Homo 74/1 (2023), 45–54
J. Comp. Hum. Biol.

Statistical classification methods for estimating sex based on five skull traits: A nonmetric assessment using 3D CT models

Yun taek Shim[1,*], Deog-Im Kim[2,*], Nahyun Aum[1], Seung gyu Choi[1], Young Seok Lee[1], Hyung Nam Koo[1], Yi-Suk Kim[3,]**

[1] Department of Forensic Medicine, National Forensic Service Institute, Seoul, Republic of Korea
[2] Department of Nursing, College of Nursing, Keimyung University, Daegu, Republic of Korea
[3] Department of Anatomy, The Catholic Institute for Applied Anatomy, College of Medicine, The Catholic University of Korea, Seoul, Republic of Korea
* These authors equally contributed to this work
** Corresponding author: jwkiss@catholic.ac.kr

With 5 figures and 3 tables

Abstract: Five cranial nonmetric traits for sex estimation for sex estimation are classified by score according to geometry. The population of origin is one of the factors influencing cranial nonmetric traits. Moreover, among the five cranial traits, the robust traits for estimating sex varied across population. The aim of this study is to suggest the most useful method for sex estimation and demonstrate the need of a suitable method for each population. One-hundred thirty-five three-dimensional skull images from 21st century Korean autopsy cadavers were evaluated using the ordinal scoring system of five cranial nonmetric traits as outlined in Buikstra & Ubelaker (1994). All scores of each trait were analyzed by linear discriminant and decision tree analyses for sex estimation. The frequency of each trait was analyzed and compared to populations from other studies. The accuracy for both sexes was 88.1% by discriminant analysis and 90.4% by decision tree. The traits with the highest accuracy were the glabella and mastoid process in both discriminant analysis and decision tree. Sex estimation in modern Korean cadavers using the cranial nonmetric method was shown to be highly accurate by both discriminant analysis and decision tree. When comparing the pattern of frequency scores in this study with those of other populations, the pattern of trait scores for estimating sex was different for each population, even among populations in the same Asian region, which suggests the need for methods suited for specific populations.

Keywords: cranial nonmetric traits; decision tree; discriminant analysis; three-dimensional image; sex estimation; population variation

Introduction

The accuracy and reliability of identification of individual human remains are affected by the condition of the remains, sex, age at death, secular trends, and population (Godde 2015; Langley et al. 2018). Visual assessment for sex estimation, which is the traditional method, usually relies on characteristics of the bone. The pelvic features are known to be the most reliable indicators of sex. However, the pelvic bone can be too poorly preserved to be useful for sex estimation due to being incomplete or having been highly modified by environmental or animal activity (Langley et al. 2018). Secondary to the pelvis, the skull is also widely used to estimate sex, since the size of the male skull is larger than that of females, and its bony markers are more prominent and more robust than females (Walker 2008; Garvin et al. 2014; Kim et al. 2015).

Identifying the sex of skeletal remains is the first and most important step for identification in the forensic and anthropological fields. Both metric and nonmetric methods are used to analyze sexual dimorphism (Kim et al. 2013a; Kim et al. 2013b). The metric method has certain challenges, such as frequent incompleteness of remains, absence of relevant anatomical landmarks, and need for special instruments for the landmark-based measurements. Cranial nonmetric traits have been criticized as being more

DOI: 10.1127/homo/2023/1632
E. Schweizerbart'sche Verlagsbuchhandlung, 70176 Stuttgart, Germany, www.schweizerbart.de

subjective than the metric method in sex estimation; however, if the difference between observers is small, subjectivity can be sufficiently represented as objectivity (Klales & Kenyhercz 2015; Godde et al. 2018). In addition, objectivity can be achieved by performing cranial nonmetric traits using an ordinal scoring system and applying enhanced statistical method to the values (Stevenson et al. 2009; Langley et al. 2018).

Many cranial features were initially proposed for sex estimation, of which five cranial traits of visual assessment received the most attention and have become the most commonly used simple ordinal scoring system by researchers (Garvin et al. 2014). Scoring system for the five cranial traits was strengthened by statistical analyses such as linear discriminant analysis (LDA), logistic regression, and decision trees. Note that the nature of cranial traits recorded on an ordinal scale can present challenges for LDA due to the violation of the normality assumption. However, the enhanced predictive performance of LDA is widely known to be very robust despite these assumption violations, particularly when the main objective is classification (Walker 2008; Nikita & Nikitas 2020).

The five cranial traits are currently the most commonly utilized nonmetric method for sex estimation by researchers, having the advantages of easy use, rapid process, and immediate results (Buikstra & Ubelaker 1994; Walker 2008; Stevenson et al. 2009; Garvin et al. 2014; Langley et al. 2018). Decision tree is one of the most widely used analysis methods in bioinformatics, and its use is currently expanding. This approach constructs and analyzes a classification system based on categorical variables, so that the results can be easily visualized and interpreted. In particular, chi-square automatic interaction detection (CHAID) emphasizes the various interaction predictor variables and maximizes the probability of making the right decision (Chen et al. 2011; Kamiński et al. 2018; Chern et al. 2019; Kim et al. 2021).

Correct sex estimation can be impacted by many factors, such as population, secular changes, disease, nutrition, health status, and environmental factors (Sakaue 2013; Garvin et al. 2014; Godde 2015). Population differences have been considered in many studies for correct sex estimation. Population differences are reflected in cranial traits; for example, the cranial traits for sexual dimorphism are different by population, which is influenced by temporal, biological, and geographic population differences (Walker 2008; Godde et al. 2018). The method of predicting sex for one population might lead to incorrect sex estimation when applied to another population. For that reason, it is important to refer to population-specific standards when predicting sex from cranial trait scores (Walker 2008; Garvin et al. 2014; Godde et al. 2018; Tallman & Go 2018).

The purpose of this study is to suggest the most useful method for estimating sex based on five skull traits in Koreans and to compare with the results from previous studies on the correct classification rates and the patterns in the frequency scores for estimating sex.

Material and methods

Samples

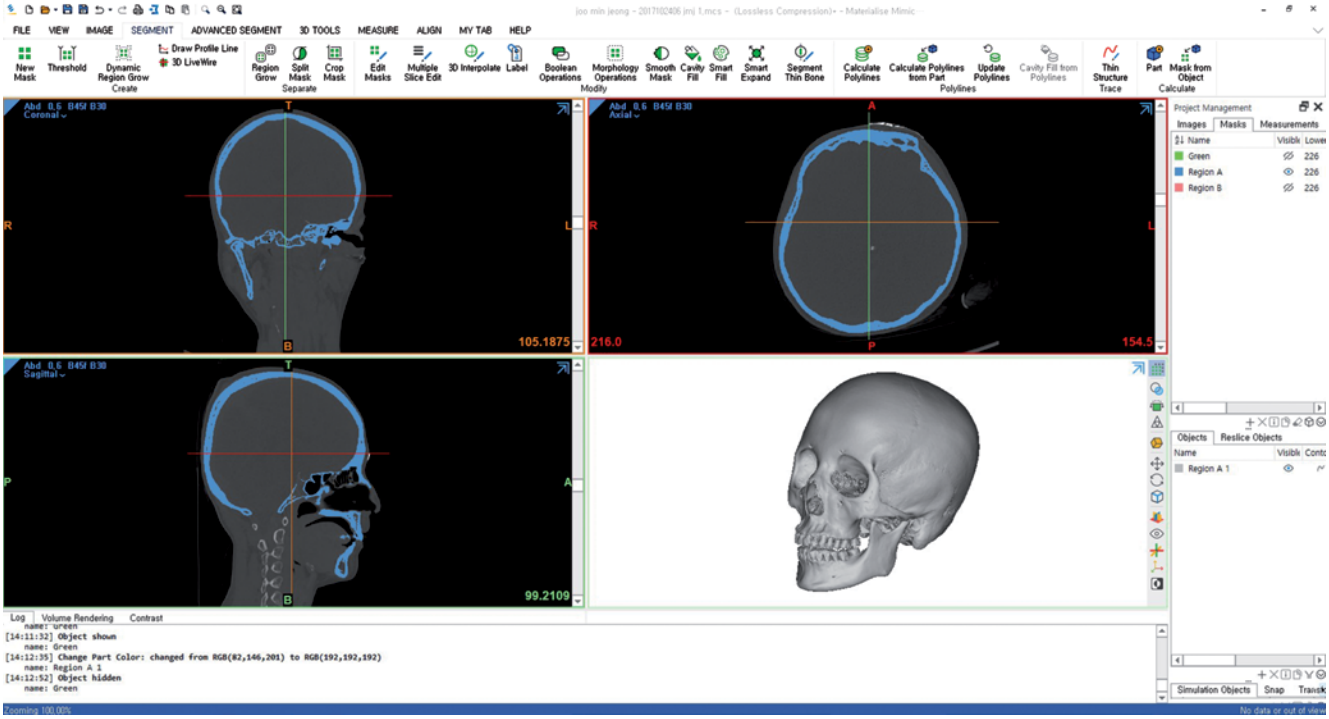
One hundred thirty-five three-dimensional (3D) skull images from Korean cadavers were obtained from the National Forensic Seoul Service Institute, randomly selected from among 8,653 cadavers collected from March 2017 to April 2020 without any injury or deformity to the skull. One researcher performed computed tomography (CT) scan of the skulls and reconstructed 3D images from them. The average age was 40.15 years (39.66 years (20–77 years) in males, 40.59 years (20–77 years) in females). The Institutional Review Board of the National Forensic Service approved this study, and informed consent was not required as it was recognized as an exempted subject, forensic autopsy (2021-05-HR).

CT examinations for forensic evaluation were performed with a SOMATOM Definition AS+ CT scanner (Siemens Healthcare, Erlangen, Germany). The imaging conditions were tube voltage 120 kVp, 210 mAs, slice thickness 0.75 mm, pitch factor 0.35, increment 0.7 mm, and rotation time 0.3 sec. The whole-body scan time was 60–75 sec depending on the height of the body. The DICOM files obtained by the CT examination were extracted using MIMICS 23.0 3D image-based engineering software (Materialize NV, Leuven, Belgium) with a threshold value of 226–3,071 Hounsfield Units (HU), and the skull imagery was extracted and converted to computer aided design (Fig. 1).

Five cranial traits

The five cranial traits were the nuchal crest, mastoid process, glabella, supraorbital margin, and mental eminence. This approach had first been published by Broca in 1875 and was further developed by Acsádi and Nemeskéri in 1970 (Walker 2008). Buikstra & Ubelaker (1994) wrote *Standards for Data Collection from Human Skeletal Remains* and developed the current scoring system in 1994. This consists of a five-point ordinal scoring system, with each cranial trait scored from 1 to 5. A score of 1 is minimal and indicates the most gracile expression of the trait associated with females (Buikstra & Ubelaker 1994).

In this study, to control for bias, data were collected with no information about the 3D skull images given to the observer. Prior to scoring, the skull images were aligned in two steps. First, each 3D reconstructed skull image on the computer monitor was enlarged to the size of the actual skull and was aligned based on the Frankfort horizontal plane. Then the aligned images were rotated from left to right and



```
    name: Green
[14:11:32] Object shown
    name: Green
[14:12:35] Change Part Color: changed from RGB(82,146,201) to RGB(192,192,192)
    name: Region A 1
[14:12:52] Object hidden
    name: Green
```

Fig. 1. Segmenting the skull and 3D CT models in the 3D image-based MIMICS software. In the study, segmentation was performed using the threshold value of Hounsfield unit.

scored based on the methodology of Buikstra & Ubelaker (1994) (Fig. 2).

Statistics

All the scores of each trait were taken twice over a period of about three months by one researcher. The statistical evaluation of the within-observer agreement included the weighted Kappa values proposed by Cohen (1960). Values of kappa greater than 0.81 indicated excellent agreement, values below 0.20 indicated poor agreement, and values between 0.21 and 0.80 indicated fair-to-good agreement (Landis & Koch 1977). The scores of each trait were analyzed by frequency, LDA, and decision tree using the statistical program SPSS 28.0 (IBM Corp., Armonk, NY, USA). The discriminant function equations were constructed in two ways: (1) using all the cranial traits and (2) using stepwise analysis.

A decision tree is usually used to construct classification systems for categorical variables and has the advantages of simplicity and good interpretability and data handling capability. In the decision tree, all samples were divided into 2 subgroups based on a cranial trait with the largest chi-square value. These groups were subdivided repeatedly by the next most significant factors (Chen et al. 2011; Kim et al. 2021). CHAID is a decision tree algorithm that examines all possible combinations of predictors and their categories. When the p-value from a chi-square test between child and parent nodes was less than 0.05, it was assumed to be statistically significant. It is also established that the parent node must have at least 10 cases and a child node at least five cases, as in Stevenson et al. (2009). CHAID analysis was used to identify the most important factors associated with sex from the five cranial traits. If the adjusted significance value was less than the 0.05 significance level, the node was separated. If not, it was considered a terminal node.

Results

The ordinal scoring system of five cranial nonmetric traits in this study was reproducible at the intra-observer level (Table 1). Using weighted Kappa coefficients, four of five cranial traits showed good agreement (> 0.61), whereas the mental eminence was considered moderate agreement with a difference of 0.01. Table 2 shows the frequency, chi-square value, and p-value for each score of the five cranial traits. Among them, the glabella had the highest chi-square value, and mental eminence had the lowest. In males, the frequency of score 4 was highest for all cranial traits. In contrast, in females, the scores with the highest frequency for differed by cranial trait but were concentrated in 1s, 2s, and 3s.

All traits were entered into discriminant analysis by the direct method. The accuracy was 88.1% for both sexes,

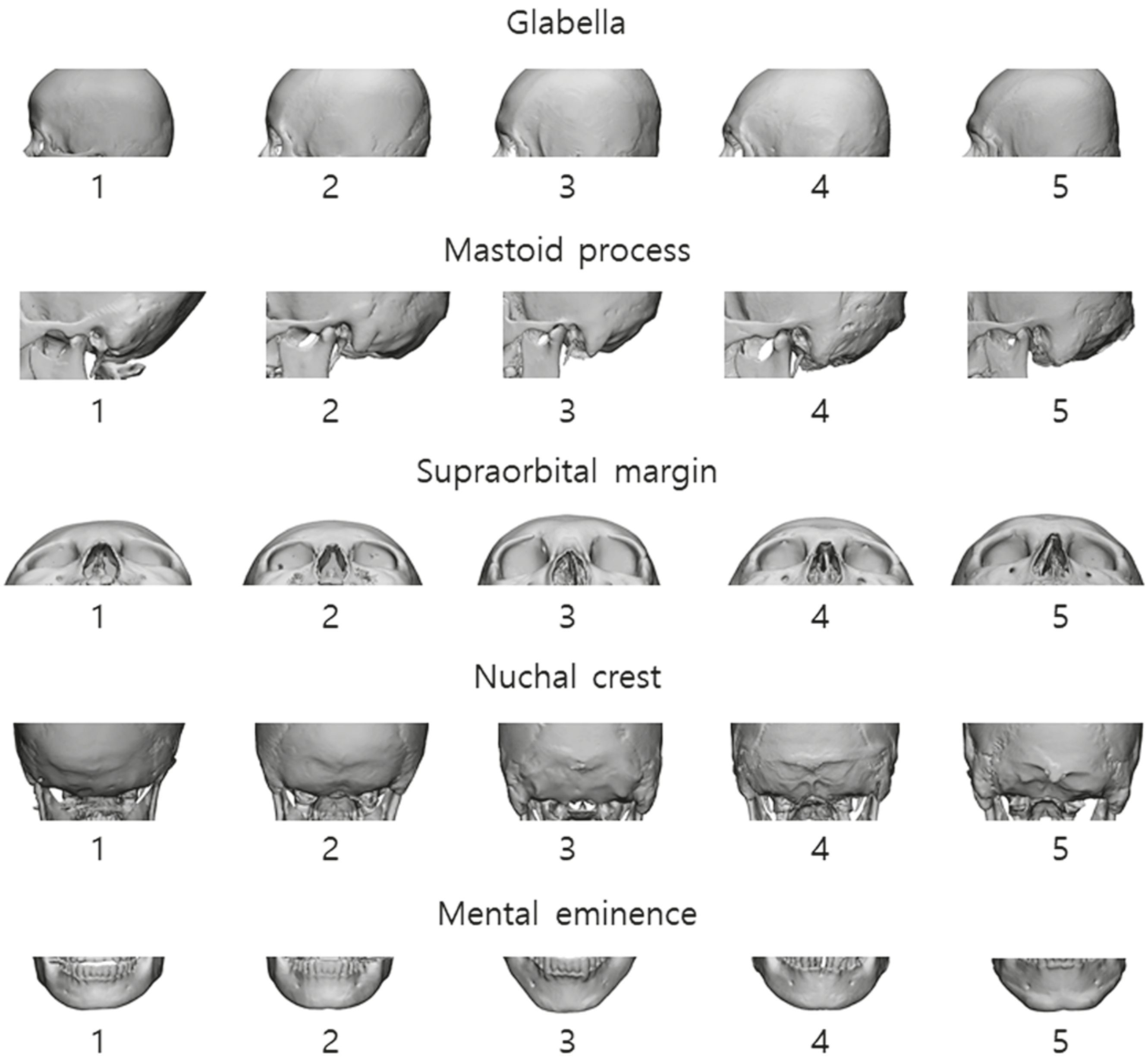


Fig. 2. Visualization of 3D skull and scores for nonmetric traits used in this study based on the methodology of Buikstra & Ubelaker (1994).

Table 1. Intraobserver assessment with weighted Kappa value.

Traits	Mean Kappa	Strength of agreement	95%CI	*p*-value
Glabella	0.79	Good	0.74–0.85	< 0.001
Mastoid process	0.61	Good	0.54–0.69	< 0.001
Nuchal crest	0.62	Good	0.55–0.69	< 0.001
Mental eminence	0.60	Moderate	0.51–0.69	< 0.001
Supraorbital margin	0.61	Good	0.54–0.69	< 0.001

CI, confidence interval.

Table 2. Results from frequency and chi-square analyses between the sexes [unit: n (%)].

Trait	Sex	Score					χ^2	*p*-value
		1	2	3	4	5		
Glabella	Males	1 (1.4)	3 (4.2)	12 (16.9)	35 (49.3)	20 (28.2)	81.024	< 0.001
	Females	25 (39.1)	24 (37.5)	10 (15.6)	5 (7.8)	0 (0.0)		
Mastoid process	Males	1 (1.4)	1 (1.4)	15 (21.1)	43 (60.6)	11 (15.5)	60.441	< 0.001
	Females	8 (12.5)	24 (37.5)	23 (35.9)	8 (12.5)	1 (1.6)		
Nuchal crest	Males	2 (2.8)	10 (14.1)	28 (39.4)	18 (25.4)	13 (18.3)	56.309	< 0.001
	Females	24 (37.5)	24 (37.5)	15 (23.4)	1 (1.6)	0 (0.0)		
Mental eminence	Males	0 (0.0)	8 (11.3)	27 (38.0)	33 (46.5)	3 (4.2)	34.608	< 0.001
	Females	6 (9.4)	27 (42.2)	23 (35.9)	8 (12.5)	0 (0.0)		
Supraorbital margin	Males	1 (1.4)	2 (2.8)	15 (21.1)	29 (40.9)	24 (33.8)	60.307	< 0.001
	Females	8 (12.5)	28 (43.8)	19 (29.7)	7 (10.9)	2 (3.1)		

Number of males was 71 and of females was 64.

93.0% in males, and 82.8% in females. The eigenvalue of this equation was 1.794, and the Wilk's lambda was 0.358. The discriminant equation (D2) using the stepwise analysis consisted of the glabella, mastoid process, and nuchal crest as follows:

$D2 = 0.703 \times \text{glabella} + 0.418 \times \text{mastoid process} + 0.385 \times \text{nuchal crest} - 4.507$

Its accuracies were 88.1%, 94.4%, and 81.3%, respectively, and were statistically significant within a 1% significance level (Table 3).

In the decision tree, there were six total nodes and four terminal nodes (Fig. 3). The most accurate tree that was generated contained the glabella and the mastoid process. The top-level node of the CHAID classification decision tree was "glabella." This decision tree, which represented the best sex predictive combinations of the traits, was accurate for both sexes combined, i.e., 90.4%, and was 94.4% in males and 85.9% in females. The risk estimate, a measure of within-node variance, was 0.096. All nodes except Node 0 had adjusted significance values < 0.05.

The decision tree starts with the researcher assessing whether the glabella has a cutoff value of 3. If the glabella was 4 or 5, it had a high probability of belonging to a male (0.92, see Fig. 3), and the researcher could stop scoring. If the glabella was 3, then the mastoid process was scored. If the mastoid process was ≤ 3, then the sample had a high probability of being female (Fig. 4).

Discussion

Sexual dimorphism is affected by environment, genetic factors, secular changes, and socio-economic situation (Godde 2015). As a result of discriminant analysis, the accuracy was 88.1% not only when all the traits were used, but also in stepwise analysis, respectively. In discriminant analysis using the stepwise method, the glabella, mastoid process, and nuchal crest were used for the equation. The first choice in the decision tree was the glabella, followed by the mastoid process. The accuracy of sex estimation using the decision tree was 90.4% and was very high at 94.4% in males. In this study, the discriminant function analysis and the decision tree for sex estimation showed an accuracy of nearly 90% (Table 3). In Walker's (2008) study, the accuracy of 88.4% for males and 86.4% for females was shown in American/English using the glabella, the mastoid process, and the mental eminence, and 69.5% for males and 82.9% for females in Native Americans using the supraorbital margin and mental eminence. Godde et al. (2018) found that, among the four cranial traits of Americans, the glabella should be weighted

Table 3. Linear discriminant analysis by three-dimensional reconstructed skull images from modern Koreans.

Functions	Coefficient		Eigenvalue	Canonical correlation	Wilk's Lambda	Sectioning point	Accuracy		
	Standardized	Unstandardized					Original		
							Female	Male	Pooled
D1. All			1.794	0.801	0.358	-18.632	93.0	82.8	88.1
Glabella	0.538	0.598							
Mastoid process	0.295	0.356							
Nuchal crest	0.345	0.367							
Mental eminence	–0.008	–0.010							
Supraorbital margin	0.239	0.258							
(Constant)		–4.759							
D2. Stepwise			1.714	0.795	0.368	-18.223	94.4	81.3	88.1
Glabella	0.632	0.703							
Mastoid process	0.347	0.418							
Nuchal crest	0.362	0.385							
(Constant)		–4.507							

more accurately to estimate sex, and the mastoid process of ancient Egyptians were superior to others. In Konigsberg & Hens (1998), the most accurate cranial trait in Native Americans was the mastoid process, with the accuracy of approximately 88%. In Stevenson's decision tree (Stevenson et al. 2009), the first choice was the glabella, and the second choice was the mastoid process, the same as in this study. However, unlike Stevenson's results, scores of 1, 2, and 3 of the mastoid process were classified as females (100%) and scores of 4 and 5 as males (75%). Consequently, the results of this study were similar to those of Americans (Walker 2008; Stevenson et al. 2009; Godde et al. 2018) in that the glabellar and mastoid processes were excellent cranial traits in order, but the results of Native Americans with the supraorbital margin (Walker 2008) and those of ancient Egyptians with the mastoid process (Godde et al. 2018) were different from this study.

In the results by the cranial nonmetric method published to date, the frequency of each trait and each score was compared by population (Fig. 5). First, the frequency of each ordinal score in Koreans was compared for Japanese, Thais, and Native Americans from Tallman & Go (2018) and Garvin et al. (2014). All other populations, including Koreans, were modern people around the 20th century, and in this study, the result of minimizing the influence of secular trends on sexual dimorphism could be predicted (Godde et al. 2018). In males, the glabella score was highest at 4 and 5 for Koreans, but in other Asian populations, this was scored as mainly 1 and 2. The frequency pattern in the mastoid process was similar across Asian populations, including Koreans, even though the Korean score of 4 was much higher than that in other Asians. The other three traits (mental eminence, nuchal crest, and supraorbital margin) in Japanese, Thais, and Native Americans had similar frequencies of scores 1 and 2, 4 and 5 based on a score of 3. The frequency pattern in females was dissimilar to that in males. There was no uniform pattern but, compared to Koreans, the glabella was almost biased toward score 1 in other Asians but score 1 and 2 in Koreans showed similar frequencies.

In U.S. White and Korean males (Fig. 5), the frequency patterns of each ordinal score were similar except for the mental eminence and supraorbital margin. In females, the frequency patterns of Koreans and U.S. Whites were similar, and it was the same as that of males. In U.S. Blacks/African Americans, the frequency patterns of each trait in males were different from of females and also different from of Koreans. More of the glabellas and nuchal crests in males were scored as 1 and 2, whereas the mental eminence and supraorbital margin had similar frequencies of 1 and 2 and of 4 and 5 based on score 3, unlike Koreans. In females, all cranial traits showed higher frequencies in scores 1 and 2. In the glabella, the score 1 was more remarkable frequent than the others. Even if the sex estimation used the same equations or methods, there was a difference in accuracy depending on the population (Garvin et al. 2014; Garvin & Klales 2017).

The frequency pattern of Koreans was closely analogous with that of U.S. Whites. That of U.S. Blacks/African Americans showed patterns similar to Japanese, Thais, and Native Americans (Fig. 5). This study found that the glabella and mastoid process performed the best, similar to other studies while the mental eminence performed the worst (Walker 2008; Stevenson et al. 2009; Garvin et al. 2014; Godde et al. 2018; Tallman & Go 2018). As has been demonstrated in other studies, the glabella is a cranial trait with fairly high reliability, while the least reliable traits in

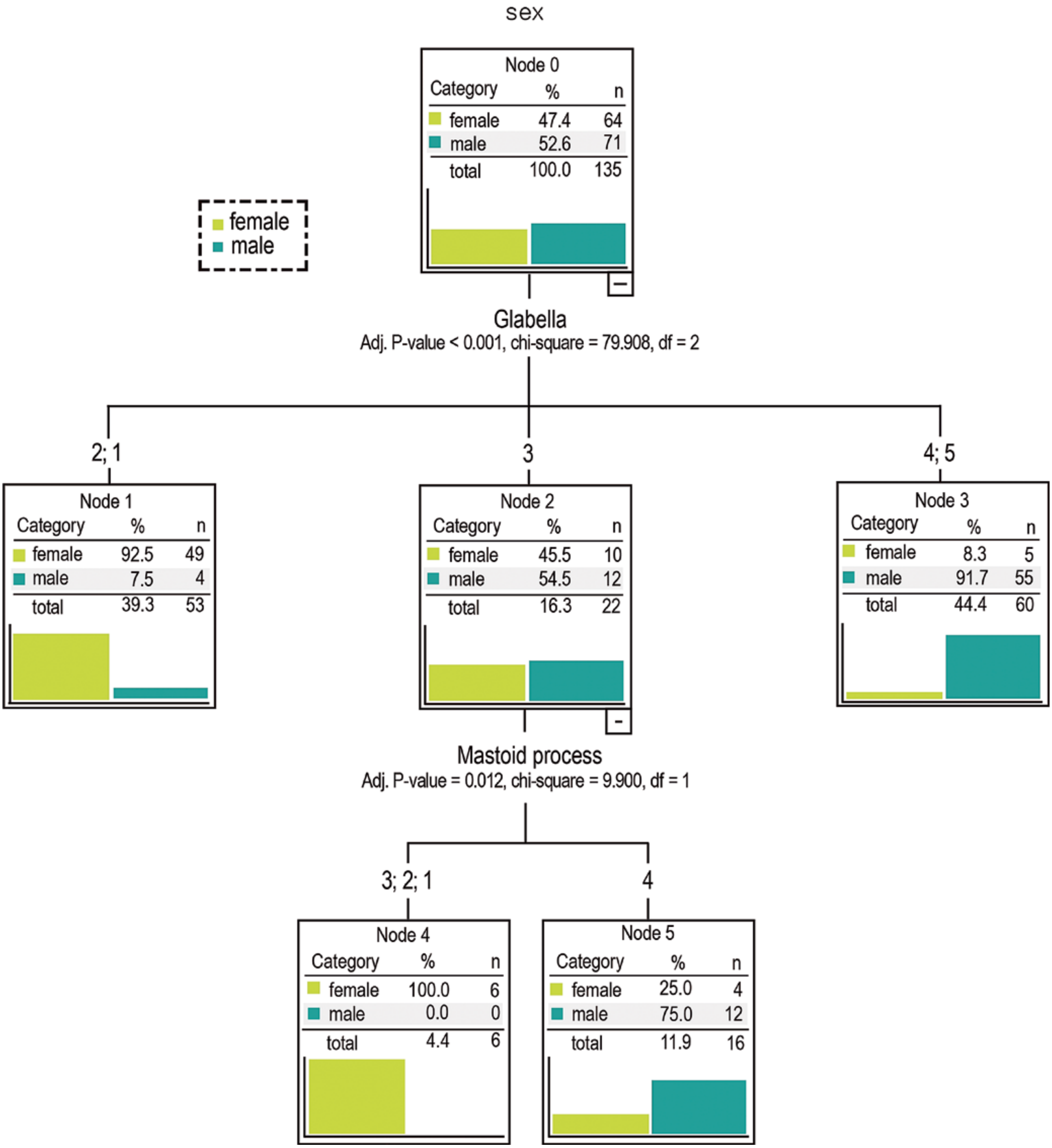


Fig. 3. A chi-square automatic interaction detection (CHAID) classification tree analysis to identify the predictors.

many studies have been the nuchal crest and mental eminence (Garvin et al. 2014; Tallman & Go 2018). These latter traits tend to be more difficult to score because they involve not only a larger area than the other traits and have a broad range of shape variation but are also unclear in terms of scoring methods. Therefore, these traits are subject to many differences in classification among researchers (Garvin et al. 2014; Godde 2015; Kim et al. 2015; Tallman & Go 2018). For example, in mental eminence, the mandible is often not excavated with the cranium (Garvin et al. 2014; Tallman & Go 2018) or is found status in edentulism or alveolar resorption may exclude this trait in analysis (Godde 2015). Tallman

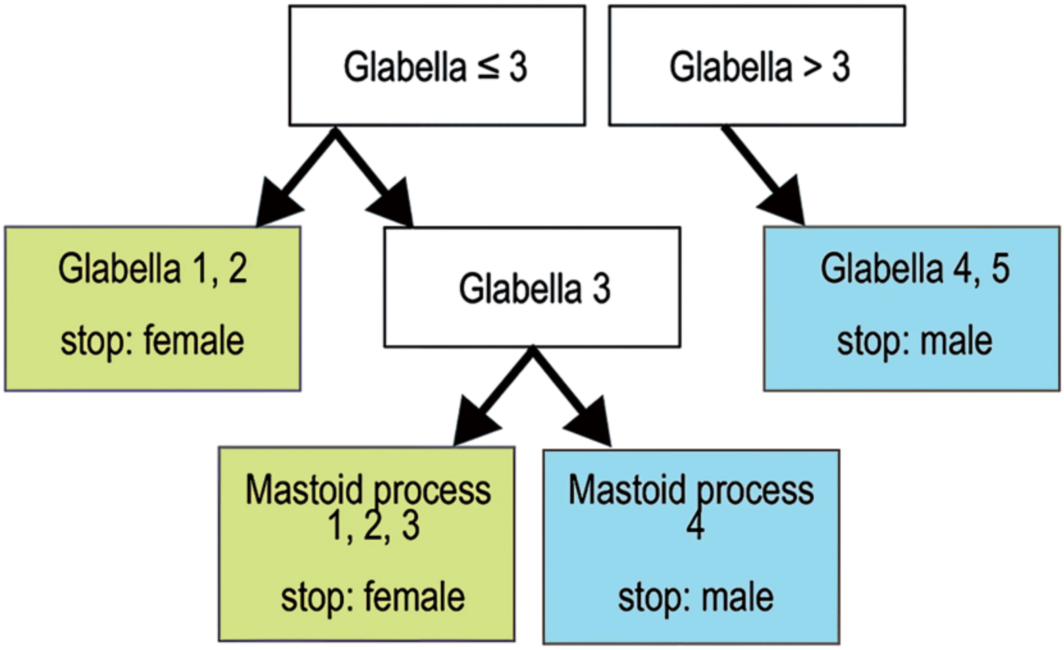


Fig. 4. Flowchart of sex determination using a decision tree (CHAID).

& Go (2018), Garvin et al. (2014), and Godde (2015) performed statistical analysis that excluded mental eminence and reported that the accuracy by this method was higher than that including mental eminence. The reason for excluding the mental eminence was that the sexual dimorphism of the mandible was likely to be obscured by the influence of dental pathology.

Walker (2008) and Stevenson et al. (2009) developed the method mostly on late 19th and early 20th century U.S. populations and 18th century British populations, and Tallman & Go (2018) used late 19th to 21st century Japanese and Thai populations. In addition, those studies analyzed real bone. However, this study analyzed 3D reconstructed contemporary Korean skull images and the accuracy of this study is believed to be slightly higher than that of previous studies (Omari et al. 2021), because it was targeted to only samples of a single population and used contemporary bodies (Godde et al. 2018). To take an extreme example, old and brittle skulls are difficult to measure physically, making them difficult to use repeatedly in studies for sex estimation. On the other hand, radiographic images, including 3D reconstructed images as in this study, have the advantage of compensating for these shortcomings of real bones and conveniences such as the handy storage location and the recent period of the research data (Omari et al. 2021).

Recent studies have documented secular changes in human bones, and geographical region and historical period also can influence the physiology affecting the expression of cranial nonmetric traits. These are affected by not only the aforementioned factors, but also by musculoskeletal and biomechanical forces, working separately or in combination (Godde 2015). In many forensic and archeological cases, an incomplete or impaired skeleton is excavated, and often only a cranium is recovered. It can be difficult to determine the sex of damaged bones using metric or nonmetric methods, and these methods might have low accuracy. Therefore, caution is required when sex estimation is attempted on unidentified bones found in a field, whereas using cranial traits with high classification rates, rather than using only one trait is a way to increase the accuracy of sex discrimination. In addition, it is necessary to use an equation suitable for each population to reduce sex determination bias, which can occur even in the same Asian regions. Therefore, considering the factors mentioned above, if the cranial trait method is used for sex estimation, it would be better in conjunction with visual assessments and also metric method such as volume of the cranial trait from 3D reconstructed images.

In conclusion, cranial nonmetric traits are usually used for sex determination in human skeletons, and they are often analyzed by discriminant function analysis or decision tree(s). Using these methods, the proportion of correct classifications is high and is more accurate in males than in females. When the cranial trait scores were compared for Korean, Japanese, Thai, Native American, and White and Black American populations, it was found that the frequency patterns were different, even among regionally proximate Asians. The pattern of frequency scores in Koreans in this study was different from that in another Asian group but was similar to Americans in most studies. Not all of the five cranial nonmetric traits have equal accuracy in assessing sex of skeletal remains. In statistical analysis, the glabella and the mastoid process gave the highest process scores.

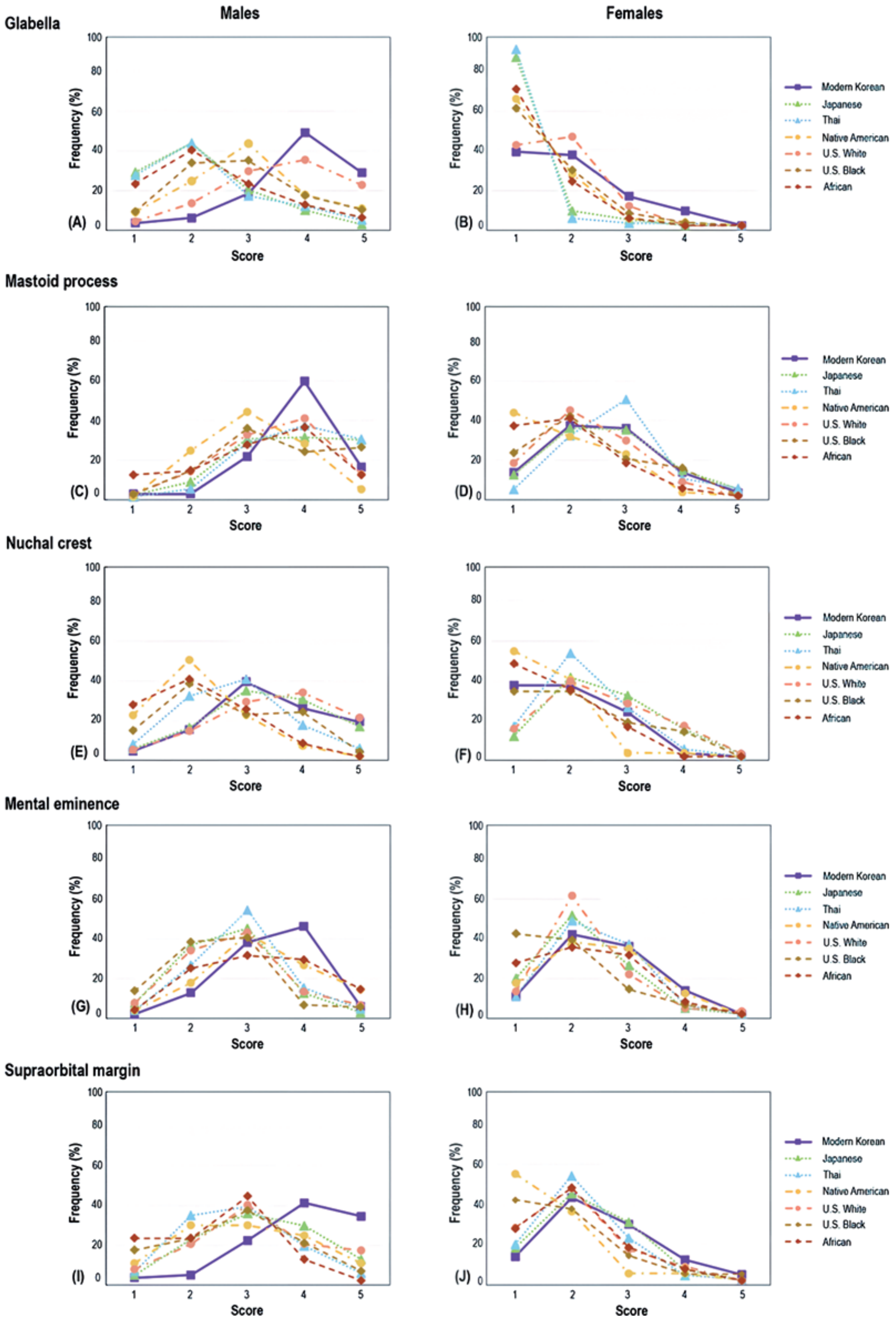


Fig. 5. Distribution of cranial trait scores by population origin from previous studies. The frequency of males by population group is (A), (C), (E), (G), and (I); and that of females is (B), (D), (F), (H), and (J). Data of Japanese and Thais from Tallman & Go (2018), and of Native Americans, U.S. Whites and Blacks, and Africans is from Garvin et al. (2014).

Conflict of interest

The authors declare no conflict of interest.

Funding

This work was supported by National Forensic Service (NFS2021MED08), Ministry of the Interior and Safety, Republic of Korea.

References

Buikstra, J. E., Ubelaker, D. H. (1994) *Standards for data collection from human skeletal remains*. Fayetteville, AK: Arkansas Archeological Survey.

Chen, X., Wang, M., & Zhang, H. (2011). The use of classification trees for bioinformatics. *Wiley Interdisciplinary Reviews. Data Mining and Knowledge Discovery, 1*(1), 55–63. https://doi.org/10.1002/widm.14 PMID:22523608

Chern, C.-C., Chen, Y.-J., & Hsiao, B. (2019). Decision tree-based classifier in providing telehealth service. *BMC Medical Informatics and Decision Making, 19*(1), 104–119. https://doi.org/10.1186/s12911-019-0825-9 PMID:31146749

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*(1), 37–46. https://doi.org/10.1177/001316446002000104

Garvin, H. M., Sholts, S. B., & Mosca, L. A. (2014). Sexual dimorphism in human cranial trait scores: Effects of population, age, and body size. *American Journal of Physical Anthropology, 154*(2), 259–269. https://doi.org/10.1002/ajpa.22502 PMID:24595622

Garvin, H. M., & Klales, A. R. (2018). A Validation Study of the Langley et al. (2017) Decision Tree Model for Sex Estimation. *Journal of Forensic Sciences, 63*(4), 1243–1251. https://doi.org/10.1111/1556-4029.13688 PMID:29148064

Godde, K. (2015). Secular trends in cranial morphological traits: A socioeconomic perspective of change and sexual dimorphism in North Americans 1849-1960. *Annals of Human Biology, 42*(3), 253–259. https://doi.org/10.3109/03014460.2014.941399 PMID:25156659

Godde, K., Thompson, M. M., & Hens, S. M. (2018). Sex estimation from cranial morphological traits: Use of the methods across American Indians, modern North Americans, and ancient Egyptians. *Homo, 69*(5), 237–247. https://doi.org/10.1016/j.jchb.2018.09.003 PMID:30269926

Kamiński, B., Jakubczyk, M., & Szufel, P. (2018). A framework for sensitivity analysis of decision trees. *Central European Journal of Operations Research, 26*(1), 135–159. https://doi.org/10.1007/s10100-017-0479-6 PMID:29375266

Kim, D.-I., Kim, Y.-S., Lee, U.-Y., & Han, S.-H. (2013a). Sex determination from calcaneus in Korean using discriminant analysis. *Forensic Science International, 228*(1-3), 177.e1–177.e7. https://doi.org/10.1016/j.forsciint.2013.03.012 PMID:23567446

Kim, D.-I., Kwak, D.-S., & Han, S.-H. (2013b). Sex determination using discriminant analysis of the medial and lateral condyles of the femur in Koreans. *Forensic Science International, 233*(1-3), 121–125. https://doi.org/10.1016/j.forsciint.2013.08.028 PMID:24314510

Kim, D.-I., & Han, S.-H. (2015). Non-metric study of the external occipital protuberance for sex determination in Koreans: Using three-dimensional reconstruction images. *Korean Journal of Physical Anthropology, 28*(4), 239–245. https://doi.org/10.11637/kjpa.2015.28.4.239

Kim, H.-J., Cho, Y., Noh, Y., Joo, J.-Y., & Park, H. R. (2021). A decision tree to identify the combinations of non-communicable diseases that constitute the highest risk for dental caries experience: A hospital records-based study. *PLoS One, 16*(10), e0257079. https://doi.org/10.1371/journal.pone.0257079 PMID:34614007

Klales, A. R., & Kenyhercz, M. W. (2015). Morphological assessment of ancestry using cranial macromorphoscopics. *Journal of Forensic Sciences, 60*(1), 13–20. https://doi.org/10.1111/1556-4029.12563 PMID:25047253

Konigsberg, L. W., & Hens, S. M. (1998). Use of ordinal categorical variables in skeletal assessment of sex from the cranium. *American Journal of Physical Anthropology, 107*(1), 97–112. https://doi.org/10.1002/(SICI)1096-8644(199809)107:1<97::AID-AJPA8>3.0.CO;2-A PMID:9740304

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159–174. https://doi.org/10.2307/2529310 PMID:843571

Langley, N. R., Dudzik, B., & Cloutier, A. (2018). A decision tree for nonmetric sex assessment from the skull. *Journal of Forensic Sciences, 63*(1), 31–37. https://doi.org/10.1111/1556-4029.13534 PMID:28508544

Nikita, E., & Nikitas, P. (2020). Sex estimation: A comparison of techniques based on binary logistic, probit and cumulative probit regression, linear and quadratic discriminant analysis, neural networks, and naïve Bayes classification using ordinal variables. *International Journal of Legal Medicine, 134*(3), 1213–1225. https://doi.org/10.1007/s00414-019-02148-4 PMID:31444553

Omari, R., Hunt, C., Coumbaros, J., & Chapman, B. (2021). Virtual anthropology? Reliability of three-dimensional photogrammetry as a forensic anthropology measurement and documentation technique. *International Journal of Legal Medicine, 135*(3), 939–950. https://doi.org/10.1007/s00414-020-02473-z PMID:33244707

Sakaue, K. (2013). Secular changes in craniofacial morphology during the Edo period of Japan. *Bulletin of the National Museum of Nature and Science Series D, 39*, 9–18.

Stevenson, J. C., Mahoney, E. R., Walker, P. L., & Everson, P. M. (2009). Technical note: Prediction of sex based on five skull traits using decision analysis (CHAID). *American Journal of Physical Anthropology, 139*(3), 434–441. https://doi.org/10.1002/ajpa.21042 PMID:19350636

Tallman, S. D., & Go, M. C. (2018). Application of the optimized summed scored attributes method to sex estimation in Asian crania. *Journal of Forensic Sciences, 63*(3), 809–814. https://doi.org/10.1111/1556-4029.13644 PMID:28940235

Walker, P. L. (2008). Sexing skulls using discriminant function analysis of visually assessed traits. *American Journal of Physical Anthropology, 136*(1), 39–50. https://doi.org/10.1002/ajpa.20776 PMID:18324631

Manuscript received: 22 March 2022
Revisions requested: 16 December 2022
Revised version received: 16 February 2023
Manuscript accepted: 28 February 2023