TECHNICAL NOTE

ANTHROPOLOGY

Alexandra R. Klales, Ph.D.; and Tesa L. Burns, B.S.

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

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 pelvic 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, respectively. While there is some variation in this fusion, the ilium typically fuses first, followed by the ischium, and the pubis last.

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

1Sociology & Anthropology, Washburn University, 1700 SW College Ave, Topeka KS 66621.
2Department of Archaeology, Dickinson College, PO Box 1773, Carlisle, PA 17013.
3Presented 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.
TABLE 1—Previous studies evaluating the use of the pelvis in subadult sex estimation.

<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Year</th>
<th>Innominate Region</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weaver</td>
<td>1980</td>
<td>Auricular surface elevation (ilium)</td>
<td>90.6% males</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>43.5% females</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>85.3% males</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>58.3% females</td>
</tr>
<tr>
<td>Mittler and Sheridan</td>
<td>1992</td>
<td>Auricular surface (ilium)</td>
<td>95.0% males</td>
</tr>
<tr>
<td>Schutkowski</td>
<td>1993</td>
<td>Sciatic notch, iliac crest, arch criterion (ilium)</td>
<td>71.4% females (sciatic notch)</td>
</tr>
<tr>
<td>Holcomb and Konigsberg</td>
<td>1996</td>
<td>Sciatic notch (ilium)</td>
<td>67.2% males</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>58.2% females</td>
</tr>
<tr>
<td>Cardoso and Saunders</td>
<td>2008</td>
<td>Arch criterion (ilium)</td>
<td>39.9% males</td>
</tr>
<tr>
<td>Wilson, MacLeod, and Humphrey</td>
<td>2008</td>
<td>Sciatic notch, iliac crest, and auricular surface (ilium)</td>
<td>100% males</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>87.5% females (sciatic notch shape)</td>
</tr>
</tbody>
</table>
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>
<thead>
<tr>
<th>Age Cohort</th>
<th>Years</th>
<th>Males (n)</th>
<th>Females (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Young Child Early</td>
<td>1.0–3.5</td>
<td>29</td>
<td>26</td>
</tr>
<tr>
<td>Young Child Late</td>
<td>3.6–6.5</td>
<td>33</td>
<td>22</td>
</tr>
<tr>
<td>Older Child Early</td>
<td>6.6–9.5</td>
<td>18</td>
<td>15</td>
</tr>
<tr>
<td>Older Child Late</td>
<td>9.6–12.5</td>
<td>30</td>
<td>24</td>
</tr>
<tr>
<td>Adolescent Early</td>
<td>12.6–15.5</td>
<td>35</td>
<td>31</td>
</tr>
<tr>
<td>Adolescent Late</td>
<td>15.6–20.5</td>
<td>40</td>
<td>31</td>
</tr>
</tbody>
</table>

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

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

<table>
<thead>
<tr>
<th>Age Cohort</th>
<th>Males</th>
<th>Females</th>
<th>Combined</th>
<th>Sex Bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Young Child Early</td>
<td>100.0</td>
<td>7.7</td>
<td>53.9</td>
<td>92.3</td>
</tr>
<tr>
<td>Young Child Late</td>
<td>100.0</td>
<td>18.2</td>
<td>59.1</td>
<td>81.8</td>
</tr>
<tr>
<td>Older Child Early</td>
<td>55.6</td>
<td>73.3</td>
<td>64.5</td>
<td>-17.7</td>
</tr>
<tr>
<td>Older Child Late</td>
<td>76.6</td>
<td>66.7</td>
<td>71.7</td>
<td>9.9</td>
</tr>
<tr>
<td>Adolescent Early</td>
<td>77.1</td>
<td>93.5</td>
<td>85.3</td>
<td>-16.4</td>
</tr>
<tr>
<td>Adolescent Late</td>
<td>97.5</td>
<td>96.8</td>
<td>97.2</td>
<td>0.7</td>
</tr>
</tbody>
</table>

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


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