J Forensic Sci, May 2005, Vol. 50, No. 3
Paper ID JFS2003385
Available online at: www.astm.org

Tracy L. Rogers,[1] *Ph.D.*

Determining the Sex of Human Remains Through Cranial Morphology

ABSTRACT: Sex determination is the keystone of a biological profile, yet few qualitative methods of cranial sex determination have been tested. This analysis examines the accuracy and precision of 17 morphological features of the skull commonly used to determine the sex of unknown skeletal remains. The sample consists of 46 identified skulls from the 19th century St. Thomas' Anglican Church Cemetery in Belleville, Canada. Nasal aperature, zygomatic extension, malar size/rugosity, and supraorbital ridge proved the most useful; of secondary value are chin form and nuchal crest; mastoid size is of tertiary consideration; nasal size and mandibular symphysis/ramus size rank fourth; forehead shape ranks fifth; and palate size/shape are sixth. Skull size/architecture provides an internal standard to assess the relative sizes of other traits. This research is a necessary step in establishing the credibility of morphological sex determination with respect to the Daubert and Mohan criteria for admissibility in a court of law.

KEYWORDS: forensic science, morphology, sex determination, skull

Sex determination is the key to establishing a biological profile of unknown human skeletal remains. Most methods of estimating age, assessing ancestry, and calculating the stature of unidentified remains are sex-dependent, becoming less accurate and precise when sexes are pooled, or unknown (1–3). Sex can be established from a gross examination of the skeleton using either metric or morphological techniques. Although there is some overlap between the two approaches, they should be used in conjunction to produce the most accurate and complete assessment of sexual dimorphism. This is particularly critical for individuals whose skeletal dimensions and characteristics do not lie at the extremes of sexual expression. Combining approaches is advocated by training manuals in Osteology and Forensic Anthropology (4–10), which report comparable levels of accuracy for both, and discuss the inherent advantages and disadvantages of each (4,6).

Metric methods of sex determination possess several benefits relative to their morphological counterparts: they are considered objective because they rely on standard landmarks; they are easier to teach, learn, and reproduce; they result in lower levels of intra- and inter-observer error; and they produce fewer indeterminate cases (11–15). Quantitative techniques can be applied to a greater range of bones than qualitative analyses. A brief review of the literature reveals metric methods of sex determination encompassing the body from head (16) to foot (17–18). Computer programs, such as FORDISC 2.0 (19), evaluate skeletal sex using measurements from a number of individual or combined bones, allowing flexibility depending on the elements available.

The accuracy, precision, and limitations of quantitative methods of sex determination have been thoroughly documented for most skeletal elements (16–18). Accuracy and precision vary by bone, ranging from success rates slightly better than chance, to those exceeding 90% (11–12,15–19). As a group, metric analyses suffer from two major limitations. They are difficult to apply on fragmentary material, because most formulae depend on measurements of several intact bones and/or dimensions that span multiple bones. Secondly, population differences in bone lengths and body proportions have traditionally restricted the application of metric techniques to the populations on which the formulae were developed and tested (20). Various solutions to this problem have been sought. Most recently, Albanese (21) has developed formulae for the pelvis based on a geographically and temporally diverse sample. By incorporating a wide range of human variability he avoids the problem of population-specific formulae. Morphological traits are typically used without concern for population affiliation, yet few have been tested to determine their applicability outside the original sample (2,22).

Skeletal sex determination relies on sexually dimorphic expressions of bony characteristics produced through different patterns, rates, and periods of adolescent growth. Males have both a longer and more intense adolescent growth spurt than females. Patterns of growth that are shared by the sexes are extended for males relative to females, creating size differences that can be measured empirically, or gauged relatively. This type of trait can, therefore, be included in either a metric or morphological analysis of the skull. Cranial characteristics, such as larger male brow ridges, eyes that appear lower in the face, and larger nasal aperatures, are the result of extending the normal downward and forward growth of the male face relative to the female face through the process of a more intense and extended male growth spurt (23).

Shape-related, sexually dimorphic features of the pelvis, such as the larger, more rectangular female pubic bone, and the position of the acetablulum, result from differences between female and male growth patterns during adolescence. While both size-related and shape-related features have been grouped together under the category "nonmetric traits", only the latter are truly morphological (more shape-based than influenced by size). Although qualitative techniques of sex determination are limited to the skull (4), pelvis (24), and distal humerus (25), they have the advantage over metric methods of being applicable to burned and fragmentary remains.

[1] Department of Anthropology, University of Toronto at Mississauga, Mississauga, Ontario, Canada.

Received 16 Nov. 2003; and in revised form 16 July and 3 Dec. 2004; accepted 6 Dec. 2004; published 6 April 2005.

Several widely used osteology and forensic anthropology texts recommend and describe a limited number of morphological traits of the skull for sex determination (4–10). The popularity of particular characteristics is difficult to understand as very few of these features have been subjected to tests of accuracy and precision (26,27). In contrast, newly developed techniques, such as mandibular flexure (28), have been tested by several independent researchers on various populations (29,30). Recent analyses of traditional features emphasize the importance of combining individual trait assessments into an overall determination of sex in an objective, unbiased manner (31,32). One of the difficulties in developing a statistical approach to decision making in this context is the absence of accuracy and precision values for each trait. Past efforts at weighting individual criteria appear to have been a matter of personal preference based largely on experience, rather than empirical evidence.

To date, metric methods of sex determination have been better suited to forensic case work than morphological, because the former more readily lend themselves to statistical testing and data manipulation. Testing and peer review are critical components of the guidelines for admissibility of expert testimony in a court of law as outlined in the Daubert (33) criteria (USA) and Mohan (34) ruling (Canada), which state the error rates and limitations of techniques and theories used in court must be known. The purpose of this analysis is to test the accuracy and precision of several morphological features of the skull commonly used to determine the sex of unknown human remains to confirm or refute their value as sex indicators. This is the first step in the process necessary to evaluate the appropriateness of an osteological technique for use in a forensic context.

Given the guidelines set out by Daubert (33) and Mohan (34) it is equally as important to understand the theory behind a technique as it is to know the reliability of the method. Familiarity with craniofacial growth patterns allows forensic anthropologists to gauge the relative importance of a particular trait, and the likely effects of age, trauma, mechanical wear, and disease on its expression. By examining the factors responsible for the final adult configuration of each characteristic included in this study, it is possible to assess the significance of their degree of dimorphism.

A review of craniofacial growth (35–37) reveals three key points relating to the expression of sexually dimorphic features of the skull: [1] the greatest relative growth from childhood to adulthood is seen in the mandible, maxilla, upper face, cranial base, and head height (37); [2] the neurocranium is the earliest growing region of the skull, followed by the mid-face and parts of the cranial base, and, lastly, structures relating to mastication (36); [3] early growing parts of the skeleton are typically less sexually dimorphic than later growing areas (36). Infant facial features, regardless of sex, can be characterized as follows: the nose is short, rounded, pug-like; the nasal bridge is low; the nasal profile is concave; the forehead is bulbous and upright; cheekbones are prominent; the face is flat and the eyes are wide set (35). As the face begins to grow changes in dimensions and proportions develop. The longer the period of growth, the more pronounced the changes.

The growth of female facial features begins to slow around the 13th year of life and maturation is completed soon afterward, while males enter a growth spurt that continues through adolescence with maturation completed in early adulthood (35,38–40). Since this is a general pattern that varies between individuals, a certain amount of overlap in the size of male and female features is inevitable. After examining stature, cranial length, and width of 73 girls aged 6–15 years and 47 boys aged 10–18 years, Baughan and Demirjian (41) concluded girls have a smaller cranium than boys, both absolutely and relatively, even before puberty and that boys experience a cranial growth spurt not seen in girls. This fundamental difference in the duration and rate of growth is the basis for sexual dimorphism of the skull, producing more extreme differences in later growing regions that experience greater relative growth, while early growing areas that experience less relative growth are less clearly sexually distinct. Age must also be considered when evaluating the sex of adult skulls, as evidence suggests the cranium continues to grow throughout adulthood, thus an older female might approximate a younger male in general size (38,41,42).

The initial impression of the sex of a particular skull is often based on its overall size and architecture (rugosity). The literature suggests there are real differences in male and female cranial growth (35,41), but the early completion of neurocranial growth relative to the rest of the skull will prevent extensive dimorphism in this area. The first hypothesis relating to the traits examined in this study contends that cranial size and architecture can be used to gain an initial impression of sex, but more importantly, they provide a baseline for comparing the relative sizes of the remaining traits.

Growth of the mandible is complex. Overall it changes from a more v-shaped appearance in children to square in adults with the development of the chin, growth of the alveolus to support the eruption of the permanent dentition, and expansion of the masticatory muscles (35). Throughout this process the ramus becomes progressively more upright (35). The characteristic male mandible (larger, higher symphysis, broader ascending ramus, and flaring goinal angle) is the result of continued male growth relative to females. Considering the greatest relative growth of the face occurs in the mandible (37) and the structures relating to mastication are later growing and, thus more likely to be sexually dimorphic (36), the mandibular criteria, including chin form, should prove among the most dimorphic of the traits examined.

According to Enlow (35), growth of the inner table of the frontal bone ceases by age five or six, as growth of the frontal lobe is completed. The outer table, however, is part of the nasomaxillary complex, which continues to remodel outward until nasal growth finishes some years later. Since the female nasal region completes its growth several years earlier than the male, there is less separation between the inner and outer tables. Greater separation in the males results in a larger frontal sinus and greater protuberance of the supraorbital ridges. The resorptive nature of growth along the anterolateral surface of the orbital roof and the depository growth of the cutaneous surface of the supraorbital ridge combine to create a protrusive superior orbital rim. As part of the nasomaxillary complex, the forehead is displaced downward and forward from the calvarium. Consequently, the male forehead changes from the bulbous, upright, infantile shape to an angled, less rounded form (35).

Growth in the nasomaxillary region lasts longer in males than females, while growth of the inner table is completed at approximately the same time (35,41). The nasal and maxillary areas undergo high levels of relative growth (37), and are mid to later growing regions of the face (36), increasing the opportunities for sexual dimorphism to develop. Since forehead slope, frontal eminence size, and supraorbital ridge size relate to the same downward and forward growth processes of the face, craniofacial growth patterns suggest these features should be considered a single point of evidence that will be a good skeletal indicator of sex.

The orbit does not require a large increase in size in order to accommodate the growing eye and its surrounding tissues (35). In contrast, the maxilla undergoes significant relative growth (35–36, 38). As the maxilla is displaced downward and the orbital floor (part of the maxilla) drifts with it, the orbit becomes unnecessarily large. To compensate for the drift and maintain proper obital size, the orbital floor deposits bone (35). The nasal floor, originally very close to the floor of the orbit in the infant, is almost twice the distance form the orbital floor by the time growth is complete (35). Since the duration of growth in this region is greater for males (41), the nasal aperture will be longer, yet the simultaneous forward projection of the nasal region will cause the aperture to appear to be situated higher on the male face than the female face. The extensive relative growth of the mid-face (37) and its pattern as a mid to later growing region of the skull (36), suggest that the size of the nasal aperature should be equally as effective as the other traits influenced by nasomaxillary growth patterns (supraorbital ridges, frontal eminence size, and forehead shape).

If one considers the palate as roughly v-shaped, its growth pattern can be described as a process of bone deposition on the inner surface and resorption on the outer surface, causing both enlargement and displacement (35). As it increases in length, it also expands in width. The duration of growth in the nasomaxillary region is extended in males (41), thus the male palate is both larger and broader. The amount of relative growth that occurs in this area and the notably longer male growth period suggest palate size and shape will be a very useful indicator of sex.

Nasal bones are usually large in males due to the extended duration of male craniofacial growth. Nasal form results from the interaction of three areas of growth, the nasals, the malars, and the maxilla (35). Growth of the malars is resorptive, causing them to become relocated posteriorly, while the adjacent nasal region of the maxilla enlarges anteriorly, producing a protrusive nose (35). Since the nasals increase in length but grow little in width (35), they gradually form a sharper angle in the midline to compensate for the divergent growth in the surrounding regions. Because the shape of the nasal bones is dependent on the growth of two different systems, nasal shape is unlikely to prove effective as a sex indicator. Nasal size, related to differences in the duration of nasomaxillary growth will exhibit some degree of sexual dimorphism, although it may be difficult to assess if the range of variation for the population is not known (as is the case in a forensic context, where the individual is the basis of analysis).

Relocation of the malars posteriorly during growth, combined with lateral growth of the zygomatic arch caused by resorption on the medial surface and deposition on the lateral surface, causes the temporal fossa to enlarge while the malar remains proportionately broad in relation to the face, jaw size, and masticatory musculature (35). Extended male growth causes the malars to be larger and the zygomatic arches to be displaced more laterally than the corresponding structures in females. Given that the upper face exhibits neither the greatest, nor least amount of relative growth in the face (37), sexual dimorphism will likely be apparent, but will demonstrate a range of variation that may be difficult to interpret in a single individual.

Keen (13) observed differences in the form of the posterior root of male and female zygomatic bones. In males the root is continuous with the supramastoid crest, which then becomes part of the temporal line. This trait is dependent on the development of the temporalis muscle (13). St. Hoyme and Işcan (40) indicate this feature is a reflection of greater male robusticity, suggesting it will be a good sex indicator only in populations that exhibit sexually dimorphic robusticity. Furthermore, its expression may vary depending on population robusticity. In more gracile populations it may be difficult to distinguish males from females, as most of the population will appear gracile, while in robust populations it may be difficult to distinguish females from males, where most of the population will demonstrate significant muscle markings.

Differences in male and female mastoid and occipital condyle size reflect differences in the duration of male growth. The cranial base exhibits less relative growth than other regions of the skull (37), and parts of the cranial base are middle growing regions (36), suggesting that some sexual differences will be evident, but these features are not likely to be among the most distinctive traits. Although both can be measured, the mastoid and occipital condyles are commonly evaluated in relative terms, scored as small, medium, or large. When evaluated relatively, the scoring of these traits becomes highly subjective, being dependent on both individual and population variation. These features can be useful in large samples where the range of variation is known, but may be problematic in individual cases.

The parietal eminence is the initial site of ossification for the parietal bone. The infant parietal is bowed outward, with the eminence being the most lateral point (35). As the brain expands the bones of the calvarium are displaced outward, decreasing the curvature of the bones as they increase. This process accounts for the larger female parietal eminence, because the male calvarium continues to grow after female growth is complete (41). Given the range of individual variation and the fact that the neurocranium is the earliest growing region of the skull (36), one can expect some overlap between males and females with only the most extreme expressions of the trait being useful for sex determination.

Despite the fact that many researchers note differences in the size of male and female teeth (4), most seem to agree with St. Hoyme and Işcan (40) that the range of overlap between males and females is simply too great to make tooth size a viable means of distinguishing male and female skeletons, thus the teeth will likely be the least effective of the traits examined in this study.

To summarize, a review of craniofacial growth patterns suggest the following: [a] cranial size and architecture can be used to gain an initial impression of sex and as a basis for comparing the relative sizes of the remaining traits; [b] the mandibular criteria, including chin form, should prove the most dimorphic; [c] forehead slope, frontal eminence size, and supraorbital ridge size relate to the same downward and forward growth processes of the face and should, therefore, be considered a single point of evidence; [d] the size of the nasal aperature and malars should be equally as effective as other traits dependent on sexual differences in the duration of nasomaxillary growth; [e] the palate will likely be useful, but not as distinctive as the nasomaxillary traits; [f] nasal size should rank 4th; [g] the orbits, parietal eminences, and teeth will all perform poorly; [h] measures of robusticity (mastoid size, zygomatic arch extension, muscle markings, and occipital condyle size) have the potential to be useful indicators when the range of variation for the population is known—this will be difficult in individual forensic cases, but could be applied by an experienced observer in conjunction with an assessment of ancestry, or by comparing these features to overall size and architecture.

Craniofacial growth patterns (35,41) also suggest the inherent direction of error for cranial sex criteria favors females, as sexual dimorphism of the skull is the result of a longer male growth period. It is likely that borderline or indeterminate individuals will be classified as females because they fail to cross the male threshold along the continuum of possible expressions of size and shape. Because the skull continues to grow with age, young males will more likely be classified female, while older females may be mistaken for males. These predictions are supported by the work of Meindl and colleagues (43) who observe that females are rarely misclassified, while males are occasionally mistaken for females. Weiss (44), on the other hand, argues that the larger/smaller criteria of several cranial features produce a temptation to call doubtful specimens male.

The accuracy and precision of each trait was evaluated to test for sex biases.

Materials and Methods

The sample consists of 245 skulls from the 19th century St. Thomas' Anglican Church cemetery in Belleville, Canada, which was excavated in 1989 to permit the construction of a new parish hall. The individuals represented in the cemetery were of European, primarily British, origin and middle to upper class (23). A total of 577 intact skeletons were excavated, representing 37% of all individuals ($N = 1564$) interred in the cemetery over a 53 year period from 1821 to 1874 (24). A subset of 46 individuals was personally identified on the basis of coffin plates recovered during excavation. Information on the coffin plates was verified using parish records maintained by the church ministers. The personally identified individuals form the basis of this analysis.

In 1955 Krogman (45) introduced a suite of 13 traits capable of distinguishing male and female skulls. These characteristics form the basis of the "traditional" cranial features used by modern forensic anthropologists and osteologists to determine the sex of unknown skeletal remains. Table 1 provides the traditional traits and current textbook authors who recommend their use. All of Krogman's (45) recommended features, and four additional characteristics: size and shape of nasal aperature; nasal bone size; extension of zygomatic arch as a crest beyond the external auditory meatus (here referred to as zygomatic extension, and in other references as suprameatal or supramastoid crest); and chin form were evaluated in this analysis. The latter four traits have been variously recommended by the authors in Table 1 (5–10). Figure 1 illustrates the data sheet used to assess each trait. Features were scored as male, female, or indeterminate. If the area was damaged, or otherwise unobservable the trait was scored as N/A (not available).

The possible expressions of the trait with the sex they are most frequently associated with are provided in Fig. 1. If a skull exhibits combinations of male and female forms, it is scored indeterminate and the specific expressions of the trait are circled. For example, if a male orbit is square, but high, small, with sharp margins, the traits square and small would be circled under the male category and high with sharp margins would be circled under the female category. Thus, the feature would be scored as indeterminate, as it exhibits combinations of male and female trait expression.

At the time of analysis, the author did not know which of the skulls were personally identifiable and which were not. All 245 skulls from the Belleville collection were examined. Groups of 80–90 were brought to the lab at one time. They were initially divided into two groups on the basis of size and architecture. Each skull was then scored individually. Six skulls, representing 3 of the largest and 3 of the smallest individuals were used as references to demonstrate the range of variability present in the population. Although available for comparative purposes throughout the examination and used as the standard by which subjective assessments of "small", "medium", or "large" were made, the sexes of these individuals were not known at the time of analysis. After all 80–90 skulls in a group were scored they were compared once again to confirm they were consistent with the group to which they had been assigned. If doubt was raised the skull was reassessed. After all 245 skulls were examined, 49 individuals (the indeterminate cases and a random selection of the remaining skulls) were re-examined to determine the maximum degree of intra-observer error. The large number of skulls examined, the time between the first and second evaluations, and the random order in which the individuals were examined made it impossible to remember the previous results, ensuring that the second assessment would not be biased by previous knowledge.

The percentage of cases for which trial one and two evaluations conflicted represents the degree of intra-observer error (precision). Accuracy was evaluated by comparing the assessments of each trait to the recorded sexes of the identified individuals. There were no discrepancies in sex estimation for the personally identified individuals included in the intra-observer error study, thus it was not necessary to decide which of the two sets of results to include in the accuracy tests. The value of each trait was a ranking that combined the least amount of intra-observer error with the highest accuracy. A Fisher's exact probability test was run to determine the effect of age on the accuracy of each trait. The age categories were: <25 years ($n = 8$), to reflect the fact that male craniofacial growth is not complete until early adulthood (35); 25–44 years ($n = 10$); and 45+ years ($n = 28$). Each age group was compared to the other two in individual 2×2 tables.

Suites of traits were also assessed for their collective effectiveness as skeletal sex indicators by calculating two by two tables, controlling for documented sex, of all possible combinations of cranial features following the procedure used by Rogers and Saunders (24). The probability of achieving a correct sex determination was calculated by dividing the number of correct assessments by the total number of estimates.

TABLE 1—*Traditional morphological cranial sex determinants.*

	Trait Recommended by					
Krogman Cranial Trait	Roberts	Byers	White	Burns	Bass	Buikstra & Ubelaker
Size & architecture	yes	yes	yes	no	yes	no
Size of supraorbital ridge	yes	yes	yes	yes	yes	yes
Size of mastoid	yes	yes	yes	yes	yes	yes
Nuchal crest	yes	no	yes	yes	yes	yes
Frontal eminences	no	no	yes	yes	yes	no
Parietal eminences	no	no	yes	no	yes	no
Orbit shape & position	no	no	yes	no	no	no
Forehead shape	yes	yes	no	no	no	no
Malar size & rugosity	no	no	no	no	no	no
Mandibular symphysis height	no	no	no	no	no	no
Mandibular ramus size	no	no	yes	yes	no	no
Palate size and shape	no	no	yes	no	yes	no
Occipital condyle size	no	no	yes	no	no	no
Size of teeth	no	no	no	no	yes	no
(See references 22–27,38)						

Group/Case:________Burial#:______________ Date:__________________

Location:____________________Condition:__________________________

Observer(s)____

	Male		**Female**	
General size/ architechture	big/rugged	_____	small/smooth	_____
Forehead	steeper	_____	rounded/full	_____
Frontal eminences	small	_____	large	_____
Supraorbital ridges	med - large	_____	small to med	_____
Orbits	square, low, small		round, high, larger	
	rounded margins	_____	sharp margins	_____
Nasal aperature	high, thin	_____	lower, wider	_____
Nasals	large	______	small	_____
Malars	big/rugged	______	smooth	_____
	more laterally arched		compressed	
Zygomatic	extends	_____	does not	_____
Parietal eminences	small	_____	large	_____
Mastoid	med- large	_____	small to med	_____
Occipital	well muscled	_____	not marked	_____
Occipital condyles	large	_____	small	_____
Palate	big, broad			
	u-shaped	_____	parabolic	_____
Mandible	large, high symph		smaller, lower	
	broad ascending		body, gonial	
	ramus, gonial angle		angle not	
	flares	______	flared	_____
Chin	square -2pt	_____	rounded-1pt	_____
Tooth size	larger	_____	smaller	_____

Comments:______________________________ **Sex Determination______**

FIG. 1—*Visual sex determination of the adult skull.*

Results

The overall intra-observer error for the 17 traits used in combination to determine sex was 12.2%, a value slightly higher than the accepted level of 10% (36), which suggests that some of the individual features are difficult to score consistently. All but four characteristics exhibited some degree of intra-observer error, but only two (nasal aperature and orbits) exceeded the acceptable level (Table 2). Comparisons of recorded sex with estimated sex revealed an overall accuracy of 89.1% when the 17 morphological features are used in combination with each trait given equal weight. The accuracy of each trait is presented in Table 2. Accuracy is lower than might be expected for many of these traits because they produced indeterminate evaluations. The percentage of indeterminate and incorrect assessments for each trait are also provided in Table 2. Five skulls were incorrectly assigned; two females and three males.

A chi square test revealed no significant difference between the levels of accuracy achieved in the three age categories ($p = 0.432$). A Fisher's exact probability test indicates that two criteria (zygomatic extension and nuchal crest) exhibit age-related patterns of accuracy. In both cases, accuracy increased with age. Rankings were created for each trait based on lowest intra-observer error and highest accuracy, giving both criteria equal weight. These are presented in Table 3.

No combination of two traits was more accurate than using all 17 traits. The most successful combination, zygomatic extension and malar size/rugosity, resulted in 88% accuracy, compared to 89.1% accuracy achieved using all 17 features. One combination of three characteristics (zygomatic extension, malar size/rugosity, and nasal aperature) produced an accuracy of 91%. All other combinations of three traits resulted in 88% accuracy or lower.

Discussion

Hrdlicka (4) stated that experienced investigators should be able to correctly identify the sex of an unknown skull in 90% of cases. Stewart (46) successfully determined the sex of 100 crania from the Terry skeletal collection with 77% accuracy. Krogman and Işcan (4) optimistically concluded that 92% accuracy could be achieved. The current research produced results closer to Hrdlicka, achieving 89.1% accuracy using 17 morphological features.

TABLE 2—*Intra-observer error rate and accuracy of each trait.*

	Intra-Observer Error $n = 49$	n	Accuracy	Indeterminate	Errors
Occipital condyle size	0.0%	43	14%	81.4%	4.6%
Tooth size	0.0%	39	10.3%	89.7%	0.0%
Size & architecture	0.0%	42	38%	54.8%	7.2%
Size of mastoid	0.0%	46	44.7%	48.9%	6.4%
Size of supraorbital ridge	2.0%	46	60.9%	34.8%	4.3%
Parietal eminences	2.0%	46	28.9%	36.2%	31.3%
Nuchal crest	2.0%	45	53.3%	33.3%	13.4%
Chin form	4.1%	46	56.3%	33.3%	10.4%
Mandibular symphysis & ramus size	4.1%	46	51.1%	40.5%	8.4%
Palate size and shape	4.1%	41	36.6%	43.9%	19.5%
Malar size and rugosity	4.1%	38	68.4%	26.4%	5.2%
Frontal eminences	4.1%	46	31.9%	36.2%	31.9%
Forehead shape	4.1%	45	44.5%	22.2%	33.3%
Zygomatic extension	6.1%	46	70.3%	25.6%	4.2%
Nasal size	6.1%	36	52.8%	13.9%	33.3%
Nasal aperature	10.3%	34	76.6%	17.6%	5.8%
Orbit shape and position	12.2%	39	43.6%	25.7%	30.7%

TABLE 3—*Trait rankings.*

Trait	Accuracy	Precision	Total	Combined Rank
Nasal aperature	1	5	6	1
Zygomatic extension	2	4	6	1
Malar size & rugosity	3	3	6	1
Size of supraorbital ridge	4	2	6	1
Chin form	5	3	8	2
Nuchal crest	6	2	8	2
Size of mastoid	9	1	10	3
Nasal size	7	4	11	4
Mandibular symphysis & ramus size	8	3	11	4
Forehead shape	10	3	13	5
Size & architecture	12	1	13	5
Palate size and shape	13	3	16	6
Orbits	11	6	17	7
Frontal eminences	14	3	17	7
Parietal eminences	15	2	17	7
Occipital condyle size	16	1	17	7
Tooth size	17	1	18	8

In general, the features of the face performed better than those of the calvarium. Of the top five criteria, four were facial and one cranial. The mean accuracy of the facial features (49%) was higher than the features of the calvarium (35%), yet the level of intra-observer error was higher for the facial traits (5%) than for the cranial (2%), suggesting some of the facial features were difficult to assess consistently. An examination of Table 2 reveals that orbit shape/position and nasal aperature size/shape were the most problematic of all the traits. One of the complicating factors in using orbit size and shape for sex determination is the fact that these characteristics also vary by population (2). Given the overall poor performance of the orbits (Table 3), this feature is not recommended for use in sex determination. Nasal aperature size and shape resulted in 10.3% intra-observer error, placing it at the borderline of acceptability (47), yet it was one of the most accurate predictors of sex (76.6%). Experience gained since the time of this analysis has resulted in increased precision with respect to scoring this trait, suggesting that intra-observer error in scoring nasal aperature can be reduced through better explanations and illustrations of the trait's expression.

The low accuracies reported for many of the traits (10.3%–76.6%) are the result of high frequencies of indeterminate assessments, the number of actual errors being quite small in most cases (Table 2). The high frequencies of indeterminate results obtained from single features emphasize the importance of relying on multiple indicators to provide the most complete analysis of sex. Traits with high frequencies of indeterminate results can still contribute to the overall assessment of sex when used in combination, providing they are correct more often than incorrect. The mandible, for example, produced indeterminate results in 40.5% of the cases, correctly predicted sex in 51.1% of the cases, and incorrectly predicted sex 8.4% of the time. Only the frontal eminences produced incorrect results as often as correct, and only the parietal eminences were more frequently wrong than right (Table 2). Given the poor performance of both the latter two traits in this analysis, they are not recommended for use in sex determination.

The importance of multiple indicators is confirmed by their individual results, as well as the accuracy of combinations of two and three traits. No single feature is capable of achieving the same level of accuracy as the traits used in combination (89.1%). The best individual trait was nasal aperature with an accuracy of 76.6%. Combinations of two traits did not exceed the accuracy of the full trait list, and only one combination of three traits achieved greater than 90%. Zygomatic extension, malar size/rugosity, and nasal aperature, the top three traits, produced an accuracy of 91%. These results suggest that both too few and too many traits can introduce error or uncertainty into a cranial assessment of sex. Relying on two or three traits, particularly the wrong two or three traits may cause the analyst to overlook valuable information, while incorporating poorly performing traits will confuse the analysis by introducing noise. This study makes it possible to devise a list of recommended traits that contribute to the analysis of skeletal sex, while reducing intra-observer error.

A final factor to consider before including a trait in the proposed list is its stability over time. Age-at-death played a role in the accuracy of only two traits, both of which were more effective as age increased. According to Meindl and colleagues [43], the success rate of assessing male skulls is lower for younger individuals, whereas in females accuracy is reduced for older individuals. It was not possible to divide the age categories by sex, due to the already small samples sizes. As a result, it is not clear whether age has no overall effect on the accuracy of morphological sex determination, or if male differences cancel female differences in pooled samples.

Returning to the predictions derived from examining patterns of craniofacial growth, the results of this analysis suggest the following: [a] cranial size/architecture exhibits low intra-observer error (0%), making it a consistent internal standard for assessing the relative sizes of the remaining traits, a factor that is of particular concern in forensic cases where skulls are assessed individually and experience provides the only other basis of comparison; [b] the mandibular criteria performed well, both chin form and mandibular symphysis/ramus ranked in the top five, but neither was the most accurate nor the most precise indicator of sex (Tables 2 and 3); [c] although the supraorbital ridges, frontal eminences, and forehead shape are all dependent upon the same growth process for their final form, the supraorbital ridges were easier to assess (2.0% intra-observer error) and more accurate (60.9%) than the other two traits; [d] the nasal aperature and malars tied with supraorbital ridges and zygomatic extension for the highest ranking; [e] palate size and shape did not perform as well as expected—it proved highly precise (4% intra-observer error), but resulted in a large proportion of indeterminate cases (44%); [f] nasal size ranked 4th as predicted; [g] the orbits, parietal eminences, and teeth all ranked low as expected; [h] of the measures of robusticity (mastoid size, zygomatic arch extension, nuchal crest, and occipital condyles), the first three were among the highest ranked, while occipital condyles ranked near the bottom. Finally, the small number of errors made it difficult to establish a particular bias in sex determination of the skull. Two females were incorrectly assessed male, while three males were incorrectly assessed female. There is too little data to comment on the significance of these results.

These predictions were developed from patterns of craniofacial growth and highlight the importance of understanding the processes responsible for the indicators we use to assess sex or ancestry. By considering the direction and duration of craniofacial growth it was possible to predict which indicators would perform well, but it was not possible to correctly rank the combined accuracy and precision of traits. Both individual and population variation play a role in the expression of many craniofacial characteristics and may be responsible for complicating the expression of sexual dimorphism in the skull. Interpreting and scoring traits is facilitated using the framework of craniofacial growth by comparing the degree of development of any given feature to its neighboring structures. This approach is particularly valuable when attempting to assess skulls that exhibit ambiguous features or combinations of male and female traits.

A review of five well-known osteology and forensic anthropology texts (5–10) reveals the most commonly recommended traits for sex determination are: [1] size of supraorbital ridge, size of mastoid; [2] nuchal crest; [3] overall size and architecture, chin form; [4] frontal eminences, zygomatic extension; [5] parietal eminences, forehead shape, mandibular ramus size, palate size/shape; [6] orbit shape and position, occipital condyle size, size of teeth. None of the texts recommended malar size and rugosity, mandibular symphysis height, nasal aperature size/shape, or nasal size. This analysis produced results that were not entirely consistent with the text recommendations, with rankings as follows: [1] nasal aperature, zygomatic extension, malar size and rugosity, size of supraorbital ridge; [2] chin form, nuchal crest; [3] size of mastoid; [4] nasal size, mandibular symphysis and ramus; [5] forehead shape, general size and architecture; (6) palate size and shape; [7] orbits, frontal eminences, parietal eminences, occipital condyle size; and [8] tooth size.

The five most commonly recommended traits (supraorbital ridges, mastoid, nuchal crest, overall size/architecture, and chin form) performed well in this analysis, yet three of the top traits identified in this study (nasal aperature, zygomatic extension, and malar size) have received very little attention in the literature. Most of the texts advocate the use of multiple features, providing instruction for scoring upwards of nine traits, with the exception of the Buikstra and Ubelaker text *Standards* (10). This resource encourages the use of five standardized analyses to ensure the comparability of skeletal analyses completed by different researchers. Limiting the number of recommended features is problematic due to variability in the condition of archaeological skeletal collections and individual forensic skeletal remains.

The results of this analysis suggest the following traits be used to determine sex based on the skull: of primary value are nasal aperature, zygomatic extension, malar size/rugosity, and supraorbital ridge; of secondary value are chin form and nuchal crest; the size of the mastoids is of tertiary consideration; nasal size and mandibular symphysis/ramus size are ranked fourth; forehead shape ranks fifth; and palate size/shape are sixth. Size/architecture should be used as an internal standard for assessing the relative sizes of the other traits.

The tests of accuracy and precision undertaken in this analysis were necessary to ensure the admissibility of this technique in a court of law. According to the Daubert (33) criteria (USA) and Mohan (34) ruling (Canada) expert testimony must be based on reliable principles and methods; the methods and theories must have been tested or be testable; the theory or technique in question must have been subjected to peer review and publication; and the potential or known error rates of the technique, as well as the standards controlling the operation of the technique, must be considered by the judge before ruling on the admissibility of the testimony. An important consideration in determining admissibility is whether experts are testifying about research conducted independent of the litigation, or they developed their opinions expressly for the purposes of the trial (33–34,49).

The principle underlying the traits examined in this analysis is normal craniofacial growth and development. The features evaluated in this study exhibit sexual dimorphism as a result of differences between male and female rates and durations of growth (35,41). The 12 recommended traits provide a reliable, accurate method of analysis. Both method and theory were tested in this study and will have undergone the process of peer review by the time of publication. An error rate of 11% was established for the full trait list of 17 features. Accuracy is expected to increase with the removal of the 5 problematic characteristics and emphasis on the recommended 12. Now that intra-observer error tests have established the precision of the traits as scored by a single investigator, the next step will be to test the reproducibility of each trait scored by different observers. It is also necessary to test the accuracy and precision of these traits on other samples, particularly modern individuals, to ensure these results are not limited to a particular population (both in time and space). Finally, this research was not completed expressly for the purpose of a specific trial, but rather to encourage the use of proven reliable and accurate techniques and traits over anecdotal favorites. Thus, the recommended morphological traits of the skull can be used in combination to accurately assess the sex of unidentified skeletal remains. This paper is the first step in establishing the admissibility of this method in a court of law under the Daubert and Mohan rulings.

References

1. Meadows L, Jantz RL. Estimation of stature from metacarpal lengths. J Forensic Sci 1992;37(1):147–54. [PubMed]

2. Rogers TL. The attribution of ancestry for European and Indian (South Asian) individuals within a forensic context [dissertation]. Burnaby (BC): Simon Fraser University, 2000.
3. Prince DA, Ubelaker D. Application of Lamendin's adult aging technique to a diverse skeletal sample. J Forensic Sci 2002;47(1):107–16. [PubMed]
4. Krogman WM, Işcan MY. The human skeleton in forensic medicine. Springfield: Charles C. Thomas, 1986.
5. Roberts CA. Forensic anthropology 1: the contribution of biological anthropology to forensic contexts. In: Hunter J, Roberts C, Martin A, editors. Studies in crime: an introduction to forensic archaeology. London: Routledge, 1996;101–21.
6. Byers S. Introduction to forensic anthropology. Boston: Allyn and Bacon, 2002.
7. White T. Human osteology (2nd edition). New York: Academic Press, 2000.
8. Burns K. Forensic anthropology training manual. New Jersey: Prentice Hall, 1999.
9. Bass WM. Human osteology: a laboratory and field manual. Columbia: Missouri Archaeological Society, 1995.
10. Buikstra JE, Ubelaker DH. Standards for data collection from human skeletal remains. Fayetteville: Arkansas Archeological Survey, 1994.
11. Giles E, Elliot O. Sex determination by discriminant function analysis of the crania. Am J Phys Anthropol 1963;21:53–68. [PubMed]
12. Henke W. On the method of Discriminant function analysis for sex determination of the skull. J Forensic Sci 1977;6:95–100.
13. Keen JA. A study of the differences between male and female skulls. Am J Phys Anthropol 1950;8:65–79. [PubMed]
14. Kelley MA. Sex determination with fragmented skeletal remains. J Forensic Sci 1979;24:154–58. [PubMed]
15. Richman EA, Michel ME, Schulter-Ellis FP, Corrunccini RS. Determination of sex by Discriminant function analysis of postcranial skeletal measurements. J Forensic Sci 1979;24:159–63. [PubMed]
16. Kalmey, JK, Rathbun TA. Sex determination by Discriminant function analysis of the petrous portion of the temporal bone. J Forensic Sci 1996;41(5):865–7. [PubMed]
17. Bidmos MA, Asala SA. Discriminant function sexing of the calcaneus of the South African Whites. J Forensic Sci 2003;48(6):1213–8. [PubMed]
18. Introna F, Jr, Di Vella G, Campobasso CP, Dragone M. Sex determination by Discriminant analysis of *calcanei* measurements. J Forensic Sci 1997;42(4):725–8. [PubMed]
19. Ousley SD, Jantz RL. Fordisc 2.0: personal computer forensic discriminant functions. Knoxville: University of Tennessee, 1996.
20. Buikstra JE, Mielke JH. Demography, diet, and health. In: Gilbert RI, Mielke JH, editors. The analysis of prehistoric diets. Orlando: Academic Press, 1985.
21. Albanese J. A metric method for sex determination using the hipbone and the femur. J Forensic Sci 2003;48(2):263–73. [PubMed]
22. Wanek V. A qualitative analysis for sex determination in humans utilizing posterior and medial aspects of the distal humerus [dissertation]. Portland (OR): Portland State University, 2002.
23. Rogers TL. Sex determination and age estimation: skeletal evidence from St. Thomas' Cemetery Belleville, Ontario [dissertation]. Hamilton (ON): McMaster University, 1991.
24. Rogers T, Saunders S. Accuracy of sex determination using morphological traits of the human pelvis. J Forensic Sci 1994;39(4):1047–56. [PubMed]
25. Rogers T. A visual method of determining the sex of skeletal remains using the distal humerus. J Forensic Sci 1999;44(1):55–9.
26. Graw M, Czarnetzki A, Haffner H. The form of the supraorbital margin as a criterion in identification of sex from the skull: investigations based on modern human skulls. Am J Phys Anthropol 1999;108:91–6. [PubMed]
27. Gülekon IN, Turgut HB. The external occipital protuberance: can it be used as a criterion in the determination of sex? J Forensic Sci 2003;48(3):513–6. [PubMed]
28. Loth SR, Henneberg M. Mandibular ramus flexure is a good indicator of sexual dimorphism. Am J Phys Anthropol 1998;105(1):91–2. [PubMed]
29. Haun SJ. Brief communication: a study of the predictive accuracy of mandibular ramus flexure as a singular morphologic indicator of sex in an archaeological sample. Am J Phys Anthropol 2000;111(3):429–32. [PubMed]
30. Hill CA. Technical note: evaluating mandibular ramus flexure as a morphological indicator of sex. Am J Phys Anthropol 2000;111(4):573–7. [PubMed]
31. Walker PL, Drayer FJ, Siefkin SK. Malibu human skeletal remains: a bioarchaeological analysis. Santa Barbara: Department of Anthropology, University of California—Santa Barbara, 1996.
32. Konigsberg L, Hens SM. Use of ordinal categorical variables in skeletal assessment of sex from the cranium. Am J Phys Anthropol 1998;107:97–112. [PubMed]
33. *United States Supreme Court in Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579 (1993).
34. *Regina v. Mohan*, [1994] S.C.R. File No. 23063.
35. Enlow DH. Handbook of facial growth (2nd edition). Toronto: W.B. Saunders Company, 1982.
36. Humphrey LT. Growth patterns in the modern human skeleton. Am J Phys Anthropol 1998;105(2):57–72. [PubMed]
37. Buschang PH, Baume RM, Nass G. A craniofacial growth maturity gradient for males and females between 4 and 16 years of age. Am J Phys Anthropol 1983;61(3):373–82. [PubMed]
38. Goldstein MS. Changes in dimensions and form of the face and head with age. Am J Phys Anthropol (old series) 1936;22:37–89.
39. Nanda RS. The rates of growth of several facial components measured from serial cephalometric roentgenograms. Am J Ortho 1955;41:658–73.
40. St. Hoyme LE, Işcan MY. Determination of sex and race: accuracy and assumptions. In: Işcan MY, Kennedy KAR, editors. Reconstruction of life from the skeleton. New York: Wiley-Liss, 1989.
41. Baughan B, Demirjian A. Sexual dimorphism in the growth of the cranium. Am J Phys Anthropol 1978;49:383–90. [PubMed]
42. Tallgren A. Neurocranial morphology and ageing—a longitudinal roentgen cephalometric study of adult Finnish women. Am J Phys Anthropol 1974;41:285–94. [PubMed]
43. Meindl RS, Lovejoy CO, Mensforth RP, Don Carlos L. Accuracy and direction of error in the sexing of the skeleton: implications for paleodemography. Am J Phys Anthropol 1985;68:79–85. [PubMed]
44. Weiss KM. On the systematic bias in skeletal sexing. Am J Phys Anthropol 1972;37:239–50. [PubMed]
45. Krogman WM. The skeleton in the forensic medicine. Postgrad Med 1955;17(2):A48–A62.
46. Stewart TD. Sex determination of the skeleton by guess and by measurement. Am J Phys Anthropol 1954;12:385–92. [PubMed]
47. Nichol CR, Turner CG. Intra- and interobserver concordance in classifying dental morphology. Am J Phys Anthropol 1986;69(3):299–315. [PubMed]
48. Rogers TL, Allard TT. Expert testimony and positive identification of human remains through cranial suture patterns. J Forensic Sci 2004;49(2):203–7. [PubMed]

Additional information—Reprints not available from author:
Tracy L. Rogers, Ph.D.
Department of Anthropology
University of Toronto at Mississauga
3359 Mississauga Road North
Mississauga Ontario L5L 1C6
Canada