

Validation of the sex estimation method elaborated by Schutkowski in the Granada Osteological Collection of identified infant and young children: Analysis of the controversy between the different ways of analyzing and interpreting the results

Javier Irurita Olivares¹ · Inmaculada Alemán Aguilera²

Received: 16 October 2015 / Accepted: 2 March 2016 / Published online: 22 March 2016
© Springer-Verlag Berlin Heidelberg 2016

Abstract Sex estimation of juveniles in the Physical and Forensic Anthropology context is currently a task with serious difficulties because the discriminatory bone characteristics are minimal until puberty. Also, the small number of osteological collections of children available for research has made it difficult to develop effective methodologies in this regard. This study tested the characteristics of the ilium and jaw proposed by Schutkowski in 1993 for estimation of sex in subadults. The study sample consisted of 109 boys and 76 girls, ranging in age from 5 months of gestation to 6 years, from the identified osteological collection of Granada (Spain). For the analysis and interpretation of the results, we have proposed changes from previous studies because we believe they raised methodological errors relating to the calculation of probabilities of success and sex distribution in the sample. The results showed correct assignment probabilities much lower than those obtained by Schutkowski as well as by other authors. The best results were obtained with the angle and depth of the sciatic notch, with 0.73 and 0.80 probability of correct assignment respectively if the male trait was observed. The results obtained with the other criteria were too small to be valid in the context of Physical or Forensic Anthropology. From our results, we concluded that Schutkowski method

should not be used in forensic context, and that the sciatic notch is the most dimorphic trait in subadults and, therefore, the most appropriate to develop more effective methods for estimating sex.

Keywords Sex estimation · Infant · Forensic anthropology · Ilium · Jaw

Introduction

The estimation of sex and age are probably the two main parameters analyzed in physical anthropology to identify or contextualize skeletal remains. In particular, the estimation of sex, in addition to its obvious usefulness for human identification, enable providing more accurate estimates of the age, given that development processes are strongly linked to sex [1–3].

However, when the skeletal remains belong to children, the estimation of sex presents great difficulties, mainly for two reasons:

1. In contrast to the estimation of age, sex estimation in children is much more imprecise than in adult individuals. This is because the appearance of features with sexual dimorphism is mostly produced as a result of a change in hormone levels at puberty, so in individuals who have not yet reached it, the discriminating bone features are minimal [4]. Therefore, the methodology used to estimate the sex of adults (mainly through features of the skull and pelvis) is not applicable to children. In addition, other indicators such as chemical differences in bone, based on levels of citrate, calcium, phosphorus, and strontium, are also only seen in individuals who have reached reproductive age [4].

✉ Javier Irurita Olivares
javier.irurita@pucese.edu.ec

¹ Centro de Investigación. Pontificia Universidad Católica del Ecuador Sede en Esmeraldas (PUCESE), Esmeraldas, Ecuador

² Laboratory of Anthropology, Department of Legal Medicine, Toxicology and Physical Anthropology, University of Granada, Granada 18012, Spain

2. It has not been possible to develop good methods due to the scarcity of samples available for research [1, 4, 5]. They must be composed of well-preserved, identified, and more contemporary skeletons. Until today, the most important osteological collections that have been used for research are the Fazekas and Kósa (1978) [6], Spitalfields [7], Lisbon [8], Argentina [9], South Africa [10], and Chile [11] collections. However, these do not fully comply with the above minimum requirements.

Despite these difficulties, numerous studies have demonstrated the existence of morphometric differences between sexes in the skeleton of children, which has allowed for designing methods that provide an estimate of sex for this age group although, for now, these are much more inaccurate compared to adults. These methods include those using the morphological differences in the humerus [12], dental [13], visual features in the jaw and the pelvis [14], cephalometric methods [15] or based on geometric morphometric analysis [16].

The method published in 1993 by Schutkowski [14], based on the macroscopic analysis of the morphological differences in the ilium and jaw, may be the most prominent method so far due to its promising results and its simplicity. To develop it, 37 boys and 24 girls were used, between 0 and 11 years, belonging to the collection of Spitalfields [7]. Subsequently, numerous studies to validate his results have been published, including those carried out by Loth and Henneberg [17] with 7 boys and 12 girls from the collection of South Africa; by Sutter et al. [11] in 30 male and 55 female mummies from Chile; by Vlak et al. [18] in 23 girls and 33 boys from the collection of Lisbon; and Cardoso and Saunders [19], who used 57 boys and 40 girls also from the collection of Lisbon.

In general, these authors obtained very different results in their publications. For example, if the angle of the sciatic notch was valued, Schutkowski [14] obtained a 95.0 % chance of success if the male trait was observed, Sutter [11] achieved 75 %, and Vlak et al. [18] 54.5 %. Some authors believed that these discrepancies were due to the small size of the samples [20], others to the difficulty of defining the different indicators [19], and others because the results were separated by age groups [3].

Perhaps the main conclusion we can draw when comparing these studies is that sex estimation in infant skeletons is still an unattained goal for research in physical anthropology. Also, it

is possible that by discussing the results featured in these publications, the contributions that have been made in identifying those constraints posed by the estimation of sex in children is more important than the validation of each of the indicators studied. In this work, we sought to contribute to this statement by describing two major errors that, in our opinion, have been observed in most studies addressing this subject. These errors prevent proper comparisons between studies being carried out and they hinder the proper interpretation of the results.

Problem statement

The following is a widespread error in the interpretation of results when methods for the estimation of sex in subadults are evaluated. The usual approach is to estimate the sex in an identified population in order to obtain the success rate in the estimate for each of the sexes; in other words, the percentage of boys and girls showing a specific indicator is quantified [11, 12, 15–17, 19–24]. This approach informs us about the distribution of these morphological features in the studied population; however, it is far from actually indicating the ability of the method to properly estimate the sex of an individual.

Briefly, we can say that to evaluate a method regarding estimating the sex, the rate of success obtained for each sex should not be calculated. Instead, we need to calculate the probability of hitting the estimate if we see a male or female indicator, which offers a very different approach.

This brief explanation, really expressing a simple idea, may be complex to understand if it is not represented numerically, so a hypothetical example to illustrate the two possible interpretations that can be performed is shown in Table 1. This example depicts the hypothetical results of a study to validate a method for estimating sex. These results have been simplified and exaggerated. The number of boys and girls of the sample with each indicator is shown.

Interpretation A: sex was correctly estimated in 100 % of boys, whereas this only occurred in 50 % of girls. In conclusion, this is a good method to estimate sex in children but not for girls. This is a common but incorrect interpretation.

Interpretation B: When the male indicator is observed, there is a 66 % chance to be a boy, while when the female indicator is observed there is a 100 % chance of being a girl. So this is a very good method for identifying girls but bad for boys. This interpretation is more rare yet correct.

Table 1 Different interpretations of the results for validation studies using methods for estimating sex. Hypothetical results

	Boys ($n = 100$)	Girls ($n = 100$)	Interpretation B (right)
Male indicator	100	50	66 %
Female indicator	0	50	100 %
Interpretation A (wrong)	100 %	50 %	

The same results show completely opposite conclusions based on the interpretation given. It is very important to note that, from a forensic perspective, specific morphological features of the skeleton are the only data that can be analyzed by the anthropologist; consequently, the method must express the probability of being a boy or a girl when that feature is observed. We are not interested in the opposite interpretation, namely, to know the probability that an individual of known sex presents a particular trait. This causes many studies to evaluate a method for estimating sex to show confusing or conflicting conclusions, and for that reason it is not possible to use them in cases of forensic identification or for comparison with similar studies if their results are not reinterpreted.

This error has already been reported by other authors [25]; however, many validation studies still use the incorrect interpretation.

Different numbers of boys and girls in the sample

This error occurs when the number of male and female individuals in the sample is not the same, which occurs in most studies [11, 12, 14–24, 26–28]. In fact, we have not found any previous work that has taken this factor into account.

As explained in the previous section, validation studies should report the probability of a correct estimate of sex when a certain feature is observed. However, to obtain a representative probability of the actual population, a similar sex distribution should be used, assuming that this will be 1:1. In other words: if the number of male individuals is greater than the females, obviously the probability of observing a feature of male individuals will be higher.

To illustrate this error, we have used the same example presented in the previous section, but now it shows the results that would have been obtained if the number of female individuals were half the number of males (Table 2).

In this example, although it also shows that 50 % of girls each have a variant of the indicator studied, only the interpretation “B” of the results is affected, which now shows a higher probability of success when the male trait is observed (80 %). This means that if the number of females in the sample is lower than that of males, the probability of

making a correct estimate if we see a masculine trait is higher than it would be in a real population. This may be particularly evident in studies published by Schutkowski [14], Sheuer [20], and Sutter [11], in which the number of boys was much higher than girls (up to three times), so the probability of properly allocating sex was highly favored for male traits.

When only small samples can be used, as in the case of osteological collections of identified infant and young children, the solution to this problem should not be to obviously reduce the number of males or females in the studied sample because useful information would be lost; instead, the results must be weighed, increasing the importance of data obtained from the underrepresented sex.

Justification and objectives

As Rösing et al. [29] stated, methods for estimating sex can be divided between complex methods but effective and simple but less effective. The choice of each depends on the specific circumstances, such as the state of preservation of the remains, the clarity of the morphological indicators, the available technical equipment, the level of precision required, etc. Because of the aforementioned problems of the current methods used in Physical Anthropology, in the forensic context, and if resources or the state of the remains permit it, a DNA test is recommended. However, this technique is not applicable in cases of burned, degraded, or contaminated remains.

Although new technologies should be used to improve the methods for estimating sex, for example using geometric morphometric techniques or chemical analysis, we consider the macroscopic analysis remains the simplest and generally useful tool in many cases, even to provide first orientation information. Therefore, the overall objective of this study was to validate and readjust, if necessary, the criteria used by Schutkowski [14] for the estimation of sex.

To carry this out, we analyzed one of the best osteological collections available for research. It was recently acquired by the Laboratory of Anthropology of the University of Granada and consists of a large number of individuals compared with other similar collections. The skeletons are in very good condition and there is wide antemortem information obtained

Table 2 Different results obtained when the sample has a different distribution of sexes. Hypothetical results

	Boys ($n = 100$)	Girls ($n = 50$)	Interpretation B (right)
Male indicator	100	25	80 %
Female indicator	0	25	100 %
Interpretation A (wrong)	100 %	50 %	

through birth, death, and burial certificates for almost all individuals [30]. Furthermore, in this case, a new perspective for the analysis and interpretation of the results has been used, different from that used by most previous studies.

The specific objectives addressed in this study were as follows:

- Validate the method by Schutkowski [14] for estimating sex in infant and juvenile skeletons in a sample of Mediterranean origin.
- Providing new hit rates for correctly assigning sex for each of the indicators studied, avoiding methodological errors in the analysis and interpretation of the results observed in previous studies.
- Compare the results with those of previous studies in different populations.

Material and methods

The sample was obtained from the osteological collection of identified infant and young children located at the Laboratory of Anthropology of the University of Granada (Spain). This collection comes from the Cemetery of San Jose located in the city. This is a relatively recent sample (mid to late twentieth century), its state of preservation is almost perfect and thanks to the existence of official documents, it has extensive ante-mortem information, such as the precise date of birth and death, immediate and fundamental causes of death, last address, months of pregnancy in the case of fetuses, etc.

For the selection of the study sample, the following exclusion criteria were used: individuals with developmental, traumatic, or taphonomic alterations, which do not allow for correctly identifying the selected indicators, cause of death involving possible developmental changes or lack of “sex” data in official documents. Once these criteria were employed for the total of 230 individuals in the osteological collection, 185 (109 boys and 76 girls) were used for this study, aged 5 months of gestation to 6 years. The sample distribution by age and sex is shown in Table 3.

Table 3 Age at death (years) and sex distribution of sample

Age	Boys	Girls	Total
Fetus	5	5	10
0–0.99	84	53	137
1–1.99	9	4	13
2–2.99	2	6	8
3–3.99	4	3	7
4–4.99	2	2	4
5–6	3	3	6
Total	109	76	185

Ilium and left jaw of each individual was evaluated macroscopically by the first author in order to assign a sex estimate for each of the seven criteria used by Schutkowski [14]. Before compiling data for a study, it is essential that the investigator is trained in the use of these criteria to become familiar with the variability that can be displayed. To facilitate identification of the different indicators, the definition was accompanied by images which clearly show male and female morphology. The slightly modified criteria used by Schutkowski [14] and the corresponding images are listed below:

- Angle of the greater sciatic notch (Fig. 1): viewed from the ventral aspect and with the auricular surface side of the angle aligned vertically. This angle will be close or similar to 90° in the male sex and more obtuse, up from 90° in female individuals.
- Depth of the sciatic notch (Fig. 2): in boys is deeper than in girls.
- “Arch” criterion (Fig. 3): the “arch” formed by drawing an extension from the inferior side of the greater sciatic notch crosses the auricular surface in girls and passes over it in boys.
- Curvature of the iliac crest (Fig. 4): seen from the top of the ilium, the iliac crest shows a pronounced S-shape in boys and softer in girls
- Morphologic features in the mandible (Fig. 5):
 - A Protrusion of the chin region: it is more prominent and square in boys than in girls.
 - B Shape of the anterior dental arcade: the alveoli of the canine protrudes with respect to molars.
 - C Eversion of the gonion region: in boys this region is protruding, while in girls is more flattened.

Data collection was repeated by the same observer and a different observer in a total of 30 individuals randomly taken



Fig. 1 Angle of the greater sciatic notch

Fig. 2 Depth of greater sciatic notch



after completion of the study, in order to calculate the intra- and interobserver errors committed to each of the criteria employees. The kappa coefficient and percentage of agreement between the original and the repeated measures were calculated.

To check if the traits investigated were actually related to the sex of the individuals, the data were analyzed using chi-square analysis.

To validate the method of Schutkowski, the positive predictive values when each trait was observed was calculated. The percentage of successful estimates for each sex was not calculated, since this information does not actually represent the reliability of the method (see “Introduction”; interpretation A in Table 1) [25].

To calculate the probability of a correct estimate when female or male characteristics were observed was necessary to weight the results according to the different distribution of boys than girls in the sample (detailed explanation in “Introduction”; Table 2). For this, an equivalent number of

individuals who would show each trait if the sample had the same number of boys as girls (0.5 probability for each sex) was calculated:

$$A^1 = \frac{A \times 50}{X}$$

$$B^1 = \frac{B \times 50}{Y}$$

Where:

- *A*: Number of boys showing each trait
- *B*: Number of girls showing each trait
- *X*: Percentage of boys in the study sample
- *Y*: Percentage of girls in the study sample
- A^1 : Equivalent number of boys that would show this feature if the sample had 50 % of each sex
- B^1 : Equivalent number of girls that would show this feature if the sample had 50 % of each sex

Fig. 3 “Arch” criterion



Fig. 4 Curvature of the iliac crest

The correct probability (weighted) was calculated using the following formula:

Probability of correctly estimating sex with male features

$$= \frac{A^1}{A^1 + B^1}$$

Probability of correctly estimating sex with female features

$$= \frac{B^1}{A^1 + B^1}$$

The results are shown jointly for all ages and also separately for individuals under 3 months of age ($n=108$) and older ($n=77$). For the choice of this age as the cutoff, changes in hormone levels in children in this age group have been taken into account: in males, circulating levels of luteinizing hormone (follicle-stimulating hormone and testosterone), which are responsible for the emergence of sex differences, begin to rise around the tenth week of gestation and decline shortly before birth; after birth, these increase again and re-stabilize around 3 months of age. Until puberty, these levels are approximately equal between boys and girls [31–33].

Results

Table 4 lists the results for intra- and interobserver error.

The results obtained for the sexual dimorphism of the traits studied using chi-square analysis is shown in Table 5.

The number of male and female individuals showing each feature, as well as the weighted probability of correctly estimating sex with Schutkowski's method [14] are shown in Table 6.

Discussion

The criteria proposed by Schutkowski [14] for the estimation of sex in subadults from the morphological features of the ilium and jaw have been evaluated. To this end, one of the largest collections of children identified [30] has been used and the results have been interpreted differently from the way most previous studies have done. The values obtained were generally much lower than those posed by Schutkowski [14] (Table 6).

The results obtained for intra- and interobserver error have shown, in general, a bad reliability for the assignment of the

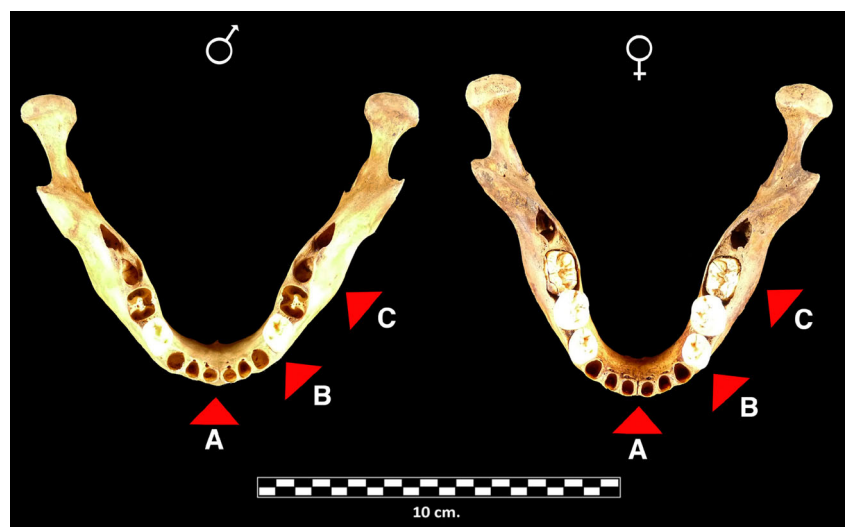
Fig. 5 Mandibular traits

Table 4 Intra- and interobserver error: Cohen's Kappa coefficient (*K*) and % agreement on traits studied (*N*=30)

	Intraobserver error		Interobserver error	
	<i>K</i>	% agreement	<i>K</i>	% agreement
Angle of the greater sciatic notch	0.79*	0.90	0.63*	0.86
Depth of the greater sciatic notch	0.87**	0.93	0.66*	0.83
Arch criterion	0.41	0.73	0.12	0.48
Curvature of the iliac crest	0.00	0.50	0.65*	0.83
Protrusion of the chin region	0.48	0.79	0.21	0.61
Shape of the anterior dental arcade	0.77*	0.86	0.63*	0.82
Eversion of the gonion region	0.56	0.79	0.45	0.72

* Good agreement; ** very good agreement

proposed traits. The best results were obtained with the angle and depth of the sciatic notch and with the shape of anterior dental arcade, which showed good or very good agreement according to the Kappa coefficient. All other traits are very difficult to identify because of the subjectivity with which they are defined. Comparing these results with those reported by Cardoso and Saunders [19], who analyzed the sex estimation by using the arc criterion, we can conclude that this trait should not be used to design methods for estimating sex in infant and young children. Unfortunately, research examining the reproducibility and repeatability of this method are very scarce, so it has not been possible to compare our results for other traits. However, this study suggests that the difficulty identifying the traits is an important factor that compromises the usefulness of this method.

A statistically significant relationship with sex for four of the traits studied was found when all age groups were analyzed, and for none of the features when only those who were younger than 3 months of age were analyzed (including fetuses) (Table 5); however, none of these groups obtained a success probability for the estimate greater than 0.67 for any criteria, values which are unacceptable in forensic contexts.

The best results were obtained in the group older than 3 months of age. In this case, the positive predictive value was 0.73 with the angle of the sciatic notch and 0.80 with sciatic notch depth, but in both cases only when masculine

traits were observed. If feminine traits were observed, the positive predictive value was lower (0.59 and 0.68, respectively). The other indicators in this age group also showed very low values.

Hormonal differences between females and males are conditioned by the presence or absence of testicles. Any individual, in the absence of androgens, will develop physiologically expressing typical female characters [34]. This is why, in our study and other previous studies that have attempted to estimate sex in children, the probability of success in the estimate is higher when male traits are observed. Another way to interpret this could be that in this age group it is easy to find boys who look like girls, but hard to find girls who look like boys.

With regard to the separation of the results by age group, there is a controversy between different authors. On one hand, it is widely known that the differences between the sexes appear gradually and accentuate the older the individuals are; for this reason, most authors have created artificial classifications of age to separate the results. Furthermore, recent studies suggested that these separations can increase the error assumed rather than reducing it [3]; instead, they proposed methods to estimate the sex that include the variable age as a continuous variable in the method itself. Indeed, these classifications by age group do not exist at the biological level because the changes occur gradually and continuously during development. When these ratings are used, they only respond

Table 5 Chi-square analysis for the sexual dimorphism of the traits studied

Trait	All ages (<i>n</i> = 185)	Under 3 months (<i>n</i> = 108)	Over 3 months (<i>n</i> = 77)
Angle of the greater sciatic notch	0.01*	0.17	0.02*
Depth of the greater sciatic notch	0.00*	0.17	0.00*
Arch criterion	0.68	0.39	0.08
Curvature of the iliac crest	0.04*	0.06	0.50
Protrusion of the chin region	0.00*	0.07	0.00*
Shape of the anterior dental arcade	0.44	0.88	0.26
Eversion of the gonion region	0.40	0.62	0.28

* Statistically significant sexual dimorphism

Table 6 Individuals with every feature and weighted positive predictive values for each sex

Trait	Estimated sex	All ages		Under 3 months		Over 3 months				
		N		Positive predictive values		N		Positive predictive values		
		♂	♀	♂	♀	♂	♀	♂	♀	
Angle of the greater sciatic notch										
≈90°	♂	38	13	0.667*	18	9	0.617	20	4	0.730*
>90°	♀	67	59	0.562	39	37	0.540	28	22	0.592*
Depth of the greater sciatic notch										
Deeper	♂	64	26	0.628*	35	22	0.562	29	4	0.797*
Shallower	♀	41	46	0.621*	22	24	0.575	19	22	0.681*
Arch criterion										
Bordering auricular surface	♂	71	47	0.511	37	34	0.472	34	13	0.586
Crosses auricular surface	♀	33	25	0.523	19	12	0.435	14	13	0.632
Curvature of the iliac crest										
Marked S-shape	♂	53	25	0.591*	26	13	0.622	27	12	0.540
Faint S-shape	♀	51	46	0.569*	30	33	0.572	21	13	0.543
Protrusion of the chin region										
Prominent	♂	61	25	0.640*	27	15	0.609*	34	10	0.665*
Non-prominent	♀	39	48	0.628*	25	30	0.581*	14	18	0.688*
Shape of the anterior dental arcade										
Protruding	♂	69	45	0.521	37	30	0.505	32	15	0.554
Not protruding	♀	31	26	0.542	15	13	0.512	16	13	0.582
Eversion of the gonion region										
Everted	♂	63	40	0.527	38	29	0.517	25	11	0.570
Not everted	♀	40	33	0.538	17	16	0.535	23	17	0.559

**p* < 0.05 (significant differences between sexes)

to a concrete objective, which is specific to each case, discipline, or area of study. Therefore, the limits of those age groups should be defined according to the characteristics of the process under study; other classifications should not be used in different contexts [35]. In this study, we have chosen to separate results by age group; however, this separation was not random: the selection of 3 months as the cutoff to separate the results was due to information relating to differences in hormone levels between boys and girls during development

[31–33], since the appearance of discriminating features depends on this.

Few studies have directly tested Schutkowski’s method, possibly because of the difficulty of accessing collections identified. To compare previous results with those obtained in this study, the data previously obtained was recalculated following the requirements outlined in this paper (Tables 1 and 2). Likewise, only the results for the group aged between 0 and 6 years was used (Table 7). Only the results

Table 7 Comparison of results with others studies

	Sample	N	Angle		Depth		Arch		Curvature		Chin		Arcade		Gonion reg.							
			♂	♀	♂	♀	♂	♀	♂	♀	♂	♀	♂	♀								
Schutkowski [14]	London	49	0.94	0.77	49	0.73	0.84	47	0.67	0.76	54	0.79	0.65	40	0.89	0.69	39	0.70	0.72	40	0.63	0.65
Sutter [11]	Chile	27	0.76	0.72	27	0.91	0.75	27	0.21	0.16	27	0.84	0.60	44	0.57	1.00	44	0.69	0.94	44	0.58	0.59
Vlak [18]	Lisbon	34	0.49	0.49	34	0.54	0.53															
This study	Granada	74	0.73	0.59	74	0.80	0.68	74	0.59	0.63	73	0.54	0.54	76	0.67	0.69	76	0.55	0.58	76	0.57	0.56

* In the case of the three other studies the ages ranged from 0 to 6 years; ** the values do not correspond to those shown in the original articles; these have been weighted according to the sex distribution in the sample; *** the values indicate the probability of success in the estimate when a particular feature is observed

published by Cardoso and Saunders [19] using the Lisbon collection were not able to be adapted due to the nature of the data. It can be seen that the positive predictive values obtained in this study was generally lower than those obtained by other authors, excluding those published by Vlak et al. [18], who obtained abnormally low data compared to the rest, even more if we consider that both the angle and depth of the sciatic notch were precisely the best criteria for the estimation of sex, as indicated by all other studies. The main difference from other studies was the results for the curvature of the iliac crest and chin, which proved to be very effective for the estimation of sex in the collections of London [5] and Chile [11], but in our study they provided very low probabilities.

The differences observed among these studies were probably due to differences in the characteristics of the samples, methods of analysis, or interpretation of results. However, examining the similarities, we can conclude that the most reliable traits were those that refer to the sciatic notch, in this case the angle and depth. Other studies did not directly analyze the method of Schutkowski, but they can also be used to corroborate our results. Garcia-Mancuso and Gonzalez [16] analyzed the geometric morphometric of the ilium of subadult individuals with a collection identified from Argentina recently acquired. In this study, they concluded, as in ours, that the sciatic notch was the most dimorphic trait between sexes in subadults, in particular the depth. However, these results should not be used to estimate sex in forensic contexts, since the positive predictive values are still small and intraobserver and interobserver error is not acceptable.

From now, it will be important to think about why these differences exist, in order to guide future studies to develop effective methods for estimating sex. In this regard, Wilson et al. (2014) suggest that this variability could be due to differences in growth rates between boys and girls, rather than morphological differences, which would be continually changing during development. Therefore, the best solution might be to propose new methods that include age as a continuous independent variable in the estimation of sex. This should be done through improved techniques of observation and statistical analysis, and using new technologies and resources available at our disposal.

In this study, the method proposed by Schutkowski [14] for the estimation of sex in subadults was tested. To do this, one of the more important collections of identified infant and young children available for research was used; however, it can still be considered insufficient to provide effective methods because it did not allow for actually knowing the effect of other factors such as population origin or age as independent variables. Although the conditions are not suitable, validation studies allow us to draw closer consensus, identify problems, and gradually perfect the methods of analysis.

Conclusions

Schutkowski's method is not acceptable for forensic purposes. The rates of correctly assigning sex in juvenile skeletons from the macroscopic analysis of the morphological differences in the ilium and jaw were lower than those published by Schutkowski [14]. Moreover, only three of the seven traits analyzed showed a reduced intra- and interobserver error: the angle and depth of the sciatic notch and the shape of the anterior dental arch. Finally, it is important to note that Schutkowski's method, as well as those that have been subsequently published to validate their results, have errors in the methodology of analysis and interpretation of results.

Acknowledgments The authors are grateful to D. Jose Antonio Muñoz, Managing Director, Maribel Martín, coordinator of services, and all EMUCESA staff at the San Jose cemetery in Granada for their invaluable assistance, and to the Magistrate Judge (Court of First Instance no. 5), responsible for the Registry Office of Granada, and to three anonymous reviewers for their insights and suggestions.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Saunders SR (2008) Juvenile skeletons and growth-related studies. In: Katzenberg MA, Saunders SR (eds) *Biological anthropology of the human skeleton*, 2nd edn. Wiley-Liss, New York, pp 117–146
2. Irurita J, Aleman I, Viciano J, De Luca S, Botella MC (2013) Evaluation of the maximum length of deciduous teeth for estimation of the age of infants and young children: proposal of new regression formulas. *Int J Legal Med*. doi:10.1007/s00414-013-0903-y
3. Wilson LAB, Ives R, Cardoso HFV, Humphrey LT (2014) Shape, size, and maturity trajectories of the human ilium. *Am J Phys Anthropol* 156(1):19–34
4. Lewis ME (2007) *The bioarchaeology of children*, 2nd edn. Cambridge University Press, Cambridge
5. Baker BJ, Drupas TL, Tocheri MW (2005) *The osteology of infants and children*. Texas A&M University Press, College Station
6. Fazekas IG, Kósa F (1978) *Forensic fetal osteology*. Ed. Akadémiai Kiadó. Budapest
7. Molleson T, Cox M, Waldron AH, Whittaker DH (1993) *The Spitalfields project volume 2—the anthropology, the middling sort*, Research Report 86. Council for British Archaeology, York
8. Cardoso HFV (2006) Brief communication: the collection of identified human skeletons housed at the Bocage Museum (National Museum of Natural History), Lisbon, Portugal. *Am J Phys Anthropol* 129:173–176
9. Salceda S, Desántolo B, García R, Plischuk M, Prat G, Inda AM (2009) Integración y conservación osteológica “Profesor Doctor Rómulo Lambre”: avances y Problemáticas. *Revista Argentina de Antropología Biológica* 11(1):133–141
10. Dayal MR, Kegley ADT, Strkalj G, Bidmos MA, Kuykendall KL (2009) The history and composition of the Raymond A. Dart Collection of human skeletons at the University of the

- Witwatersrand, Johannesburg, South Africa. *Am J Phys Anthropol* 140:324–335
11. Sutter RC (2003) Nonmetric subadult skeletal sexing traits: I. A blind test of the accuracy of eight previously proposed methods using prehistoric known-sex mummies from Northern Chile. *J Forensic Sci* 48(5):927–935
 12. Rogers TL (1999) A visual method of determining the sex of skeletal remains using the distal humerus. *J Forensic Sci* 44(1):57–60
 13. Viciano J, López-Lázaro S, Alemán I (2013) Sex estimation based on deciduous and permanent dentition in a contemporary Spanish population. *Am J Phys Anthropol* 152:31–43
 14. Schutkowski H (1993) Sex determination of infant and juvenile skeletons: I. morphognostic features. *Am J of Phys Antropol* 90:199–205
 15. Hsiao TH, 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
 16. García-Mancuso R, González PN (2013) Reconocimiento de rasgos dimórficos en ilion infantil mediante el uso de morfometría geométrica. *Cs Morfol* 15(1):1–11
 17. Loth SR, Henneberg M (2001) Sexually dimorphic mandibular morphology in the first few years of life. *Am J Phys Anthropol* 115:179–186
 18. Vlák D, Roksandic M, Schillaci MA (2008) Greater sciatic notch as a sex indicator in juveniles. *Am J Phys Anthropol* 137:309–315
 19. Cardoso HFV, Saunders SR (2008) Two arch criteria of the ilium for sex determination of immature skeletal remains: a test of their accuracy and an assessment of intra- and inter-observer error. *Forensic Sci Int* 178:24–29
 20. Scheuer L (2002) Brief communication: a blind test of mandibular morphology for sexing mandibles in the first few years of life. *Am J Phys Anthropol* 119:189–191
 21. Weaver DS (1980) Sex differences in the ilia of a known sex and age sample of fetal and infant skeletons. *Am J Phys Anthropol* 52:191–195
 22. Falys CG, Schutkowski H, Weston DA (2005) The distal humerus. A blind test of Rogers' sexing technique using a documented skeletal collection. *J Forensic Sci* 50(6):1289–1293
 23. Gapert R, Black S, Last J (2009) Sex determination from the occipital condyle: discriminant function analysis in an eighteenth and nineteenth century British sample. *Am J Phys Anthropol* 138:384–394
 24. Rogers TL (2009) Sex determination of adolescent skeletons using the distal humerus. *Am J Phys Anthropol* 140:143–148
 25. Klepinger LL (2006) Assignment of sex. In: Editor A (ed) *Fundamentals of forensic anthropology*. Wiley, Hoboken, NJ, pp 25–42
 26. Mittler DM, Sheridan SG (1992) Sex determination in subadult using auricular surface morphology: a forensic science perspective. *J Forensic Sci* 37(4):1068–1075
 27. Veroni A, Nikitovic D, Schillaci MA (2010) Brief communication: sexual dimorphism of the juvenile basicranium. *Am J Phys Anthropol* 141:147–151
 28. Gama I, Navega D, Cunha E (2015) Sex estimation using the second cervical vertebra: a morphometric analysis in a documented Portuguese skeletal sample. *Int J Legal Med* 129:365–372
 29. Rösing FW, Graw M, Marré B, Titz-Timme S, Rothschild MA, Röttscher K, Schmeling A, Schröder I, Geserick G (2007) Recommendations for the forensic diagnosis of sex and age from skeletons. *HOMO-J Comp Hum Biol* 58:75–89
 30. Alemán I, Irurita J, Valencia AR, Martínez A, Lopez-Lazaro S, Viciano J, Botella MC (2012) Brief communication: the Granada osteological collection of identified infants and young children. *Am J Phys Anthropol* 149:606–610
 31. Mann DR, Fraser HM (1996) The neonatal period: a critical interval in male primate development. *J Endocrinol* 149:191–197
 32. Andersson AM, Toppari J, Haavisto AM, Petersen JH, Simell T, Simell O, Skakkebaek N (1998) *J Clin Endocrinol Metab* 83(2):675–681
 33. Main KM, Toppari J, Skakkebaek NE (2006) Gonadal development and reproductive hormones in infant boys. *Eur J Endocrinol* 155:51–57
 34. Haseltine FP, Ohno S (1981) Mechanisms of gonadal differentiation. *Sci* 211:1272–1278
 35. Halcrow SE, Tayles N (2008) The bioarchaeological investigation of childhood and social age: problems and prospects. *J Archaeol Method Theory* 15:190–215