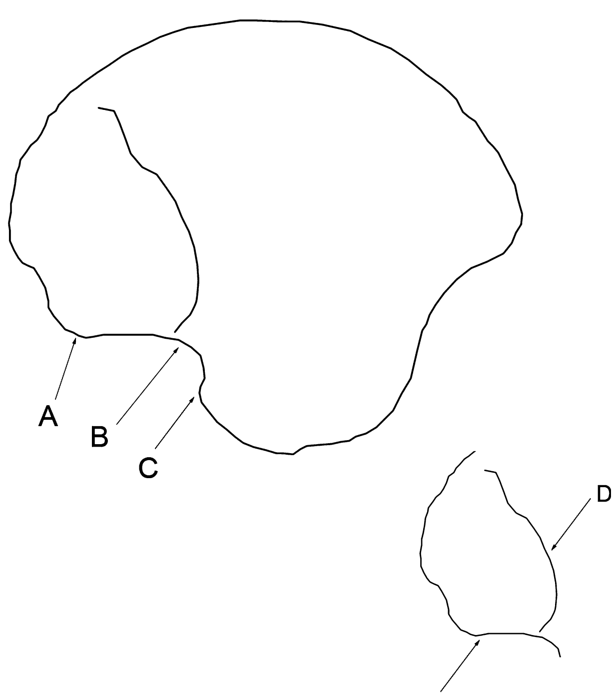

FIG. 3—*The angle of the greater sciatic notch was defined using points A, B, and C. The outline of the greater sciatic notch was defined to begin at point A and finish at point C, a trace tool was used to record 50 evenly spaced points along the outline of the bone. The auricular surface morphology was defined to begin at point A and finish at point D; a trace tool was used to record 50 evenly spaced points along the outline of the bone.*

Shape of Greater Sciatic Notch

The outline of the greater sciatic notch area was recorded using the Image-Pro tracing tool (Media Cybernetics, Inc.,) with 50 evenly spaced coordinate point measurements taken along the trace outline. For each sample the trace began at the point at which the notch shape was distinct from the curvature of the crest, and ended at the point of transition to the basal attachment area (Fig. 3). To test the stability and reproducibility of the discriminant function axes, a Jack-knife test (25) was performed on the dataset. This involved sequentially removing one individual and repeating the eigenshape analysis after each removal to generate 25 new discriminant functions. The original discriminant function score for each sample, a single value representing the location of the sample along a line defined by the linear discriminant function, was then recalculated to verify accuracy. A t-test was performed on the discriminant function scores to enable comparison with the nonlinear angle measurement data. The t-test was used to test the null hypothesis that the mean value representing the female group was equal to the mean value representing the male group.

Curvature of the Iliac Crest (as a Closed Curve)

To record the curvature of the iliac crest the ilium was viewed from the top and oriented at 90° to the photograph platform with the dorsal surface aligned vertically (Fig. 2) such that the tracing tool could be used to record 100 evenly spaced coordinate measurements along the outline of the crest.

Curvature of the Upper Plane and Lower Plane of the Iliac Crest

Two landmark points were used to divide the crest, as considered when the bone was viewed in the orientation displayed in Fig. 4, in order to quantify variation in curvature of the upper and lower portions of the bone independently. These landmark points were defined as the most distal point of the crest on the right of the ilium and the most distal point on the left side of the ilium

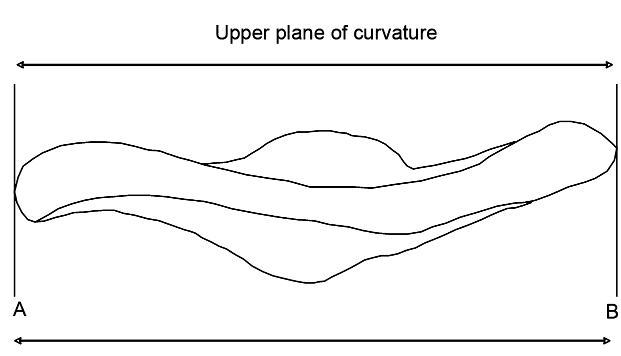

FIG. 4—*Outline of the iliac crest illustrating the division of the crest into two separate planes. Point A was defined as the most distal point on the left of the crest and point B was defined as the most distal point on the right of the crest. A trace tool was used to record 50 evenly spaced coordinate points from point A to point B for the upper plane and from point B to point A for the lower plane, following the outline of the bone.*

(Fig. 4). This enabled the upper and lower portions of the crest to be captured separately using the tracing tool to record 50 evenly spaced coordinate measurements on each part.

Auricular Surface Morphology

The trace measurement was used to record 50 evenly spaced coordinate points along the outline of the auricular surface to capture the shape. The auricular surface is the roughened area located on the lateral surface of the sacrum; this surface articulates with the ilium in the sacroiliac articulation. The sacrum in females is wider at the sacral ala than in males, and females have a wider sacro-vertebral angle than males (26). This study considers, for the first time, whether the auricular surface displays sexual dimorphism analogous to that exhibited by the attached bone, the sacrum. The auricular surface was defined for each sample to begin at the point closest to the iliac crest where the auricular morphology was distinct from the surface of the bone. An outline was then traced until the sacro-iliac surface was no longer clearly distinguishable; this second landmark point was closest to the greater sciatic notch region of the ilia (Fig. 3).

Age distribution within the sample was considered a potential source of error within the discriminant function produced from the eigenshape analysis. In particular, the identification of landmark points required to accurately constrain the trace outline was considered to be more problematic with samples younger than 0.5 years. To evaluate the impact of very young (<0.5 years) individuals on the accuracy of the discriminant function, the eigenshape analysis was recalculated for each of the areas of measurement excluding individuals less than 0.5 years old, irrespective of sex. This exclusion was then repeated a second time, excluding individuals younger than 0.1 years old. Eigenshape results, along with the resultant discriminant functions, were compared with those produced from analysis of the full sample.

Results

Intra-observer error arising from data collection methods was quantified using the *r*-statistic, applicable to circular data (23). A correlation (*r*) of 1.0 would indicate maximum concentration of data about the mean value and hence minimum error. The reverse would be true for an *r*-value of 0.0. For both error samples the *r*-value was 0.999, which is significantly higher than the critical value of 0.569 ($p < 0.001$) for this sample.

Correct sex identification of juvenile ilia varied from 83% to 100% for males and 25% to 88% for females using the five different techniques proposed in this study (Table 2). Of the five study areas subjected to eigenshape analysis, the greater sciatic notch criterion (Table 2) is the best discriminator for sex; 96% of the individuals in the sample were correctly identified when the eigenshape scores were subjected to canonical variates analysis to generate a discriminant function. The misidentified individual was female and aged 0.1 years; the discriminant function analysis determined this individual to be male with a 79.3% probability of membership to the male group. The discriminant function generated from the auricular surface morphology data enabled 84% of individuals to be allocated to the correct sex group. This criterion was less successful than the shape of the greater sciatic notch, but was considerably more effective, especially with female identification (62.5%) than any of the criteria associated with the curvature of the iliac crest. Correct identification of females using the curvature of the iliac crest criteria varied from 25% to 37.5%, less than half that achieved for the other two criteria displayed in Table 2. Females were also considerably less well identified than males. Accurate identification using the curvature of the iliac crest criteria ranged from 82.4% to 88.2% for males (Table 2).

To test the reproducibility of the 96% accuracy achieved with the greater sciatic notch criterion, a Jack-knife test was performed on the data set. Results are detailed in Table 3. Correct identification varied across the tests from 87.5% to 100% with an average accuracy of 91.7%. In some cases removal of the youngest individuals improved the accuracy of the discriminant function, whereas the accuracy fell below 92% when males and females older than 2 years were removed from the data set. The recomputed F1 values listed in Table 4 illustrate the stability of the discriminant function axes; each recomputed F1 value from the reduced data set remained close to the group centroid generated from the canonical analysis of the complete data set. Hence, on the basis of the recomputed F1 values, none of the test samples would be reclassified as belonging to the incorrect sex group. Table 4 also lists the probability associated with the classification of each sample to a sex group, based upon the original discriminant function, generated with the complete data set. When correctly identified, a female individual was assigned to the female group with an average certainty of 91%. Similarly, on average a male individual was assigned to the male group with a certainty of 93%.

The discriminatory capacity of each criterion altered when the data set was adjusted to exclude samples younger than 0.1 years and also samples younger than 0.5 years. For all variables, apart from the curvature of the lower plane of the iliac crest, exclusion of younger individuals reduced the error margin, and hence increased the separation of the male and female groups along the discriminant function (Table 5). The data for the curvature of the lower plane of the iliac crest demonstrated an identification accuracy of 72% using the full sample set of 25 individuals (Table 2). Accuracy was reduced slightly (to 70%) by the removal of individuals younger than 0.1 years (three males and two females) and increased slightly (to 70.6%) when the test was repeated without individuals younger than 0.5 years (five males and three females).

TABLE 2—*Percentage of juvenile ilia correctly identified as male or female out of a sample of 17 males and eight females, using discriminant function analysis on five anatomical areas. Chance of correct identification calculated according to Bayes Theorem, probability calculated for an average population, assuming equal prevalence of males and females (*p *= 0.5).*

Anatomical Region	Full Sample	% Correctly Identified		% Chance of Correct Identification	
		Males	Females	Males	Females
Auricular surface morphology	84	94.1	62.5	75	76
Greater sciatic notch shape	96	100	87.5	100	90
Curvature of iliac crest (closed curve)	64	82.4	25	54	57
Curvature of iliac crest (upper plane)	68	88.2	25	59	64
Curvature of iliac crest (lower plane)	72	88.2	37.5	65	60

TABLE 3—*Result of Jack-knife analysis for the shape of the greater sciatic notch; each individual was removed from the sample set sequentially and the discriminant function was then recalculated for the removed individual.*

	Age of Sample Excluded (years)	Sex of Sample Excluded	Correctly Identified (%)	No. of Males Misidentified	No. of Females Misidentified
All data	–	–	96.0	0	1
Test 1	0.04	Female	91.7	1	1
Test 2	0.04	Female	91.7	1	1
Test 3	0.1	Female	95.8	1	0
Test 4	1.1	Female	87.5	1	2
Test 5	1.14	Female	87.5	1	2
Test 6	1.24	Female	95.8	0	1
Test 7	1.64	Female	91.7	1	1
Test 8	3.21	Female	87.5	2	1
Test 9	0	Male	91.7	1	1
Test 10	0.06	Male	100.0	0	0
Test 11	0.06	Male	87.5	1	2
Test 12	0.12	Male	95.8	1	0
Test 13	0.42	Male	87.5	1	2
Test 14	0.83	Male	91.7	1	1
Test 15	0.92	Male	91.7	1	1
Test 16	1.29	Male	91.7	1	1
Test 17	1.5	Male	100.0	0	0
Test 18	2.46	Male	91.7	1	1
Test 19	2.56	Male	87.5	2	1
Test 20	2.61	Male	91.7	1	1
Test 21	4.17	Male	87.5	2	1
Test 22	4.5	Male	91.7	1	1
Test 23	4.55	Male	91.7	1	1
Test 24	5.4	Male	91.7	1	1
Test 25	7.88	Male	91.7	1	1

TABLE 4—*Results of Jack-knife test; distances to the female and male centroid values are based on the discriminant function generated from the complete data set and are included to enable the movement between the original F1 values (Canonical discriminant function coefficients) and those recomputed after the Jack-knife test, to be evaluated in context.*

	Age of Sample Removed (years)	Sex of Sample Removed	Before Jack-knife F1 (n = 25)	After Jack-knife F1 (n = 24)	Probability Female	Probability Male
Test 1	0.04	Female	2.45	−0.775	**0.992**	0.008
Test 2	1.24	Female	1.725	0.797	**0.939**	0.061
Test 3	0.04	Female	1.18	−1.197	**0.76**	0.24
Test 4	1.1	Female	3.702	0.267	**1**	0
Test 5	1.14	Female	2.658	1.403	**0.996**	0.004
Test 6	1.64	Female	1.053	−1.875	**0.687**	0.313
Test 7	3.21	Female	2.685	−0.007	**0.996**	0.004
Test 8	0.1	Female	0.32	0.611	0.207	**0.793**
Test 9	0.06	Male	−0.789	1.204	0.01	**0.99**
Test 10	7.88	Male	0.36	1.416	0.227	**0.773**
Test 11	0.42	Male	−0.593	0.68	0.018	**0.982**
Test 12	4.5	Male	−1.027	0.303	0.005	**0.995**
Test 13	0.92	Male	−1.32	0.403	0.002	**0.998**
Test 14	0.83	Male	0.663	0.366	0.415	**0.585**
Test 15	2.61	Male	−1.608	0.93	0.001	**0.999**
Test 16	2.46	Male	−1.177	−0.453	0.003	**0.997**
Test 17	1.5	Male	−1.022	−2.777	0.005	**0.995**
Test 18	4.55	Male	−0.931	1.348	0.007	**0.993**
Test 19	4.17	Male	−2.864	0.095	0	**1**
Test 20	2.56	Male	−0.659	−1.33	0.015	**0.985**
Test 21	0.06	Male	−2.067	−1.224	0	**1**
Test 22	0.12	Male	0.192	3.999	0.153	**0.847**
Test 23	0	Male	−0.815	−0.212	0.01	**0.99**
Test 24	1.29	Male	−2.303	−1.003	0	**1**
Test 25	5.4	Male	0.186	−0.89	0.151	**0.849**

Numbers in bold indicate the probability of membership to the male or female group for the correctly identified individuals.

These results may be contrasted with those obtained for the other criteria, most markedly the curvature of the iliac crest. The initial analysis of these closed curve data revealed identification accuracy to be 64% (Table 2). This result improved to 100% with removal of the individuals under 0.5 and 0.1 years. The removal of younger individuals from the data for greater sciatic notch shape had a comparable effect upon accuracy, which increased from 96% to 100% (Table 2). There was also a marginal increase in accuracy for the auricular surface morphology data from 84% to 85% when individuals younger than 0.1 years were excluded from the analysis (Table 2). Accuracy was further improved (to 94%) with the additional exclusion of individuals younger than 0.5 years. On average,

TABLE 5—*Comparison of error rate generated when the sample was analyzed excluding individuals younger than 0.5 years and individuals younger than 0.1 years. Three females and five males were under the age of 0.5 years, reducing the sample size to seventeen. Three males and two females were under the age of 0.1 years reducing the sample size to twenty. The error rate was generated from the learning sample used to create the discriminant function.*

	All Data Excluding <0.5 Years Old			All Data Excluding <0.1 Years Old		
	% Correctly Identified			% Correctly Identified		
Anatomical Region	Full Sample	Male	Female	Full Sample	Male	Female
Auricular surface morphology	94.1	100.0	80.0	85.0	100.0	50.0
Greater sciatic notch shape	100.0	100.0	100.0	100.0	100.0	100.0
Curvature of iliac crest (closed curve)	100.0	100.0	100.0	100.0	100.0	100.0
Curvature of iliac crest (upper plane)	70.6	91.7	20.0	70.0	85.7	33.3
Curvature of iliac crest (lower plane)	70.6	91.7	20.0	70.0	92.9	16.7

across all five shape variables, 91% of males and 47.5% of females were correctly identified using the full sample. In comparison, 95.7% of males and 60% of females were accurately identified when individuals younger than 0.1 years were excluded. The accuracy with which both sexes were correctly assigned to a group further improved with the exclusion of individuals less than 0.5 years old, with average correct identifications of 96.7% for males and 64% for females. The covariance between the first and second eigenshape axes was analyzed in relation to age (Fig. 5). The eigenshape-based ordination of the data for greater sciatic notch shape reveals a clustering of males older than 3.5 years while females appear not to display any age related clustering within the group (Fig. 5*a*). Clustering of males older than 3.5 years is also evident within the data for the curvature of the iliac crest (Fig. 5*b*).

The results of the *t*-test (Table 6) indicate greater sciatic notch shape data and auricular surface morphology data provide the greatest level of confidence ($p < 0.001$) with which to reject the null hypothesis that the mean of the male group is equal to the mean of the female group. A nonparametric Watson–Williams U^2 test was used to test the same null hypothesis for the nonlinear greater sciatic notch angle data, producing a value of 0.22 indicating the null hypothesis can be rejected with a confidence of $p < 0.05$. This was slightly less significant than the confidence level calculated for the curvature of the iliac crest ($p < 0.0184$). For the other two variables differences between males and females were not statistically significant (Table 6).

The greater sciatic notch angle measurements taken for the two randomly selected ilia, to test the error margin inherent within the data collection methods, were subject to angular dispersion testing. Results (Table 7) indicate the angle measurements for both Error sample 1 and Error sample 2 were highly concentrated ($r = 0.999$) for each. This is significantly higher than the critical value of 0.569 ($p < 0.001$) for the sample. Since an *r*-statistic value of 1.0 represents maximum concentration, the value of 0.999 can be associated with minimum error within the sampling methods as error free sampling should, in theory, generate a series of identical results that would have an *r*-value of 1.0. The value of 0.987 (Table 7) for the angle measurements from the complete data set reflects the relatively small amount of dispersion in the data set with the smallest greater sciatic notch angle measuring 108.6° and the largest 143.6°, irrespective of sex. Male angle measurements ranged from 108.6° to 143.6° whereas females had a comparatively smaller range of 123.4° to 142° (Fig. 6). Hence, the standard deviation of 8.2° about a mean angle of 132.5° for females is smaller than the standard deviation for males of 8.6° about a mean angle of 126.0° (Table 8).

Rayleigh's *R* was used to test the null hypothesis of random distribution in the population. For all tests an *R*-value of 24.697 (Table 7) indicates the population as a whole does not have a random distribution, thus rejecting the null hypothesis with a confidence level of $p < 0.001$ (12.806). The application of Rayleigh's test to the error sample data provides an assessment of confidence in the mean angle, calculated from the 20 tests for each sample. Even if there is no mean direction in the population, a random sample may still display a calculable mean (21). Consequently, a significant Rayleigh's *R* would indicate the mean represented the true direction of the population. For both error samples the *R* values are significantly higher than the $p < 0.001$ value of 11.375 (Table 7). The Watson–Williams test, a parametric two-sample test, was used to consider the null hypothesis that the mean angle for females was equal to the mean angle for males. The Watson–Williams test combines the *R*-values from the two populations (male and female) to generate an *F*-statistic. This was calculated as 2.574 which indicates the male angle measurements are different from the female angle measurements such that the two sexes represent separate populations; thus the null hypothesis may be rejected with a confidence level of $p < 0.1$ (Table 7).

The greater sciatic notch angle for males displays a trend towards becoming narrower with age (Fig. 7). ANCOVA analysis was performed on the full data set to assess the relationship between angle measurement (dependent variable), age (quantitative variable), and sex (qualitative variable). The result of the ANCOVA test produced an *F*-statistic value of 2.583; this was not statistically significant ($p = 0.098$), but there may still be a significant trend found in males. The failed ANCOVA test likely reflects the effects of a small sample size.

Discussion

Sex identification using the criteria proposed in this study has resulted in varying levels of success. The highest level of accuracy was achieved with the greater sciatic notch shape criterion (96%). The Jack-knife test performed on this data set indicated a high level of reproducibility; analysis of the recalculated F1 values indicated the original discriminant function axes were completely stable, such that none of the individuals were reclassified into the opposite sex group. Moreover, the level of certainty with which an individual could be assigned to a group was consistently high for both males (93%) and females (91%), thus highlighting the significant level of separation, along the discriminant axis, between male and female shape. The individual misidentified (accounting for the 4% error) in the original sample was a 0.1-year-old female, which improves the confidence that may be held in the result. If the misidentified individual had been much older then the greater sciatic notch shape criteria would have contrasted with the trend, displayed by several of the criteria, towards improved accuracy with age. Schutkowski

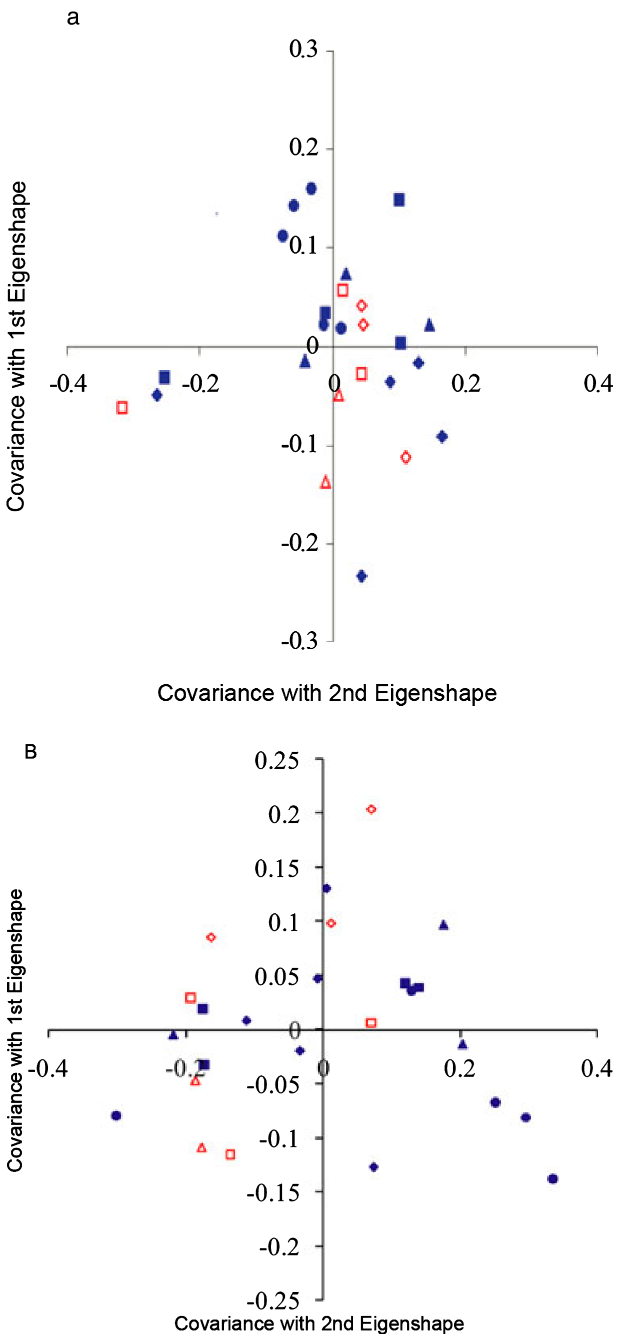

FIG. 5—*Eigenshape space plots of shape difference relations existing within the data collected for the greater sciatic notch shape (a) and the curvature of the iliac crest (b). Each plot was created by determining the covariance between the shape functions of each outline and the first two eigenshapes, determined from a singular value decomposition of the inter-object shape function covariance matrix. Males are represented by filled shapes, females are represented by open shapes. For each sex, where applicable, the data is grouped into age categories as follows: <0.5 years (◇), <1.5 years (□), <3.5 years (Δ), and >3.5 years (○).*

achieved 81.2% accuracy with males and 76.5% accuracy with females when using an angular scoring method to reflect the shape of the greater sciatic notch (13). This may be compared to 100% male accuracy and 88% female accuracy produced using the shape based techniques discussed in this study. Shape models derived from the greater sciatic notch eigenshape data indicate the first

TABLE 6—*The t-test values for each criterion studied, the data for the greater sciatic notch angle was subject to a directly equivalent test, the Watson–Williams* U^2 *test, used for circular data (24). The* t*-test is used to assess the null hypothesis that the male group mean is not statistically different from the female group mean, the* p *(two-tailed) value gives the probability with which this hypothesis can be rejected.*

Anatomical Region	*t*-value	*p* (two-tailed)
Auricular surface morphology	5.74	<0.001
Angle of greater sciatic notch	0.22*	<0.05
Greater sciatic notch shape	6.76	<0.001
Curvature of iliac crest (closed curve)	2.44	>0.05
Curvature of iliac crest (upper plane)	1.77	>0.05
Curvature of iliac crest (lower plane)	1.80	>0.05

*This value is a U^2 value however is directly comparable to the linear *t*-value.

TABLE 7—*Circular distribution analysis for greater sciatic notch angle measurement data (24). Angular dispersion (*r*) was calculated to describe the dispersion of the data,* r *varies inversely within the parameters of 0–1.0, with the dispersion of the data. The circular standard deviation (*s*) is analogous to the standard deviation for linear data. Rayleigh's test (*R*) was used to test the hypothesis of a random distribution in the population; when considering the error samples this value can be used to assess whether the mean represents the population direction accurately since a random sample may still display a calculable mean. The* F*-value is the critical value used by the Watson–Williams test to consider the null hypothesis that the mean angle for males is equal to the mean angle for females, and hence angle measurement can not be used as a basis to sex an individual.*

	All Data	Error Sample 1	Error Sample 2
Sample size (*n*)	25	20	20
r	0.987	0.999	0.999
$p < 0.001$ value	0.512	0.569	0.569
s	8.947	1.478	2.348
R	24.697	19.993	19.983
$p < 0.001$ value	12.806	11.375	11.375
F	2.574	–	–

r = measure of concentration, varies from 0 (maximum dispersion) to 1.0 (maximum concentration).
s = circular standard deviation.
R = Rayleigh's—measure to test random distribution within a population.
F = Critical value for Watson–Williams two sample parametric test.
p = Probability.

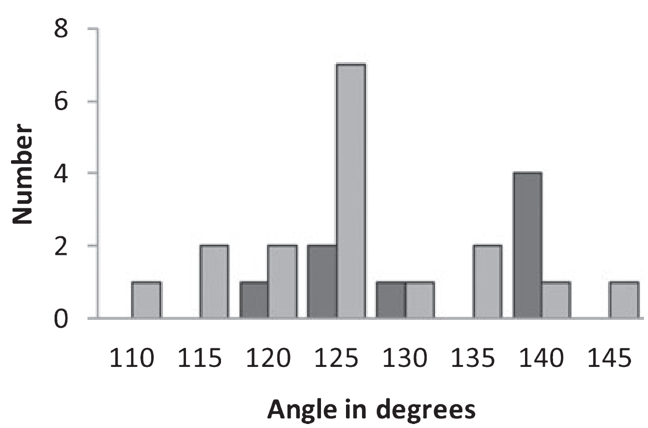

FIG. 6—*Histograms illustrating the comparative distribution of male and female greater sciatic notch angle measurements. Males are indicated by dark gray and females are represented by light gray bars. See text for discussion.*

eigenshape axis is associated with a progressive widening of the notch with shape change focused upon the central region of the notch whilst the second eigenshape axis delineates localized shape

TABLE 8—*Circular distribution analysis of male and female greater sciatic notch angle measurements. The circular standard deviation statistic* (s) *illustrates the spread about the mean angle for each population. The angular dispersion* (r) *describes the dispersion of the data;* r *has no units and varies inversely within the parameters of 0–1.0, with the dispersion of the data (24).*

	Male (n = 17)	Female (n = 8)
Mean angle (degrees)	126.0	132.5
s	8.631	8.177
r	0.989	0.990

s = circular standard deviation.
r = measure of concentration, varies from 0 (maximum dispersion) to 1.0 (maximum concentration).

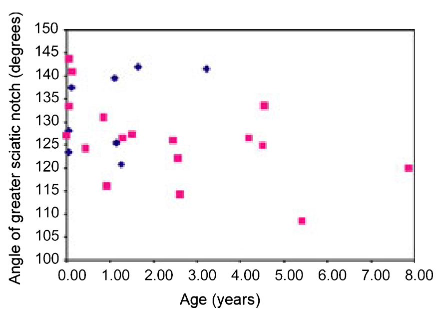

FIG. 7—*Illustration of how male (□) and female (◇) greater sciatic notch angle measurements vary with age. See text for discussion.*

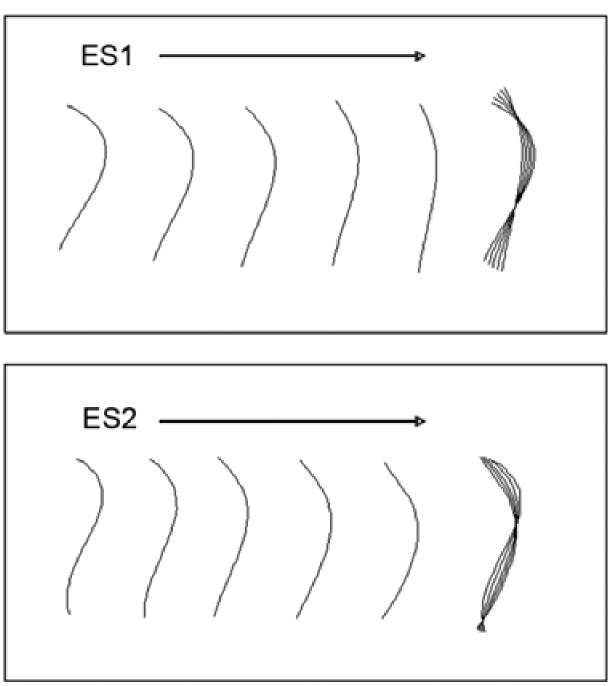

FIG. 8—*Shape models generated from the first two eigenshape axes of a covariance-based eigenshape analysis of the greater sciatic notch. These two axes represent 66% of the shape variation. See text for discussion.*

change occurring at either end of the notch outline (Fig. 8). Furthermore the mean shapes generated from the eigenshape analysis enable the shape difference between males and females to be clearly visualized (Fig. 9). Since the initial eigenshape analysis used a mean-centered shape comparison metric, the Phi functions generated from the decomposition of the covariance matrix were used to calculate mean male and mean female shapes. In comparison to the mean male greater sciatic notch shape, the mean female greater sciatic notch appears shallower in the center, with less steeply sloping edges (see Fig. 9).

Schutkowski considered the discriminatory capacity of the greater sciatic notch angle using a simple method of scoring an individual to be male if the angle was *c.* 90° and female if the angle was greater, or much greater, than 90° (14). The present study adopted a different methodology than Schutkowski (14) to define the greater sciatic notch angle. Coordinate data were extracted from three landmark points, this was used to calculate an angle. Thus, rather than grouping the angles into two discrete categories that reveal little in terms of population shape trends, this study utilized the continuous shape variation present in the population. Based on the landmarks used in this study the mean angle for females was 132.5° and the mean angle for males was 126.0°. All the angle measurements for males produced in this study are greater than 90° reflecting the significantly different methodology used to collect angle measurement data in this study compared to Schutkowski (14). Schutkowski (14) identified 95% males accurately and 71.4% of females using the angle criterion, supported by a p (two-tailed) value of much less than 0.001. The analyses of angle data from this study produced a comparably less significant result; the Watson–Williams test indicated that the mean angle for females was statistically different to that for males, but only to a critical level of $p < 0.1$. Sutter (27) performed a blind test of Schutkowski's technique and achieved an accuracy of 69.2% for males and 78.6% for females. The smaller sample size for this study (27 individuals) compared with that used by Schutkowski (63 individuals) may account for the reduced clarity of separation between the sexes in terms of this character.

The curvature of the iliac crest, as a whole, was shown by Schutkowski (14) to display either a marked S-shape (males) or a faint S-shape (females). Using this qualitative scoring technique Schutkowski (14) achieved 81.2% correct identification of males and 62.1% correct identification of females. This study built upon the geometry inherent in Schutkowski's approach by identifying the shape variation present in both the upper and lower plane separately, as well as viewing the iliac crest as a whole. Out of the

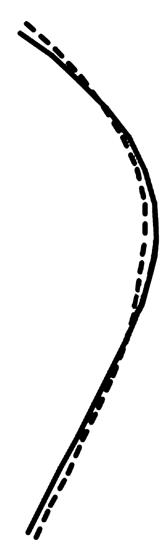

FIG. 9—*Meanshape models of the greater sciatic notch generated from mean-centered eigenshape analysis. Male meanshape represented by a solid line, female meanshape represented by a dashed line. See text for discussion.*

three criteria proposed in this study the lower plane of the iliac crest proved the best discriminator of sex with an accuracy of 88.2% for males, but only 37.5% for females.

Males were identified to a higher level of accuracy (82.4–88.2%) than that achieved by Schutkowski (14) using the three iliac crest criteria. Females, though, were less well identified than by previous comparable studies (14,27). Sutter reported an overall accuracy of 92.9% for females when replicating Schutkowski's qualitative techniques, and this increased to 100% when using individuals aged 2–5 years.

The poor accuracy achieved for curvature of the iliac crest may suggest male and female shape variation in this anatomical region displays a greater amount of overlap than for other variables, for which an accuracy of 62.5–87.5% was obtained. The small female sample size ($n = 8$) may also have contributed to the low resolution, though this could have only acted to limit the identification of clustering since the other criteria displayed a comparatively much higher accuracy for females despite the small sample size. Earlier studies considering dimorphism in the iliac crest using nonmetric criteria were more successful than the results produced here suggesting this trait may be assessed more accurately in this way (14,27).

The results of this study are consistent with previous studies of sex determination using the juvenile ilium that have reported less accurate results for females than for males. Mittler & Sheridan (3) reported female identification accuracy at only slightly better than chance (58.3%) when retesting the technique used by Weaver (8) for identification of auricular surface elevation. Similarly Holcomb & Konigsberg achieved 58.2% accuracy with female identification using the fetal sciatic notch (16). Sutter suggests reduced accuracy for female identification may be due to the expression of sex-related characteristics being associated with differential growth among females (27). Moreover from their study of the adult pelvis, Rogers & Saunders propose that a larger proportion of individuals characterized by an intermediate level of expression are likely to be classified as male because of the lack of female-trait expression (2).

In contrast, male-trait expression commonly leads to poor male identification in forensic cases (28) because whilst large numbers of males are correctly identified as male (true positives), a high quantity of females are falsely identified as male (false positives). The sample studied here exemplifies this trend; when applying Bayes theorem of conditional probability (29) the accuracy of female identification improved for each of the five criteria studied whilst the accuracy of male identification reduced marginally (Table 2). The marked increase in accuracy of female identification reflects a smaller number of false negatives (males falsely identified as females) compared with a relatively higher number of false positives (females falsely identified as males) within the data set. False identifications are included within the probability metric (Table 2) to enable the chance, rather than the incidence, of correct identification to be assessed; this provides a more relevant statistic for forensic casework (27). Moreover, since each criterion exhibits a high level of specificity, as indicated by male identification accuracy of 82.4–100% (Table 2), a forensic caseworker may have more confidence that the shape exhibited by females for each of the criteria reflects the true female shape for that trait.

Results for the data set with younger individuals removed indicate a marked improvement in accuracy. For the curvature of the iliac crest criterion, the shape difference between males and females is so pronounced that there is no error for sex identification with juveniles older than 0.5 years. The improvement of accuracy is most acute for females, with the original sample generating only 25% accuracy. This result may indicate one of several relationships between age and sex.

1. Females may exhibit a shape of curvature analogous to males and subsequently deviate from the male shape after 6 months of age.
2. Males may exhibit a shape of curvature analogous to females and subsequently deviate from the female shape after 6 months of age.
3. Males and females may share an original shape up to the age of 6 months, after which both develop a new shape, different from one another and distinct from the original.

The scope of this study did not enable these proposed explanations to be assessed fully, though the eigenshape space plot excluding the younger individuals (Fig. 5*b*) appears to indicate a possible "original" shape from which both males and females diverge on separate trajectories. The use of qualitative methodology (14,27) does not allow for age variation to be considered in terms of shape, other than to deduce that the individual becomes "more" or "less" analogous to the criterion defined at the beginning of the study. Indeed, the clarity synonymous with quantitative measurement enables the generation of hypotheses providing a more detailed reflection of relationships between sex, age, and shape than is possible with qualitative techniques.

Sutter also assessed the affect of age on the accuracy of sex identification techniques, though not using any form of shape interpretation, and found the correct identification of males improved from 0–50% for newborn to 1 year olds, to 71–100% for individuals 2–5 years old (27). The correct identification of females also improved, although less markedly, from 71–86% in individuals aged less than 1 year to 86–100% for 2–5 year olds (27). Weaver (8) proposed that discrimination of sex diagnostic characteristics would be most obvious at the fetal stage, typically from the 15th week onwards, with the appearance of fetal testosterone and the resultant onset of major sexual differentiation. Moreover, Weaver argued sex diagnostic characteristics would display a decreased amount of differentiation later in development as they were masked by growth (8). The criteria studied here indicate the opposite to be true. Only female identification using the upper and lower planes of curvature of the iliac appears to be relatively static in terms of accuracy with the exclusion of younger individuals, indicating a lack of sensitivity to sample age.

Summary

The results presented here indicate the greater sciatic notch shape criterion to be most useful for sex determination of juvenile material. Given the accuracy achieved with the greater sciatic notch shape criterion, compared to the other regions studied, the authors recommend this criterion should be used to assess juvenile sex. The curvature of the iliac crest criterion displays a relationship with sample age such that "typical male" shape and "typical female" shape diverge after 0.5 years of age. The two shapes are so markedly different from one another that there is no error for sex identification with samples older than 0.5 years. Sample age also appears directly related to identification accuracy with three of five criteria displaying improved identification accuracy with increased age. Previous hypotheses regarding the use of the greater sciatic notch angle for sex determination (14) are supported by the statistically significant difference between male and female sciatic notch angle measurements produced in this study.

In view of the morphological and metric variation that exists between and within human populations (30), caution should be applied when attempting to use the techniques discussed in this study on a different population. The discriminant functions and accuracy calculations generated in this study are unique, and thus inherently limited, to the population of Christ Church, Spitalfields. The expression of skeletal indicators of sex is influenced by the age of onset and the quantity of androgen production (27). These two factors are, in turn, influenced by a variety of factors including genetic variation and fetal environment. As a result techniques developed using one population may not always be applied successfully to a different population. Nevertheless, the results obtained in this study illustrate the potential for improvement upon previous more qualitative and indirect approaches to sex determination of juvenile bones that can be achieved by adopting a more objective, geometric approach. If developed further such an approach could prove highly useful for palaeo-demographic studies and the identification of children in forensic cases. Further exploration of these techniques using a different population would allow the general applicability of the methods used in this study to be evaluated. The use of a larger sample size, with a greater age distribution, may also enable the further investigation of several interesting relationships with age that have emerged from the study of this limited sample set, in particular those associated with the curvature of the iliac crest.

Acknowledgments

We thank Jonathan Krieger and Graeme Smith for valuable suggestions and thoughtful discussions during the preparation of this manuscript and two anonymous reviewers, whose useful comments helped improve this paper.

References

1. MacLaughlin SM, Bruce MF. The sciatic notch/acetabular index as a discriminator of sex in European skeletal remains. J Forensic Sci 1986;31:1380–90.
2. Rogers T, Saunders S. Accuracy of sex determination using morphological traits of the human pelvis. J Forensic Med 1994;39(4):1047–56.
3. Scheuer L, Black S. Developmental juvenile osteology. London: Academic Press, 2000.
4. Humphrey LT. Patterns of growth in the modern human skeleton. Am J Phys Anthropol 1998;105:57–72.
5. Mittler DM, Sheridan SG. Sex determination in subadults using auricular surface morphology: a forensic science perspective. J Forensic Sci 1992;37(4):1068–75.
6. Washburn SL. Sex differences in the pubic bone. Am J Phys Anthropol 1948;6:199–207.
7. Phenice TW. A newly developed visual method of sexing the os pubis. Am J Phys Anthropol 1969;30:297–302.
8. Weaver DS. Sex differences in the ilia of a known sex and age sample of fetal and infant skeletons. Am J Phys Anthropol 1980;52(2):191–5.
9. MacLaughlin SM, Bruce MF. Morphological sexing of the os pubis. An anatomical approach. Am J Phys Anthropol 1980;81:260–1.
10. Hunt DR. Sex determination in the subadult ilia: an indirect test of Weavers non-metric sexing method. J Forensic Sci 1990;35:881–5.
11. Cox M, Mays S, editors. Human osteology in archaeology and forensic science. London: Greenwich Medical Media Ltd, 2000;118.
12. White TD, Folkens PA. Human osteology. 2nd ed. San Diego, CA: Elsevier Science, 2000;365.
13. Boucher BJ. Sex differences in the foetal pelvis. Am J Phys Anthropol 1957;15:581–600.
14. Schutkowski H. Sex determination of infant and juvenile skeletons: I. Morphognostic features. Am J Phys Anthropol 1993;90(2):199–205.
15. Krogman WM, Iscan MY. The human skeleton in forensic medicine. Springfield, IL: Charles C Thomas, 1986.
16. Holcomb SMC, Konigsberg LW. Statistical study of sexual dimorphism in the human fetal sciatic notch. Am J Phys Anthropol 1995;97:113–25.
17. Bookstein FL. Morphometric tools for landmark data: geometry and biology. Cambridge: Cambridge University Press, 1991;434–5.
18. Lohmann GP. Eigenshape analysis of microfossils: a general morphometric method for describing changes in shape. Math Geol 1983;15:659–72.
19. Lohmann GP, Schweitzer PN. On eigenshape analysis. In: Rohlf FJ, Bookstein FL, editors. Proceedings of the Michigan morphometrics workshop. Ann Arbor, MI: The University of Michigan Museum of Zoology, Special Publication No. 2,1990;145–66.
20. MacLeod N. Generalizing and extending the eigenshape method of shape space visualization and analysis. Paleobiology 1999;25(1):107–38.
21. MacLeod N. Geometric morphometrics and geological form-classification systems. Earth Sci Rev 2002;59:27–47.
22. Molleson T, Cox M, Waldron AH, Whittaker DH. The spitalfields project volume 2—the anthropology, the middling sort. Research Report 86. York: Council for British Archaeology, 1993.
23. Zahn CT, Roskies RZ. Fourier descriptors for plane closed curves. IEEE Trans Comput 1972;21:269–81.
24. Zar JH. Biostatistical analysis. Upper Saddle River, NJ: Prentice Hall, 1974;101–17, 310–26.
25. Good P. Resampling methods. 3rd ed. Boston: Birkheuser, 2006.
26. France DL. Observational and metric analysis of sex in the skeleton. In: Reichs KJ, editor. Forensic osteology: advances in the identification of human remains. Springfield, IL: Charles C Thomas, 1998;164.
27. Sutter RC. Nonmetric subadult skeletal sexing traits: I. A blind test of the accuracy of eight previously proposed methods using prehistoric known-sex mummies from Northern Chile. J Forensic Sci 2003;48(5):1–9.
28. Klepinger L, Giles E. Clarification or confusion: statistical interpretation in forensic anthropology. In: Reichs KJ, editors. Forensic osteology: advances in the identification of human remains. Springfield, IL: Charles C Thomas, 1998;435–9.
29. Bayes T. An essay towards solving a problem in the doctrine of chances. Biometrika 1958;45:296–315.
30. MacLaughlin SM, Bruce MF. A simple univariate technique for determining sex from fragmentary femora: its application to a Scottish short cist population. Am J Phys Anthropol 1985;67:413–7.

Additional information and reprint requests:
Laura Wilson, B.Sc.
Universität Zürich
Palaontologisches Institut und Museum
Karl Schmid-Strasse 4
CH-8006 Zürich
Switzerland
E-mail: laura.wilson@pim.uzh.ch