

# The application of traditional and geometric morphometric analyses for forensic quantification of sexual dimorphism: preliminary investigations in a Western Australian population

Daniel Franklin · Andrea Cardini · Ambika Flavel ·  
Algis Kuliukas

Received: 16 December 2011 / Accepted: 15 February 2012 / Published online: 8 March 2012  
© Springer-Verlag 2012

**Abstract** A current limitation of forensic practice in Western Australia is a lack of contemporary population-specific standards for biological profiling; this directly relates to the unavailability of documented human skeletal collections. With rapidly advancing technology, however, it is now possible to acquire accurate skeletal measurements from 3D scans contained in medical databases. The purpose of the present study, therefore, is to explore the accuracy of using cranial form to predict sex in adult Australians. Both traditional and geometric morphometric methods are applied to data derived from 3D landmarks acquired in CT-reconstructed crania. The sample comprises multi-detector computed tomography scans of 200 adult individuals; following 3D volume rendering, 46 anatomical landmarks are acquired using OsiriX (version 3.9). Centroid size and shape (first 20 PCs of the Procrustes coordinates) and the inter-landmark (ILD) distances between all

possible pairs of landmarks are then calculated. Sex classification effectiveness of the 3D multivariate descriptors of size and shape and selected ILD measurements are assessed and compared; robustness of findings is explored using resampling statistics. Cranial shape and size and the ILD measurements are sexually dimorphic and explain 3.2 to 54.3 % of sample variance; sex classification accuracy is 83.5–88.0 %. Sex estimation using 3D shape appears to have some advantages compared to approaches using size measurements. We have, however, identified a simple and biologically meaningful single non-traditional linear measurement (glabella–zygion) that classifies Western Australian individuals according to sex with a high degree of expected accuracy (87.5–88 %).

**Keywords** Forensic anthropology · 3D landmarks · Size measurements · Sexual dimorphism · Geometric morphometrics · Population standards · Sex discrimination

**Electronic supplementary material** The online version of this article (doi:10.1007/s00414-012-0684-8) contains supplementary material, which is available to authorized users.

D. Franklin (✉) · A. Cardini · A. Flavel · A. Kuliukas  
Centre for Forensic Science, The University of Western Australia,  
M420, 35 Stirling Highway,  
Crawley 6009, Australia  
e-mail: daniel.franklin@uwa.edu.au

A. Cardini  
Dipartimento di Biologia, Università di Modena e Reggio Emilia,  
via Campi 213,  
41100 Modena, Italy

A. Cardini  
Centre for Anatomical and Human Sciences,  
The Hull York Medical School,  
Heslington,  
York YO10 5DD, UK

## Introduction

Recent research trends in the discipline of forensic anthropology indicate a growing awareness of the importance (and requirement) for contemporary population-specific standards; the development, refinement and testing of alternative methodological approaches have also been given due consideration. In relation to assessing sex in human skeletal remains, the former issue is listed as a recommended practice in the ‘Best Practice’ guidelines published by the Scientific Working Group for Forensic Anthropology. Additional objectives of this working group involve the dissemination of research and technology relating

to the discipline and to identify best practice and future directions [1].

With regard to forensic practice in Western Australia (and Australia generally), there is a relative paucity of contemporary population-specific skeletal standards for the estimation of sex; in attempting to fortify the capability of forensic anthropologists in this country, we are developing standards for estimating sex that are applicable to a number of different skeletal elements, complete and/or fragmentary [e.g. 2, 3]. The formulation of accurate, robust (e.g. statistically quantified) and “useable” (e.g. able to be applied to incomplete bones) standards is clearly a desirable and necessary aim, but it is equally important that alternate methodological approaches be given due consideration; this ensures that only the most robust and accurate methods end up being disseminated into the wider forensic community. Ideally, this will result in the “end-user” (forensic scientist/investigator) being better prepared, such that they can select from a suite of available methods that most appropriately suit a given situation.

An important reason for developing and/or applying standards specific to the geographical region in which a forensic investigator is working is the well-established notion that the application of non-population-specific standards (e.g. for estimating sex) results in a reduction in classification accuracy and/or an unacceptably large sex bias [e.g. 4, 5]. In the forensic anthropological discipline, it is now (more than ever) crucial that methods are empirically tested and the associated degree of uncertainty in arriving at any final estimation is statistically quantified because this is a core requirement for the admittance of expert evidence testimony [6, 7].

Repositories of documented human skeletons have traditionally provided the necessary biological data for the development of anthropological standards. Such collections, however, are not globally available. In recent years there has been an increasing movement towards utilizing radiological imaging data that more appropriately represent contemporary individuals from a variety of geographically diverse populations. In terms of forensic validation studies, it has been demonstrated that 3D CT data are suitable for the reliable quantification of traditional osteometric measurements [e.g. 8, 9]. It has also been established that radiological (or digital) skeletal data can be used to formulate sexing standards that have a high degree of expected accuracy—above 80 % and even beyond 90 % (e.g. cranial CT scans [9], humeral digital radiographs [10], lateral skull cephalographs [11] and digital photographs of the glenoid cavity [12]). Such accuracy rates are comparable to, if not higher than, results achieved in traditional morphometric studies of the same bones [e.g. 13–15].

With regard to the refinement and testing of both novel and established traditional methodological approaches, the main aim of the present study is to explore how accurately cranial form predicts sex in a sample of adult Australians by

applying traditional morphometric and geometric morphometric methods to data derived from three-dimensional anatomical landmarks acquired from CT-reconstructed crania. From an original dataset of more than 1,000 pairwise ILD measurements, a series of traditional measurements that most accurately estimate sex in this population was selected. We concurrently assess and compare the effectiveness of three-dimensional multivariate descriptors of cranial size and shape for estimating sex. The robustness of all findings is then explored using resampling statistics.

## Materials and methods

### Subjects

The present study examines cranial multi-detector computed tomography (MDCT) scans of 100 male and 100 female adult individuals who were patients during 2010 and/or 2011 that presented at various Western Australian hospitals for clinical cranial evaluation. The scans were anonymous when received by the authors, with only sex and age data retained. The mean age of the male individuals is 36.4 years (range 18–63) and for the females it is 40.3 years (range 18–61). Specific information regarding individual ethnicity of participants is unavailable, but the sample is overall taken as being representative of a “typical” contemporary Western Australian population (e.g. primarily Caucasian—see [2]). CT scans presenting acutely abnormal morphologies resulting from trauma and/or pathology (e.g. serious fractures) were not included if it affected our ability to accurately locate the required cranial landmarks (see the following discussion). Research ethics approval was granted by the Human Research Ethics Committee of the University of Western Australia (RA/4/1/4362).

### Radiological data acquisition

The selection criteria were very specific: only scans with a slice thickness  $\leq 1$  mm and which included the entire region of interest (cranial vault to alveolar margins) were accepted. Cranial patient CT imaging was performed using MDCT on a Phillips Brilliance 64 scanner (Philips Healthcare, North Ryde, Australia). Following 3D volume rendering, the 3D coordinates of 46 anatomical landmarks were acquired (by AF) using OsiriX<sup>®</sup> (version 3.9–64 bit). The landmarks are accordingly defined and illustrated in Table 1 and Fig. 1, respectively.

### Geometric morphometrics

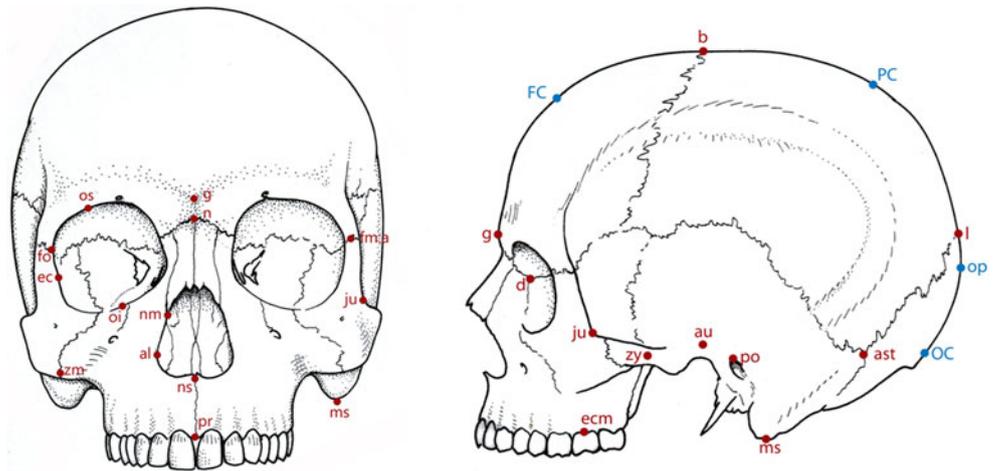
Raw landmark coordinates were imported in MorphoJ [16] as an ascii text file in the format as specified by the

**Table 1** Definitions of the landmarks used in the present study (see Fig. 1 for illustration)

Number	Landmark	Definition
Bilateral points (right and left)		
3 and 8	Ectoconchion (ec) <sup>a</sup>	The intersection of the most anterior surface of the lateral border of the orbit and a line bisecting the orbit along its long axis (parallel to the superior margin).
4 and 9	Superior orbital margin (os) <sup>b</sup>	The point on the superior orbital margin that meets with the short axis of the orbit, perpendicular to the superior orbital margin. [Used for orbital breadth and biorbital breadth measurements.]
5 and 10	Inferior orbital margin (io) <sup>b</sup>	The point on the inferior orbital margin that meets with the short axis of the orbit, perpendicular to the superior orbital margin. For an inside measurement. [Used for orbital height measurement]
6 and 11	Fronto-zygomatic orbitale (fo) <sup>b</sup>	The point where the frontozygomatic suture crosses the lateral orbital margin.
7 and 12	Fronto-malar anterior (fma) <sup>a</sup>	The most anterior point on the fronto-malar suture.
13 and 14	Alare (al) <sup>c, d</sup>	The most lateral point on the nasal aperture in a transverse plane.
17 and 18	Zygomaxillare anterior (zm:a) <sup>a</sup>	The most anterior point on the zygomaticomaxillary suture, on the lateral and inferior border of the bone. The intersection of the suture and the limit of the attachment of the masseter muscle.
19 and 33	Dacyron (d) <sup>c</sup>	Point on the medial border of the orbit that marks the junction of the sutures between the frontal, maxillary and lachrymal bones.
20 and 34	Nasomaxillary (nm) <sup>b</sup>	The most inferior point on the nasomaxillary suture.
21 and 32	Ectomolare (ecm) <sup>c</sup>	The outer surface of the alveolar margins, usually opposite the middle of the maxillary second molar tooth.
22 and 30	Jugalia (ju) <sup>a</sup>	The deepest points in the curvature between the frontal and temporal processes of the malars.
23 and 29	Auriculare (au) <sup>d</sup>	A point on the lateral aspect of the root of the zygomatic process at the deepest incurvature.
24 and 31	Porion (po) <sup>c</sup>	The highest point on the superior margin of the external auditory meatus.
25 and 28	Asterion (ast) <sup>a</sup>	The junction of the lambdoidal, parietomastoid and occipitomastoid sutures.
37 and 38	Foramen magnum lateral (fml) <sup>b</sup>	The point of greatest lateral curvature of the foramen magnum. [Used for foramen magnum breadth measurement]
39 and 40	Mastoidale (ms) <sup>c</sup>	The most inferior point on the mastoid.
41 and 42	Zygion (zy) <sup>c, d</sup>	Instrumentally determined as the most lateral point on the zygomatic arch.
Midline points		
1	Nasion (n) <sup>a, f</sup>	The junction of the internasal suture with the nasofrontal suture, in the median plane.
2	Glabella (g) <sup>c, e</sup>	The most forward projecting point of the forehead in the midline at the level of the supraorbital ridges and above the nasofrontal suture.
15	Nasospinale (ns) <sup>d</sup>	The point where a line drawn between the inferiormost points of the nasal aperture crosses the midsagittal line.
16	Prosthion (pr) <sup>a</sup>	The most anterior point in the midline on the alveolar processes of the maxilla.
26	Lambda (l) <sup>c</sup>	The point of intersection of the sagittal and lamdoidal sutures, in the midline.
27	Bregma (b) <sup>c</sup>	The point of intersection of the coronal and sagittal sutures, in the midline.
35	Opisthion (o) <sup>f</sup>	The midpoint of the posterior margin of the foramen magnum in the mid-sagittal plane.
36	Basion (ba) <sup>c</sup>	The midpoint of the anterior margin of the foramen magnum in the mid-sagittal plane
43	Frontal convexity (fc) <sup>b</sup>	The highest point on the convexity of the frontal bone in the midplane, relative to the nasion-bregma chord. [Used for nasion-bregma subtense and fraction measurements.]
44	Parietal convexity (pc) <sup>b</sup>	The highest point on the convexity of the parietal bones in the midplane, relative to the bregma-lambda chord. [Used for bregma-lambda subtense and fraction measurements.]
45	Occipital convexity (oc) <sup>b</sup>	The most prominent point on the basic contour of the occipital bone in the midplane, relative to the occipital chord. [Used for lambda-opisthion subtense and fraction measurements.]
46	Opisthocranion (op) <sup>f</sup>	Instrumentally determined most posterior point of the skull not on the external occipital protuberance.

<sup>a</sup>Howells [46]<sup>b</sup>Not a traditionally defined landmark; the definition presented here was devised and is used by authors at the Centre for Forensic Science (UWA)<sup>c</sup>Bass [47]<sup>d</sup>Buikstra and Ubelaker [48]<sup>e</sup>DeVilliers [49]<sup>f</sup>Wood and Lynch [50]

**Fig. 1** Anterior and lateral views of the cranium showing the landmarks used in the present study. Only selected landmarks are illustrated (see Table 1 for definitions)



programme user's guide. In this software, the shape coordinates were obtained using a Procrustes superimposition of the original raw landmark configurations [17]. This results in size being standardized and any differences in translation and rotation are minimized using a least square method. Centroid size (CS)—the square root of the sum of squared distances between the landmarks and their centroid (i.e. the mean of the  $X$ ,  $Y$  and  $Z$  raw coordinates)—was computed for each specimen; this was used as a biologically meaningful expression of the overall size of the landmark configuration and thus of the cranium [18].

Quantification of differences between bilateral landmarks was performed in MorphoJ using a hierarchical ANOVA to assess whether asymmetry was negligible compared to individual variation. Shape analyses were performed on the first 20 principal components (PCs) of the Procrustes coordinates, which were shown to provide an accurate summary of total shape variation using the approach of Fadda and Corti [19] and Cardini et al. [20]. Analyses designed to test the “robustness” of the results of the discriminant analysis data (see below) were performed as described in Viscosi and Cardini [21]. More detailed information about this and other statistical analyses can be found in the “[Electronic supplementary material](#)”.

#### Traditional morphometrics

The percentage of variance explained by sex was computed in PAST as the squared correlation coefficient ( $R^2$ ) between each ILD among all possible pairs of landmarks and a dummy variable for sex. The ILDs were ordered according to increasing  $R^2$  values; those measurements  $\geq R^2$  values of centroid size (which summarizes size differences, simultaneously taking all landmarks into account—see the preceding discussion) were selected and individually used to predict sex. This approach is a crude (albeit simple) method of exploring whether simple cranial linear distance measurements can be used to accurately estimate sex in this

particular population. Confidence intervals (CI) for the  $R^2$  estimates were computed using 1,000 bootstraps in *NTSYS-pc* [22] following the approach of Manly [23].

#### Statistical analyses of sexual dimorphism

Sexual dimorphism was assessed in PAST using a test for mean group differences with 10,000 permutations [24, 25]. Sex prediction accuracy was estimated by performing discriminant analyses (in PASW 18 [26]) using both a 50 % holdout-sample cross-validation and a jackknife leave-one-out approach.

For centroid size, shape and the linear measurement selected as an example of a potentially simple trait for a rapid and accurate assessment of sex, the proportion of individuals whose sex was estimated with a posterior probability higher than 0.80 was also computed, as well as the proportion of individuals with typicality probabilities lower than 0.05.

Following Sanfilippo et al. [27], the predictive accuracy of shape and size across a series of cutoff points was compared using a receiver operator characteristic curve (ROC) and the corresponding area under the curve (AUC); both were computed in PASW 18 [26] from the discriminant analysis posterior probabilities. The 95 % confidence intervals for AUC were also estimated; females were coded 0 and males were 1.

#### Results

Overall differences between the left and the right side of the cranium were small and negligible compared to the variation among individuals controlling for sex ( $P < 0.0001$  for both size and shape); the sides were thus accordingly averaged. Further justification of this approach is provided by the correlations of averaged

and original distances, which were always higher than 0.9 (range 0.92–0.98).

The percentages of variance ( $R^2$ ) explained by sex with 95 % confidence intervals (1,000 bootstraps) for centroid size, shape and 13 ILD measurements are shown in Table 2. Sex was always highly significant ( $P < 0.0001$  for all size and shape measurements). It is apparent that sex explains a much higher proportion of variance in size (49 %) than shape (3.2 %). When sampling error is estimated, the variance explained by centroid size ranges between 35.8 and 64.2 %; for shape, the corresponding figures are between 2.7 and 4.8 % (Table 2). The  $R^2$  values for the 13 ILD measurements are also shown in Table 2; only those measurements with  $R^2$  values  $\geq$  centroid size  $R^2$  are presented (see “Materials and methods”). The percentages of variance explained by the ILD measurements are between 49.3 and 54.3 %; the range of the corresponding low to high confidence interval figures is 36.1 to 69.1 % (Table 2).

The cross-validated sex prediction accuracy of centroid size, shape and the 13 ILD measurements is presented in Table 2, respectively. With regard to evaluating the predictive power of the multivariate cranial shape data (summarized by the first 20 PCs), it is evident that it is powerful enough to produce a relatively high cross-validated accuracy (83.5 % using the leave-one-out method) despite the large amount of variance unrelated to sex differences (Table 2). The associated mean male and female shape differences are illustrated in Fig. 2; females (relative to

males) have a flatter and narrower face, reduced anterior projection of the glabella, proportionally larger orbits, less inferior projection of the mastoid process and a more brachycephalic neurocranium.

In consideration of the sex prediction effectiveness of centroid size and the ILD measurements, it appears that size is in absolute terms a more accurate predictor of sex, with jackknife classification accuracies ranging between 82.5 and 88 % (Table 2). In further considering the ILD measurements that most accurately predict sex, the distance between glabella and zygion (GZ, mean of ILD 2-41 and 2-42) provides an example of the accuracy achievable (87.5–88 %) using a potentially simple linear measurement to rapidly quantitatively estimate sex in Australian adult crania. The selection of GZ is based on the fact that glabella and zygion are two of the most common anatomical landmarks in the ILD measurements outlined in Table 2 and that it has an  $R^2$  value higher than that of centroid size. Furthermore (and importantly) the distances between these two landmarks are relatively easy to measure in a standardized manner. With either shape, centroid size or GZ, very few specimens (less than 5 %) have a typicality probability smaller than 0.05 and at least 60 % (CS and GZ) and up to 80 % (shape) of individuals have posterior probabilities higher than 0.8.

ROC curves were computed for shape, centroid size and GZ (Fig. 3). The three curves are very similar and their AUC confidence intervals largely overlap (Table 3). These analyses suggest, therefore, that differences in accuracy may be

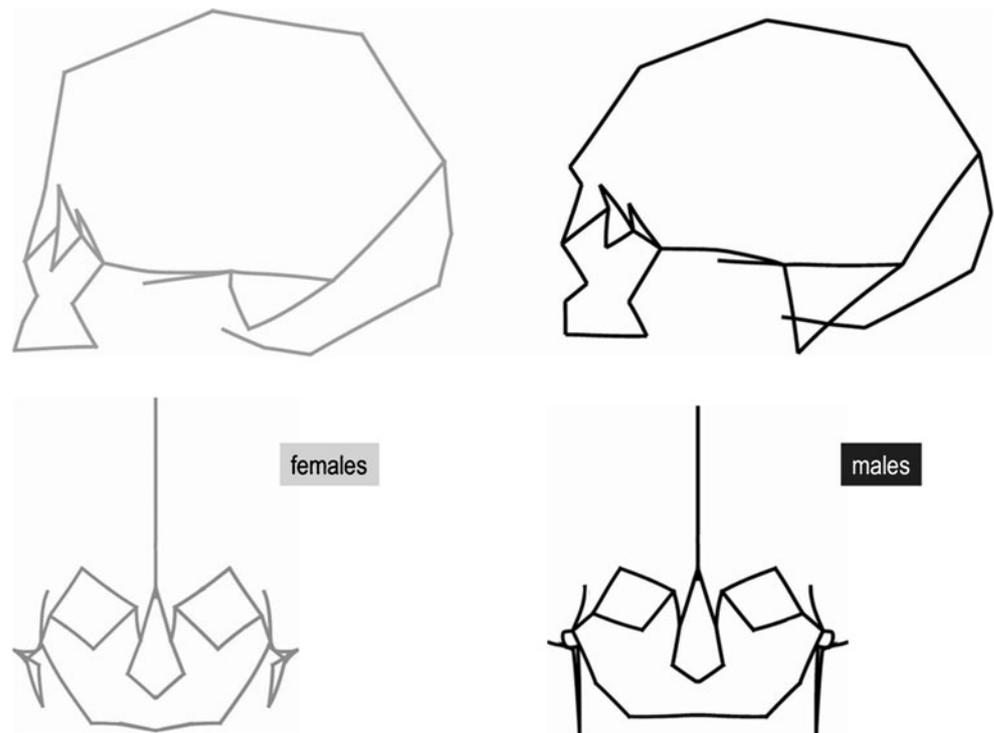
**Table 2** Percentages of variance explained ( $R^2$ ) by sex with 95 % confidence intervals estimated using 1,000 bootstraps and cross-validated sex-predictive accuracy for multivariate shape, centroid size, and the selected ILD measurements

Cranial morphology	$R^2$	Confidence interval		Correctly classified	
		Observed (%)	Lower (%)	Higher (%)	Holdout (%)
First 20 PCs of shape	3.2	2.7	4.8	73.0	83.5
Centroid size	49.0	35.8	64.2	84.0	85.5
ILD 12-25, 7-28	(fma-ast) 49.3	36.5	64.5	87.0	85.5
ILD 2-24, 2-31	(g-po) 49.6	36.1	64.9	83.0	82.5
ILD 7-40, 12-39	(fma-ms) 49.6	36.9	65.3	88.0	84.5
ILD 19-42, 33-41	(d-zy) 50.1	36.7	64.2	87.0	85.5
ILD 9-39, 4-40	(os-ms) 50.5	37.4	65.5	89.0	85.0
ILD 15-41, 15-42	(ns-zy) 50.9	37.4	66.0	84.0	84.5
ILD 1-41, 1-42	(n-zy) 51.0	37.4	65.6	86.0	83.5
ILD 4-28, 9-25	(os-ast) 51.0	38.2	65.5	88.0	86.5
ILD 2-25, 2-28	(g-ast) 51.9	38.8	67.4	83.0	85.0
ILD 34-42, 20-41	(nm-zy) 52.0	38.6	66.2	83.0	86.5
ILD 2-41, 2-42 GZ	(g-zy) 53.9	40.6	68.7	88.0	87.5
ILD 2-40, 2-39	(g-ms) 54.3	40.9	69.5	88.0	85.5
ILD 20-42, 34-41	(nm-zy) 54.3	41.2	69.1	86.0	88.0

Key to measurements is shown in Table 1

GZ glabella–zygion measurement

**Fig. 2** Lateral and anterior images showing the female (light grey) and the male (black) mean shape differences at  $\times 5$  magnification. In Procrustes-based shape analysis, the size of the landmark configuration is conventionally standardized to unit. For this reason, the female cranium appears to be as large as that of the male but with a comparatively smaller face and larger vault



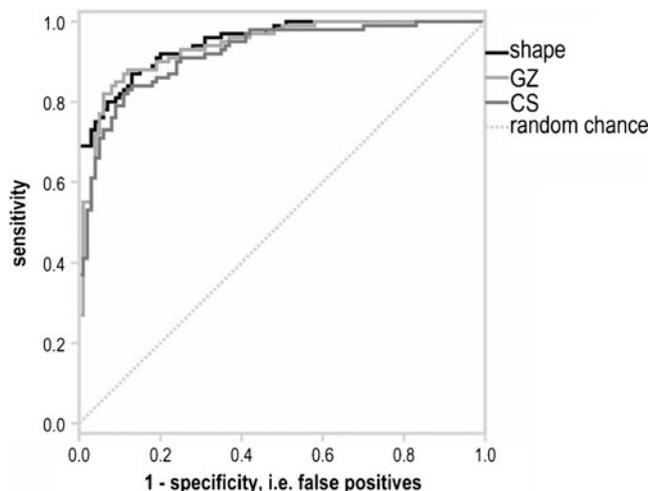
negligible. However, it is interesting to note that in terms of absolute cross-validated accuracy, GZ is more accurate than centroid size, which is in turn more accurate than shape; in relation to the AUC statistics, shape is more accurate than GZ, which is more accurate than centroid size. It is evident, however, that the overall difference between the highest and lowest AUC values is small (0.948 to 0.921; Table 3).

Further analyses of the GZ ILD distance were performed to fully quantify the statistical robustness of this simple linear measurement for predicting sex. These analyses indicate that: (1) sex is correctly predicted in 87.5–88 % of individuals (Table 2), (2) 65 % of individuals are not only

correctly sexed but have a high posterior probability ( $\geq 80\%$ ), (3) the thresholds for reaching the 0.8 posterior probability of correct classification in the discriminant analysis is  $GZ \leq 92$  mm for females and  $GZ \geq 97$  mm for males (Fig. 4), (4) the smallest GZ for males is 88 mm—therefore, all individuals with  $GZ \leq 88$  mm were females and (5) misclassified individuals are evenly distributed between males (12) and females (13), which indicates that there is no real bias towards one or the other sex in relation to the predictive accuracy of this measurement.

## Discussion

It is imperative in the forensic anthropological discipline that alternative (non-traditional) approaches to quantifying biological form are given due consideration for their potential forensic applications; it is equally essential that refinements to established approaches are explored. It is also

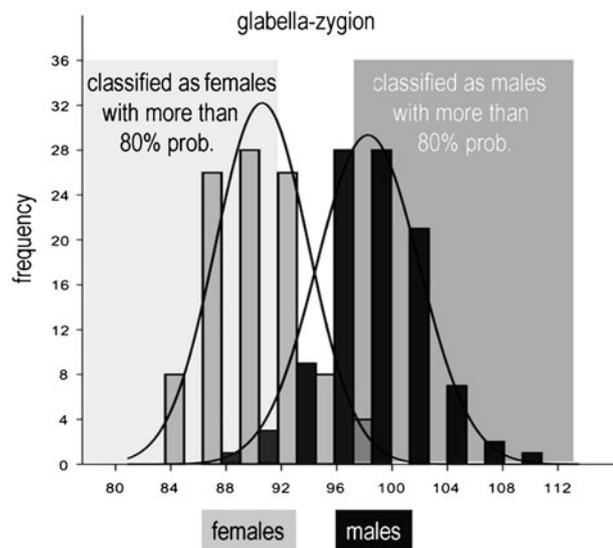


**Fig. 3** A comparison of ROC curves for shape, centroid size and the GZ ILD

**Table 3** AUC of ROC curves computed from the DA posterior probabilities for shape, centroid size and the GZ ILD

Cranial morphology	AUC	Confidence interval	
		Lower	Upper
Shape	0.948	0.922	0.975
GZ	0.939	0.907	0.970
CS	0.921	0.884	0.958

**Fig. 4** Glabella–zygion ILD measurement (shown on the lateral and anterior skull images) and associated frequency histograms with estimated normal curves for females and males. The approximate range of measurements where sex is estimated with a posterior probability higher than 0.8 (80 %) is emphasized with a *light grey* (females) and *dark grey* (males) background



important to consider that statistical quantification of any methodology and/or standard is necessary because such data are inherently intertwined with evaluating the admissibility of forensic expert testimony based on the methods and practice of science. To this end, this study is a proof of concept, whereby we show that using a sample of Australians, and despite missing information on its composition (specific individual ethnicities), we can effectively extract accurate shape and size sex predictors and have statistically quantified the accuracy of using cranial form to predict sex in a Western Australian population. A multifactorial approach was applied, whereby both traditional morphometric and geometric morphometric methods were applied to data acquired in CT-reconstructed crania.

Although in the present study we have focused on evaluating the sex prediction power of three-dimensional landmark configurations representing cranial size and shape, in addition to a single simple linear measurement (GZ), many other ILD size variables were shown to have comparable  $R^2$  values (see Table 2); therefore, they can also be potentially useful to estimate sex. A further justification of this is demonstrated by the high degree of sex classification accuracy for those ILD measurements (Table 2). It is important to note, however, that as all of those variables estimate the same quantity (e.g. size), they will be similarly affected by sampling; measurement choice, therefore, is likely to be a matter of practicality as some of the measurements may be easier to accurately acquire using traditional instruments (e.g. spreading and sliding calipers).

Other measurements (among those that have a high correlation with sex) might also be considered; they potentially

include some of the variables with  $R^2$  smaller than the arbitrary threshold (see “Traditional morphometrics”) we applied. For instance, using an ILD between nasion and asterion (a measurement that has an  $R^2$  just below our threshold but may be more likely preserved in fragmentary material) correctly classify sex in more than 82 % of individuals and has an AUC of 0.920, which is virtually identical to centroid size. Indeed that threshold was simply used as an expedient method of reducing the total number of predictors that we examined in this exploratory study from an original set of more than 1,000 variables. Using the same threshold in a different sample (or a different threshold in our study sample) might elucidate other interesting predictors that are simple to measure and thus enabling sex to be accurately estimated with a single linear measurement. The predictive accuracy in a linear discriminant analysis, however, will likely be similar because, as we already stated, these are all size measurements and therefore capture the same aspect of cranial variation. Resampling statistics suggest that these size estimators are reasonably stable in terms of their association to sexual dimorphism as they all produced comparably accurate predictions and similar confidence intervals around the observed  $R^2$  values.

With regard to the alternative simple linear GZ (and the other ILD) measurement(s), it is important to note that the data presented here represent a proof of concept; they are clearly non-traditional linear measurements that are practical to acquire, are biologically meaningful, and can be used to accurately predict sex. The sex classification accuracy of our single ILD measurements (range 82.5–88 %; Table 2) falls

within the range of (if not higher than) selected published research using established single traditional linear measurements (e.g. South African White—80.2 % [28], South African Black—78 % [5], Cretan—81.9 % [29], Japanese—84.6 % [30]).

Further exploration of the forensic utility of the alternative ILD measurements presented in the present study will ideally require quantification of their intra- and inter-observer error and evaluation of congruence between digital and caliper measurements in addition to estimates of generalizability to other samples within and outside Australia. Our previous research [3] using CT-reconstructed data and the same data acquisition methods, however, demonstrated that for 33 traditional linear measurements, the level of intra- and inter-observer error is well within accepted standards (e.g.  $rTEM < 5\%$ ; low TEM and high  $R$  values [31–33]). Furthermore, based on our analyses of purposely scanned skulls from the Centre for Forensic Science teaching collection, we have established that there are no significant differences in measurements acquired in CT scans compared to the same measurements taken in dry skulls using traditional instruments.

The present study has demonstrated in this population highly significant sexual dimorphism in cranial shape; the visualization of the associated mean male and female shape differences also elucidated a series of sexually dimorphic features, including the shape of the forehead contour, anterior projection of glabella and inferior projection of the mastoid (amongst others; Fig. 2). Such shape features, which are well established for distinguishing male and female crania (e.g. 34–36), are related to basic biomechanical and functional differences between the sexes (e.g. 37–40). It is, however, the relative expression of these morphological features, under the influence of different genetic and environmental influences, which varies between global populations [18, 41, 42]. This is in essence why contemporary population-specific standards are a necessary requirement in the formulation of a forensic biological profile.

In considering the cross-validated sex prediction accuracy of the three-dimensional configurations of the 46 anatomical landmarks analysed in the present study, it is evident that multivariate cranial shape and size both produced a relatively high accuracy rate (83.5–85.5 %; Table 2). Previous research using multivariate cranial shape data yielded a similar degree of expected classification accuracy (87 %) in a South African population [43]. Other published research in a variety of populations have indicated a wide range of prediction accuracy using cranial size and shape data (e.g. Portuguese—77.86 % [44], American Black and White—89.65 and 86.65 % [4], Bohemian—100 % [45]); a point of commonality between all those studies is that the inclusion of centroid size in the analysis considerably increased classification accuracy.

The ROC curves and AUC confidence intervals calculated for shape, centroid size and the GZ ILD measurement demonstrated that differences in overall accuracy are relatively negligible (Fig. 3; Table 3). It is evident, however, that consideration of the ROC and AUC data combined seems to imply that complex measurements like shape might, on the whole, lead to higher confidence in statements about sex discrimination compared to simple size measurements. Obviously, however, the use of three-dimensional classifiers is a specialization that may not always be available in many forensic laboratories. To that end, the use of straightforward linear measurements still yields accurate and statistically robust classifications of sex (Table 2); accordingly, a series of cranial linear standards are currently being devised based on the statistical analyses of an expanded Western Australian sample.

## Conclusions

In the present study, significant size and shape cranial sexual dimorphism was elucidated in a Western Australian population and the accuracy of using cranial form to predict sex was statistically quantified. It is evident that the cranium is a useful element for estimating sex in this population, with classification accuracy using geometric morphometric data, specifically three-dimensional multivariate descriptors of size and shape, are well above 80 %. The use of a single traditional linear measurement (glabella–zygion) is also able to classify Western Australian individuals according to sex with a high degree of expected accuracy (87.5–88 %). The present study has demonstrated the forensic potential of alternative and established traditional approaches for sex estimation in the Western Australian milieu.

**Acknowledgements** The authors would like to thank A/Prof. Rob Hart, Frontier Medical Imaging International, Western Australia, for assistance with obtaining and interpreting the CT scans. We also thank Dr. Paul Sanfilippo, University of Melbourne, for assistance with statistical applications and interpretations. The authors also thank the reviewers and editor for their helpful comments. Funding is provided by ARC Discovery Grant (DP1092538).

## References

1. SWGANATH (2011) The Scientific Working Group for Forensic Anthropology. <http://www.swganth.org>. Accessed 13 December 2011
2. Franklin D, Flavel A, Kuliukas A, Cardini A, Marks MK, Oxnard CE, O'Higgins P (2011) Estimation of sex from sternal measurements in a Western Australian population. *Forensic Sci Int*. doi:10.1016/j.forsciint.2011.11.008
3. Franklin D, Flavel A, Kuliukas A, Hart R, Marks MK (2012) The development of forensic anthropological standards in Western Australia. 64th annual meeting of the American Academy of Forensic Sciences. Atlanta, Georgia. February 20, 2012

4. Kimmerle EH, Ross A, Slice D (2008) Sexual dimorphism in America: geometric morphometric analysis of the craniofacial region. *J Forensic Sci* 53:54–57
5. Franklin D, Freedman L, Milne N (2005) Sexual dimorphism and discriminant function sexing in indigenous South African crania. *HOMO-J Comp Hum Biol* 55:213–228
6. Grivas CR, Komar DA (2008) Kumho, Daubert and the nature of scientific inquiry: implications for forensic anthropology. *J Forensic Sci* 53:771–776
7. Christensen AM, Crowder CM (2009) Evidentiary standards for forensic anthropology. *J Forensic Sci* 54:1211–1216
8. Verhoff MA, Ramsthaler F, Krahahn J, Deml U, Gille RJ, Grabherr S, Thali MJ, Kreutz K (2008) Digital forensic osteology—possibilities in cooperation with the Virtopsy® project. *Forensic Sci Int* 174:152–156
9. Ramsthaler F, Kettner M, Verhoff MA (2010) Digital forensic osteology: morphological sexing of skeletal remains using volume-rendered cranial CT scans. *Forensic Sci Int* 195:148–152
10. Kranioti EF, Nathana D, Michalodimitrakis M (2011) Sex estimation of the Cretan humerus: a digital radiometric study. *Int J Legal Med* 125:659–667
11. Hsiao JH, Tsai SM, Chou ST, Pan JY, Tseng YC, Chang HP, Chen HS (2010) Sex determination using discriminant function analysis in children and adolescents: a lateral cephalometric study. *Int J Legal Med* 124:155–160
12. Macaluso PJ (2011) Sex discrimination from the glenoid cavity in black South Africans: morphometric analyses of digital photographs. *Int J Legal Med* 125:773–778
13. Spradley MK, Jantz RL (2011) Sex estimation in forensic anthropology: skull versus postcranial elements. *J Forensic Sci* 56:289–296
14. Robinson MS, Bidmos MA (2009) The skull and humerus in the determination of sex: reliability of discriminant function equations. *Forensic Sci Int* 186:86.e1–86.e5
15. Dabbs GR, Moore-Jansen PH (2010) A method of estimating sex using metric analysis of scapula. *J Forensic Sci* 55:199–152
16. Klingenberg CP (2011) MorphoJ: an integrated software package for geometric morphometrics. *Mol Ecol Resour* 11:353–357
17. Rohlf FJ, Slice DE (1990) Extensions of the Procrustes method for the optimal superimposition of landmarks. *Syst Zool* 39:40–59
18. Viðarsdóttir US, O’Higgins P, Stringer C (2002) A geometric morphometric study of regional differences in the ontogeny of the modern human facial skeleton. *J Anat* 201:211–229
19. Fadda C, Corti M (2000) Three dimensional geometric morphometric study of the Ethiopian *Myomys–Stenocephalemys* complex (Murinae, Rodentia). *Hystrix* 10:131–143
20. Cardini A, DinizFilho JAF, Polly PD, Elton S (2010) Biogeographic analysis using geometric morphometrics: clines in skull size and shape in a widespread African arboreal monkey. *AMT Elewa* (ed) *Morphometrics for nonmorphometricians. Lecture notes in Earth Sciences* 124, Springer, Heidelberg. doi:10.1007/978-3-540-95853-6\_8
21. Viscosi V, Cardini A (2011) Leaves, taxonomy and geometric morphometrics: a simplified protocol for beginners. *PLoS One* 6(10):e25630. doi:10.1371/journal.pone.0025630
22. Rohlf FJ (2011) NTSYSpc, version 2.21k. Exeter Software, Setauket, New York
23. Manly BFJ (2001) Randomisation, bootstrap and Monte Carlo methods in biology. CRC, Boca Raton
24. Hammer Ø, Harper DAT, Ryan PD (2001) PAST: Paleontological Statistics software package for education and data analysis. *Palaeontologia Electronica* 4:9
25. Hammer Ø (2011) PAST ver. 2.09. <http://norges.uio.no/past/download.html>. Accessed 8 June 2011
26. SPSS (2009) PASW Statistics 18. Release version 18.0.0. SPSS, Inc., Chicago, IL. [www.spss.com](http://www.spss.com)
27. Sanfilippo P, Cardini A, Sigal IA, Ruddle J, Chua B, Hewitt A, Mackey DA (2010) Geometric morphometric assessment of the optic cup in glaucoma. *Exp Eye Res* 91:405–414
28. Steyn M, İşcan MY (1998) Sexual dimorphism in the crania and mandibles of South African whites. *Forensic Sci Int* 98:9–16
29. Kranioti EF, İşcan MY, Michalodimitrakis M (2008) Craniometric analysis of the modern Cretan population. *Forensic Sci Int* 180:110.e1–110.e5
30. Nagaoka T, Shizushima A, Sawada J, Tomo S, Hoshino K, Sato H, Hirata K (2008) Sex determination using mastoid process measurements: standards for Japanese human skeletons of the medieval and early periods. *Anth Sci* 116:105–113
31. Reynolds M, Franklin D, Raymond MA, Dadour I (2008) Bloodstain measurement using computer fitted theoretical ellipses: a study in accuracy and precision. *J Forensic Ident* 58:469–484
32. Ward RE, Jamison PL (1991) Measurement precision and reliability in craniofacial anthropometry: implications and suggestions. *J Craniofac Genet Dev Biol* 11:156–164
33. Weinberg SM, Scott NM, Neiswanger K, Marazita ML (2005) Intraobserver error associated with measurements of the hand. *Am J Hum Biol* 17:368–371
34. Workshop of European Anthropologists (1980) Recommendations for age and sex diagnoses of skeletons. *J Hum Evol* 9:517–549
35. Buikstra JE, Ubelaker DH (1994) Standards for data collection from human skeletal remains. Arkansas Archaeological Survey, Arkansas
36. Walker PL (2008) Sexing skulls using discriminant function analysis of visually assessed traits. *Am J Phys Anthropol* 136:39–50
37. Hylander WL, Picq PG, Johnson KR (1991) Masticatory-stress hypotheses and the supraorbital region of primates. *Am J Phys Anthropol* 86:1–36
38. Lahr MM, Wright RVS (1996) The question of robusticity and the relationship between cranial size and shape in *Homo sapiens*. *J Hum Evol* 31:157–191
39. Rosas A, Bastir M (2002) Thin-plate spline analysis of allometry and sexual dimorphism in the human craniofacial complex. *Am J Phys Anthropol* 117:236–245
40. Bastir M, Godoy P, Rosas A (2011) Common features of sexual dimorphism in the cranial airways of different human populations. *Am J Phys Anthropol* 146:414–422
41. Johnson DR, O’Higgins P, Moore WJ, McAndrew TJ (1989) Determination of race and sex of the human skull by discriminant function analysis of linear and angular dimensions. *Forensic Sci Int* 41:41–53
42. Kemkes A, Gobel T (2006) Metric assessment of the “mastoid triangle” for sex determination: a validation study. *J Forensic Sci* 51:985–989
43. Franklin D, Freedman L, Milne N, Oxnard CE (2006) A geometric morphometric study of sexual dimorphism in the crania of indigenous southern Africans. *S African J Sci* 102:229–238
44. Gonzalez PN, Bernal V, Perez SI (2011) Analysis of sexual dimorphism of craniofacial traits using geometric morphometric techniques. *Int J Osteoarch* 21:82–91
45. Bigoni L, Velemínska J, Bruzek J (2010) Three-dimensional geometric morphometric analysis of cranio-facial sexual dimorphism in a Central European population of known sex. *HOMO-J Comp Hum Biol* 61:16–32

46. Howells WW (1973) Cranial variation in man: a study by multivariate analysis of patterns of difference among recent human populations. Peabody Museum of Archaeology and Ethnology, Massachusetts
47. Bass WM (2005) Human osteology: a laboratory and field manual of the human skeleton. Missouri Archaeological Society, Missouri
48. Buikstra JE, Ubelaker DH (1994) Standards for data collection from human skeletal remains. Arkansas Archaeological Survey, Arkansas
49. de Villiers H (1968) The skull of the South African Negro: a biometrical and morphological study. Witwatersrand University Press, Johannesburg
50. Wood CG, Lynch JM (1996) Sexual dimorphism in the craniofacial skeleton of modern humans. In: Marcus LF, Corti M, Loy A, Naylor GLP, Slice DE (eds) Advances in morphometrics. Plenum, New York, pp 407–414