Sue M. MacLaughlin,[1] Ph.D. and Margaret F. Bruce,[2] Ph.D.

The Accuracy of Sex Identification in European Skeletal Remains Using the Phenice Characters

ABSTRACT: Three documented European skeletal series were examined to assess the accuracy and reliability of the pubic variables described by Phenice for correctly identifying the sex of adult human skeletal remains. The accuracy and objectivity of these variables, as reported by Phenice, Kelley, Sutherland and Suchey, and Lovell, could not be confirmed on this European material. In general, the subpubic concavity feature, when used alone, proved to be the most reliable variable for sex identification.

In this study, the level of correct sex identification that could be achieved using the Phenice variables was shown to be significantly affected by the previous experience of the observer.

KEYWORDS: physical anthropology, human identification, Phenice method, osteology

The identification of sex from adult human skeletal remains is more reliable if the complete skeleton is available for analysis. In a sample of 750 intact skeletons, Krogman [1] reported close to 100% accuracy of sex identification. In the same sample, he achieved some 80% accuracy when using only the long bones and around 90% accuracy when only the skull was examined. Using the pelvis alone yielded an accuracy approaching 95%. This high accuracy of discrimination has its basis in the unique form of sexual dimorphism exhibited by the adult human pelvis and, in particular, in its innominate component. It is for this reason that the innominate is the skeletal element most preferred for sex determination.

There are essentially two approaches to sex identification. The first is based on a visual assessment of the shape or relative proportions of sexually dimorphic features. The second is a metric approach, which offers several advantages over the visual approach in that (1) it is inherently more objective, (2) replicability is high, (3) it is less dependent on previous observer experience, and (4) it is more readily amenable to statistical analysis and thus facilitates between-sample and between-study comparisons. However, it does depend on readily identifiable and unambiguous osteometric landmarks.

Phenice [2] described a visual approach to sex identification which he considered to be "accurate, rapid, highly objective and does not require years of experience for accurate application." The method involved noting the presence or absence of three morphological

features of the os pubis—a ventral arc, a subpubic concavity, and a ridge on the medial border of the ischiopubic ramus. Each feature was reported to be present in a "typically female" innominate and absent in a "typically male" innominate. Using this method, Phenice [2] correctly assigned sex to almost 96% of a sample of innominates selected from the Terry dissecting-room skeletal collection ($n = 275$).

Kelley [3] tested the reliability of the Phenice approach on an archaeologically derived American Indian skeletal collection and reported the method to be "highly consistent with other pelvic indicators of sex." From that paper [3] it could be calculated that the "Phenice estimated" sex was in agreement with the "pelvic sex" in 100% of the males and 90% of the females in Kelley's sample ($n = 362$).

Lovell [4] tested the reliability of the Phenice approach on a small sample of 36 individuals of documented sex but uncertain ethnic origin and concluded that close to 83% of her sample could be correctly identified.

Using only the ventral arc component, Sutherland and Suchey [5] found that 96% of a large sample ($n = 1284$) of documented American autopsy material could be correctly sexed.

Because of these high levels of accurate sex identification and the reported ease of application, the use of the three pubic features described by Phenice [2] is advocated by many North American physical anthropologists [6–13], and the approach is now beginning to be reported in European literature [14–15]. Yet, to our knowledge, the accuracy and reliability of the pubic variables, as reported by Phenice [2], have not been rigorously assessed on any large sample of human skeletal material of documented sex which is not of North American origin. Thus, the general applicability of the approach to forensic science and archaeological identification is still uncertain.

The purpose of this investigation was to examine the accuracy with which sex could be assessed from adult *European* skeletal remains of documented sex using the Phenice pubic variables and to compare these results with those obtained by Phenice [2], Kelley [3], Sutherland and Suchey [5], and Lovell [4].

Materials and Methods

A total of 273 skeletons of documented sex and age at death was selected from three European skeletal collections.

The first is a 17th to 18th century English cemetery collection housed in the crypts of St. Bride's Church in the City of London. Although some 250 skeletons were available, documentation could be regarded as secure for only some 160 individuals. After exclusion of nonadult material and of bones showing evidence of gross pathological conditions, innominates representing 85 individuals were available for study (52 males and 33 females).

The second collection is housed in the Department of Anatomy and Embryology at the University of Leiden, the Netherlands, and is of modern dissecting-room origin. After exclusion of individuals, using the same criteria as for the first collection, innominates representing 136 individuals were available for study (72 males and 64 females).

The third collection, also of modern dissecting-room origin, was made available by the Department of Anatomy, University of Aberdeen, Scotland. After exclusion of individuals on the grounds already stated, innominates representing 52 individuals were available for study (28 males and 24 females).

For each individual, only the right innominate was examined, but where this was absent or fragmentary, the left was substituted.

For convenience, the collections will be referred to hereafter as "English," "Dutch," and "Scottish," respectively.

Morphological Features of the Os Pubis

Phenice [2] defined the three morphological features of the os pubis as follows:

> *Ventral Arc*—"A slightly elevated ridge of bone which extends from the pubic crest and arcs inferiorly across the ventral surface to the lateral most extension of the sub-pubic concavity where it blends with the medial border of the ischio-pubic ramus."
> *Subpubic Concavity*—"A lateral recurve in the ischio-pubic ramus a short distance below the lower margin of the pubic symphysis."
> *Ischiopubic Ridge*—"A ridge on the ischio-pubic ramus immediately below the symphyseal surface."

Each feature was reported to be present in the above form in a "typically female" innominate and absent in a "typically male" innominate.

However, during a pilot investigation of the material, it became apparent that each feature was not exclusively present or absent, as defined, in either a typical male or a typical female form. For this reason, an ambiguous category of scoring was introduced. The inclusion of this category was considered appropriate as Phenice had noted that ambiguity occasionally arose, since not every os pubis presented each feature in either a "typical" male or "typical" female form. In this study, the ambiguous category was scored only if there was uncertainty as to whether a feature was present or absent. In this way, the use of this category was selected against.

For every innominate in this study, the ventral arc, subpubic concavity, and ischiopubic ridge features were independently scored, as follows:

Score 1 = typical female form,
Score 2 = ambiguous form, and
Score 3 = typical male form.

No weight was given to any one feature and at least two out of three characters were required to be in "agreement of sex" before sex was assigned; that is, for an innominate to be scored as female, a minimum of two scores of "1" was required, or, conversely, to be scored as male, a minimum of two scores of "3" was required. Where more than one ambiguous character (Score 2) or where three different scores were recorded, the innominate was assigned ambiguous sex status and, in the course of the subsequent analysis, was considered not to have been correctly sexed.

The discriminatory capacity of each component of the Phenice method was then assessed independently. Where an "ambiguous" score occurred, the innominate was considered *not* to have been correctly sexed.

The above tests were carried out by a single observer (SMM) without prior knowledge of the documented sex of the specimens.

Multiple Observer Test

A multiple-observer test of reliability was performed to test Phenice's statement (supported by Lovell [4]) that the level of accuracy achieved was independent of the experience of the observer.

In this study, 34 observers independently tested the method. The results were examined in the following three categories, according to the experience of the observer:

Group 1—experienced observers ($n = 12$);
Group 2—undergraduate students of human anatomy with moderate experience ($n = 20$); and
Group 3—human anatomy technical staff with no previous experience in this field ($n = 2$).

Each observer was requested to examine 34 innominates (22 male and 12 female), randomly selected from the Scottish series, and to assign sex solely on the basis of the three Phenice characters. The documented sex of the material was not made available to any observer.

Results

Single-Observer Test

While some 83% of the English innominates were correctly sexed by the Phenice method, only 68% of the Dutch and 59% of the Scottish innominates were sexed in agreement with the documented sex (Table 1). Females of both the English and Scottish groups were assigned to the correct sex more frequently than males (by some 22% in the English and 25% in the Scottish).

The discriminatory capacity of each feature was assessed independently (Table 2). In all three groups, the subpubic concavity proved to be the single most reliable indicator of sex and, in general, proved to be more reliable than the three features combined in the original approach. Used in isolation, both the ventral arc and the ischiopubic ridge features were found to be relatively poor indicators of sex in all three groups.

Multiple-Observer Test

The results of the multiple-observer test suggested that the Phenice variables could provide some 74% accuracy in sex determination in the sample of Scottish innominates

TABLE 1—*Accuracy of sex identification using the Phenice (1969) variables.*

Collection	Male + Female		Male		Female	
	n	%	*n*	%	*n*	%
English	85	82.7	52	71.7	33	93.7
Dutch	136	68.3	72	68.0	64	68.7
Scottish	52	58.6	28	46.4	24	70.8

TABLE 2—*Accuracy of sex identification for each feature of the Phenice (1969) approach, examined as an independent predictor of sex.*

Feature	Collection	Total, *n*	Accuracy of Sex Prediction, %		
			Male + Female	Male	Female
	English	85	64.9	58.8	71.0
	Dutch	136	55.0	68.1	41.9
Ventral arc	Scottish	52	45.4	42.8	48.0
	English	85	87.1	82.7	91.4
	Dutch	136	72.1	76.4	67.7
Subpubic concavity	Scottish	52	75.3	71.4	79.2
	English	85	61.0	52.8	69.2
	Dutch	136	59.2	45.1	73.4
Ischiopubic ridge	Scottish	52	56.2	42.8	69.6

(Table 3). The subpubic concavity, again, proved to be the single most reliable indicator of sex. The ventral arc and ischiopubic ridge features proved to be somewhat less reliable.

A statistically significant difference was found between experienced and inexperienced observers in the levels of correct sex identification (Table 4). Only the difference between Groups 1 and 2 was tested because of the small number of observers in the third group ($n = 2$). For the Phenice three-feature approach, Group 1 (experienced observers) achieved significantly higher levels of accuracy than Group 2 ($P < 0.001$), the mean accuracy of correct sex identification being 79% in Group 1 and 70% in Group 2. When each feature was examined independently, the experienced observers still tended to show significantly higher levels of correct sex identification than the inexperienced observers. The greatest difference between the results for the two groups was for the subpubic concavity feature, where experienced observers correctly assigned sex to some 81% of

TABLE 3—*Multiple-observer test of accuracy (Scottish sample).*

Sex	No. of Innominates	Accuracy of Phenice Approach, %	Accuracy of Each Feature as an Independent Predictor of Sex, %[a]		
			VA	SPC	IPR
Males + females	34	74	66	74	67
Males	22	74	69	73	65
Females	12	74	62	75	71

[a]Key to abbreviations.
  VA = ventral arc.
  SPC = subpubic concavity.
  IPR = ischiopubic ridge.

TABLE 4—*Accuracy of sex determination using the Phenice approach and of each feature as an independent predictor—experienced versus inexperienced observers (Scottish sample).*

	Accuracy of Sex Prediction, %		
	Males + Females ($n = 34$)	Males ($n = 22$)	Females ($n = 12$)
Phenice method			
Experienced ($n = 12$)	78.9***[a]	76.1*	84.0***
Inexperienced ($n = 20$)	70.0	70.2	69.6
Ventral arc			
Experienced ($n = 12$)	71.8**	72.3*	70.8***
Inexperienced ($n = 20$)	63.4	66.4	57.9
Subpubic concavity			
Experienced ($n = 12$)	80.6***	78.8***	84.0***
Inexperienced ($n = 20$)	68.8	68.6	69.2
Ischiopubic ridge			
Experienced ($n = 12$)	69.8	66.3	76.4***
Inexperienced ($n = 20$)	64.6	63.0	67.5

[a]The asterisks indicate the following probability levels:
  * $P < 0.05$.
  ** $P < 0.01$.
  *** $P < 0.001$.

the sample, whereas inexperienced observers were correct for only 69% of the innominates (a difference of 12%).

Discussion

From a sample of 275 innominates (72 Negro and 203 Caucasian) selected from the Terry skeletal collection, and using only three morphological features of the os pubis, Phenice [2] reported an accuracy of sex determination in excess of 95%. Kelley [3] tested the method on an archaeologically derived skeletal series of predominantly American Indian origin, and reported a high degree of concordance between "Phenice estimated sex" and sex estimated by other pelvic indicators. A further test by Lovell [4] on a small sample of 36 individuals of known sex but uncertain ethnic origin yielded results indicating approximately 83% accuracy of sex identification.

Dittrick and Suchey [12] found that, when the Phenice method was used in conjunction with "the relative width of the pubic bone and the configuration of the ventral rampart," 99% accuracy of correct sex prediction could be achieved.

In contrast to these results, the only test of the method on non-American material [15] reported that the Phenice method was useful only as corroborative evidence of sex. This was based on an autopsy sample of 100 pubic bones of Asian origin.

In this study of three European skeletal collections of documented sex and age at death, the high discriminatory capacity of the method reported by Phenice [2], Kelley [3], Sutherland and Suchey [5], and Lovell [4] could not be confirmed (Table 5). Although some 83% of English innominates were assigned to the correct sex, at best, only two thirds of the innominates in the Dutch and Scottish groups were sexed in agreement with the documented sex. In general, the subpubic concavity feature proved to be the single most reliable feature, and, in fact, the inclusion of the ventral arc and ischiopubic ridge features did not appear to increase the level of correct sex identification significantly.

The difference between the results obtained in this study and those obtained on North American skeletal material may be the result of a number of factors.

A more rigorous procedure for sex estimation was adopted in this study than was used

TABLE 5—*The results of this study compared with those of Phenice (1969), Kelley (1978), Sutherland and Suchey (1987), and Lovell (1989).*

Study	Sample Size, n		Mean Accuracy of Phenice Method, %	Accuracy of Individual Components, %[a]		
	Male	Female		VA	SPC	IPR
Phenice						
Caucasian	160	43	96%	. . .	. . .	. . .
Negro	20	52	94%	. . .	. . .	. . .
Kelley	191	171	95%	. . .	. . .	. . .
Sutherland and Suchey	703	581	. . .	96%	. . .	70%
Lovell	13	23	84%	. . .	. . .	. . .
This study						
English	52	33	83%	65%	87%	61%
Dutch	72	64	68%	55%	72%	59%
Scottish	28	24	59%	45%	75%	56%

[a]Key to abbreviations:
VA = ventral arc.
SPC = subpubic concavity.
IPR = ischiopubic ridge.

in the original evaluation by Phenice [2] or in the reliability tests of Kelley [3] and Lovell [4]. For this investigation, at least two out of the three features were required to be in agreement before sex was assigned. The other studies relied on only one feature if it was considered to be "obviously indicative of sex" [2]. In our view, this introduces a heavily subjective element into the technique and brings the claimed objectivity of the approach into serious question. To date, it appears that no other study has attempted to evaluate the relative contribution of each of the three features to successful sex identification.

In this study, all three features were given equal weighting. An intermediate or ambiguous category was included to take account of the cases in which there was some uncertainty over the presence or absence of a trait. The inclusion of an intermediate score does reduce the chance of correct sex prediction. On the other hand, because of the nature of the, albeit arbitrary, weighting procedure in the original Phenice approach, intermediate values would be less likely to result in misclassification if at least one of the three features was "obviously indicative of sex."

Phenice reported that ambiguous or intermediate forms were not a serious consideration, and Kelley noted that only occasionally were intermediate forms of each feature encountered. In this study, intermediate or ambiguous forms were not uncommon. High levels of intermediacy may arise from (1) difficulties in interpretation of the original Phenice description among observers or (2) sample differences in the degree to which each feature is expressed, or from both causes.

The description of each morphological feature of the os pubis is not unequivocal. Phenice described the ridge of bone on the ventral surface of the pubic body in females as being in the shape of an arc and stated that this did not occur in males. However, he reported that, in some cases, a ridge of bone may occur in males but then stated that its course differed from that of the female arc. Sutherland and Suchey [5] added one addendum to Phenice's description of the ventral arc in that the female arc should be "palpable." It appears that a ridge of bone may therefore exist in both sexes, the difference lying principally in the shape of the ridge, its subsequent course, and its elevation. The critical criterion for the separation of the sexes, based on this feature, may not therefore be the objective assessment of the *presence* or *absence* of the feature but rather a subjective appraisal of its shape. Subjective assessments are open to more variability in interpretation and produce higher levels of interobserver error, which, in turn, may result in differing levels of accuracy by different workers.

Sutherland and Suchey [5] also reported that the identification of the ventral arc may be age dependent, with a less obtrusive precursor arc occurring in younger individuals (less than 30 years of age). The age at death was not considered to be a complicating factor in this study as the mean age at death for the three groups ranged from 56 to 67 years, with only 8% of the English, 2% of the Dutch, and 5% of the Scottish individuals younger than 30 years of age at death. It has been suggested that the arc should in fact be more readily discernible with advanced age, and therefore the dominance of older individuals in this study should, in fact, have increased the accuracy with which the ventral arc character was identified and thus have minimized confusion over a precursor arc condition.

Lovell [4] reported that "previous experience in human osteological analysis was shown to have no effect on accuracy in this test, confirming Phenice's assertion that the technique does not require extensive experience to yield accurate results." The results of this study show that the experienced observers did indeed show a higher level of correct sex identification than the inexperienced observers. This was true for all three Phenice features when assessed independently as well as for the combined feature approach. Interestingly, the greatest difference between the results for the two observer groups occurred for the subpubic concavity feature, where 81% accuracy was achieved by the experienced group and only 69% accuracy by the inexperienced group. This difference between the groups

arose from the fact that, of the three Phenice features, this trait was more likely to be assigned correctly by experienced observers, whereas the inexperienced observers achieved similar levels of accuracy for each of the three features. This implies that the experienced observers were incorporating more than a simple objective "present" or "absent" observation into their decisions. Therefore, the previous experience of the observer will play a considerable role in the accuracy with which this method can be used.

The lower levels of correct sex identification reported in this study may, of course, reflect population differences. Previous work on the three skeletal series studied in this investigation [16] showed that the English sample was significantly more sexually dimorphic in pelvic dimensions than either the Dutch or the Scottish sample. This higher level of sexual dimorphism may, of course, account in part for the higher level of accuracy of sex determination found in the English sample. It might be expected that a population with a high degree of sexual dimorphism would show a higher incidence of the characteristically male and female forms of the pubic traits, leading to higher levels of correct sex identification.

Variation in population levels of sexual dimorphism may therefore account, in part, for the disparity between the North American results and those found in this study. Unfortunately, no indication of the levels of sexual dimorphism in the samples studied by Phenice [2], Kelley [3], Sutherland and Suchey [5], or Lovell [4] is available for comparison.

Before population differences can be explored, and understanding of the functional or developmental basis, or both, of the sexual dimorphism in the three features is required. No attempt was made by Phenice [2], Kelley [3], Sutherland and Suchey [5], or Lovell [4] to explain fully the developmental or functional basis of the dimorphism of the three morphological traits of the os pubis. Phenice suggested a connection between the ischiopubic ridge and the attachment of the genital crura. Buikstra and Mielke [17] also linked the origins of the dimorphism of all three features to the attachment of the penile or clitoral crura. Bass [7], on the other hand, associated all three features with the attachment of the arcuate ligament of the pubis. Neither the ventral arc, subpubic concavity, or ischiopubic ridge is described in any detail in any standard anatomical reference text.

Clearly, further studies are required to determine the basis of the dimorphism of these three features. In the meantime, caution must be urged in the application of this method to collections which differ from the one on which the technique was developed. As a preliminary to its use, it would be advisable to determine the following in advance.

(a) the level of sexual dimorphism in the sample.
(b) a protocol for the relative weighting of each feature, and
(c) the previous experience of the observer.

References

[1] Krogman, W. M., *The Human Skeleton in Forensic Medicine*, Charles C Thomas, Springfield, IL, 1973.

[2] Phenice, T. W., "Newly Developed Method of Sexing the *Os Pubis*," *American Journal of Physical Anthropology*, Vol. 30, 1969, pp. 297–302.

[3] Kelley, M. A., "Phenice's Visual Sexing Technique for the *Os Pubis*: A Critique," *American Journal of Physical Anthropology*, Vol. 48, 1978, pp. 121–122.

[4] Lovell, N. C., "Test of Phenice's Technique for Determining Sex from the Os Pubis," *American Journal of Physical Anthropology*, Vol. 79, 1989, pp. 117–120.

[5] Sutherland, L. D. and Suchey, J. M., "Use of the Ventral Arc in Sex Determination of the *Os Pubis*," paper presented at the 39th Annual Meeting of American Academy of Forensic Sciences, San Diego, CA, 16–21 Feb. 1987.

[6] Ubelaker, D. H., *Human Skeletal Remains: Excavation, Analysis, Interpretation, Taraxacum Manuals on Archaeology*, No. 2, Smithsonian Institution, Washington, DC, 1978.
[7] Bass, W. M., *Human Osteology: A Laboratory and Field Manual of the Human Skeleton*, 2nd ed., Missouri Archaeological Society, Columbia, MO, 1981.
[8] Ortner, D. J. and Putschar, W. G. J., *Identification of Pathological Conditions in Human Skeletal Remains, Contributions to Anthropology*, No. 28, Smithsonian Institution Press, Washington, DC, 1981.
[9] Skinner, M. and Lazenby, R. A., *Found! Human Remains,* Archaeology Press, Simon Fraser University, Barnaby, British Columbia, Canada, 1983.
[10] Rathbun, T. A. and Buikstra, J. E., *Human Identification: Case Studies in Forensic Anthropology*, Charles C Thomas, Springfield, IL, 1984.
[11] Meindl. R. S. Lovejoy, C. O. Mensforth, R. P. and Don Carlos, D., "Accuracy and Direction of Error in the Sexing of the Skeleton: Implications for Paleodemography," *American Journal of Physical Anthropology*, Vol. 68, 1985, pp. 79–85.
[12] Dittrick, J., and Suchey, J. M., "Sex Determination of Prehistoric Central California Skeletal Remains Using Discriminant Analysis of the Femur and Humerus," *American Journal of Physical Anthropology*, Vol. 70, 1986, pp. 3–9.
[13] Krogman, W. M. and Iscan, M. Y., *The Human Skeleton in Forensic Medicine,* Charles C Thomas, Springfield, IL, 1986.
[14] Atkinson, S. "Variation in Rates of Age-Related Changes of Skeletal Tissue in a Romano-British Population," Ph.D. thesis, University of Durham, Durham, England, 1986.
[15] Suri, R. K. and Tandon, J. K., "Determination of Sex from the Pubic Bone," *Medicine, Science and the Law*, Vol. 27, No. 4, 1987, pp. 294–296.
[16] MacLaughlin, S. M. and Bruce, M. F., "Population Variation in Sexual Dimorphism in the Human Innominate," *Sexual Dimorphism in Living and Fossil Primates,* M. Pickford and B. Chiarelli, eds., Il Sedicesimo, Florence, Italy, 1986.
[17] Buikstra, F. E. and Mielke, J. H., "Demography, Diet and Health," *The Analysis of Prehistoric Diets,* R. I. Gilbert and J. H. Mielke. eds. Academic Press, New York, 1985.

Address requests for reprints or additional information to
Dr. S. M. MacLaughlin
Department of Anatomy
Guy's Hospital
London Bridge, London
SE1 9RT England