International Journal of Osteoarchaeology
Int. J. Osteoarchaeol. **21**: 82–91 (2011)
Published online 14 October 2009 in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/oa.1109

Analysis of Sexual Dimorphism of Craniofacial Traits Using Geometric Morphometric Techniques

P. N. GONZALEZ,* V. BERNAL AND **S. I. PEREZ**

División Antropología, Facultad de Ciencias Naturales y Museo, Universidad Nacional de La Plata. Paseo del Bosque s/n, La Plata 1900, Argentina

ABSTRACT This work deals with the assessment of cranial sexual dimorphism in human skeletal samples applying geometric morphometric techniques. The purpose of this research is to apply such techniques to quantitatively describe in craniofacial traits the degree and pattern of shape and size sexual dimorphism. Likewise, we evaluate the precision and accuracy of semilandmark-based techniques for sex estimation. We employ a sample of 125 adult skulls of known sex from the Coimbra collection. A set of coordinate points was selected to describe glabella, mastoid, frontal and zygomatic processes. The results of intra-class correlation coefficient (ICC) show excellent intra- and inter-observer agreement (ICC > 0.96) in the location of the coordinates of points employed. The principal component analysis (PCA) performed on shape variables shows a large superposition of both sexes, suggesting a relatively low degree of dimorphism in shape. As a consequence, the average percentages of correct sex estimations based on these variables were of 60.12 and 68.90%, obtained by discriminant analysis with leave-one-out cross validation and k-means clustering respectively. Conversely, when centroid size is included in PCA, females and males exhibit large separation along the first component. The highest values of correct assignment (77.86 and 72.15%) were found using shape–size variables with discriminant and k-means clustering analysis, indicating that the traits analysed display marked sex differences related to the larger size and more robust features of males. Finally, the advantages of geometric morphometric techniques are discussed.

Key words: cranial sex; semilandmarks; discriminant analysis; k-means clustering

Introduction

To identify sex from skeletal samples correctly is very important in bioarchaeological research. In this context, the studies generally aim to establish the sexual composition of large samples as well as to compare the degree and pattern of sexual dimorphism in different populations. Because sexual dimorphism is not uniformly expressed in the skeleton, the accuracy of estimations varies considerably between the different osteological elements (Buikstra & Ubelaker, 1994; Meindl & Russell, 1998). Although the highly dimorphic pelvic traits are considered to be the most reliable sex indicator, the skull is frequently used in archaeological contexts, due to its better preservation (Novotny *et al.*, 1993). In addition, numerous osteological collections formed in the last century are composed exclusively by skulls and, as a consequence, a great effort has been made to find criteria capable of distinguishing male and female skulls either suggesting new suites of traits or applying different morphometric and statistic approaches to register and analyse the cranial traits with acceptable levels of precision and accuracy.

The protocol for sex estimation by visual assessment of non-metric traits usually consists on seriating each feature of the skull and then sorting them into categories previously defined based on shape and size differences (Buikstra & Ubelaker, 1994; Konigsberg & Hens, 1998; Graw *et al.*, 1999; Graw *et al.*, 2005; Rogers, 2005; Williams & Rogers, 2006). To obtain the final estimation of sex, the traits used are ranked according to their accuracy and precision (Rogers, 2005). However, such approach has been largely criticized for being highly subjective, and because quantitative analysis of visual data is less developed than for metric variables (Konigsberg & Hens, 1998; Williams & Rogers, 2006). In addition, the seriation used to assign

* Correspondence to: Facultad de Ciencias Naturales y Museo, Museo de La Plata, Paseo del Bosque s/n, La Plata (1900), Argentina.
e-mail: pgonzalez@museo.fcnym.unlp.edu.ar

Received 4 February 2009
Revised 8 June 2009
Accepted 23 June 2009

sex turns to be problematic when dealing with large series, because it is doubtful that crania could be seriated in any replicable manner, and each cranial character must be examined independently (Konigsberg & Hens, 1998).

All that being said, it is often more desirable to perform a metric analysis when working with morphological data, since it has proven to be more objective. Because measurements rely on standard landmarks, results exhibit lower levels of intra- and inter-observer error. Likewise there are more powerful statistical methods for the analysis of continuous data. Linear measurements are commonly evaluated by uni- and multivariate statistical analyses, like discriminant functions, which allow to separate sexes with a rather high accuracy (Giles & Elliot, 1963; Franklin *et al.*, 2005). However, traditional linear measurements are not able to capture the shape differences of some complex and rounded structures, e.g. orbit shape, or prominence of the glabella or chin. For that reason, with a greater emphasis put on shape rather than size, visual assessment methods provided the most valuable tool to assess shape differences, at least until recently.

Because of these difficulties, researchers have explored alternative methods of quantifying and analysing morphological shape. A fundamental change started in the 1980s, regarding the nature of the data gathered and analysed (Rohlf & Marcus, 1993). Data that captured the geometry of the morphological structure employing 2D and 3D coordinates of anatomical landmarks and semilandmarks (points along homologous surfaces, curves or outlines) were of particular interest, and methods to analyse such data were developed (Bookstein, 1991, 1997; Adams *et al.*, 2004). This approach, now called geometric morphometrics, presents some advantages in relation to linear measurements. Researchers can preserve geometric information about the relative positions of coordinates of points, visualise results of multivariate analyses as configurations of landmarks back in the original space of the organism, and assess variation in structures with few or no landmarks (Adams *et al.*, 2004). Thus, geometric morphometrics would be more suitable to describe subtle differences in sexually dimorphic structures. Although some recent papers used landmark-based techniques to analyse the sexual dimorphism of some cranial traits, the accuracy attained by these techniques has not yet been evaluated (Rosas & Bastir, 2002; Franklin *et al.*, 2004; Katina *et al.*, 2004; Pretorius *et al.*, 2006; Kimmerle *et al.*, 2008).

The purpose of this research is to apply geometric morphometric methods to quantitatively describe the degree and pattern of sexual dimorphism in craniofacial traits. We evaluate the precision and accuracy of semilandmark-based techniques to describe and quantify shape and size differences between sexes in a sample of adult skulls of known sex.

Materials and methods

The sample consists of 125 adult skulls of known sex (Table 1) randomly selected from the collection of documented skeletons housed at the Museu Antropologico de Coimbra (University of Coimbra, Coimbra, Portugal). The Collection consists of 266 males and 239 females, with ages at the time of death that range from 7 to 96 years old. The individuals used in this study were born in Portugal between 1822 and 1921, and died between 1904 and 1936 (Santos, 2000). The majority of both males and females were from a lower socio-economic class (Santos, 2000).

The specimens were photographed with an Olympus SP-350 digital camera. The skulls were positioned in the Frankfurt plane and the camera lens was oriented parallel to the sagittal plane using a spirit level (Buikstra & Ubelaker, 1994). The images from lateral view were taken at a 300 mm distance from the euryon point. Coordinates for 12 landmarks (■) and 25 semilandmarks (●) were obtained (Figure 1) to describe the glabella, malar, mastoid process and frontal and zygomatic processes.

These structures were chosen because: (a) they exhibit high levels of preservation in archaeological contexts (Paiva & Segre, 2003; Peterson & Dechow, 2003); (b) the influence of artificial cranial deformation, a common cultural practice in several populations, is smaller than in traits of cranial vault and therefore generally negligible (Perez, 2006); (c) previous studies suggest that these structures are sexually dimorphic (Buikstra & Ubelaker, 1994; Rogers, 2005; Williams & Rogers, 2006).

Digital images were processed with MakeFan6 (Sheets, 2003). MakeFan6 places alignment 'fans' at

Table 1. Age and sex composition of the Coimbra sample

Age (years)	Female (*n*)	Male (*n*)
15–19.9	3	3
20–24.9	7	7
25–29.9	10	16
30–34.9	8	9
35–39.9	10	14
40–44.9	4	10
45–49.9	4	8
>50	4	8
Total	50	75

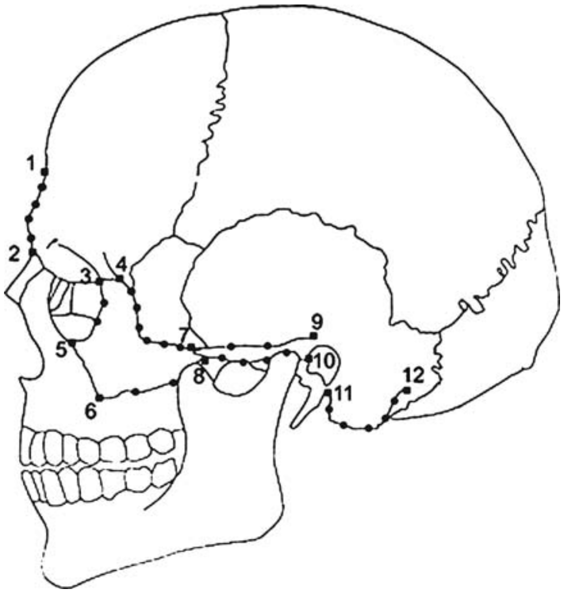

Figure 1. Allocated landmarks (■) semilandmarks (●) on craniofacial structures. Landmarks: (1) frontex: most inferior posterior midline point above glabella; (2) nasion; (3) frontomalare anterior; (4) frontomalare temporale; (5) infraorbitale; (6) zygomaxillare anterior; (7) the most superior point on the suture between zygomatic process of the temporal bone and the temporal process of the zygomatic bone; (8) the most inferior point on the same suture that landmarks 7; (9) auriculare; (10) it is defined as the point on the lateral aspect of the inferior border of the root of the zygomatic process; (11) it is defined as the anterior point on the root of mastoid process; (12) it is defined as the posterior point on the root of mastoid process. Landmarks 2, 4 and 9 were digitised following Buikstra & Ubelaker (1994), landmarks 3 and 6 following Howells (1973).

equal angular spacing along a curve to ensure a consistent placement of semilandmark coordinates. The landmark and semilandmark coordinates were digitised by means of tpsDIG 1.40 software (Rohlf, 2007).

Within geometric morphometrics the shape is defined as the information remaining after the effects of position, orientation and scale have been held constant (Rohlf & Slice, 1990). In this study, the Generalised Procrustes analysis (Gower, 1975; Rohlf, 1990; Rohlf & Slice, 1990) was used to remove these effects in landmark and semilandmark configurations, and centroid size was employed as size measurement (Bookstein, 1991). To convert the evenly distributed points along contours into semilandmarks, they were aligned by means of the perpendicular projection or minimum Procrustes distance criteria (Sheets *et al.*, 2004). This operation extends the Generalised Procrustes analysis (Gower, 1975; Rohlf & Slice, 1990) by sliding the semilandmarks until they match as well as possible the positions of corresponding points along an outline in a reference specimen, minimising the Procrustes distance (Sheets *et al.*, 2004). This results in an alignment of the semilandmarks along the curve so that the semilandmarks on the target form lie along the lines perpendicularly to a curve passing through the corresponding semilandmarks on the reference form (Sampson *et al.*, 1996; Sheets *et al.*, 2004).

Shape differences between specimens were studied using the aligned coordinates to perform a principal component analysis (PCA) to describe major trends in shape between males and females (Bookstein, 1991; Rohlf, 1993). The principal components obtained from these variables are known as relative warps (RW; Bookstein, 1991; Rohlf, 1993). An important aspect of this analysis is that variation along the relative warp axes can be expressed as intuitive deformation grid diagrams showing the difference from the mean form or reference. To visualise sexual dimorphism, graphical representations of shape differences were generated as deformation grids of female and male individuals relative to the reference configuration (i.e. consensus configuration). In addition, a PCA based on a matrix that includes shape coordinates and an additional column with log centroid size was made to describe the differences in the shape–size space (Mitteroecker *et al.*, 2004).

To estimate the precision of geometric morphometric techniques, intra- and inter-observer error associated with the placement of point coordinates were evaluated independently. A general consensus on methodologies suitable for assessing landmark and semilandmark error has not been reached yet (see von Cramon-Taubadel *et al.*, 2007 for a discussion about available methods). Whereas some methods evaluate the individual landmark precision (Corner *et al.*, 1992; von Cramon-Taubadel *et al.*, 2007), others emphasise the analysis of the 'overall effect of landmark error' (O'Higgins & Jones, 1998; Lockwood *et al.*, 2002; Viðarsdóttir *et al.*, 2002; Gonzalez *et al.*, 2007). The first follows the same approach as traditional morphometrics where the effect of discrepancy in data recording is tested for each variable. However, in geometric morphometrics, it is more difficult to test the extent of observer-differences in data acquisition until the configurations of landmarks are compared (von Cramon-Taubadel *et al.*, 2007). The second approach does not provide information about the differential precision of each coordinate, but it is a practical approach to the overall effect of intra- and inter-observer error on 'individual specimen affinity' (Lockwood *et al.*, 2002). In this study, we follow the last approach because observer error in semilandmarks cannot be evaluated individually, and it allows us to estimate whether observer error confounds the discrimination of individuals in the sample.

In order to assess intra-observer error, P.N.G. digitised the landmarks and semilandmarks on photographic images of 30 crania randomly selected from the skeletal collections hosed at the La Plata Museum. Two sets of variables were obtained for each skull with a week's interval between the scoring sessions. Then, to assess inter-observer error, the other authors (V.B. and S.I.P) digitised the landmarks and semilandmarks on the same sample. The sets of point coordinates obtained each time were used to perform a relative warps analysis, and the ordinations obtained were compared to evaluate both intra- and inter-observer error by comparing the score of each individual on the first two relative warps using an intra-class correlation coefficient (ICC; Shrout & Fleiss, 1979). Intra-class correlation assesses rating reliability by comparing the variability of different ratings by the same subject to the total variation across all ratings and all subjects. The ICC values range from 0 to 1, corresponding to lowest and highest agreement, respectively. Significance level was established as $p < 0.05$ and indicates if the agreement is greater than the one expected by chance.

The accuracy of geometric morphometric techniques was assessed by estimating the sex of individuals using two statistical methods, discriminate analysis with leave-one-out cross validation and k-means clustering. Their performance was examined by comparing the percentage of cases which estimated sex correctly matched the actual sex of the skull (i.e. the percentage of correct classification).

Discriminant analysis is a method used to find a set of axes that grants the greatest ability possible to discriminate between two or more groups (Manly, 1994). The main purpose of discriminant analysis is to achieve a predictive classification of individuals. The first step is to estimate the discriminant functions that best discriminate between groups, computing the classification scores for the individuals. The next step is to classify the individuals according to the group for which they have the highest classification score. Finally, the accuracy of the classifications is evaluated using a cross-validation analysis. In cross-validation, each case is classified by the functions derived from all cases other than that case. Therefore, the analysis is performed several times, excluding one individual at a time, as a way to establish whether or not it is well classified. This gives an unbiased estimate of the percentage of individuals that were wrongly classified. The significance of discriminant functions can be tested using Wilks' lambda and F value.

Discriminant analysis makes some assumptions about the variables employed (i.e. less predictor variables than the sample size of the smallest group, lack of high multicollinearity, normality and homogeneity of covariance matrices) that must be tested before performing the analysis. Because discriminant analysis requires more individuals than variables per group, the use of outline methods poses difficulties due both to the large number of semilandmarks per individual needed to describe outlines and to the representation of semilandmark points by two coordinates (x- and y-) when there is only one degree of freedom per point (Sheets *et al.*, 2006). Therefore, principal components analysis is used to reduce the dimensionality of the data by analysing a limited number of scores instead of the original data. In addition, the use of principal components avoids multicollinearity. In this study, the discriminant analysis was based on the score of individuals along the first two axes of PCA obtained for shape and shape–size (i.e. form) variables. We tested the normality employing Kolmogorov–Smirnov test (Hair *et al.*, 1998). The results show that principal components of shape (RW1: $d = 0.04889$; RW2: $d = 0.04598$) and form (PC1: $d = 0.05778$; PC2: $d = 0.02919$) do not differ significantly from the normal distribution. The homogeneity of variance–covariance matrices was checked using Bartlett's test (D'Agostino & Russell, 1998). The results indicate that homogeneity of variance–covariance matrices for shape ($X^2 = 4.3852$, $p = 0.2228$) and form ($X^2 = 3.8485$, $p = 0.2783$) variables cannot be rejected.

In k-means clustering analysis, a set of specimens is divided into k-groups fixed *a priori* in such a way that the specimens within the k-groups are more similar to one another than to specimens in the other clusters, thereby minimising within-group variation (MacQueen, 1967). The first step in this analysis is to define k centroids, one for each cluster. The next step is to take each specimen and associate it to the nearest centroid. Then, k new centroids are re-calculated and the specimens are associated to the nearest new centroid. Consequently, a loop is generated, where k centroids will change their location step by step until no more changes are done. The k-means clustering is different from discriminant analysis because no information about the specimens is required, so the clusters are generated based solely on the morphological similarity among specimens. In this study, we classify the individuals in two groups representing both sexes. Then, we can assess grouping accuracy for individuals in comparison with their actual sex.

Finally, four indices of sexual size dimorphism using centroid size were examined: the mean distance index

$(\mathrm{MDI} = \overline{X}_{\mathrm{m}} - \overline{X}_{\mathrm{f}}/\overline{X}_{\mathrm{m}}100)$; the ratio male/female $(\overline{X}_{\mathrm{m}}/\overline{X}_{\mathrm{f}})$; the Student's *t*-test and the *F*-test. These measurements were selected since their simplicity of application permits comparisons with cases from the literature and because of their widespread use in physical anthropology (Marini *et al.* 1999).

Geometric morphometric analyses were performed using tpsRelw 1.44 (Rohlf, 2007) and Semiland6 software (Sheets, 2003). All statistical analyses were performed using R 1.9.1 (R Development Core Team, 2008).

Results

The values obtained by ICC show excellent agreement ($ICC > 0.99$; $p < 0.01$) (Fleiss, 1981) between the two series registered by the same observer (P.N.G.). Likewise, high and significant values of ICC were found between the three observers, both for RW1 and RW2 ($ICC > 0.96$ and $p < 0.01$ for all comparisons).

Figure 2a is a plot of the first two relative warps calculated from the landmarks and semilandmarks of the craniofacial traits which account for 43.2% of the explained variance. Male morphologies—which locate at the most negative values of first relative warp—display a wider frontal and zygomatic processes as well as a more developed glabella and mastoid process, when compared to morphologies located at more positive values (Figures 2b and c). In the shape space represented by the first two relative warps, there is a great superposition of males and females. Conversely, in the shape–size space, both sexes are more differentiated along the first principal component (Figure 3). The first two components explain 42.29% of the total variance.

Figures 4a and b are box-plots of the discriminant scores obtained using shape and form. These figures show a great superposition between sexes when shape variables are used, while they tend to separate when size is incorporated to the analysis. The test of significance also shows these differences. Although both discriminant analyses were highly significant, Wilks' lambda for form was smaller ($\lambda = 0.650$; $F = 34.530$; $p = 0.0000$) than for shape variables ($\lambda = 0.863$; $F = 10.217$; $p = 0.0001$), indicating greater differences between group means when shape and size are analysed together.

The average accuracy obtained by discriminant analysis based on shape variables was of 60.12%, and slightly higher for females (64.15%) than for males (60.75%) (Table 2). When multivariate statistical analyses were performed using shape and size, the percentages of correct estimations increased (Table 3). The average accuracy obtained was 77.86%, with similar percentages for females and males (76.9 and 78.48%, respectively).

The results obtained by k-means clustering based on shape variables were 68.90% for average accuracy, 61.1% and 74.35% for females and males, respectively (Table 2). The percentage of correct allocation by k-means clustering of shape–size variables was 72.51%, and was similar for both sexes (73.07% for females and

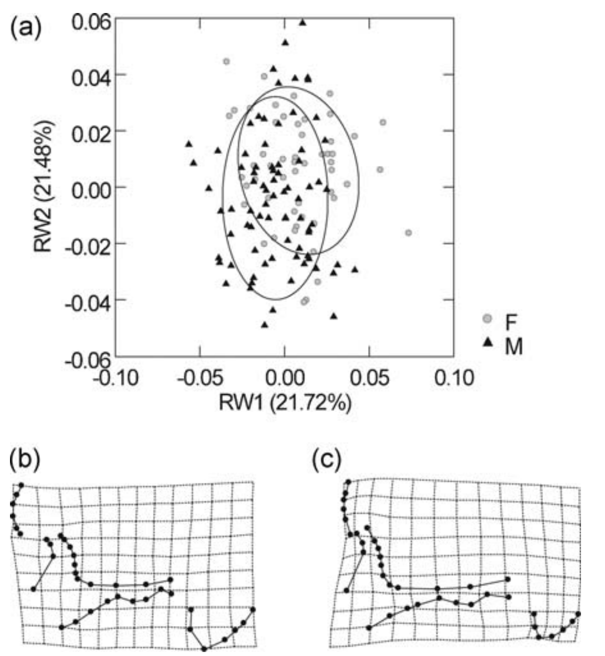

Figure 2. Relative warps of the craniofacial traits (a). The deformation grids represent the variation along the first relative warp's axis, showing typically male (b) and female morphology (c). F: female; M: male. The elipses represent 68% confidence intervals for males and females.

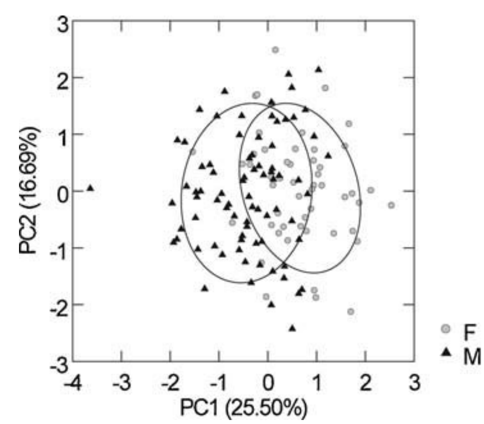

Figure 3. Principal component analysis obtained with shape and size variables. F: female; M: male. The elipses represent 68% confidence intervals for males and females.

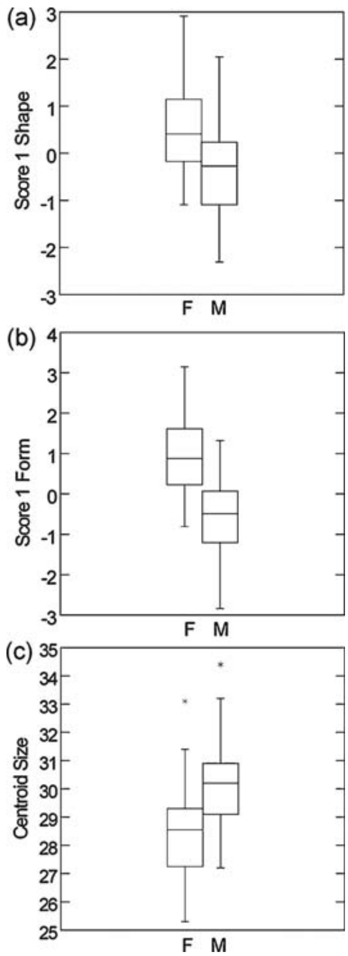

Figure 4. Box-plots of the discriminant scores obtained using variables of shape (a), form (b) and centroid size (c). F: female; M: male.

72.15% for males, Table 3). These results show that k-means clustering analysis yield similar results than discriminant analysis.

Figure 4c shows the differences in centroid size between males and females. The values obtained for size dimorphism using the mean distance index and the ratio male/female were 6.196 and 1.066, respectively. The values of t and F were 6.231 ($p = 0.000$) and 38.82 ($p = 0.000$). This F value is greater than the one obtained using shape variables but yet is similar to the value obtained using form, showing the contribution of size variation to cranial sexual dimorphism in this sample.

Discussion

The results obtained in this study indicate a high intra- and inter-observer agreement in the location of the coordinates of points employed to describe craniofacial traits. According to ICC values, the individuals were consistently ordered along the first two relative warps using the data digitised in different sessions. Even though slight differences can exist in the location of some landmarks and semilandmarks, they have a small influence on the overall description of the cranial structures, which contributes to enhance the agreement among observations. Therefore, the classification of individuals as males or females is minimally influenced by the observer. Consequently, the location of the coordinates of points employed to describe craniofacial traits can be excluded as a source of error in sex estimation based on these geometric morphometric techniques.

Table 2. Percentages of correct estimations obtained with discriminant analysis and k-mean clustering based on shape variables. F: female; M: male

Sex	Discriminant analysis			k-means clustering		
	Correctly assigned	Incorrectly assigned	% correctly assigned	Correctly assigned	Incorrectly assigned	% correctly assigned
F	34	19	64.15	33	21	61.1
M	31	48	60.75	20	58	74.35
Total	65	67	62.12	53	79	68.90

Table 3. Percentages of correct estimations obtained with discriminant analysis and k-mean clustering based on shape (Procrustes coordinates) and size (centroid size) variables. F: female; M: male

Sex	Discriminant analysis			k-means clustering		
	Correctly assigned	Incorrectly assigned	% correctly assigned	Correctly assigned	Incorrectly assigned	% correctly assigned
F	40	12	76.9	38	14	73.07
M	17	62	78.48	22	57	72.15
Total	57	74	77.86	60	71	72.51

The techniques employed to estimate sex should achieve high levels of precision and accuracy. As in any protocol, repeatable procedures are required to ensure the reproducibility of the measurements taken by the same observer in different times or by different observers (Yezerinac *et al.*, 1992; Arnqvist & Martensson, 1998; von Cramon-Taubadel *et al.*, 2007). The existing differences, if any, should be quantified to evaluate their impact on the data that is being interpreted. But despite the importance of obtaining accurate and unbiased estimations of sex from osteological remains, studies of sex assessment were not concerned, until recently, with ascertaining the consistency among observers (Walrath *et al.*, 2004; Rogers, 2005; Walker, 2005; Williams & Rogers, 2006; Gonzalez *et al.*, 2007). The results of these studies indicate that some features of the skull present high levels of error, and as a consequence they are not as useful to estimate sex as it was previously thought (Walrath *et al.*, 2004; Rogers, 2005; Williams & Rogers, 2006). For instance, Walrath *et al.* (2004) found levels of concordance among observers that range between 0.21 and 0.92. The authors state that the degree of error is highly influenced by the ambiguity in the definition of variables and the availability of reference pictures or graphics to the assigned individuals. Williams & Rogers (2006) also found that some traits, such as the orbital shape and the size and roughness of malar and parietal bones, display extremely high levels ($> 10\%$) of intra-observer error. It must be noted that due to the characteristics of non-metric data, measurement error is tested for each trait. Conversely, our study considered error on specimen similarity assessing the overall effect of intra- and inter-observer error.

These results suggest that the combination of landmark and semilandmark-based techniques with multivariate analyses permit different observers to consistently order large sets of skulls. This remarks the relevance of the geometric morphometric techniques to overcome one of the problems related to morphoscopic techniques, i.e. seriating large samples in a replicable manner. When non-metric data are used, the main difficulties arise when dealing with intermediate morphologies, since subtle differences are hard to assess visually, while individuals with extreme morphologies can be rated more consistently (Walker, 2005).

Specific patterns of sexual dimorphism related to shape found in the sample studied are illustrated in the deformation grids (Figures 2b and c). These grids show that supraorbital region exhibit the greatest dimorphism, followed by mastoid and frontal processes, while the zygomatic process is the least dimorphic trait. The large superposition of both sexes along the first two relative warps (Figure 2a) indicates that, within this sample, the degree of dimorphism in shape is relatively low. As a consequence, the percentage of correct sex estimations based on shape variables was low, with average accuracies of 60.12 and 68.90% obtained by discriminant and k-means clustering analysis, respectively.

Conversely, when centroid size is included in the PCA, females and males exhibit large separation along the first component (Figure 3). The highest values of correct assignment were found when shape–size variables were used, 77.86 and 72.15% with discriminant and k-means clustering analysis, indicating that this structure displays marked sex differences in morphology, related to the larger size and robust features of males. This result indicates a greater degree of size over shape dimorphism in the Coimbra sample. The values of F for size ($F = 38.83$) and shape ($F = 10.217$) also show that males and females differ more in size than in shape. Our results agree with a recent study of sexual dimorphism in American Whites and Blacks based on landmarks of the whole skull, which also found an increase in accuracy when centroid size is included to perform the discriminant analysis (Kimmerle *et al.*, 2008). In these samples the multivariate analysis of landmarks and semilandmarks of craniofacial traits allows distinguishing between males and females when shape variables and centroid size are analysed together.

Although morphometric techniques allow precise and accurate sex estimations, one of the main problems with them is that discriminant functions are population specific (Franklin *et al.*, 2004). Several authors have called attention to the applicability of discriminant functions derived from one sample to others, and then specific standards for each population are developed to optimize the accuracy of identification (Giles & Elliot, 1963; Steyn & Iscan, 1998; Franklin *et al.*, 2004). However, the pattern of sexual dimorphism is unknown in many cases due to the lack of documented references, and prehistoric populations are an example of this. The results obtained in this study suggest that internal variation could be used to sex undocumented skeletal material, since similar levels of accuracy were obtained with discriminant analysis and k-means clustering, which do not require a reference sample of known sex (Tables 2 and 3). Consequently, the individuals could be grouped according to their morphological affinity without requiring previous information about their sex.

In summary, the results obtained indicate that the application of geometric morphometric techniques to estimate sex has some advantages over visual techniques and traditional morphometrics. Among them we can state that they are considerably more objective, since they are based on anatomical points which are well defined; result in low levels of intra- and inter-observer error; allow to adequately describe anatomical structures with few landmarks by including coordinates of points that describe contours (semilandmarks); have the ability to obtain consistent results in large series of skulls and finally, are suitable to discern and objectively compare the degree and pattern of sexual dimorphism in different populations. For example, this technique has been applied to large samples of populations from South America to analyse the interpopulation variation in sexual dimorphism in size and shape, as well as to estimate the sex using cranial traits (Perez, 2006; Gonzalez, 2008). Considering that the use of morphometric analyses has a wide tradition in bioarcheological and forensic research, and the recognised advantages of quantifying morphological variation, geometric morphometric techniques have high probabilities of being largely adopted in the analysis of sexual dimorphism. This is supported by an increasing number of studies published in the last few years which employ these techniques to analyse sexual variation in morphological traits (Rosas & Bastir, 2002; Franklin *et al*., 2004; Katina *et al*., 2004; Pretorius *et al*., 2006; Gonzalez *et al*., 2007, 2009; Kimmerle *et al*., 2008; Kranioti *et al*., 2009). Finally, it is recommended to expand the use of these techniques to other samples in order to evaluate population variation in sexual dimorphism, as well as to determine the accuracy of sex determination in populations of different ancestry.

Acknowledgements

The authors are grateful to Eugenia Cunha, Ana Luisa Santos, Nuno Porto and the staff of the Museu Antropologico de Coimbra, and Héctor Pucciarelli and the staff of the Museo de La Plata for access to the material under their care. We also thank Amelia Barreiro for assisting us with the proofreading of the manuscript.

References

Adams DC, Rohlf FJ, Slice DE. 2004. Geometric morphometrics: ten years of progress following the 'revolution'. *Italian Journal of Zoology* **71**: 5–16.

Arnqvist G, Martensson T. 1998. Measurement error in geometric morphometrics: empirical strategies to assess and reduce its impact on measures of shapes. *Acta Zoologica Academiae Scientiarum Hungaricae* **44**: 73–96.

Bookstein FL. 1991. *Morphometric Tools for Landmark Data: Geometry and Biology*. Cambridge University Press: Cambridge.

Bookstein FL. 1997. Landmark methods for forms without landmarks: localizing group differences in outline shape. *Medical Image Analysis* **1**: 225–243.

Buikstra J, Ubelaker D. 1994. *Standards for data collection from human skeletal remains*. Arkansas Archaeological Survey Research Series 44: Fayetteville.

Corner BD, Lele S, Richtsmeier JT. 1992. Measuring precision of three-dimensional landmark data. *Journal of Quantitative Anthropology* **3**: 347–359.

D'Agostino RB, Russell HK. 1998. Multivariate Bartlett test. *Encyclopedia of Biostatistics.* **4**: 2843–2844.

Fleiss JL. 1981. *Statistical Methods for Rates and Proportions*. 2nd edn. John Wiley: New York.

Franklin D, Milne N, Freedman L. 2004. Geometric morphometric study of cranial sexual dimorphism in selected indigenous populations of South Africa. *American Journal of Physical Anthropology* **123**(S38): 96.

Franklin D, Freedman L, Milne N. 2005. Sexual dimorphism and discriminant function sexing in indigenous South African crania. *Homo: Journal of Comparative Biology* **55**: 213–228. DOI: 10.1016/j.jchb.2004.08.001

Giles E, Elliot O. 1963. Sex determination by discriminant function analysis of crania. *American Journal of Physical Anthropology* **21**: 53–68. DOI: 10.1002/ajpa.1330210108

Gonzalez PN. 2008. Morfometría geométrica aplicada al análisis del dimorfismo sexual en restos óseos humanos de individuos adultos y subadultos. *Unpublished PhD Thesis*, Universidad Nacional de La Plata: La Plata.

Gonzalez PN, Bernal V, Perez SI, Barrientos G. 2007. Analysis of dimorphic structures of the human pelvis: its implications for sex estimation in samples without reference collections. *Journal of Archaeological Science.* **34**: 1720–1730.

Gonzalez PN, Bernal V, Perez SI. 2009. Geometric morphometric approach to sex estimation of human pelvis. *Forensic Science International.* In press. DOI: 10.1016/j.forsciint.2009.04.012

Gower JC. 1975. Generalized procrustes analysis. *Psychometrika* **40**: 33–51.

Graw M, Czarnetzki A, Haffner HT. 1999. The form of the supraorbital margin as a criterion in identification of sex from the skull: investigations based on modern human skulls. *American Journal of Physical Anthropology* **108**: 91–96.

Graw M, Wahlb J, Ahlbrecht M. 2005. Course of the meatus acusticus internus as criterion for sex differentiation. *Forensic Science International* **147**: 113–117. DOI: 10.1016/j.forsciint.2004.08.006

Hair JF, Tatham RL, Anderson RE, Black W. 1998. *Multivariate Data Analysis*, 5th edn, Prentice Hall: New Jersey.

Howells WW. 1973. Cranial variation in man: a study by multivariate analysis of patterns of difference among recent human populations. *Papers of the Peabody Museum no. 67*. Harvard University Press: Cambridge, MA.

Katina S, Šefčáková A, Velemínská J, Brůžek J, Velemínský P. 2004. A geometric approach to cranial sexual dimorphism in fossil skulls from Předmostí (Upper Palaeolithic, Czech Republic) Časopis Národního muzea, Řada přírodovědná. *Journal of Natural Museum* **173**: 133–144.

Kimmerle EH, Ross A, Slice D. 2008. Sexual dimorphism in America: geometric morphometric analysis of the craniofacial region. *Journal of Forensic Science* **53**: 54–57. DOI: 10.1111/j.1556-4029.2007.00627.x

Konigsberg LW, Hens SM. 1998. Use of ordinal categorical variables in skeletal assessment of sex from the cranium. *American Journal of Physical Anthropology* **107**: 97–112.

Kranioti EF, Bastir M, Sánchez-Meseguer A, Rosas A. 2009. A geometric-morphometric study of the cretan humerus for sex identification. *Forensic Science International.* In press. DOI: 10.1016/j.forsciint.2009.04.0

Lockwood CA, Lynch JM, Kimbel WH. 2002. Quantifying temporal bone morphology of great apes and humans: an approach using geometric morphometrics. *Journal of Anatomy* **201**: 447–464. DOI: 10.1046/j.1469-7580.2002.00122.x

MacQueen J. 1967. Some methods for classification and analysis of multivariate observations. *Proceedings of the 5th Berkeley Symposium on Mathematics, Statistics and Probability*. Vol 1, University of California Press: Berkeley; 281–297.

Manly BFJ. 1994. *Multivariate Statistical Methods: A Primer*. 2nd edn. Chapman & Hall: London.

Marini E, Racugno W, Borgognini Tarli SM. 1999. Univariate estimates of sexual dimorphism: the effects of intrasexual variability. *American Journal of Physical Anthropology.* **109**: 501–508.

Meindl RS, Russell KS. 1998. Recent advances in method and theory in paleodemography. *Annual Review of Anthropology* **27**: 375–399. DOI: 10.1146/annurev.anthro.27.1.375

Mitteroecker P, Gunz P, Bernhard M, Schaefer K, Bookstein FL. 2004. Comparison of cranial ontogenetic trajectories among hominoids. *Journal of Human Evolution* **46**: 679–698. DOI: 10.1016/j.jhevol.2004.03.006

Novotny V, Iscan M, Loth S. 1993. Morphologic and osteometric assessment of age, sex, and race from the skull. In *Forensic Analysis of the Skull*, Iscan MY, Helmer RP (eds). Wiley-Liss: New York; 71–88.

O'Higgins P, Jones N. 1998. Facial growth in *Cercocebus torquatus*: an application of three dimensional geometric morphometric techniques to the study of morphological variation. *Journal of Anatomy* **193**: 251–272. DOI: 10.1046/j.1469-7580. 1998.19320251.x

Paiva LAS, Segre M. 2003. Sexing the human skull through the mastoid process. *Revista do Hospital das Clínicas* **58**: 15–20. DOI: 10.1590/S0041-87812003000100004

Peterson J, Dechow PC. 2003. Material properties of the human cranial vault and zygoma. *The Anatomical Record Part A*: **274**: 785–797. DOI: 10.1002/ar.a.10096

Perez SI. 2006. Influencia de la deformación artificial del cráneo sobre la morfología facial: implicancias para los estudios de relaciones evolutivas inter-poblacionales. *Revista Argentina de Antropología Biológica* **8**: 37–55.

Pretorius E, Steyn M, Scholtz Y. 2006. Investigation into the usability of geometric morphometric analysis in assessment of sexual dimorphism. *American Journal of Physical Anthropology* **129**: 64–70. DOI: 10.1002/ajpa.20251

R Development Core Team. 2008. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org

Rogers TL. 2005. Determining the sex of human remains through cranial morphology. *Journal of Forensic Science* **50**: 493–500. DOI: 10.1520/JFS2003385

Rohlf FJ. 1990. Rotational fit (Procrustes) methods. In *Proceedings Michigan Morphometrics Workshop*, Rohlf FJ, Bookstein FL (eds). Museum of Zoology. University of: Michigan: Michigan; 227–236.

Rohlf FJ. 1993. Relative warps analysis and an example of its application to Mosquito wings. In *Contributions to Morphometrics*, Marcus LF, Bello E, García-Valdecasas A (eds). Monografías del Museo Nacional de: Ciencias Naturales: Madrid; 132–159.

Rohlf FJ. 2007. tps serie softwares. Available at http//life.bio.sunysb.edu/morph/.

Rohlf FJ, Slice DE. 1990. Extensions of the procrustes method for the optimal superimposition of landmarks. *Systematic Zoology* **39**: 40–59.

Rohlf FJ, Marcus LF. 1993. A revolution in morphometrics. *Trends in Ecology and Evolution* **8**: 129–132.

Rosas A, Bastir M. 2002. Thin plate spline analysis of allometry and sexual dimorphism in the human craniofacial complex. *American Journal of Physical Anthropology* **117**: 236–245. DOI: 10.1002/ajpa.10023

Sampson PD, Bookstein FL, Sheehan H, Bolson EL. 1996. Eigenshape analysis of left ventricular outlines from contrast ventriculograms. In *Advances in Morphometrics*, Marcus LF, Corti M, Loy A, Naylor GJP, Slice DE (eds). Nato ASI Series, Series A: Life Science Vol. 284, Plenum: New York; 131–152.

Santos ALC. 2000. A skeletal picture of tuberculosis: macroscopic, radiological, biomolecular, and historical evidence from the Coimbra identified skeletal collection. *Unpublished PhD Thesis*, Universidade de Coimbra.

Sheets HD. 2003. IMP-Integrated Morphometrics Package. Department of Physics, Canisius College, Buffalo, New York.

Sheets HD, Covino KM, Panasiewicz JM, Morri SR. 2006. Comparison of geometric morphometric outline methods in the discrimination of age-related differences in feather shape. *Frontiers in Zoology* **3**: 15. DOI: 10.1186/1742-9994-3-15

Sheets HD, Keonho K, Mitchell CE. 2004. A combined landmark and outline-based approach to ontogenetic shape change in the Ordovician Trilobite *Triarthrus becki*. In *Applications of Morphometrics in Paleontology and Biology*, Elewa A (ed.). Springer: New York; 67–81.

Shrout PE, Fleiss JL. 1979. Intraclass correlations: uses in assessing rater reliability. *Psychological Bulletin* **2**: 420–428.

Steyn M, Iscan MY. 1998. Sexual dimorphism in the crania and mandibles of South African whites. *Forensic Science International* **98**: 9–16. DOI: 10.1016/S0379-0738(98)00120-0

Viðarsdóttir SU, O'Higgins P, Stringer C. 2002. The development of regionally distinct facial morphologies: a geometric morphometric study of population-specific differences in the growth of the modern human facial skeleton. *Journal of Anatomy* **201**: 211–229. DOI: 10.1046/j.1469-7580.2002.00092.x

von Cramon-Taubadel N, Frazier BC, Lahr MM. 2007. The problem of assessing landmark error in geometric morphometrics: theory, methods, and modifications. *American Journal of Physical Anthropology* **134**: 24–35. DOI: 10.1002/ajpa.20616

Walker PL. 2005. Greater sciatic notch morphology: sex, age, and population differences. *American Journal of Physical Anthropology* **127**: 385–391. DOI: 10.1002/ajpa.10422

Walrath DE, Turner P, Bruzek J. 2004. Reliability test of the visual assessment of cranial traits for sex determination. *American Journal of Physical Anthropology* **125**: 132–137. DOI: 10.1002/ajpa.10373

Williams BA, Rogers TL. 2006. Evaluating the accuracy and precision of cranial morphological traits for sex determination. *Journal of Forensic Science* **51**: 729–735. DOI: 10.1111/j.1556-4029.2006.00177.x

Yezerinac SM, Lougheed S, Handford P. 1992. Measurement error and morphometric statistical power and observer experience. *Systematic Biology* **41**: 471–482.

