Int J Legal Med
DOI 10.1007/s00414-012-0769-4

ORIGINAL ARTICLE

Reliability of Schmeling's stages of ossification of medial clavicular epiphyses and its validity to assess 18 years of age in living subjects

R. Cameriere · S. De Luca · D. De Angelis · V. Merelli · A. Giuliodori · M. Cingolani · C. Cattaneo · L. Ferrante

Received: 21 May 2012 / Accepted: 27 August 2012

Abstract Nowadays, due to the global increase in migration movements, forensic age estimation of living young adults has become an important focus of interest. Minors often have no identification documents providing their correct birth dates. Establishing the age of majority is therefore fundamental in order to determine whether juvenile penal systems or penal systems in force for adults are to be applied. Radiological examination of the clavicles is one of the methods recommended by the Study Group on Forensic Age Diagnostics. In this retrospective study, a sample of chest radiographs of 274 subjects, aged between 12 and 25 years, was studied according to Schmeling's method in order to examine the ossification of both medial clavicular epiphyses. All stage classifications were evaluated by five examiners. Intra- and inter-examiner reliability was analysed by Cohen's K statistic. Intra-examiner agreement was insufficient for two of the experts. Inter-examiner agreement, among the other three operators, was moderate (K= 0.509). Study of reliability highlighted difficulties in interpretation, the need to select qualified personnel and choice of the best radiographic image in order to reduce any anatomic overlaps. Although ossification of the medial clavicular epiphyses is recommended to assess whether an individual has already reached the age of majority or not, these results suggested that it is very difficult to clearly identify the five stages of ossification by using conventional chest radiography.

Keywords Forensic sciences · Age estimation · Schmeling's method · Medial clavicular epiphyses · Italy

R. Cameriere · S. De Luca · A. Giuliodori · L. Ferrante
AgEstimation Project, Institute of Legal Medicine,
University of Macerata,
Macerata, Italy

S. De Luca (✉)
Department of Legal Medicine, Toxicology and Physical Anthropology, Faculty of Medicine, University of Granada,
Granada, Spain
e-mail: sluca@ugr.es

D. De Angelis · V. Merelli · C. Cattaneo
Laboratorio di Antropologia e Odontologia Forense (LABANOF),
Institute of Legal Medicine, University of Milan,
Milan, Italy

M. Cingolani
Institute of Legal Medicine, University of Macerata,
Macerata, Italy

L. Ferrante
Department of Biomedical Sciences and Public Health,
Faculty of Medicine, Polytechnic University of Marche,
Ancona, Italy

Introduction

Forensic Age Estimation constitutes a field of expertise in forensic medicine which aims at defining as accurately as possible the chronological age of unidentified individuals involved in legal proceedings [1, 2]. In the last two decades, Europe has received a flood of immigrants from countries which are not members of the European Union (EU). In many cases, these non-EU citizens have no documents which indicate their chronological age, or there are doubts regarding the correctness of the alleged chronological age. The situation is made even more complex in cases when immigrants are presumed to be minors. European legislation and international treaties ensure special treatment of unaccompanied minor immigrants because they are considered to be under the protection of European authorities [1]. This special treatment must also be ensured in cases of criminal proceedings in which minors have special protection as victims or as those responsible for criminal activities. In

Published online: 12 September 2012

all these cases, the various courts of law and public institutions require expert reports from medical forensic specialists [1–4].

Although numerous age estimation techniques have been developed in the fields of forensic anthropology and forensic odontology, there is still no consensus in European countries on which methods should be applied when the age of a presumed minor is to be estimated [1, 5]. In most countries, age assessment includes an interview in some form. In some cases, this is no more than a brief interview, and a visual evaluation of the person's age is made by a police or migration officer at the port of entry. In other cases, the interview is part of a more elaborate process carried out by trained personnel or social workers and may involve behavioural observations. Some countries rely on medical examinations, primarily in the form of radiographs of the skeleton and/or teeth. Anthropometric measurement and examination of sexual development are occasionally included [2–4, 6]. The choice of a particular method depends on the specific conditions of each case and mainly on the accuracy achieved [5–8].

The two main anatomic aspects, which are traditionally examined, are teeth [7–10] and the hand-wrist bones [11–15]. In most situations, these techniques were developed for biological purposes, which means that their applicability in the forensic field is relatively limited, both on grounds common to all methodologies (biological variability, characters of the selected sample, racial, social and health differences, etc.) and due to the limited amount of information at key ages. Particularly for the purpose of assessing the 18th year of age, such features are either fully developed, as is the case in both males and females for the hand-wrist bones [15], and as regards teeth, only the third molars are still developing. Although several studies have reported that the third molar development is a reliable predictor for defining the age of majority [16–19], variability in its development has been recognised, and its effect on forensic age estimations has been investigated with variable results [20]. In addition, as third molars are sometimes missing for reasons of agenesis or extraction, estimating the age of majority based on this tooth is impossible [1, 2, 9]. Therefore, and to confirm the dental evaluation, skeletal parameters are also considered.

One of the most extensively studied bones is the clavicle which, of all long bones in the human skeleton, has the slowest and most prolonged development [21–28]. From the embryological point of view, the clavicle consists of laminated bone [21]. The first mesenchymal differentiation centres arise between the 7th and 12th weeks of embryogenesis, followed by lateral and medial bone apposition via chondral ossification. However, bony fusion only occurs during early adulthood [22, 23].

Several studies have been conducted on the time frame for ossification of the medial clavicular epiphyseal cartilage in the age group of concern for forensic age diagnostics. One group of studies adopted an anatomical perspective, assessing ossification by means of autopsy or direct skeletal inspection [24–30]; the other group took a radiological approach. Radiological methods allow the medial clavicular epiphysis in living individuals to be examined and consist of conventional radiography [31–36] and computed tomography (CT) [37–41], as well as new approaches with magnetic resonance (MR) imaging [42, 43] and ultrasound [44, 45].

Both the Study Group on Forensic Age Diagnostics (AGFAD) and Forensic Anthropology Society of Europe, a subsection of the International Academy of Legal Medicine, have also introduced in their guidelines for forensic age estimation the possibility of recourse to X-rays, CT scans or MR imaging of the clavicle to establish the threshold of the age of majority [2, 6, 46–48]. The latest version of these recommendations can be found on the AGFAD website at http://agfad.uni-muenster.de/german/start.htm.

From biological and forensic standpoints, studies relying on CT and MR have proved to be extremely helpful in estimating age in subjects still undergoing the process of growth. CT scan, especially multislice acquisition, significantly improves spatial definition [41]. Progress in computer technology offers two- and three-dimensional reconstructions of high quality, with no spatial deformation of the features studied [41, 46]. Several studies have confirmed the value of CT scan for the assessment of changes in the medial epiphysis of the clavicle in relation to the chronological age of the subjects examined [49]. The disadvantages of CT scans may be, in some cases, use of contrast dye injections, like low osmolar contrast material and a disproportionately higher radiation dose than other diagnostic X-rays methods [50, 51]. In addition, in CT examination of the medial clavicular epiphysis, it has been noted that, if slice thickness is increased above a certain maximum value, this might lead to errors when assessing ossification status [52]. On the other hand, the MR examination can considerably decrease the radiation exposure using the necessary imaging procedures [42, 43, 50]. However, MR scanning may be expensive and uses contrast materials, like gadolinium, and its closed-in structure may create a claustrophobic sensation for some persons [50]. In addition, the real availability of such significant sample out of hospital environment is often slightly probable. Questions also exist regarding reliability and reproducibility, together with the same concerns over the effects of ethnicity, coupled with the range of normality for bone fusion. Unfortunately, to date, there is no good information on the application of these techniques to age assessment of asylum seekers [1–3].

In several cases, the need to estimate the age of majority is not limited to subjects involved in criminal proceedings but extends to individuals waiving their right to asylum or non-deportation [1, 5, 46]. The prohibition against sending, expelling, returning or, otherwise, transferring (refoulement) a refugee to "territories where his life or freedom would be threatened on account of his race, religion, nationality or membership of a particular social group" is recognised in Article 33 of the 1951 UN Convention on the Status of Refugees and its 1967 protocol and is enshrined in the texts of many other treaties [1, 4, 53–55].

Traditional radiography has been used as a tool to estimate the 21st year of age by Schmeling et al. in a study published in IJLM in 2004 [34]. Whereas traditional classification systems differentiate four stages of clavicle ossification (stage 1, ossification centre is not ossified; stage 2, ossification centre is ossified, and epiphyseal plate is not ossified; stage 3, epiphyseal plate is partly ossified; stage 4, epiphyseal plate is fully ossified), Schmeling et al. [34] divided the stage of total epiphyseal fusion into two additional stages (stage 4, epiphyseal plate is fully ossified, and epiphyseal scar is visible; stage 5, epiphyseal plate is fully ossified, and epiphyseal scar is no longer visible). The results of their study indicate that stage 4 includes subjects with a minimum age of 20 years in women and 21 years in men. Stage 5 comprises ages which are very far from legally relevant thresholds, as the minimum age is 26 for women and 26.7 for men. Subsequently, when the existence of stage 5 development is verified, subjects may be considered of legal age beyond reasonable doubt. This stage 5 is characterised by total fusion of the epiphyseal cartilage and the disappearance of the epiphyseal scar. As Schmeling et al. [34] noted, forensic age assessment of living subjects can assume that, when this stage is observed, the subject must have attained the age of 21 at least 5 years prior to the examination and the age of 18 at least 8 years prior to the examination. This stage becomes therefore relevant in cases where there is a period of several years between the crime and the court proceedings.

The main aim of this work is twofold: first, to determine the intra- and inter-observer agreement of clavicle X-rays in blind trials in order to obtain data on the reliability of this method; and second, to test the validity of Schmeling's method [34] in a sample of conventional radiographs in order to assess the suitability of clavicular development in discriminating whether or not an individual has reached the age of 18 years.

Materials and methods

A sample of 300 conventional postero-anterior radiographs of the chest was evaluated retrospectively to determine the ossification stage of the medial clavicular epiphyses. All images were taken at the Department of Radiology of Macerata Hospital (Italy) between 2010 and 2011. Protocols to collect radiographs for human subjects were approved by the Ethics Committee for Research Involving Human Subjects of the University of Macerata (Italy), and the study was conducted in accordance with the ethical standards laid down by the Declaration of Helsinki (Finland). The World Medical Association developed the Declaration of Helsinki as a statement of ethical principles for medical research involving human subjects, including research on identifiable human material and data.

The initial number of subjects recruited to the study was 300. All examinations regarding traumatologic or neoplastic changes of the clavicles were not included in the sample, although indications were not statistically evaluated. All examinations with poor image quality were also excluded. In 26 radiographs, reliable assessment of the degree of ossification of the medial clavicular epiphyseal cartilage on both sides proved impossible due to graphic overlaps or superimposition effects by the lungs, first rib and/or vertebrate column. The final number of study subjects was 274. Table 1 shows sample size by age and sex group for all subjects examined, 159 males and 115 females, aged between 12 and 25 years, in which reliable assessment of ossification status was possible. The ossification stages of the medial clavicular epiphyses were defined according to the staging system proposed by Schmeling et al. [34], as described above. If developmental differences between the left and right side were observed, which was the case in 47 cases (15.6 %), the more advanced side is chosen to assess the degree of ossification [40].

All stage classifications were analysed by five examiners: two of them, specialised forensic anthropologists, had slight experience in the radiological staging of clavicular ossification, and the other three, two forensic odontologists and one

Table 1 Age and sex distribution of studied sample

Age	Boys	Girls
12	2	6
13	5	2
14	14	4
15	11	7
16	7	1
17	10	8
18	13	13
19	18	13
20	21	12
21	11	8
22	19	22
23	10	15
24	11	3
25	7	1
Total	159	115

physician, are experienced in Schmeling's method, and they had also many years of education in the fields of radiology, anatomy and forensic research.

All participants were given a brief explanation of the purpose of the work, which was to assess intra- and inter-examiner repeatability in scoring ossification of the medial clavicular epiphyses. Each observer worked independently, and throughout the comparison phase, the observers were blinded to the identity of the subjects. Prior to and during analysis of radiographic images, the age of the individuals was not known to the examiners.

Statistical analysis

Statistical analyses were performed with the R statistical program [56]. The significance threshold was set at 5 %.

In order to evaluate intra-examiner reliability (repeatability), five examiners made randomly repeated measures of 63 plain chest radiographs after an interval of 3 weeks. Repeatability was studied with the weighted Cohen's K statistic [57]. Disagreements were weighted according to their squared distance from perfect agreement, i.e. quadratic weights were proportional to the square of the number of categories apart.

The same 63 plain chest radiographs scored above were compared between all the examiners. Inter-examiner reliability (reproducibility) was studied with Light's K statistic [58]. The 95 % confidence intervals of K values were based on 1,000 bootstrap replicates.

After analysis of intra- and inter-examiner agreement, sensitivity and specificity were used to analyse the validity of the data generated by the examiners [56]. Sensitivity, also called the true positive fraction (TPF), was defined as the proportion of correct estimates of age as 18 years or older. Specificity was the proportion of cases correctly classified as younger than 18 years of age. The false positive fraction (FPF), i.e. the proportion of cases incorrectly classified as older than 18 years, was the complement of specificity (1− specificity). The examiner's responses were expressed as the degree to which the observer believes the case to be a person aged more than 18 years as judged from ossification of the medial clavicular epiphysis, and each degree was used as a cutoff point to create an array of TPF/FPF pairs. Although it is a fundamental requirement of forensic identification that a test yield extremely high true positive and true negative decisions, it is also important that the test does not produce a high proportion of false positive identifications. False negatives are the less critical error since there is an opportunity to retest or use alternative methods to lead to a correct identification.

Taking into account, only the examiners' ratings showed very good repeatability (i.e. examiners whose K values were higher than 0.8), and three tests were therefore performed to discriminate between individuals who were or were not aged 18 years or more, considering clavicular ossification stages 2, 3 or 4 as cutoffs. As a result, the subjects were classified according to whether they were older or younger than 18 and their clavicular ossification stage. In addition, for each test, sensitivity (Se), specificity (Sp) and correct classification (accuracy of the test, i.e. the proportion of all correct results), E, were evaluated.

Results

Table 2 lists the mean age distribution (in years) and standard deviation at each developmental stage according to each examiner, for both sexes, in the study sample. Mean age gradually increased with each stage of ossification and also varied between boys and girls, but the differences were not significant. A statistically significant difference in mean age was recorded between stages when compared with the previous stage ($P<0.05$) for the five examiners, excluding examiner III, who found a non-significant difference between stages 3 and 4 (Fig. 1).

Table 2 Descriptive statistics for age by sex and ossification stage for each rater

Stage	Sex	Rater (mean ± SD)				
		I	II	III	IV	V
1	Girls	12.8 (1.8)	13.7 (2.3)	14.1 (2.2)	13.1 (1.8)	15.1 (3.2)
	Boys	13.6 (0.8)	14.7 (1.6)	15.4 (3.1)	13.9 (0.9)	16.1 (3.1)
2	Girls	17.3 (2.4)	18.5 (2.3)	18.7 (2.8)	17.4 (2.0)	18.6 (3.2)
	Boys	16.5 (2.6)	17.8 (3.2)	18.4 (2.2)	16.9 (2.2)	18.5 (3.8)
3	Girls	20.8 (2.0)	20.2 (2.7)	20.6 (2.5)	20.6 (1.8)	19.8 (2.5)
	Boys	20.4 (2.2)	20.4 (2.5)	20.3 (3.1)	20.2 (2.1)	19.9 (2.4)
4	Girls	22.4 (0.8)	22.0 (1.2)	21.1 (4.0)	22.4 (1.4)	22.2 (1.4)
	Boys	23.9 (1.0)	22.4 (2.2)	20.8 (4.7)	23.1 (1.5)	21.6 (3.0)
5	Girls	–	–	–	24 (−)	22.3 (0.5)
	Boys	–	–	24 (−)	24.5 (0.6)	22.3 (2.5)

SD standard deviation

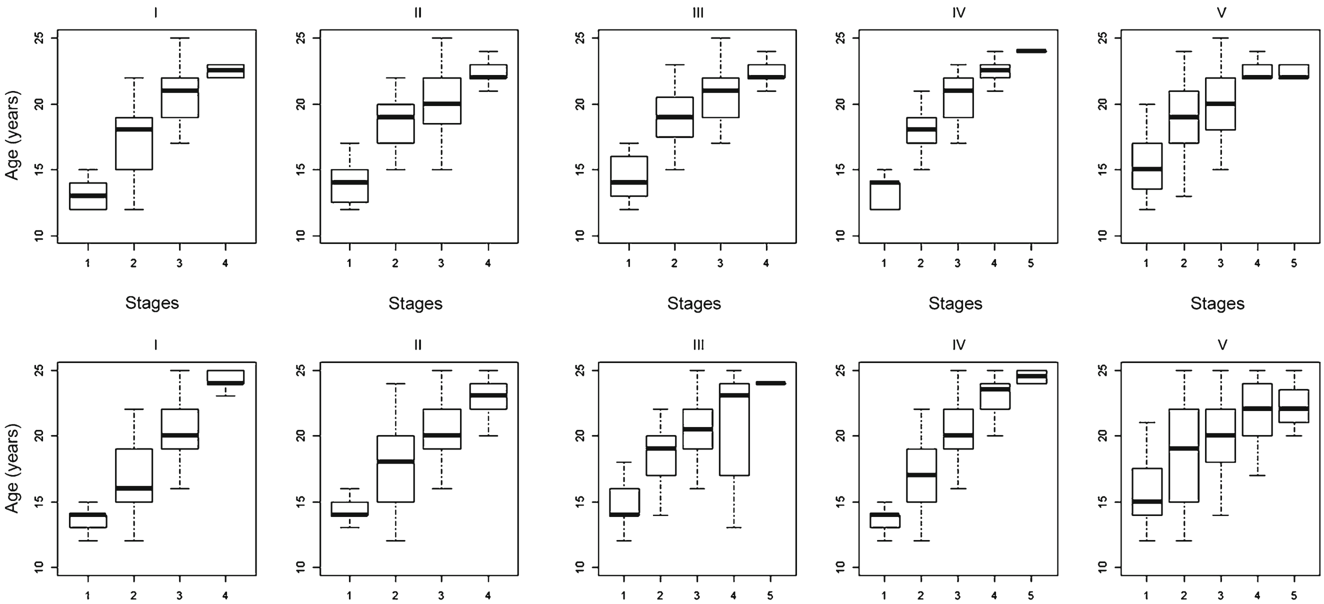

Fig. 1 Age distribution of girls (*upper row*) and boys (*lower row*) vs stages, evaluated for each observer

For intra-examiner repeatability of Schmeling' stages of clavicular ossification, good agreement was obtained between the two measurements made by the examiner "II" using the K statistic (95 % confidence interval), K=0.898 (0.823, 0.950), but moderate agreement between the two measurements made by the examiner "V", with K=0.523 (0.240, 0.734), was also observed.

The inter-examiner reproducibility of Schmeling's stages was not good with Light's K statistic (95 % confidence interval), K=0.482 (0.434, 0.520), indicating a moderate degree of homogeneity of evaluation among operators (Table 3). Also, when inter-examiner reproducibility was evaluated by excluding examiners who did not show very good repeatability (i.e. examiners V and I, with K values lower than 0.8), the K statistic persisted in being moderate at K=0.509 (0.451, 0.578).

The results of these three tests obtained considering clavicular ossification stages 2, 3 or 4 as cutoffs and performed to discriminate between individuals, who were or were not adult, are listed in Table 4, in which the sensitivity, specificity and overall accuracy of the observers were recorded. Observers II, III and IV are the forensic odontologists and the physician experienced in Schmeling's method and with several years of education and practice in radiology and anatomy.

The highest values of the hypothetical probability of chance agreement, E, were obtained at stage 2 (minors judged correctly), although the differences between stages 2 and 3 were not significant when examiners II and IV made the calculations (Table 4). When stage 3 was chosen as cutoff, their performance measures showed high variability, as sensitivity and specificity ranged from 63 to 86 % and from 61 to 90 %, respectively. In addition, as for forensic purposes, it is important for the test to show a low proportion of individuals younger than 18 whose test is positive (i.e. a high specificity level), and it seemed appropriate to pay more attention to the chance of a false positive than to that of a false negative. The choice of stage 2 to characterise adult age always yielded specificity lower than that obtained when stage 3 was chosen (Table 4). Finally, in case of choosing stage 4 as cutoff, the specificity increased up to a maximum of 100 %, but the overall sensitivity decreased (as judged subjects of age and older).

Discussion

In several countries, the need to ascertain 18 years of age is crucial in determining the procedures to be followed in cases of persons without definite identification. This problem is associated in particular with the rights of hospitality and with asylum seekers from non-EU countries [1, 5].

Table 3 Weighted Cohen's K and its bootstrapped 95 % confidence interval evaluated for each examiner

Examiner	K value	95 % confidence interval	
		Lower	Upper
I	0.711	0.569	0.807
II	0.898	0.823	0.950
III	0.815	0.662	0.910
IV	0.859	0.766	0.924
V	0.523	0.240	0.734

Table 4 Percentages of sensitivity, specificity and correct classification (95 % confidence interval) of test of adult age when stages 2, 3 and 4 of clavicular ossification are used

Stage	Sex	Examiner								
		III			II			IV		
		Se[a]	Sp[a]	E	Se[a]	Sp[a]	E	Se[a]	Sp[a]	E
2	M	92 (86–96)	37 (24–52)	75 (68–82)	96 (91–99)	40 (26–55)	79 (72–85)	97 (92–99)	45 (31–60)	81 (74–87)
	F	99 (94–100)	38 (21–58)	84 (76–90)	97 (90–99)	50 (31–69)	85 (78–91)	98 (92–100)	48 (29–67)	86 (78–91)
3	M	64 (55–73)	61 (46–74)	63 (55–70)	73 (63–81)	82 (69–91)	75 (68–82)	86 (78–92)	86 (74–94)	86 (80–91)
	F	63 (52–73)	76 (56–90)	66 (56–74)	70 (59–79)	82 (63–94)	73 (64–80)	83 (74–90)	90 (73–98)	85 (77–91)
4	M	12 (7–20)	90 (79–97)	36 (29–44)	21 (14–30)	96 (86–100)	44 (36–52)	26 (18–35)	96 (87–100)	48 (40–55)
	F	11 (6–20)	93 (77–99)	32 (23–41)	36 (26–47)	100 (82–100)	51 (42–61)	26 (17–36)	97 (82–100)	43 (34–53)

[a] Results are expressed as means (minimum values–maximum values) at 95 % confidence intervals

The number of unaccompanied children seeking asylum in Europe has increased in recent years. According to United Nations High Commissioner for Refugees (UNHCR) statistics for 2009, 15,100 unaccompanied children sought asylum in Europe in that year [1]. The increase was particularly high in Italy and Spain, which together received one third of all unaccompanied minors in Europe in 2009 [1, 4, 5]. Worldwide, UNICEF [61–63] estimates that about one third of all children are not registered with their countries' authorities before their fifth birthday. The great majority of unaccompanied children currently seeking asylum in Europe come from places (e.g. Africa, Eastern Europe) where the basic services of the state, such as birth registration, have not been functioning for many years. This makes it impossible for many unaccompanied minors to prove their age with official documents [1, 61–66]. The UNHCR has also produced documents which provide guidance on the use of age assessment processes for refugee and migrant children [64–66].

"Objective" methods to estimate the age of unaccompanied children and adolescents were therefore required, and in many countries, the medical profession has been asked to assist in this. A great variety of methods for age assessment has thus been integrated into immigration procedures across Europe, sometimes involving medical professionals [3, 4, 54, 55].

Regarding the responsible use of X-rays in forensic age estimations, it is indispensable not only to use dose-lowering techniques but also to take the presumed age of the person being examined into greater account [50]. Established doses for X-ray examinations in forensic age estimations vary from less than 0.1 mSv (left hand-wrist X-ray) to more than 800 mSv (CT) per examination [50, 69, 70]. Ramsthaler et al. [50] referred to a publication by Jurik et al. [71] who compared the exposure doses for spiral CT and conventional tomography of sternoclavicular joints, the values being 0.6 and 0.8 mSv, respectively. As regards chest radiography, the responsible use of X-rays demands critical selection of methods which are suitable for the specific age range. From this point of view, the use of conventional X-rays is not appropriate for individuals under 18, and like the collar bone CT, it should be recommended as a method only to answer the question of whether the subject is over 18/21 years [1, 4, 46, 50]. Article 3 of the European Council Directive 97/43/Euratom states that, when exposing an individual to ionising radiation, the net benefit to the individual must outweigh the risks [50, 72, 73]. Nowadays, more recent data [74, 75] are also available that compare radiation exposure from medical X-rays with the hazards of everyday living. Based on these studies, the resulting risks from using X-rays in age estimation procedures (with the exception of CT scan on sternoclavicular joints) is very low in comparison to other life risks [50]. However, the need for the moment is to minimise all radiation as far as possible and to evaluate risks from radiation when performing X-ray examinations in age estimation procedures [50].

Some authors suggest that care should be taken not to rely too much on the results of one single examiner, pointing out the importance of obtaining a second opinion [67, 68]. The same is true of age estimation methods. Some methods may not be as reliable as they are claimed to be, and the conscientious forensic researcher should always use several age estimation methods in order to check findings [67]. As regards the first aim of this study, to our knowledge, no previous study has been performed to determine intra- and inter-observer variations in analysis of the ossification of medial clavicular epiphyses according to Schmeling et al. [34].

The final results of this study showed that analysis of radiographic images of the medial clavicular epiphysis is neither simple nor easy to reproduce. The radiographs were analysed by five different observers, each one with different degrees of experience and education in assessing and interpreting chest X-rays. They had practised using the method before they studied these radiographies. However, when inter-observer reproducibility was evaluated, excluding observers who showed not very good repeatability (i.e. observers V and I, with K values lower than 0.8), the K statistic persisted in being moderate at K=0.509 (0.451, 0.578). This reveals considerable variation between observers, with moderate agreement. In inter-observer agreement studies, varying experience of examiners might influence the level of reproducibility [60]. In this work, the different levels of experience could have influenced the results.

Therefore, these results indicate that, at the present time, evaluation of conventional postero-anterior radiographs of the chest for age assessment has to be carried out carefully and/or by observers with ample specific experience and practice. It is commonly held that almost any physician is qualified to "read" X-rays, and that any dentist can evaluate dental X-rays. Still, there is opinion that simply being a physician does not in and of itself qualify one as a competent interpreter of a diagnostic imaging study. However, the more highly qualified and specialised the expert, the more weight his opinion will carry. Most juries will appreciate that the diagnostic images must be interpreted and explained to them by a competent expert familiar with both medical and technical factors applicable to the evidence. A physical anthropologist, for instance, may qualify as an expert to testify on radiographs of the skeleton but, probably, will be required to explain or document special training or experience in the use of this tool. Therefore, an expert in the field of radiology should be preferred. One previous study on experience level and its effect on radiological identification have shown, as expected, that those with the most medical and anatomy experience performed better than a group with less experience [76]. Koot et al. [77], in their study of hand radiographs, confirms the role of experience and training in radiographic human identification.

Misinterpretation of recognised findings may be due to a number of factors which can be grouped into several categories: e.g., *technical*, *perception* (search) and *analysis* (judgement/cognition/decision) [78, 79]. Technical errors are those related to the creation of the radiographic image and include: number and type of radiographic views, patient positioning, parameter selection, inappropriate use of collimation and grids, processing errors and other image artefacts (such as fog, extraneous objects). Although some types of technical error can be reduced with digital radiography, they are not eliminated, and several errors may be inherent in digital radiography. They include, among others, underexposure (insufficient X-ray exposure; quantum mottle or "graininess") and overexposure (excessive X-ray exposure; saturation and planking). In addition, consideration must be given to the medium on which the image is being viewed. For digital images, a properly calibrated, high-luminance, high-resolution grey-scale screen is essential. As regards patient positioning, X-rays of each clavicle should be taken separately. Care must be taken to ensure that the individual being examined is bending the shoulder so that the X-ray beams are perpendicular to the medial clavicle and the zone of ossification.

Perception is defined as *awareness* of the content of a radiographic image. For example, perception errors may occur when a specific feature is being sought. They include visual perception errors in the translation of a structure to three dimensions and also summation, silhouetting, magnification and distortion, and they often result in false positive analysis. Search errors are common and most often result from incomplete evaluation of radiographs [78, 79].

Probably, owing to the fact that the X-rays used in this study were not made for the purpose of examining the clavicle, in some of them, the (nearly fused) epiphyseal disc was more difficult to explore than in others. Consequently, this sometimes hindered correct classification of the clavicle. In addition, the medial epiphysis of the clavicle, like most epiphyses in the shoulder girdle region, is subjected to pronounced morphological variation ranging in appearance from small nuclei in fossa-like depressions to disk-shaped structures [80]. Normal anatomy and certain abnormalities are not well visualised on either overexposed or underexposed radiographs [80]. Therefore, a deficient knowledge of normal morphological variants becomes crucial while differentiating between normal and abnormal shape. The morphology of attachment area of costo-clavicular ligament on the clavicle, as well as the first rib, is a relatively neglected district that could potentially cause diagnostic errors [80].

In view of this and taking into account the different levels of experience of all five examiners, it is concluded that, even with extensive experience, Schmeling's protocol does not perform well if good-quality chest radiographs are not used. In addition, because of the ethical limitation in using X-rays, the use of non-ionising radiation methods is intuitively attractive. Future studies could show whether different imaging methods, such as computed tomography and MR, can be applied as reliable methods in age estimation practice. To our knowledge, both the thin-slice CT scans and the MR examination would lead to a considerable improvement of the diagnostic accuracy of age diagnosis [40–43]. However, as Cunha et al. [46] noted, sometimes the best methods are not those with the best published standard error, but those which have been tested by many researchers on many different populations, which are suitable for a specific forensic scenario, practical, user-friendly, relatively quick and economical. Further studies must be done to validate the MR and CT approaches to assessing age in normal populations before considering their use as a routine method for asylum seekers. Furthermore, the technology demands expensive equipment and specialist expertise limited to few locations.

In the forensic field, as what interests judges is whether the individuals in question have reached a specific threshold, some borderline results may make decisions difficult for judges [1, 2, 46]. For forensic practice purposes, their errors during application may be classified into two types: technically unacceptable errors and ethically unacceptable ones. Errors due to a forensic age estimate indicating that a subject actually over 18 is in fact a minor fall into the first class. However, errors indicating that a minor is over 18 may be classified as ethically unacceptable, since they lead to a violation of minors' rights. If they are judged to be under 18, they will be treated as children, with the full benefit of safeguarding through being children in care; a child incorrectly judged to be over 18 may be at risk of abuse or exploitation if placed with adult migrants [1, 5, 59]. There are also financial consequences for local authorities in looking after children and practical implications for deportation if an asylum claim is rejected [63]. Consequently, in forensic age diagnosis, test methods must reduce technically unacceptable errors to a minimum, but it is even more important for ethically unacceptable errors to disappear, especially in cases involving the possible criminal responsibility of the presumed minor [1, 5, 77].

In this study, after analysis of intra- and inter-examiner reliabilities, the validity of the methods under study was assessed by analysing the sensitivity and specificity values and the likelihood ratios for positive and negative results, according to different cutoff points. In fact, as Schmeling et al. [2] noted, in the case of living persons, an age diagnosis usually has to assess the probability that the examined subject has reached a specific legally relevant age threshold. However, the results of this study clearly showed that significant variability, showing fair agreement, existed between examiners. Due to this kind of inter-observer agreement and the values of sensitivity, specificity and accuracy (i.e. the proportion of all correct classifications), the results of these tests did not allow the authors to provide a cutoff statistically significant for estimation of the legal threshold of 18 years in living people.

Conclusions

The present study is the first to analyse intra- and inter-examiner agreement in order to eliminate technical inconsistencies and to standardise methods for assessing epiphyseal ossification. The major conclusion of this study is that, at the present time, there is an evident lack of reliability in assessing the medial epiphyseal ossification of the clavicle by X-rays for the purposes of estimating chronological age.

The results clearly indicate the following:

1. According to the low-to-moderate level of agreement for the final scoring of this sample, it is concluded that, although the instructions for scoring morphological traits and giving final scores may be clear [34], summation, silhouetting, magnification and distortion of the X-rays images often result in false positive analysis. Therefore, with standardisation of image acquisition procedures, further improvement of performance may be achieved. Instead of using conventional chest radiography, an alternative approach could be the application of thin-slice CT scans or MR imaging in order to try to improve the reliability of Schmeling's method. However, serious effort should be put also into research into refining the "non-ionising methods" approach and documenting its day-to-day applicability in the context of the enormous burden on immigration control caused by soaring rate of influx of migration.
2. The professional background and the degree of experience of the examiner are very important in age estimation in the living. In fact, the inter-observer variability may have been largely influenced by the differences in image acquisition or interpretation during the procedure. It is therefore imperative in the future to select qualified personnel with ample, specific experience in providing and analising radiographic images. Similarly, it would be desirable that specialised radiologists were involved in the process in order to generate accurate and less biased age estimates. However, this depends on rigorous training and use of auditable protocols.
3. Although Schmeling's stages of ossification [34] of the medial clavicular epiphysis for age estimation seem to produce results which can be used for forensic purposes, further research efforts on specific problems, such as inter-observer agreement and imaging methods, are required. Standardisation of this technique remains problematic, especially with regard to the correct evaluation of X-ray images. Given the above, it would be most valuable if minimum image quality requirements were to be drawn up for the different radiographic procedures.

Even so, until more rigorous definitions of such elements have been agreed upon, forensic researchers are recommended to use the conventional radiographic technique with extreme care or to perform it using high-quality images in order to maximise the accuracy in the assessment of age of majority. There is a real need for accurate education of staff in age assessment in order to improve capacity, consistency and competence.

Acknowledgments The authors would like to thank the staff of the Department of Radiology at Macerata Hospital (Italy) for their assistance on this project and also Ms. Gabriel Walton for editing the English text. The authors are also grateful to anonymous reviewers for their comments and suggestions which greatly improved the manuscript.

References

1. Schmeling A, Garamendi PM, Prieto JL, Landa MI (2011) Forensic age estimation in unaccompanied minors and young living adults. In: Duarte NV (ed) Forensic medicine—from old problems to new challenges. InTech, Rijeka, pp 77–120
2. Schmeling A, Geserick G, Reisinger W, Olze A (2007) Age estimation. Forensic Sci Int 165:178–181
3. Schmeling A, Reisinger W, Geserick G, Olze A (2006) Age estimation of unaccompanied minors. Part I. General considerations. Forensic Sci Int 159S:S61–S64
4. Hjern A, Brendler-Lindqvist M, Norredam M (2011) Age assessment of young asylum seekers. Acta Paediatrica; doi:10.1111/j.1651-2227.2011.02476.x
5. Garamendi PM, Landa MI, Ballesteros J, Solano MA (2005) Reliability of the methods applied to assess age minority in living subjects around 18 years old. A survey on a Moroccan origin population. Forensic Sci Int 154:3–12
6. Schmeling A, Grundman C, Fuhrman A, Kaatsch HJ, Knell B, Ramstahler F, Reisinger W, Riepert T, Ritz-Timme S, Rosing FW, Rotzscher K, Geserick G (2008) Criteria for age estimation in living individuals. Int J Legal Med 122(6):457–460
7. Foti B, Lalys L, Adalian P, Giustiniani J, Maczel M, Signoli M, Dutour O, Leonetti G (2003) New forensic approach to age determination in children based on tooth eruption. Forensic Sci Int 132(1):49–56
8. Demirjian A, Goldstein H, Tanner JM (1973) A new system of dental age assessment. Hum Biol 45:211–227
9. Mincer HH, Harris EF, Berryman HE (1993) The A.B.F.O. study of third molar development and its use as an estimator of chronological age. J Forensic Sci 38:379–390
10. Cameriere R, Ferrante L, De Angelis D, Scarpino F, Galli F (2008) The comparison between measurement of open apices of third molars and Demirjian stages to test chronological age of over 18 year olds in living subjects. Int J Legal Med 122:493–497
11. Tanner M, Healy MJR, Goldstein H, Cameron N (2001) Assessment of skeletal maturity and prediction of adult height (TW3 method). Saunders, London
12. Roche AF, Cameron Chumlea W, Thissen D (1988) Assessing the skeletal maturity of the hand-wrist: FELS method. Charles C. Thomas, Springfield
13. Cameriere R, Ferrante L (2008) Age estimation in children by measurement of carpals and epiphyses of radius and ulna and open apices in teeth: a pilot study. Forensic Sci Int 174:60–63
14. Thevissen PW, Fieuws S, Willems G (2010) Human dental age estimation using third molar developmental stages: does a Bayesian approach outperform regression models to discriminate between juveniles and adults? Int J Legal Med 124(1):35–42
15. Tise M, Mazzarini L, Fabbrizzi G, Ferrante L, Giorgetti R, Tagliabracci A (2011) Applicability of Greulich and Pyle method for age assessment in forensic practice on an Italian sample. Int J Legal Med 125(3):411–416
16. Thevissen PW, Alqerban A, Asaumi J, Kahveci F, Kaur J, Kim YK, Pittayapat P, Van Vlierberghe M, Zhang Y, Fieuws S, Willems G (2010) Human dental age estimation using third molar developmental stages: accuracy of age predictions not using country specific information. Forensic Sci Int 201:106–111
17. Gunst K, Mesotten K, Carbonez A, Willems G (2003) Third molar root development in relation to chronological age: a large sample sized retrospective study. Forensic Sci Int 136:52–57
18. Rozkovcova E, Markova M, Mrklas L (2005) Third molar as an age indicator in young individuals. Prague Med Rep 106(4):367–398
19. Liversidge HM, Herdeg B, Rosing FW (1998) Dental age estimation of non adults. A review of methods and principles. In: Alt KW, Rosing FW, Teschler-Nicola M (eds) Dental anthropology fundamentals, limits, and prospects. Springer, New York, pp 419–442
20. Knell B, Ruhstaller P, Prieels F, Schmeling A (2009) Dental age diagnostics by means of radiographical evaluation of the growth stages of lower wisdom teeth. Int J Legal Med 6:465–469
21. Gardner E (1968) The embryology of the clavicle. Clin Orthop Relat Res 58:9–16
22. Neer CS (1960) Non-union of the clavicle. J Am Med Assoc 172:1006–1011
23. Todd TW, D' Errico J (1928) The clavicular epiphyses. Am J Anat 41:25–50
24. McKern TW, Stewart TD (1957) Skeletal age changes in young American males. Analysed from the standpoint of age identification. In: Technical report EP 45. Quartermaster Research and Development Center, Environmental Protection Research Division. Natick, Massachusetts, pp. 89–971
25. Owings Webb PA, Myers Suchey J (1985) Epiphyseal union of the anterior iliac crest and medial clavicle in a modern multiracial sample of American males and females. Am J Phys Anthropol 68:457–466
26. MacLaughlin SM (1990) Epiphyseal fusion at the sternal end of the clavicle in a modern Portuguese skeletal sample. Antropol Port 8:59–68
27. Ji L, Terazawa K, Tsukamoto T, Haga K (1994) Estimation of age from epiphyseal union degrees of the sternal end of the clavicle. Hokkaido Igaku Zasshi 69:104–111
28. Black SM, Scheuer JL (1996) Age changes in the clavicle: from the early neonatal period to skeletal maturity. Int J Osteoarcheol 6:425–434
29. Shirley NR (2009) Age and sex estimation from the human clavicle: an investigation of traditional and novel methods. Dissertation. University of Tennessee, Knoxville
30. Singh J, Chavali KH (2011) Age estimation from clavicular epiphyseal union sequencing in a Northwest Indian population of the Chandigarh region. J Forensic Legal Med 18:82–87
31. Flecker H (1933) Roentgenographic observations of the times of appearance of epiphyses and their fusion with the diaphyses. J Anat 67:118–164
32. Galstaun G (1937) A study of ossification as observed in Indian subjects. Indian J Med Res 25:267–324
33. Jit I, Kullkarni M (1976) Times of appearance and fusion of epiphysis at the medial end of the clavicle. Indian J Med Res 64:773–782
34. Schmeling A, Schulz R, Reisinger W, Mühler M, Wernecke KD, Geserick G (2004) Studies on the time frame for ossification of medial clavicular epiphyseal cartilage in conventional radiography. Int J Legal Med 118:5–8
35. Garamendi PM, Landa MI, Botella MC, Alemán I (2011) Forensic age estimation on digital X-ray images: medial epiphyses of the clavicle and first rib ossification in relation to chronological age. J Forensic Sci 56:S3–S12
36. Kreitner K-F, Schweden F, Schild HH, Riepert T, Nafe B (1997) Die computertomographisch bestimmte Ausreifung der medialen Klavikulaepiphyse—eine additive Methode zur Altersbestimmung im Adoleszentenalter und in der dritten Lebensdekade? Fortschr Röntgenstr 166:481–486
37. Schulz R, Mühler M, Mutze S, Schmidt S, Reisinger W, Schmeling A (2005) Studies on the time frame for ossification of the medial epiphysis of the clavicle as revealed by CT scans. Int J Legal Med 119:142–145
38. Schulze D, Rother U, Fuhrmann A, Richel S, Faulmann G, Heiland M (2006) Correlation of age and ossification of the medial clavicular epiphysis using computed tomography. Forensic Sci Int 158:184–189
39. Bassed RB, Drummer OH, Briggs C, Valenzuela A (2010) Age estimation and the medial clavicular epiphysis: analysis of the age of majority in an Australian population using computed tomography. Forensic Sci Med Pathol. doi:10.1007/s12024-010-9200-y

40. Kellinghaus M, Schulz R, Vieth V, Schmidt S, Pfeiffer H, Schmeling A (2010) Enhanced possibilities to make statements on the ossification status of the medial clavicular epiphysis using an amplified staging scheme in evaluating thin-slice CT scans. Int J Legal Med 124:321–325
41. Kellinghaus M, Schulz R, Vieth V, Schmidt S, Schmeling A (2010) Forensic age estimation in living subjects based on the ossification status of the medial clavicular epiphysis as revealed by thin-slice multidetector computed tomography. Int J Legal Med 124:149–154
42. Schmidt S, Mühler M, Schmeling A, Reisinger W, Schulz R (2007) Magnetic resonance imaging of the clavicular ossification. Int J Legal Med 121:321–324
43. Hillewig E, De Tobel J, Cuche O, Vandemaele P, Piette M, Verstraete K (2011) Magnetic resonance imaging of the medial extremity of the clavicle in forensic bone age determination: a new four-minute approach. Eur Radiol 21:757–767
44. Schulz R, Zwiesigk P, Schiborr M, Schmidt S, Schmeling A (2008) Ultrasound studies on the time course of clavicular ossification. Int J Legal Med 122:163–167
45. Quirmbach F, Ramsthaler F, Verhoff MA (2009) Evaluation of the ossification of the medial clavicular epiphysis with a digital ultrasonic system to determine the age threshold of 21 years. Int J Legal Med 123:241–245
46. Cunha E, Baccino E, Martrille L, Ramsthaler F, Prieto JL, Schuliar Y, Lynnerup N, Cattaneo C (2009) The problem of ageing human remains and living individuals: a review. Forensic Sci Int 193:1–13
47. Cattaneo C, Baccino E (2002) A call for forensic anthropology in Europe. Int J Legal Med 116:N1–N2
48. Baccino E (2005) Forensic Anthropology Society of Europe (FASE): a subsection of IALM. Int J Legal Med 119:N1
49. Kreitner KF, Schweden FJ, Riepert T, Nafe B, Thelen M (1998) Bone age determination based on the study of the medial extremity of the clavicle. Eur Radiol 8:1116–1122
50. Ramsthaler F, Proschek P, Betz W, Verhoff MA (2009) How reliable are the risk estimates for X-ray examinations in forensic age estimations? A safety update. Int J Legal Med 123:199–204
51. Brenner DJ (2002) Estimating cancer risks from pediatric CT: going from the qualitative to the quantitative. Pediatr Radiol 32:228–233
52. Mühler M, Schulz R, Schmidt S, Schmeling A, Reisinger W (2006) The influence of slice thickness on assessment of clavicle ossification in forensic age diagnostics. Int J Legal Med 120:15–17
53. Galbraith KS (1999) Moving people: forced migration and international law. 13 Geo Immigr LJ 597
54. Cranston A (2000) Refugees in crisis. 3 Alt LJ 12
55. Feller E (2001) The UN and the protection of human rights: the evolution of the International Refugee Protection Regime. Washington J Law P 5:129–143
56. R Development Core Team (2008). R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org (Last access 12/03/2012)
57. Cohen J (1960) A coefficient of agreement for normal scales. Educ Psychol Meas 20:37–46
58. Light RJ (1971) Measures of response agreement for qualitative data: some generalizations and alternatives. Psychol Bull 76(5):365–377
59. Martín-de las Heras S, García-Fortea P, Ortega A, Zodocovich S, Valenzuela A (2008) Third molar development according to chronological age in populations of Spanish and Magrebian origin. Forensic Sci Int 174:47–53
60. Altman DG (1991) Practical statistics for medical research. Chapman & Hall, London
61. UNHCR (2011) Birth registration. UNICEF, Geneva. http://www.unhcr.org/4ed34f1c9.html (Last access 07/03/2012)
62. UNICEF (2005) The ‘rights’ start to life: a statistical analysis of birth registration. UNICEF, New York, pp 5–11
63. Kvittingen AV (2011) Negotiating childhood: age assessment in the UK asylum system. Working paper series Oxford: Refugee Studies Centre, Oxford Department of International Development, University of Oxford
64. UNHCR (2009) Guidelines on international protection: Child Asylum Claims under Articles 1(A)2 and 1(F) of the 1951 Convention and/or 1967 Protocol relating to the Status of Refugees
65. UNHCR (1997) Guidelines on policies and procedures on dealing with unaccompanied children seeking asylum, p 5
66. UNHCR (1994) Refugee children: guidelines on protection and care preface, Geneva
67. Willems G, Moulin-Romsee C, Solheim T (2002) Non-destructive dental-age calculation methods in adults: intra- and inter-observer effects. Forensic Sci Int 126:221–226
68. Thevissen PW, Pittayapat P, Fieuws S, Willems G (2009) Estimating age of majority on third molar developmental stages in young adults from Thailand using a modified scoring technique. J Forensic Sci 54(2):428–432
69. Huda W, Atherton JV, Ware DE, Cumming WA (1997) An approach for the estimation of effective radiation dose at CT in pediatric patients. Radiology 203:417–422
70. Maher MM, Kalra MK, Toth TL, Wittram C, Saini S, Shepard J (2004) Application of rational practice and technical advances for optimizing radiation dose for chest CT. J Thorac Imaging 19:16–23
71. Jurik AG, Jensen LC, Hansen J (1996) Radiation dose by spiral CT and conventional tomography of the sternoclavicular joints and the manubrium sterni. Skeletal Radiol 25:467–470
72. Schmeling A, Reisinger W, Geserick G, Olze A (2006) Age estimation of unaccompanied minors. Part I. General considerations. Forensic Sci Int 159(Suppl 1):S61–S64
73. Aynsley-Green A (2009) Unethical age assessment. Brit Dent J 206:337
74. Hall EJ (2009) Radiation biology for pediatric radiologists. Pediatr Radiol 39(suppl 1):57–64
75. Hall EJ, Brenner DJ (2008) Cancer risks from diagnostic radiology. Br J Radiol 81:362–378
76. Hogge JP, Messmer JM, Doan QN (1994) Radiographic identification of unknown human remains and interpreter experience level. J Forensic Sci 39(2):373–377
77. Koot MG, Sauer NJ, Fenton TW (2005) Radiographic human identification using bones of the hand: a validation study. J Forensic Sci 50(2):1–6
78. Pescarini L, Inches I (2006) Systematic approach to human error in radiology. Radiol Med 111:252–267
79. Robinson PJ (1997) Radiology’s Achilles’ heel: error and variation in the interpretation of the Rontgen image. Br J Radiol 70:1085–1098
80. Schmidt H, Köhler A, Zimmer EA, Freyschmidt J, Holthusen W (1993) Borderlands of normal and early pathologic findings in skeletal radiography. Georg Thieme, Stuttgart, pp 305–309