JOURNAL OF FORENSIC SCIENCES

J Forensic Sci, 2017
doi: 10.1111/1556-4029.13332
Available online at: onlinelibrary.wiley.com

TECHNICAL NOTE

ANTHROPOLOGY

Alexandra R. Klales,[1] Ph.D.; and Tesa L. Burns,[2] B.S.

Adapting and Applying the Phenice (1969) Adult Morphological Sex Estimation Technique to Subadults*

ABSTRACT: This research evaluated whether adult morphological sex estimation methods of the innominate could be adapted and applied to subadults. The subpubic concavity, described by Phenice (1969) and revised by Klales et al. (2012), was modified for use with subadults. Two observers scored radiographic images from the PATRICIA database of 334 individuals of both sexes aged between 1.19 and 20.47 years. Score frequencies shifted from score 2 (straight) to higher frequencies of score 3 (convexity) in males and score 1 (concavity) in females with increasing age. Using ordinal logistic regression, sex classification was highest for the oldest age cohort at 97.2% and then decreased by age cohort. Interobserver error rendered a high level of agreement (0.806) using the intraclass correlation coefficient. Results indicate that the Klales method can be modified and applied to subadults to accurately estimate sex following the onset of puberty with a high degree of reliability and validity.

KEYWORDS: forensic science, forensic anthropology, pelvic nonmetrics, biological profile, sex estimation, subadults, Phenice

One of the primary goals of physical and forensic anthropologists is the estimation of a person's biological profile to determine the identity of unknown human remains. Sex estimation is an important component of the biological profile, and many methods have been developed, tested, and utilized using different regions of the human skeleton. Generally, methods of adult sex estimation using the pelvis, long bones, and skull are well accepted and utilized within physical and forensic anthropology. Sex classification in adults using morphological traits, metric measurements, and geometric morphometrics of the pelvis has been high, which is likely why the pelvis is considered by most to be the best indicator for sex estimation (1).

Phenice (2) devised a method of sexing the adult pelvis, specifically the pubis bone, using three traits: the ventral arc, the subpubic concavity, and the medial aspect of the ischio-pubic ramus. The high classification rate reported by Phenice (2), over 95%, likely contributes to why this method remains the most popular morphological sex estimation method used today (1). In light of Daubert (3), Klales and colleagues (4) revised the Phenice (2) method to include a standardized ordinal recording scheme and a logistic regression model to predict the probability of sex membership for an unknown individual based on pubic trait scores. Mean classification accuracy of the method is 94.5%. In validation studies of the Klales et al. (4) revised method, accuracy rates range from 86.2% to 99.2% (4–7) depending on the sample. Metric studies using interlandmark distances reported even higher classification than morphological methods [e.g., 99.0% combined percent correct in Klales et al. (8) and 97.5% in Baumgarten and Ousley (9)]. Geometric morphometric analyses examining shape also produced equally high sex classification accuracy between adult males and females [e.g., 98.5% combined percent correct by Bytheway and Ross (10) and by Vollner and colleagues (11)].

While there are several reliable and valid methods for adult sex estimation using the pelvis, the same is not true for subadults. In fact, DiGangi and Moore (12:106) note, "sub-adult sex estimation has been perhaps the most problematic area of sex estimation." To date, there are no generally agreed upon methods within the field to accurately sex unknown subadult individuals. Further, the Scientific Working Group for Forensic Anthropology describes sex estimation for individuals under 12 years of age as an unacceptable practice. Because many sexually dimorphic parts of the human skeleton do not show a difference between males and females until after puberty, sexing subadult individuals remains challenging. New research on subadult sexing is further hindered by a lack of large, known subadult skeletal collections.

Developmentally speaking, the innominate is composed of three bones, the ilium, the ischium, and the pubis, which begin to fuse around the age of four (13). The ischium and the pubis fuse first, between four and eight years of age, and then, the ilium fuses to the other two bones between 11 and 14 years of age in females and between 14 and 17 years of age in males (13). Scheuer and Black (14) suggest a slightly later fusion of the ischio-pubic ramus at between five and eight years of age,

---

[1] Sociology & Anthropology, Washburn University, 1700 SW College Ave, Topeka KS 66621.

[2] Department of Archaeology, Dickinson College, PO Box 1773, Carlisle, PA 17013.

*Presented at the 67th Annual Meeting of the American Academy of Forensic Sciences, February 16–21, 2015, in Orlando, FL.

Received 6 April 2016; and in revised form 6 Aug. 2016; accepted 20 Aug. 2016.

while also suggesting that commencement of acetabular fusion, where the three bones fuse together, can extend in females until the age of 15 years. As the bones of the pelvis develop, the greatest sex difference is in the subpubic concavity/contour as the pelvic inlet begins to enlarge in females (15). Rapid hormonal changes are responsible for this enlargement during the pubertal growth spurt (16). Very generally speaking, puberty begins between the ages of 10 and 14 years, with onset often occurring earlier in females than in males; however, the onset of puberty is extremely variable between populations and between individuals (17). Most growth leading to sexual dimorphism takes place during adolescence, when males show greatest growth in the acetabulum and females show greatest growth and change in the pelvic cavity, the length of the pubis, and the development of the subpubic concavity (15,18). Because the sexually dimorphic features of the pubis do not fully develop until puberty, sexing the pelvis in prepubescent individuals is difficult, if not impossible.

Previous studies have addressed the sexing of the subadult pelvis using different morphological and metric methods (e.g., 19–25) (Table 1). Currently, morphological traits have proven better than metric methods in subadult sexing and in many past studies, male subadults have classified better than females (13). Weaver (19) evaluated six metric traits, converted into three indices, and one nonmetric trait of the ilia in fetal (from 6 months *in utero*) and infant (birth to six months) remains. No significant differences between the sexes were found in the metric measurements; however, the one nonmetric trait, auricular surface elevation, showed more promising results. Males classified more accurately using this method (fetal 91.7%; newborn 73.1%; six months 90.6%). Females showed less reliable classification (75.0%, 54.2%, and 43.5% in the fetal, newborn, and six months age groups, respectively), suggesting a bias in sex classification. Two studies by Hunt (20) and Mittler and Sheridan (21) subsequently tested Weaver's method. Hunt (20), using an unknown sample, compared the sex ratio of his sample to that found by Weaver. Hunt's results were biased toward the raised auricular surface trait (5.6 times more prevalent in his sample), which suggests that auricular surface elevation should not be considered an accurate indicator of sex in subadults in all populations. Mittler and Sheridan (21) also tested the usefulness of auricular surface elevation in sex estimation of subadults, and the authors found results similar to Weaver's (19) original study: 85.3% accuracy for males and 58.3% accuracy for females. They note that the accuracy of the method improved with increasing age and that the method had a strong male bias, particularly in younger individuals. Based on their validation of the trait, Mittler and Sheridan (21) suggest that the method is only an accurate predictor of sex for forensic contexts in individuals older than nine years of age.

Schutkowski (22) examined both the ilium and the mandible of subadult individuals from birth to 5 years of age. On the ilium, the angle and depth of the sciatic notch, the iliac crest, and arch criterion were evaluated. The angle of the sciatic notch proved most reliable of the methods, correctly classifying 95.0% of males and 71.4% of females. Holcomb and Konigsberg (23) focused on fetal remains and used morphometric techniques to study the shape of the greater sciatic notch for sexual dimorphism. The authors found that, although there is no significant difference in sciatic notch depth between male and female fetuses, the "anterior to posterior location of the maximum depth of the sciatic notch" does show dimorphism (23:122). More recently, Cardoso and Saunders (24) examined the arch criterion of the ilium in subadults. While this feature is considered by the authors as "somewhat successful" for sex estimation in adults, classification accuracy was poor (26.7–52.6%) when applied to subadults (24). Agreement within and between researchers was also found to be poor, indicating that the method is invalid and unreliable for sexing subadult remains at this time. Wilson et al. (25) analyzed geometric measurements of the sciatic notch, iliac crest, and auricular surface of subadult remains and were able to estimate sex with 96% accuracy (in a sample of juveniles from birth to 7.88 years) using the greater sciatic notch shape.

Although all of these studies utilized the subadult pelvis, most of them focused on the ilium and few have attempted to use the pubic bone. Because most studies to date have ignored the pubic bone, the aim of this research was to test the applicability of the subpubic concavity/contour, used by Phenice (2) and Klales et al. (4) in adults, for sex estimation in subadult individuals. Phenice (2:300) cautions that the subpubic concavity is "not well developed until the female has reached about 20 years of age" and suggests that his method and the traits included should be limited to adult material until further tests have been conducted in subadults. To the current knowledge of the authors, neither method has been applied to sex estimation of subadult pelves. A preliminary study by the primary author using subadult pubic bones from the Hamann-Todd Osteological Collection has shown that it is possible to use a modified version of the Klales et al. (4) method to estimate sex in subadult age categories younger than previously believed. The results from this preliminary study were encouraging; however, a larger sample size was needed for the results to be statistically significant. The present research examines the subadult pubis using a revision of the

TABLE 1—*Previous studies evaluating the use of the pelvis in subadult sex estimation.*

Author(s)	Year	Innominate Region	Accuracy
Weaver	1980	Auricular surface elevation (ilium)	90.6% males 43.5% females
Mittler and Sheridan	1992	Auricular surface (ilium)	85.3% males 58.3% females
Schutkowski	1993	Sciatic notch, iliac crest, arch criterion (ilium)	95.0% males 71.4% females (sciatic notch)
Holcomb and Konigsberg	1996	Sciatic notch (ilium)	67.2% males 58.2% females
Cardoso and Saunders	2008	Arch criterion (ilium)	39.9% males 45.4% females
Wilson, MacLeod, and Humphrey	2008	Sciatic notch, iliac crest, and auricular surface (ilium)	100% males 87.5% females (sciatic notch shape)

Klales et al. (4) subpubic contour scoring. The first goal of the study was to determine whether this revised technique can be used to differentiate sexes in subadult remains. The second goal was to evaluate the timing or onset of sexual dimorphism in the subpubic contour. The third goal of the study was to test the level of observer agreement for trait scoring.

Materials and Methods

Sample and Scoring

The sample used in this research was derived from the PATRICIA Radiographic Data Bank (26). The database contains radiographs from a "geographically and ethnically diverse" sample of modern American subadults (born after 1990) with known demographic information collected from various coroner and medical examiner offices throughout the United States (26:183). A query was run in PATRICIA to include radiographs with specific features: *Quality* (fair, good, or very good), *Orientation* (anterior–posterior), and *Body Areas Represented* (left and right ilium, ischium, and pubis).

For this research, a total of 334 individuals of both sexes were included (Table 2). Individuals in the sample ranged in age from 1.19 to 20.47 years. The sample was divided into six age cohorts that were slightly modified versions of the age categories presented in Baker et al. (13): Young Child Early (1.0–3.5 years), Young Child Late (3.6–6.5 years), Older Child Early (6.6–9.5 years), Older Child Late (9.6–12.5 years), Adolescent Early (12.6–15.5 years), and Adolescent Late (15.6–20.5 years).

Two observers, one experienced observer (ARK) and one upper level anthropology student (TLB), scored each radiograph using a revision of the Klales et al. (4) adult subpubic contour scale/figures. For the present research, the original adult SPC scores were reduced from five ordinal scores to three (Fig. 1). Score 1 (pronounced SPC/concavity) for the subadults corresponds closely to score 1 in the original adult method devised by Klales et al. (4). Score 2 (straight SPC) for the subadults corresponds to score 3 in the original adult method devised by Klales et al. (4). Score 3 (convex SPC) for the subadults is a modified version of score 5 from the original adult method devised by Klales et al. (4). Radiograph examples of each score are presented in Fig. 2.

Prior to scoring, the radiographs were anonymized by creating a new unique identification number for each of the radiograph

TABLE 2—*Age cohorts and sample size used in the current research. Age cohorts were modified from Baker et al. (13).*

Age Cohort	Years	Males (*n*)	Females (*n*)
Young Child Early	1.0–3.5	29	26
Young Child Late	3.6–6.5	33	22
Older Child Early	6.6–9.5	18	15
Older Child Late	9.6–12.5	30	24
Adolescent Early	12.6–15.5	35	31
Adolescent Late	15.6–20.5	40	31

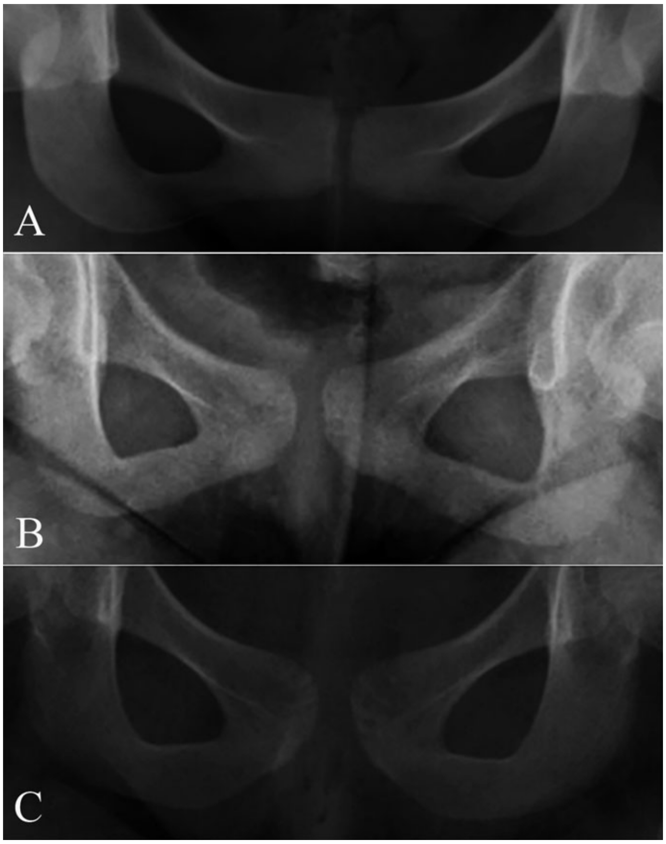

FIG. 2—*Radiographic examples of subpubic contour scores: A) score 1, concavity; B) score 2, straight; and C) score 3, convexity.*

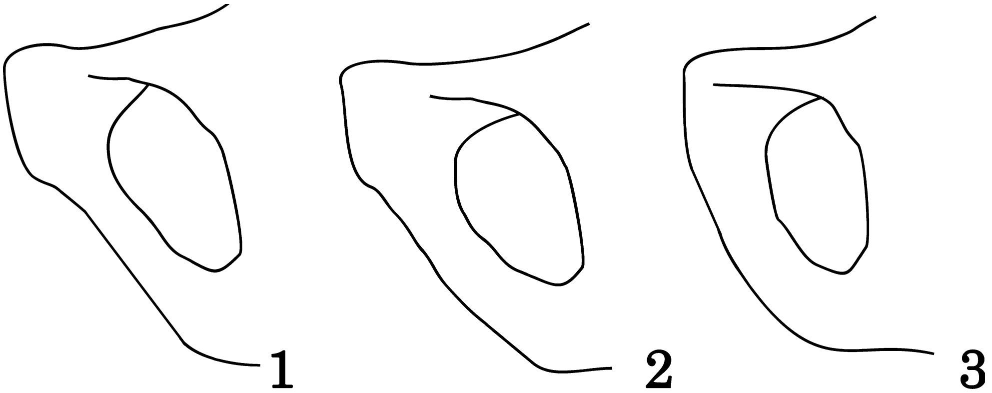

FIG. 1—*Subadult ordinal scores used in the current research. Modified from Klales et al. (4).*

file links provided online in PATRICIA. The links were then used to directly view the radiographs (without any demographic data) in random order, so that both of the two observers could blindly score each individual. In some cases, external genitalia were visible in the original radiographs. These individuals were included in the sample, and the images were enlarged in an attempt to exclude the genitalia from view when possible prior to scoring. The radiographs were used “as is” with no manipulation of contrast and brightness; however, some images were enlarged to view the area of interest. Following scoring, the unique identification numbers associated with the file links were linked with the known demographic and scan data for later analyses.

Statistical Methods

All statistical analyses were performed in SPSS (27) and followed the same methodology used in the original adult sex estimation article by Klales and colleagues (4). First, trait frequency distributions for each ordinal score by sex were calculated for each of the six age cohorts. Next, sex classification accuracy using SPC ordinal scores was tested using ordinal logistic regression (OLR). Lastly, the intraclass correlation coefficient (ICC) was utilized to test the degree of agreement for trait scoring between the experienced and inexperienced observer.

Results

Score frequencies varied by age cohort (Fig. 3). In the Young Child cohorts, score 2 (straight) was most prevalent for both males and females, followed by score 3 (convexity) then score 1 (concavity). In the Older Child cohorts, all males were scored as either a score 3 (most prevalent) or score 2. In females of the Older Child cohorts, score 2 was the most prevalent; however, scores were more variable than for the males in the same cohorts. In the Adolescent cohorts, score frequencies clearly differed by sex, with females predominantly being scored as a score 1 (pronounced concavity), while males were predominantly scored as score 3 (convexity present) (Fig. 3). As age increased after the Young Child Late phase (6.5 years), the frequency of score 2 continuously decreased for both males and females. Using ordinal logistic regression, sex classification accuracy was highest for the oldest age cohort (Adolescent Late) at 97.2% combined correct and then decreased in order of age cohort (Table 3). The ICC (0.806) indicated a high level of agreement in scoring between the two observers.

Discussion

While 85% accuracy has generally been considered the minimum standard for adult sex estimation methods, DiGangi and Moore (12:107) suggest “the goal for accuracy should be at least 75% (which is 50% better than chance)” for subadult sex classification. Most studies examining the subadult innominate have failed to meet that criterion, especially for fetal and infant age groups. Based on the 75% standard, only the Adolescent Early (12.6–15.5 years) and Adolescent Late (15.6–20.5 years) cohorts in this study classified accurately enough for use in subadult sex

TABLE 3—*Sex classification accuracy (%) for each age cohort using ordinal logistic regression.*

Age Cohort	Males	Females	Combined	Sex Bias
Young Child Early	100.0	7.7	53.9	92.3
Young Child Late	100.0	18.2	59.1	81.8
Older Child Early	55.6	73.3	64.5	−17.7
Older Child Late	76.6	66.7	71.7	9.9
Adolescent Early	77.1	93.5	85.3	−16.4
Adolescent Late	97.5	96.8	97.2	0.7

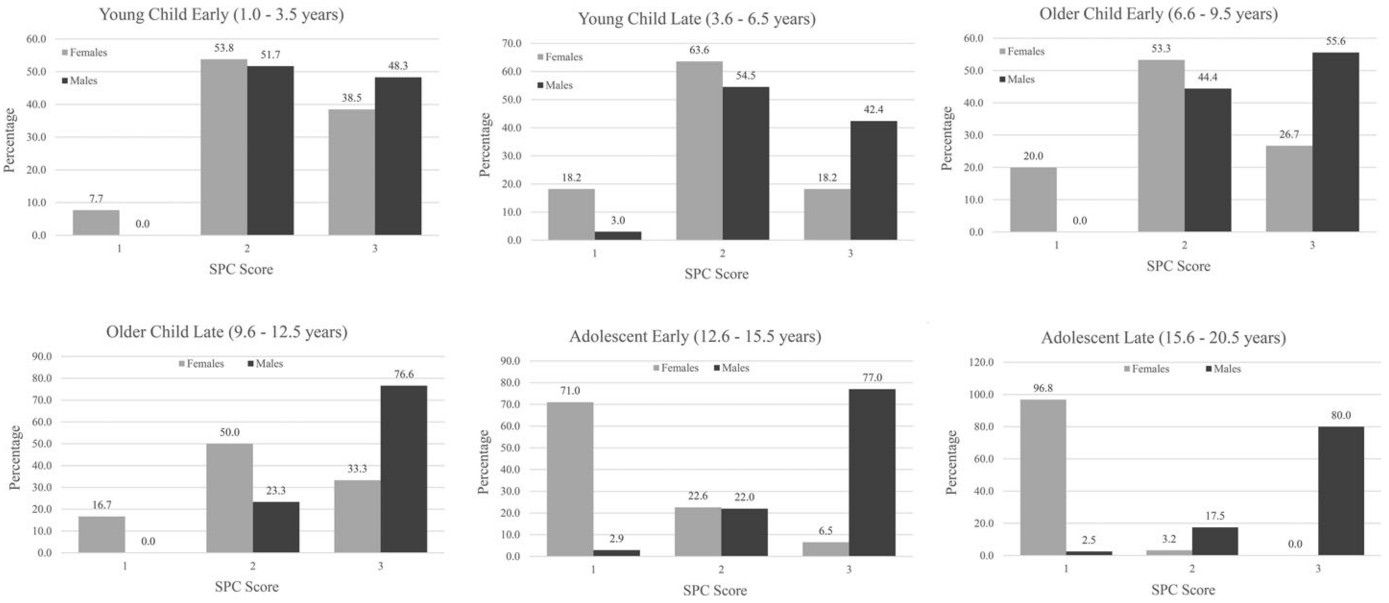

FIG. 3—*Score frequencies by sex for each age cohort.*

estimation (Table 3); however, sex bias was fairly high (−16.4%) for the Adolescent Early cohort. The Adolescent Late cohort showed the best classification accuracy, with only one specimen from each sex being classified incorrectly using OLR (97.2% accuracy), closely followed by Adolescent Early (85.3%). These results show that this method can be used with confidence to classify individuals aged from 12.6 to 20.5 years, which is considerably earlier than previously stated by Phenice (2). In the Older Childhood Late (9.6–12.5 years) age cohort, males classified (76.6%) above the standard but a male bias resulted in a combined sex classification accuracy just below (71.7%) the 75% standard. Based on these findings, this method for subadult sex estimation can only be reliably used in adolescents.

Conversely, the method proved invalid for the younger cohorts and is generally consistent with previous literature suggesting that sex estimation fails to produce accurate results prior to puberty. Combined sex results for the Young Child Early and Young Child Late age cohorts were only slightly better than chance (53.9% and 59.1%, respectively), and there was a significant male bias in classification (92.3% and 81.8%, respectively). Almost all of the individuals from these age groups classified as male (100% of males and 94.3% of females in the Young Child Early cohort and 100% of males and 81.5% of females in the Young Child Late cohort). Sex differentiation improved with age in the Older Child age cohorts, but was only slightly higher than chance and were also not at or above the 75% threshold suggested by DiGangi and Moore (12). For these ages, results show a male bias due to the differential development of the pelvis in males and females. The female pelvis diverges from the male form during puberty as they begin to adapt to childbirth (12). Scheuer and Black (14) suggest that only after this development can accuracy be achieved using the pelvis: “from around mid-puberty, the secondary sexual differentiation of the pelvis is probably sufficiently advanced in females to permit reliable estimation of that sex” (14:343).

As Scheuer and Black (14:16) caution, “any new method for sex determination needs a rigorous standard against which to test its validity and reliability and this can only be achieved on a sample of known biological identity.” By quantifying and assigning ordinal scores to morphological traits and then analyzing those using statistical methods, it is possible to test the validity and reliability of this method in accordance with the *Daubert* criteria (3,4). This study has shown validity, or accuracy, in the Adolescent cohorts. It also tested reliability (repeatability or observer agreement) using the intraclass correlation coefficient. These statistical results are important to consider in the application of sex estimation techniques in forensic contexts. Because the sample used for this study was taken from a pool of radiographs of individuals of known sex and age, the method was testable against the known sample. However, DiGangi and Moore (12:107) suggest that the use of radiographs may be problematic, as they “do not necessarily compare well with dry bones.” Future research, using perhaps a larger sample size from a skeletal collection, would solidify the usefulness and reliability of this method for use by forensic anthropologists.

Acknowledgments

Thanks go to Dr. Stephen Ousley, Dr. Kyra Stull, and Katie Frazee for their work creating the PATRICIA Database and making it publically accessible for research purposes. Thanks also go to the anonymous reviewers for their suggestions and comments that improved the manuscript.

References

1. Klales AR. Current practices in forensic anthropology for sex estimation in unidentified, adult individuals. Proceedings of the 65th Annual Meeting of the American Academy of Forensic Sciences; 2013 Feb 18–23; Washington, DC. Colorado Springs, CO: American Academy of Forensic Sciences, 2013;19(H81):439–40.
2. Phenice TW. A newly developed method of sexing the os pubis. Am J Phys Anthropol 1969;30:297–302.
3. Daubert V. Merrell Dow Pharmaceuticals. US Supreme Court 509.U.S.579,113S.Ct.2786, 125L. Ed.2d 469, 1993.
4. 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.
5. Kenyhercz MK. Sex estimation using pubic bone morphology in a modern South African sample: a test of the Klales et al. method. Am J Phys Anthropol 2012;147(S54):179–80.
6. Kenyhercz MW, Fredette SM, Klales AR, Dirkmaat DC. Metric and non-metric assessment of sex: accuracy, correlation and corroboration. Proceedings of the 64th Annual Meeting of the American Academy of Forensic Sciences; 2012 Feb 20–25; Atlanta, GA. Colorado Springs, CO: American Academy of Forensic Sciences, 2012;18(H85):406–7.
7. Stull KE, Kenyhercz MW, L’Abbe EN. Non-metric cranial and pelvic traits as a measure of sexual dimorphism in a modern South African population. Am J Phys Anthropol 2013;150(S56):266.
8. Klales AR, Vollner JM, Ousley SD. A new metric procedure for the estimation of sex and ancestry from the human innominate. Proceedings of the 61st Annual Meeting of the American Academy of Forensic Sciences; 2009 Feb 16–20; Denver, CO. Colorado Springs, CO: American Academy of Forensic Sciences, 2009;15:311.
9. Baumgarten S, Ousley SD. Estimating sex from the innominate utilizing new measurements. Proceedings of the 67th Annual Meeting of the American Academy of Forensic Sciences; 2015 Feb 16–21; Orlando. FL. Colorado Springs, CO: American Academy of Forensic Sciences, 2015;21(A73):142.
10. Bytheway JA, Ross AH. A geometric morphometric approach to sex determination of the human adult os coxa. J Forensic Sci 2010;55:859–64.
11. Vollner JM, Klales AR, Ousley SD. Geometric morphometric analysis of the human innominate for sex estimation. Am J Phys Anthropol 2010;141(S50):237.
12. DiGangi EA, Moore MK, editors. Research methods in human skeletal biology. Oxford, U.K.: Elsevier Academic Press, 2013.
13. Baker BJ, Tocheri MW, Dupras TL. The osteology of infants and children. College Station, TX: Texas A&M University Press, 2005.
14. Scheuer L, Black S. Developmental juvenile osteology. New York, NY: Elsevier Academic Press, 2000.
15. Coleman WH. Sex differences in the growth of the human bony pelvis. Am J Phys Anthropol 1969;31(2):125–52.
16. Greulich WW, Thomas H. The growth and development of the pelvis of individual girls before, during, and after puberty. Yale J Biol Med 1944;17(1):91–8.
17. Stanford C, Allen JS, Antón SC. Biological anthropology, 3rd rev. edn. Boston, MA: Pearson, 2013.
18. LaVelle M. Natural selection and developmental sexual variation in the human pelvis. Am J Phys Anthropol 1995;98(1):59–72.
19. 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.
20. Hunt DR. Sex determination in the subadult ilia: an indirect test of Weaver’s nonmetric sexing method. J Forensic Sci 1990;35(4):881–5.
21. Mittler DM, Sheridan SG. Sex determination in subadults using auricular surface morphology: a forensic science perspective. J Forensic Sci 1992;37(4):1068–75.
22. Schutkowski H. Sex determination of infant and juvenile skeletons: I. Morphognostic features. Am J Phys Anthropol 1993;90(2):199–205.
23. Holcomb SMC, Konigsberg LW. Statistical study of sexual dimorphism in the human fetal sciatic notch. Am J Phys Anthropol 1996;97(2):113–25.
24. Cardoso HFV, Saunders SR. Two arch criteria of the ilium for sex determination of immature skeletal remains: a test of their accuracy and an assessment of intra- and inter-observer error. Forensic Sci Int 2008;178:24–9.

25. Wilson LA, MacLeod N, Humphrey LT. Morphometric criteria for sexing juvenile human skeletons using the ilium. J Forensic Sci 2008;53(2):269–78.
26. Ousley SD, Frazee K, Stull K. Assessing bone growth and development in modern American children. Proceedings of the 79th Meeting of the American Association for Physical Anthropology; 2010 Apr 14–17; Albuquerque, NM. Am J Phys Anthropol 2010;141(S50):183.
27. SPSS. SPSS for Windows, Rel. 14.0. Chicago, IL: SPSS Inc., 2005.

Additional information and reprint requests
Alexandra R. Klales, Ph.D.
Sociology & Anthropology HLRC #218
Washburn University
1700 SW College Avenue
Topeka, KS 66621
E-mail: alexandra.klales@gmail.com