Available online at www.sciencedirect.com
SCIENCE DIRECT®
Forensic Science International 147 (2005) 159–164
Forensic Science International
www.elsevier.com/locate/forsciint

The reliability of sex determination of skeletons from forensic context in the Balkans

Marija Đurić[a,*], Zoran Rakočević[b], Danijela Đonić[a]

[a]*Laboratory of Anthropology, Department of Anatomy, Medical Faculty, University of Belgrade, 4/2 Dr Subotica, 11000 Belgrade, Serbia*
[b]*Department of Radiology, Dental School, University of Belgrade, Serbia*

Abstract

In this study we have tested the applicability of morphological methods for sex assessment, based on seven pelvic and nine cranial traits, using contemporary Balkans population. The material involved in the study comprises 262 pelvic bones and 180 skulls of male individuals from two mass graves in Serbia. The material was examined separately by an experienced and an inexperienced physical anthropologists.

Sex was correctly estimated by the experienced anthropologist in 100% of individuals using all of the 16 pelvic and cranial criteria. In fact, sex differences in pelvic morphology were large enough to allow sexing the individuals with 100% accuracy. Among seven features observed on the pelvic bones, the least reliable single sex indicator was the width of the great sciatic notch (with accuracy of 79.15%). Looking at the skull alone, sex was correctly determined in 70.56% cases. It was shown that the most accurate single indicators among cranial methods was the robustness of the mandible (with accuracy of 70.93%), while the sharpness of the supraorbital margins was the least reliable indicator demonstrating accuracy in only 28.75% of crania.

Examination of the sample by an individual with training in physical anthropology, but no case experience, suggests that experience is likely to contribute moderately to the accuracy of the sex determination. Namely, the inexperienced anthropologist accurately assessed the sex of the sample 95.04% of the time; 4.06% less accurate than the experienced anthropologist. The two anthropologists showed the least agreement in scoring the ventral arc and composite arc on the pelvic bones.

Keywords: Forensic identification; Sex assessment; Pelvis; Skull; Skeletal remains

1. Introduction

The anthropological analysis for sexing skeletal material from a forensic context is an integral part of the identification process as it provides relatively fast and accurate data, which could narrow the police investigator's field of search. Although many bones of the skeleton present sexually related differences [1–5], a number of studies have demonstrated that the pelvis is the single most reliable area for sex determination [4–9]. Therefore, sexing of skeletons based on pelvic morphology has been widely used in past decades. However, despite the improvements of the methods and the introduction of different scoring systems [10–12], there is still a need for establishing population specific standards. Anthropologists participating in the identification process in the region of former Yugoslavia are faced with this problem particularly when dealing with body parts and poorly preserved individuals. This article presents practical issues

* Corresponding author. Tel.: +381 11 686 172; fax: +381 11 686 172.
E-mail address: marijads@eunet.yu (M. Đurić).

0379-0738/$ – see front matter © 2004 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.forsciint.2004.09.111

resulting from assessments of pelvic and cranial sex indicators of skeletal remains of Albanian populations from Kosovo.

2. Materials and methods

To examine the usefulness and accuracy of various morphological traits of the pelvis and cranium in estimating sex of the Balkan population, 262 pelvic bones and 180 skulls from a forensic context were examined. Donor individuals (262) were males excavated during 2001 and 2002 from mass graves in Batajnica near Belgrade. These mass graves were created in the spring of 1999 at the height of the NATO bombings of Serbia (including Kosovo). For this study we selected individuals who were recovered with documentation of their sex, and bodies with well preserved soft tissue and obvious sexual characteristics. Only materials from male individuals were used, because the mass graves contained only several female individuals. The excavation and identification was performed by the forensic team (including the authors of this paper) of the School of Medicine, University of Belgrade with the help of International Commission on Missing Persons (ICMP). The number of crania examined is less than number of pelvic bones because some of skulls were missing (or recovered separately), or crashed. All individuals identified at the time of the study (70 out of 262 examined) were Albanians from Kosovo.

The pelvis and crania of individuals were assessed and scored for each morphological sex indication separately, but the final judgement for each case was based on complete pelvis and skull visual assessment independently. Seven pelvic variables were considered: robustness of the pelvic bones, presence of the preauricular sulcus, presence of the ischiopubic ramus ridge, shape of the subpubic angle, shape of the ventral arc, shape of the great sciatic notch, and shape of the composite arch. The crania were scored for nine common sex indicators: size of the mastoid process, size of the occipital protuberance, nuchal cresting, sharpness of the supraorbital margin, supercilliary arch form, prominence of the supramastoid ridge, robustness of the mandible, size of the mental eminence, and size of the frontal tuber. The skeletal material was not always completely preserved, so there is a lack of information on each trait of some specimens. The five-point scale, recommended by the Workshop of European Anthropologists [10], and later in Standards for Data Collection from Human Skeletal Remains [12] was initially applied in scoring each trait. However, in order to simplify further analysis we condensed the scoring into the simplified three-point scale [13] running from male to female, i.e. male, ambiguous, and female characteristic.

All scorings were made by two anthropologists: first being the principle author and the second a person trained in physical anthropology, but less experienced (second co-author). Skull and pelvic sex were judged independently in two phases: recording each variable and visual assessment based on overall morphology. The pathologists performing the autopsies did not pass on information regarding the sex of the individuals to the authors at the time of examination, in essence making the assessments 'blind'. Although 262 pelvic bones and 180 crania were available many of these items were subject to trauma and could not be completely assessed. Thus, the sample sizes used in the study vary from 227 to 262 for pelvis, and from 144 to 180 for skulls. The observed agreement beyond the chance and interobserver variability were tested by kappa statistic and interpreted following the guidelines introduced by Byrt [18].

3. Results

The experienced anthropologist, using all of the pelvic criteria observed, correctly estimated sex in 100% of

Table 1
Observed classification of pelvic morphological traits (experienced anthropologist)

Trait	Male		Ambiguous		Female		Total
	n	%[a]	*n*	%[b]	*n*	%[c]	
Robustness	246	93.89	16	6.11	0	0	262
Composite arc	257	98.47	2	0.77	2	0.77	261
Subpubic angle	255	99.22	2	0.78	0	0	257
Ventral arc	256	99.22	0	0	2	0.78	258
Ischiopub. r. ridge	211	92.95	0	0	16	7.05	227
Preauricular sulc	237	90.11	4	1.52	22	8.36	259
Great sciatic notch	205	79.15	51	19.69	3	1.16	259
Total of all characteristics	1667	93.49	75	4.21	41	2.30	1783
Total individuals	262	100.00	0	0	0	0	262

[a] Sensitivity of a test in % (true positive rate in %).
[b] Percentage of false negative (due to ambiguous).
[c] Percentage of false negative (due to female).

Table 2
Observed classification of cranial morphological traits (experienced anthropologist)

Trait	Male		Ambiguous		Female		Total
	n	%[a]	*n*	%[b]	*n*	%[c]	
Mastoid process	81	51.59	53	33.75	23	14.65	157
Occip. protuberan.	67	41.90	57	35.60	36	22.50	160
Nuchal crest	61	38.12	64	40.00	35	21.88	160
Supraorbit. margin	46	28.75	49	30.60	65	40.63	160
Supercilliary arch	69	46.00	53	33.13	28	18.67	150
Supramastoid ridg	72	40.00	68	37.78	40	22.22	180
Robustness mand.	122	70.93	38	22.90	12	6.98	172
Mental eminence	77	45.03	72	42.10	22	12.86	171
Frontal tuber	67	46.53	47	32.64	30	20.83	144
Total of all characteristics	662	45.53	501	34.46	291	20.01	1454
Total individuals	127	70.55	27	15.00	26	14.44	180

[a] Sensitivity of a test in % (true positive rate in %).
[b] Percentage of false negative (due to ambiguous).
[c] Percentage of false negative (due to female).

individuals. Looking at the skull alone, sex was correctly inferred in 127 (70.56%) cases, 26 (14.44%) skulls were identified as females, while 27 (15.0%) were considered ambiguous. The mean ratio of correct sexing of the pelvis (Table 1) was 93.49%, the least reliable pelvic sex indicators was the shape of the great sciatic notch with accuracy of 79.15%, while subpubic angle, ventral arc and composite arc showed correct sexing in over 98% of cases.

The mean ratio of correctly sexing of the skulls was only 45.53% (Table 2). The most accurate single sex indicator among cranial traits was the robustness of the mandible (with accuracy of 70.93%), followed by the size of the mastoid process (with accuracy of 51.59%). The sharpness of the supraorbital margins was the least reliable indicator showing the correct sexing in only 28.75% of crania. Applying the criteria proposed by Solowiej [14] and Novotný et al. [15], in which very reliable traits must correspond to actual sex in more than 60% of cases and misclassify less than 10%, to our sample material robustness of the mandible was the only trait that could be considered very reliable. Traits designated as reliable, which allow correct assessment in more than 50% of specimens and misclassify less than 15%, were included of only the size of the mastoid process, while the rest of the traits belong to the group of unreliable sexing traits.

The influence of the investigator's experience on sex assessment was tested on a separate occasion between an experienced and inexperienced anthropologist. The overall assessment of the inexperienced anthropologist achieved correct diagnosis in 95.04% (249 cases); the accuracy was 91.22% (239 cases) when the pelvis was used, and 54.44% (98 cases) when skulls were examined. The distribution of sex assessment of each characteristic is given in Tables 3 and 4. The results suggest similar distribution and frequency rate between the observers. Statistical analysis showed good agreement between anthropologists in total observation of pelvis (0.718), and very good agreement in observation of skulls (0.9035). Excellent agreement was

Table 3
Observed classification of pelvic morphological traits (inexperienced anthropologist)

Trait	Male		Ambiguous		Female		Total
	N	%[a]	*n*	%[b]	*n*	%[c]	
Robustness	218	83.21	40	15.27	4	1.53	262
Composite arc	234	89.66	2	0.77	25	9.58	261
Subpubic angle	256	99.61	1	0.39	0	0	257
Ventral arc	240	93.02	14	5.42	4	1.55	258
Ischiopub. r. ridge	213	93.83	4	1.76	10	4.41	227
Preauricular sulc	242	93.43	4	1.52	17	6.56	259
Great sciatic notch	183	70.66	72	27.80	4	1.54	259
Total of all characteristics	1586	88.95	137	7.68	64	3.59	1783

[a] Sensitivity of a test in % (true positive rate in %).
[b] Percentage of false negative (due to ambiguous).
[c] Percentage of false negative (due to female).

Table 4
Observed classification of cranial morphological traits (inexperienced anthropologist)

Trait	Male		Ambiguous		Female		Total
	n	%[a]	*n*	%[b]	*n*	%[c]	
Mastoid process	67	42.67	63	40.12	27	17.20	157
Occip. protuberan.	60	37.5	66	41.25	34	21.25	160
Nuchal crest	45	28.13	73	45.63	42	26.25	160
Supraorbit. margin	40	25.00	54	33.75	66	41.25	160
Supercilliary arch	60	40.00	65	43.33	25	16.67	150
Supramastoid ridg	68	37.78	72	40.0	40	22.22	180
Robustness mand.	120	69.77	40	23.25	12	6.98	172
Mental eminence	78	45.61	70	40.93	23	13.45	171
Frontal tuber	55	38.19	54	37.5	35	24.31	144
Total of all characteristics	593	40.78	557	38.31	304	20.91	1454

[a] Sensitivity of a test in % (true positive rate in %).
[b] Percentage of false negative (due to ambiguous).
[c] Percentage of false negative (due to female).

reached in estimation of robustness of the mandible, mental eminence and supramastoid ridge, while the ventral arc and composite arc showed poor or slight agreement (Table 5).

In addition, when we started scoring sex characteristics of the left pelvic bone of a case no. 7/88 (while the right bone was still being cleaned by the assistant), we noticed that the left bone showed mainly female characteristics (robustness, width of sciatic notch, preauricular sulcus, shape of ventral arc). Later comparison with the right pelvis bone revealed remarkable differences and obvious male characteristics (Figs. 1 and 2). After careful observation of the whole skeleton, complete asymmetry in the size of the left and the right leg were detected. Radiography revealed bony underdevelopment of the left leg followed by osteoporosis.

4. Discussion

The results of this study have shown that the observed aspects of skull morphology, which are usually considered as reliable sex indicators [14,15,17], were not successful in sexing this skeletal sample. It indicates that interpopulation variations, i.e. anthropological characteristics of certain populations, could considerably affect the accuracy of sex assessment. The crania of Albanian males investigated in the study are often characterised by: slight to moderate prominence of the glabella and supercilliary arches, no expressed rugosity of the occipital squama, sharp supraorbital margins, and smooth supramastoid relief. These are features that are generally considered as female characteristics. Although the accuracy of each single cranial trait does not exceed 52% (except for robustness of the mandible), by using a combination of all cranial criteria observed we were able to estimate sex with much better accuracy (70.56%). The probable reason for this phenomenon is that when there was ambiguity concerning one or more criteria, there were some other criteria obviously indicative of a male. However, the experience in sexing individuals from this population has most likely contributed to the general assessment being better than the analysis of particular traits, i.e. general visual

Fig. 1. Asymmetry of the pelvic bones that could confuse the sex determination.

Fig. 2. Radiography in frontal view showing morphological differences between affected and non affected side.

Table 5
Differences between frequencies of male characteristics recorded by two observers

Trait	Sample size (*n*)	Observed agreement (in %)	Kappa statistics
Composite arc	257	91.19	0.252
Subpubic angle	255	99.61	0.664
Ventral arc	256	93.79	0.189
Ischiopub. r. ridge	211	99.12	0.929
Preauricular sulc	237	94.21	0.585
Great sciatic notch	205	91.51	0.776
Total of pelvic characteristics	1783	95.46	0.718
Mastoid process	81	91.08	0.822
Occip. protuberan.	67	95.63	0.909
Nuchal crest	61	90.00	0.777
Supraorbit. margin	46	96.25	0.905
Supercilliary arch	69	94.00	0.878
Supramastoid ridg	72	97.78	0.953
Robustness mand.	122	98.84	0.972
Mental eminence	77	99.42	0.988
Frontal tuber	67	91.67	0.831
Total of cranial characteristics	1454	95.25	0.9035

assessment based on recognition of population specific characteristics was of crucial importance for the final judgement in each case.

The mean ratio of correct sexing of the pelvis was 93.49% in our study confirming that the pelvis is indeed the most reliable sex indicator, as it was pointed out in previous studies [6–9]. However, comparison between the reliability of pelvis and crania in sexing skeletal remains showed less discrepancy in some other investigations: pelvis 95%, skull 90% [11], pelvis 96%, skull 92% [16].

A good overall agreement between observers on sex assessment could be explained by the facts that both the experienced and the inexperienced anthropologists were trained in the same anthropological school, and that the majority of sex-related characteristics are easily observable even without long-term experience. However, indicators that showed poor agreement should probably not be recommended as sex indicators in routine forensic practice.

In addition, there is a little attention paid to the side to side difference in sex indicators of skeletal material. Although we did not examine this item systematically, our long-term experience with bones from both forensic and archaeological contexts, also suggested that there are not side to side differences important for sex diagnosis. However, as demonstrated with the body with asymmetric pelvic girdle and lower extremities (case no. 7/88) suggest the importance of investigating the entire skeleton.

In summary, when applied to Albanian population, cranial sex indicators appear to be less reliable because morphological characteristics of the male skulls could easily be attributed to female indicators. When dealing with pelvic sex indicators, recording of ventral arc and composite arc showed significant influence of the investigator's experience, so these indicators probably require better training of anthropologists. It should be also noted that asymmetry of the skeleton in some instances could lead to misinterpretation of sex, so it is useful to observe the entire skeleton whenever possible.

Acknowledgments

The authors wish to acknowledge all the members of the Institute of Forensic Science, University of Belgrade, and International Commission on Missing Persons, who supported our anthropological investigation during exhumation of mass graves in Batajnica.

References

[1] E. Giles, Sex determination by discriminant functional analysis of the mandible, Am. J. Phys. Anthrop. 22 (1963) 129–135.
[2] D.G. Steele, The estimation of sex on the basis of the talus and calcaneum, Am. J. Phys. Anthrop. 45 (1976) 581–588.
[3] M.Y. İşcan, P. Miller-Shaivitz, Determination of sex from the tibia, Am. J. Phys. Anthrop. 64 (1984) 53–58.
[4] W.M. Krogman, M.Y. İşcan, The Human Skeleton in Forensic Medicine, Charles C. Thomas, Springfield, IL, 1986.
[5] M.Y. İşcan, Rise of forensic anthropology, Yrbk. Phys. Anthropol. 31 (1988) 203–230.
[6] T.D. Stewart, Sex determination in the skeleton by guess and by measurement, Am. J. Phys. Anthropol. 12 (1954) 385–392.
[7] S.T. Genoves, Diferencias sexuales en el hueso coxal. Public del Instituto de Historia, Universidad Nacional Autonoma de Mexico, Mexico City, 1959.
[8] T.W. Phenice, A newly developed visual method of sexing the os pubis, Am. J. Phys. Anthropol. 30 (1969) 297–302.
[9] D. Uberlaker, C. Volk, A test of Phenice method for the estimation of sex, J. Forensic Sci. 47 (1) (2002) 19–24.
[10] D. Ferembach, L. Schwidetzky, M. Stloukal, Recommendation for age and sex diagnoses of skeletons, J. Hum. Evol. 9 (1980) 517–549.

[11] W.M. Krogman, M.Y. İşcan, The human skeleton in forensic medicine, second ed. Charles C. Thomas, Springfield, 1986.
[12] J. Buikstra, D. Uberlaker (Eds.), Standards for data collection from human skeletal remains, Arkansas archaeological survey research series no. 44, 1994, pp. 16–19.
[13] L. Konigsberg, S. Hans, Use of ordinal categorical variables in skeletal assessment of sex from the cranium, Am. J. Phys. Anthropol. 107 (1998) 97–112.
[14] D. Solowiej, Určováni pohlavi podle lebkz—overeni metod doporučovanzch rùznymi autory, Universitas Carolina, Prague, 1982.
[15] V. Novotný, M.Y. İşcan, S. Loth, in: Işcan, Helmer (Eds.), Forensic Analysis of the Skull—Morphologic and Osteometric Assessment of Age, Sex, and Race from the Skull, Wiley–Liss, New York, 1993, p. 72.
[16] R. Meindl, O. Lovejoy, R. Mensforth, L. Carlos, Accuracy and direction of error in the sexing of the skeleton: implication for paleodemography, Am. J. Phys. Anthrop. 68 (1985) 79–85.
[17] G. Ascádi, J. Nemeskéri, History of Human Life Span and Mortality, Akademiai Kiado, Budampest, 1970.
[18] T. Byrt, How good is that agreement? Epidemiology 7 (1996) 561–562 (Letter to Editor).