A8 The Accuracy and Reliability of the Klales et al. Non-Metric Pelvic Sexing Method

Kate M. Lesciotto, JD, MS, Pennsylvania State University, Dept of Anthropology, 409 Carpenter Bldg, State College, PA 16802; and Lily J. Doershuk, BS, Pennsylvania State University, Dept of Anthropology, 409 Carpenter Bldg, State College, PA 16802*

After attending this presentation, attendees will better understand the accuracy and inter-observer and intra-observer error rates for the Klales et al. method of sex estimation utilizing three non-metric traits of the pelvis.[1]

This presentation will impact the forensic science community by contributing an external validation study on the accuracy and reliability of using pelvic non-metric traits to estimate sex for unknown individuals.

When attempting to estimate sex from a set of unknown skeletal remains, the pelvis is considered to be the most reliable element. In 1969, Phenice created a now widely used method that evaluates three nonmetric traits of the pelvis: the ventral arc, subpubic concavity, and medial aspect of the ischiopubic ramus.[1] In 2012, Klales et al. created an ordinal scoring method for the Phenice traits in an attempt to provide robust statistical analysis, posterior probabilities, and error rates, as required by *Daubert*.[2] As non-metric pelvic traits are commonly used in estimating sex in forensic cases, it is imperative that this method be externally validated for both accuracy and reliability.

The Klales et al. method was used to score three nonmetric pelvic traits in United States White and United States Black individuals from the Hamann-Todd Collection (N=279).[2] Two different observers scored each of the innominates to evaluate overall accuracy rates using the ordinal scores and logistic regression equation provided in the 2012 publication. The entire sample was also utilized to examine inter-observer error rates, while intra-observer error was evaluated on a subsample of 50 innominates that were rescored by each observer. Additionally, after rescoring the traits on this subsample, each observer also blindly provided an overall "gestalt" evaluation of sex based on the traits, other morphological characteristics, size, and personal experience.

Observers A and B attained similar accuracy rates of 73.1% and 72.4% with the Klales et al. regression formula, with each observer's results exhibiting a large sex bias. Females were correctly identified most of the time by both observers (95.7% and 97.1%, respectively), while males were correctly identified with accuracy approximately equal to chance (50.7% and 47.9%, respectively). Fifty-three individuals were incorrectly sexed by the Klales regression formula by both observers, including 51 males and only 2 females. In contrast to the regression equation, both observers achieved high accuracy based on their gestalt estimation of the intra-observer error subsample (98.0% and 90.2%, respectively).

Pairwise comparisons completed using Cohen's weighted kappa were used to evaluate observer error. Both the inter-observer and Observer A's intra-observer comparison exhibited moderate agreement for the ventral arc and subpubic concavity ($kappa$=0.539 and 0.549, respectively) and substantial agreement for the medial ischiopubic ramus ($kappa$=0.651 and 0.669, respectively). Observer B's intra-observer comparison revealed moderate agreement for the ventral arc (0.572) and substantial agreement for the subpubic concavity and the medial ischiopubic ramus (0.634 and 0.645, respectively) (all $p < 0.001$). The intra-class coefficient correlation tests found similar patterns, with agreements ranging from 0.693-0. 832 (all $p < 0.001$).

Results demonstrate that each of the three non-metric pelvic traits can be scored consistently both between and within observers, with a low incidence of systematic observer bias; however, the high accuracy rates originally reported by Klales et al. were not reached. These results, combined with the high sex bias in accuracy, reveal

problems with the method that will likely impact the method's use in forensic contexts when sex estimation of an unknown individual is considered to be the most important aspect of the biological profile. Additional research should be directed toward utilizing less subjective techniques to better quantify the traits and creating a more representative set of descriptions and visual guides for scoring.

Reference(s):

1. Phenice T. A newly developed visual method for sexing the os pubis. *Am J Phys Anthropol.* 1969;30:297-302.
2. Klales A.R., Ousley S.D., Vollner J.M. A revised method of sexing the human innominate using Phenice's nonmetric traits and statistical methods. *Am J Phys Anthropol.* 2012;149:104-114.

Sex Estimation, Innominate, Observer Error