Reliability of the methods applied to assess age minority in living subjects around 18 years old
A survey on a Moroccan origin population

P.M. Garamendi[a,*], M.I. Landa[a], J. Ballesteros[b], M.A. Solano[a]

[a]*Basque Legal Medicine Institute, Vizcaya Division, Forensic Clinical Department, Spain*
[b]*University of the Basque Country, U.P.V.-E.H.U., Neuroscience Department, Spain*

Abstract

We present a review of a population of 114 immigrant Moroccan males for which an age estimation was requested. The subjects' real chronological age was confirmed by the Moroccan Embassy in Spain. The confirmed age range was between 13 and 25, with an average age of 18.1 years and a standard deviation of 2.03. The following tests were performed to arrive at the forensic estimation of age: general physical examination, carpus X-ray (Greulich and Pyle method) and dental orthopantomography to determine the degree of maturity of the third inferior molars (Demirjian's method). Carpus X-ray (skeletal age) was the most useful method, followed by Demirjian's method (dental age), as prediction factors of a chronological age of over or under 18. The combination of skeletal and dental age variables represented a significant improvement in the prediction of the chronological age of the subjects in this population, reducing the number of ethically unacceptable test errors to a minimum.
© 2004 Elsevier Ireland Ltd. All rights reserved.

Keywords: Age determination by skeleton; Bone age; Dental age; Age minority; Morocco

1. Introduction

Spain is the second country in the European Community with the largest number of illegal undocumented immigrants. Minors are subject to special Spanish and community regulations which are different from those applied to subjects over 18. According to these regulations, under 14-year-old subjects are exempt from criminal liability, minors between 14 and 18 are subject to special criminal standards, and persons over 18 but under 21 could be subject in the future in Spain to the criminal standards now applied to minors under 18. Illegal immigrants under 18 may be placed under the guardianship of the authorities.

Both in a clinical setting and in the specific field of forensic medicine, there is a growing demand by the courts for appropriate medical tests aimed at estimating the approximate age of supposed minors without documentation.

In the year 2000, the Arbeitsgemeinschaft für Forensische Altersdiagnostik der Deutschen Gesellschaft für Rechtsmedizin published its guidelines for the forensic estimation of the chronological age of living individuals subject to criminal proceedings [1]. These guidelines recommended the performance of the following tests to determine majority or minority of age (18 years) for criminal purposes, in living subjects:

1. Physical examination: anthropometric measurements (weight, height, build); inspection of signs of sexual maturity; identification of diseases which could alter maturity development.
2. X-ray examination of the left hand.
3. External examination of the condition of teeth and dental X-ray.
4. X-ray examination of the *medial clavicular epiphyseal cartilage* [2], to confirm if the chronological age is over or under 21.

When interpreting the results, the guidelines themselves recommend that data from the tests above should be compared with reference studies relevant to the specific individual in question. They finally recommend that, when the final expert assessment has been made, the results of each of the tests performed should be recorded separately and that the age estimated should be identified as the most probable, specifying the degree of probability of each estimated result.

In Spain, there are several studies on our national population according to which the recommended tests can be interpreted with the pertinent adjustments in assumed minors of Spanish origin. Most illegal immigrants without documentation are from Morocco. In this case, the forensic estimation of their age is difficult because no systematic studies have been conducted in that country that could inform us of the maturity parameter variations applicable to its population.

This article presents a study conducted in a population of immigrant males, supposedly minors and of Moroccan origin. The purpose of the study was centred on analysing the efficacy of the tests available to estimate approximate chronological age and, particularly, the reliability of an age estimation of 18 or more, attempting to reduce the number of ethically unacceptable errors to a minimum.

2. Material and methods

The population sample on which the study was conducted consisted of a total of 114 males of Moroccan origin, supposedly minors, and illegal immigrants in Spain. The original studies were requested by the Juvenile Division of the Public Prosecutor's Office of the Basque Country.

Confirmation of the subject's date of birth was obtained from the Moroccan Embassy in Spain while processing their residency applications. Along with this information, the chronological age for this study was calculated in relation to the date of the X-ray examinations performed on the supposed minors.

The study conducted on this population sample consisted on the following examinations and additional tests:

1. Physical examination: basic anthropometric measurements (weight, height and body mass index, or BMI), inspection of signs of sexual maturity (stage of evolution of pubic hair P1–P6 using Tanner's method; Table 1),

Table 1
Stages of pubic hair development, according to Tanner [10]

P1 No hair
P2 Some pubic hair around the root of the penis or on the labia majora, not recognizable on a full body photograph
P3 Dense hair within defined limits, visible on a photograph
P4 Dense hair like that of adults, but less extended
P5 Dense and widespread hair with a horizontal upper limit and spreading sideways towards the thighs
P6 Triangular growth up to the navel

general physical examination and personal interview to rule out the existence of disease.
2. X-ray study of the hand and wrist.
3. Dental maturity study: this included a general examination of the oral cavity and an X-ray examination of the stage of maturity based on dental orthopantomography.

The data from the physical examination and the external examination of the oral cavity was obtained by the forensic physician on duty, who was responsible for this examination.

The hand and wrist X-rays and the dental orthopantomography were performed in the Basurto Hospital's radiodiagnosis department. The data obtained from the hand and wrist X-rays were interpreted using the Greulich and Pyle method [3].

The data from the X-ray examination of dental maturity was centred on assessing the stage of evolution of the third inferior molars in both dental arches. The degree of maturity was quantified by Demirjian's stage system, unaltered, stages A–H (Table 2) [4]. Although not originally devised by Demirjian et al. to quantify third molars maturation, the method has been previously applied for this purpose in several scientific papers. Demirjian's stage system has been applied in this research following A.B.F.O. (American Board of Forensic Odontologists) recommendations published in 1993 by Mincer et al. [5]. The quantification of the third superior molars was ruled out since it was impossible to interpret them on the orthopantomography plates because bone contours overlapped the molars, making it difficult to see them entirely.

The two X-ray studies were interpreted by three of the raters (Greulich and Pyle's method) or by two (Demirjian's method). Two of them (PMG, MAS) have been regularly responsible for interpreting these tests in the Basque Legal Medicine Institute since 1999. The third author (MIL) is a forensic doctor unaccustomed to performing this kind of X-ray assessment. When making his assessments, the third author had no more than the instructions provided in the original text of the Greulich and Pyle system (1959 edition). She received no specific instructions on how to apply the methods from the other two authors. This study has been based on the interpretations of the carpus X-rays by investigators PMG, MIL and MAS and the interpretations of the orthopantomographies by investigators PMG and MIL.

The results of the above studies were analysed by the fourth author, a specialist in statistical research in biome-

Table 2
Reproduction of the dental maturation system of Demirjian et al. [4] recommended by Mincer et al. [5] and applied to the third molars in this research

A	The points of the cusps are mineralised, but not yet joined	E	The interradicular bifurcation has started. The root is shorter than the crown
B	The cusps are joined and the crown morphology is well defined	F	The root is as long as the crown. The root terminations are funnel-shaped
C	The crown is half formed; the pulp chamber and dentin apposition are evident	G	Root walls are parallel, but the root apices remain open
D	The crown is complete up to the amelo-cemental junction. The pulp chamber is trapezoidal	H	The root apices are completely closed. The width of the periodontal membrane around the roots is constant

dical science (JB). Agreement between observers in their X-ray interpretations was analysed by the intraclass correlation coefficient (ICC) [6] for the Greulich and Pyle method, and with the quadratic weighted kappa index for the dental orthopantomography [7]. The validity of the criterion of the methods under study in relation to chronological age at the time of examination was evaluated by analysing the sensitivity and specificity values and the likelihood ratios for positive and negative results, according to different cut-off points [8]. Finally, the forensic age determination diagnosis methods for which the best validity of criterion values had been obtained, were compared by calculating the area under the ROC curve [9].

3. Results

3.1. General description

The study was conducted on a population of 114 male subjects of Moroccan origin, of known chronological age (mean age 18.1 years, S.D. = 2; range: 13–25). Of these, 62 were 18 or over at the time of these assessments.

3.2. Agreement between observers

For the Greulich and Pyle method (114 subjects assessed by 3 observers), the intraclass correlation coefficient (ICC) is 0.93 (CI 95%: 0.91–0.95), whereas for Demirjian's technique (84 subjects assessed by 2 observers) the value of the weighted kappa estimator is the same for both third inferior molars (kappa = 0.88; CI 95%: 0.66–1.0). In both cases, the level of agreement is very high.

The validity of the chronological age of 18 or over criterion in relation to the results obtained by the Greulich and Pyle method (GP skeletal age) is analysed in Table 3.

The validity of the chronological age of 18 or over criterion in relation to the results obtained by Demirjian's method (dental age) for both third inferior molars, is analysed in Table 4.

3.3. Comparison of the validity of the Greulich and Pyle method and Demirjian's technique

In the 80 subjects assessed by both methods, a comparison of the area under the ROC curves (see Fig. 1) showed

Table 3
Criterion validity (chronological age ≥ 18)

Cut-off point (estimated age, years)	TP	FN	FP	TN	Sensitivity (CI 95%)	Specificity (CI 95%)	LR+ (CI 95%)	LR− (CI 95%)
16/17	52	10	23	28	0.84 (0.72–0.92)	0.55 (0.40–0.69)	186 (1.39–2.64)	0.29 (0.16–0.53)
17/18	42	20	11	40	0.68 (0.55–0.79)	0.78 (0.65–0.89)	3.14 (1.88–5.55)	0.41 (0.27–0.59)
18/19	28	34	7	44	0.45 (0.32–0.58)	086 (0.74–0.94)	3.29 (1.64–6.93)	0.64 (0.48–0.81)

Greuiich and Pyle method (N = 113). Area under the ROC curve = 0.77 (CI 95%: 0.68–0.85). TP: true positives; FN: false negatives; FP: false positives; TN: true negatives; LR+: likelihood ratio for a positive result; LR−: likelihood ratio for a negative result.

Table 4
Criterion validity (chronological age ≥ 18)

Cut-off point (Demirjian states)	TP	FN	FP	TN	Sensitivity (CI 95%)	Specificity (CI 95%)	LR+ (CI 95%)	LR− (CI 95%)
Right third inferior molar ($N = 80$)								
E/F	33	9	17	21	0.79 (0.63–0.90)	0.55 (0.38–0.71)	1.76 (1.23–2.67)	0.39 (0.20–0.72)
F/G	25	17	9	29	0.60 (0.43–0.74)	0.76 (0.60–0.89)	2.51 (1.40–4.77)	0.53 (0.34–0.78)
G/H	10	32	2	36	0.24 (0.12–0.39)	0.95 (0.82–0.99)	4.52 (1.22–17.73)	0.80 (0.64–0.96)
Left third inferior molar ($N = 87$)								
E/F	36	11	19	21	0.77 (0.62–0.88)	0.53 (0.36–0.68)	1.61 (1.15–2.39)	0.45 (0.24–0.79)
F/G	30	17	10	30	0.64 (0.49–0.77	0.75 (0.59–0.87)	2.55 (1.49–4.65)	0.48 (0.31–0.72)
G/H	13	34	2	38	0.28 (0.16–0.43)	0.95 (0.83–0.99)	5.33 (1.53–21.25)	0.76 (0.61–0.91)

Demirjian method. Area under the ROC curve, both molars = 0.72 (CI 95%: 0.61–0.83). TP: true positives; FN: false negatives; FP: false positives; TN: true negatives; LR+: likelihood ratio for a positive result; LR−: likelihood ratio for a negative result.

that there were no significant differences between them ($\chi^2 = 0.08$ in 2 d.f., $P = 0.96$).

The validity of the chronological age of 18 or over criterion in relation to the evolution of the pubic hair variation parameter (Tanner's method) is specified in Table 5.

3.4. BMI results

These results were evaluated in 112 subjects in relation to its variation from known chronological age. The validity of the criterion is low, with an area under the ROC curve of 0.62 (CI 95%: 0.52–0.73). Pearson's correlation between BMI and age was also low ($r = 0.27$).

The BMI by age group, as an indicator of under or over nutrition, indicates an acceptable distribution. The WHO considers the following as normal BMI values: from 16 to 23 for minors of 14 and under, 18 to 23 for subjects between 14

and 19 years of age, and 18 to 25 for subjects of 19 or over. In the group of minors under 14 there was only one case with a BMI of 17.07. For ages from 14 to 19 (77 cases), the mean BMI was 20.6, with a standard deviation of 2.05, a median of 20.5 and maximum and minimum values of 28.69 and 16.18, respectively. Among the subjects over 19 (36 cases), the mean BMI was 20.8 with a standard deviation of 2.0, a median of 20.8 and maximum and minimum values of 27.4 and 18.0, respectively.

Table 6 analyses the results of the validity of the chronological age of 18 or over criterion in relation to a joint assessment by Demirjian's method for both molars, and the Greulich and Pyle method.

Fig. 2 presents the frequency distribution of differences in each subject under study between actual chronological age and the age estimated using the Greulich and Pyle method.
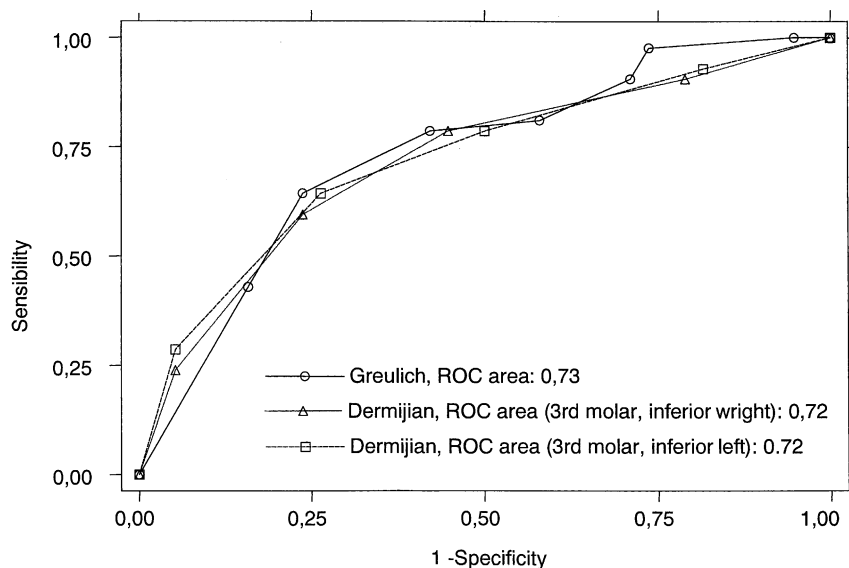

Fig. 1. Comparison of method validity for Greulich and Pyle and Demirjian's technique in right and left third inferior molars. A comparison of the areas under the ROC curve shows no significant validity differences for majority of age (18 or over) status (χ^2 test = 0.08 in 2 d.f., $P = 0.96$).

Table 5
Criterion validity (chronological age $\geq$ 18)

Cut-off point (pubic hair states)	TP	FN	FP	TN	Sensitivity (CI 95%)	Specificity (CI 95%)	LR+ (CI 95%)	LR− (CI 95%)
P3/P4	47	3	29	8	0.94 (0.83–0.99)	0.22 (0.10–0.38)	1.20 (1.02–1.51)	0.28 (0.08–0.90)
P4/P5	32	18	19	18	0.64 (0.49–0.77)	0.49 (0.32–0.66)	1.25 (0.87–1.86)	0.74 (0.45–1.22)
P5/P6	5	45	5	32	0.10 (0.03–0.22)	0.87 (0.71–0.95)	0.74 (0.25–2.25)	1.04 (0.89–1.27)

Tanner method (N = 87). Area under the ROC curve = 0.57 (CI 95%: 0.45–0.69). TP: true positives; FN: false negatives; FP: false positives; TN: true negatives; LR+: likelihood ratio for a positive result; LR−: likelihood ratio for a negative result.

Table 6
Criterion validity (chronological age $\geq$ 18) for a combination of the Greulich method (17/18) and Demirjian's technique (states G/H) (N = 80)

	Criterion +	Criterion −	Sensitivity (CI 95%)	Specificity (CI 95%)	LR+ (CI 95%)	LR− (CI 95%)
Test +	9	1	0.21 (0.10–0.37)	0.97 (0.86–0.99)	8.14 (1.45–48.74)	0.81 (0.66–0.94)
Test −	33	37	0.21 (0.10–0.37)	0.97 (0.86–0.99)	8.14 (1.45–48.74)	0.81 (0.66–0.94)

LR+: likelihood ratio for a positive result; LR−: likelihood ratio for a negative result.
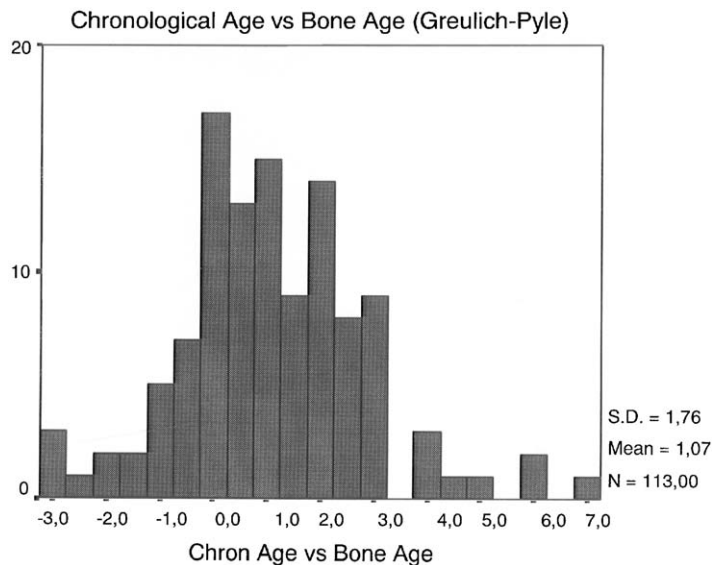

Fig. 2. Histogram of distribution of frequency of results of differences between known age and bone age estimated by examiner 1 using the Greulich and Pyle method. >0 difference results indicate that known chronological age is greater than the bone age estimated by the GP atlas. <0 difference results indicate that known chronological age is lower than the age estimated by the GP atlas. Differences of 0 indicate that both ages coincide.

4. Discussion

The application of the guidelines of the Arbeitsgemeinschaft für Forensische Altersdiagnostik der Deutschen Gesellschaft für Rechtsmedizin in forensic practice, not only implies a need to consider the results of the tests performed in the light of prior scientific studies, but also involves the difficulty of deciding which results interpretation method is the most appropriate, and its particular degree of reliability.

4.1. Physical examination, anthropometrical variables

The maturation of secondary sexual features is usually quantified according to the system proposed by Tanner [10]. This system quantifies the following parameters: development of axillary hair (A1–A3), pubic hair (P1–P5), mammary gland development (B1–B5), external genital development (G1–G5), testes volume (using Prader's orchidometer), length of the flaccid penis, menarche, sperm

production and voice features. Irrespective of the difficulties involved in interpreting results with inter- and intra-observer differences, which have not been well studied, there are few series analysing the evolution of these parameters with chronological age in different populations, and the few that there are, are fundamentally centred on developed countries.

Anthropometrical variables are of little use as predictors of estimated forensic age, but they may be useful as factors suggesting the existence of pathological conditions which, since they delay or enhance maturity, may significantly alter a comprehensive interpretation of the results.

4.2. X-ray study of the left carpus

The skeletal age parameter obtained from an X-ray examination of the left hand is a statistical concept derived from clinical experience that is useful for strictly clinical purposes when estimating a subject's individual rate of maturity, and predicting phenomena such as expected height. In different population studies it has been seen that this parameter is not absolutely correlated to the subjects' actual chronological age. Nevertheless, this parameter is considered to be more physiologically stable than dental maturation and is therefore the main parameter at our disposal for clinical and forensic purposes to estimate chronological age at the end of adolescence with some accuracy.

The choice of the method for assessing the skeletal age of the carpus that is most appropriate for estimating approximate age is subject to controversy. The three models available (atlas systems, such as the GP method; numerical systems, such as the TW2 (Tanner–Whitehouse) or CASAS method; and mixed systems such as FELS) have their supporters and detractors. The TW2 method is more widely used in Europe and the GP technique on a world-wide scale. Their results, for the purposes of this type of study, can be applied together, but GP has the advantage that it is more economical, since it takes less time [11]. It could therefore be considered as the method of choice for this type of study.

The interpretation of results in relation to the population under study reveals significant differences related to the population sample. Differences have been detected, according to how the sample is affected by racial, socio-economic and pathological demographic factors.

There is a large number of studies in the context of ethnic and racial impact, some of them based upon dubious sample selection and often contradictory results, mainly conducted on populations of Caucasian Europeans, Caucasian North Americans, other North American ethnic groups, different Mongoloid and Caucasian populations from Asia and some incomplete studies on central and south African populations.

The most recent studies in Europe appear to indicate that European Caucasians are close to the maturity rates marked by the GP and the TW2 system (Italy [12], Belgium [13], Holland [14], Finland [15], Denmark [16], Sweden [17], Spain [18], Austria [19]). In some cases, the differences from the original methods were small but so statistically significant that the need for specific tables and atlases has been proposed for these populations.

The classical studies in the 1970s and 1980s on Mongoloid Asian populations showed that the Chinese and Japanese are delayed in skeletal age in relation to chronological age in the pre-puberal period, but that the process speeded up in the post-puberal phase and that they complete maturation at a similar age to Caucasian Europeans and Americans. The most current series on modern populations with better socio-economic conditions show a trend to adjust the results even further to the rate of bone age maturity of western populations [20,21]. Findings have been similar in India and Pakistan, with an advance in bone age in relation to chronological age during the post-puberal stage, especially in subjects belonging to a high social class [22].

In the U.S., studies indicate that Caucasian subjects are well adjusted to the GP and TW2 standards, or show a slightly advanced maturation rate [23]. On the other hand, results for Negroid subjects are contradictory. For Gross et al. [24] the black race is better adjusted to the GP standard than the white. In the series conducted by Ontell et al. [26] and Lodler et al. [25] the black race is ahead of the GP standards. The series obtained by Marshall et al. [27] in black Jamaicans compared with the TW2 UK60, indicates a delay in bone age from 13 years of age, irrespective of socio-economic factors. Finally, Gilsanz et al. [28], in a study with black and white subjects in the U.S., find no significant differences between skeletal age and chronological age in the two races when socio-economic factors are equal.

As far as we know, the inhabitants of Moslem countries in the Near East and Northern Africa, and the populations of these countries emigrated to developed countries, have not been systematically studied and it is not known whether the rate of bone maturation follows the same evolution as other populations. The only systematic study identified in our bibliographical review, conducted by Koc et al. [29] on a modern Turkish population, shows a bone age delay up to the age of 13 and a discreet advance after that age for the GP. The main problem with this study is that the population sample is only up to 17 years of chronological age, and it therefore has no absolute utility for the purposes of our study.

Some authors consider that the socio-economic characteristics of each population are the most important factors in bone age maturity rate variation [30]. Other authors, on the other hand, do not appear to confirm this hypothesis. In a black Jamaican population, Marshall observed no variations depending on the subjects' high or low social class [27]. Oestreich [31] also observes no differences between the rural and urban German population in the Thiemann-Nitz method series. However, modern and specific studies on this variable by authors such as Jahari et al. [32], on an Indonesian population sample, Fleshman [33], on an African population sample, or Melsen et al. [34] on a group of adopted foreign minors in Denmark, clearly identify socio-economic factors and poverty as producing significant delays in bone maturation during the pre-puberal stage.

The pathological factors clearly identified as bone age evolution rate alteration factors include, among others, nocturnal enuresis, GH deficit, the practice of competitive sports, skeletal malformations or the effects of physical agents, such as cold.

4.3. Dental examination

With regards to external and X-ray dental examinations, there are different studies and result interpretation variants. Initially, since this maturation parameter is less influenced by environmental circumstances than other maturation parameters, it could be especially useful for the purpose of our research.

The external examination, centred on the eruption or not of the third molars, is too inaccurate, due to the highly variable appearance of this physiological phenomenon. It is therefore more appropriate to evaluate the development of the third molars, which is a more stable phenomenon within its high degree of variability. There are several methods to quantify the degree of maturity, principally Demirjian's original graphic method [4] or its variants, and some numerical methods, like the technique proposed by Kulman [35]. Although not originally devised to test third molars development, eight-stage Demirjian's method is one the most practical, easy to use and world-wide used [5,36]. It has been previously applied on forensic purposes in different scientific papers to establish third molar's maturation degree. Nevertheless, there are many other methods to quantify dental maturity of third molar most of them based on X-ray findings similar as in Demirjian's.

All the series studied appear to suffer from the same defect, in that they do not guarantee with a sufficient degree of reliability, that a subject with a stage of less than H (crown mineralization and root development incomplete) is not over 18 years of age, or that a subject in stage H (crown mineralization and root development complete) is not under 18 years of age [5,36–38] (Table 7).

4.4. Discussion

Our study population consisted of 114 male supposed minors of Moroccan origin, with a confirmed chronological age of between 13.75 and 25.75. Demographic data places

Morocco in WHO nutritional status group 3 (5–19% of the population undernourished), similar to other Northern African countries like Algiers, eastern countries like Indonesia or the present Russian Federation. As for its ethnic distribution, in our sample all the subjects had a Caucasoid appearance and none of them presented a nutritional deficit. Since we do not know the minors' prior socio-economic status, the BMI was used as an indicator of their nutritional status, following WHO recommendations. None of the subjects under study referred to either a disease identified as a bone and dental maturation altering factor, or competitive sports or intense physical activities.

The primary objective of our study was to attempt to verify the degree of validity of estimated forensic age results based on the tests described, compared with the known chronological age of an undocumented immigrant population from Morocco. Our interest was focused on confirming the specific degree of reliability when estimating whether forensic age is 18 or over.

In the interpretation of the results, our research provides several conclusions of practical interest.

The degree of agreement between the results obtained by the different investigators, irrespective of their prior experience, is high. This conclusion leads us to consider that the direct assessment of the radiographs performed can be carried out both by doctors experienced in such an interpretation and by non-experienced physicians.

In our series, the degree of maturity of secondary sexual features, in particular the evolution of pubic hair, showed that subjects with a stage of less than P3 are usually under 18. However, the statistical significance of this result is insufficient, as revealed by the likelihood ratios both for positive and negative results (Table 5). It can therefore not be considered a valid age diagnosis method.

The Greulich and Pyle method's X-ray study of the carpus (skeletal age) appears to be the most powerful prediction factor for chronological age. The results show that in this sample there is a high coincidence between the age estimated by this method and the known chronological age. The mean difference between the two is 1.07 with a standard deviation of 1.76 (chronological age older than age estimated by GP); the median difference was 0.9 year and the mode was −1.0 year. These differences are due to the fact that a significant group of subjects (36 cases) chronologically exceed the maximum age considered for the method in question (19 years) (Fig. 2). The likelihood ratios for positive results of the confirmation of the hypothesis that chronological age is effectively at 18 years in relation to the age of 18 and 19 estimated by the Greulich and Pyle method are acceptable (values of 3.00 or over), but the likelihood ratios for negative results are too high (over 0.20) (Table 3). The latter shows that as an isolated test, the Greulich and Pyle method is a good, but not optimal, way of diagnosing age.

The study of the degree of dental maturity (dental age), although it is statistically a less powerful predictor of chronological age, is however statistically significant.

Table 7
Results of the study from Mincer et al. conducted in a population of 853 subjects between 14.1 and 24.9 years of age [5]

	Group				
	D	E	F	G	H
Maxilla	15.9	27.8	44.0	46.8	85.3
Mandible	6.1	69.4	40.5	56.0	90.1

Empirical probabilities (%) of an individual being at least 18 years of age based on the grade of third molar formation using eight grade scheme developed by Demirjian et al. [4] (white males sample).

The likelihood ratios for a positive result confirming the hypothesis that chronological age is effectively over 18 for the results in both third inferior molars are also acceptable, but once again, the likelihood ratios for negative results are higher than they should be (Table 4). These results show that as an isolated test, Demirjian's technique applied to the assessment of third molar development is also a good, but not optimal, way of diagnosing age.

The difference in the efficacy of both methods when confirming or ruling out a chronological age of 18 or over is minimal. The areas under the ROC curve are 0.73 for the Greulich and Pyle method, and 0.72 for Demirjian's technique in either of the two third inferior molars, and the χ^2 test shows no significant differences between the two methods (Fig. 1). These results show that both methods are comparable in their efficacy confirming a chronological age of 18 or over.

Nevertheless, the statistical results show that both methods on their own have too high an error rate, and that they should therefore not be considered as optimal diagnostic methods. The errors occurring in their application can be classified into two types on forensic practice purposes: technically unacceptable errors and ethically unacceptable errors. The errors derived from a forensic age estimation indicating that a subject actually over 18 is a minor fall into the first class. Incorrect estimates of this kind lead to a more benevolent criminal treatment of these subjects, normally reserved for minors, and they also generate social expenses derived from the need for special protective measures. However, errors indicating that minors are over 18 can be classified as ethically unacceptable, since they lead to a violation of minors' rights. Consequently, in forensic age diagnosis, the tests methods used have to reduce technically unacceptable to a minimum, but it is even more important for ethically unacceptable errors to disappear, especially in cases involving the possible criminal liability of the supposed minor.

Finally, our series shows that the combination of the two chronological age prediction factors represented a significant increase in the efficacy of the prediction that a subject was under the 18-year-old age limit or not. Table 6 shows that the combination of a Greulich and Pyle result of 18 or 19 years of age with an H result in both third inferior molars by applying Demirjian's technique limits the number of false positive results (ethically unacceptable errors) to 1 subject in a series of 80; however, this is at the expense of an important increase in the number of false negative results (technically unacceptable errors).

From the practical perspective applied to forensic medicine in Spain and most European countries, the above results are of special importance if we consider the fact that, both from a criminal viewpoint and according to undocumented immigrant acceptance policies, the age limit for the need to adopt urgent criminal or acceptance measures is 18 years.

Our main purpose was centred on analysing the reliability of a forensic age estimation of over or under 18 in a Moroccan population basing age estimations on the tests recommended by the Arbeitsgemeinschaft für Forensische Altersdiagnostik der Deutschen Gesellschaft für Rechtsmedizin. So, we needed to apply a statistical analysis methodology different from that applied to prior studies. The statistical analysis methodology was aimed at evaluating the relative validity of the diagnostic methods used for the specific purpose established in our research, not at providing a mere statistical description of the findings.

In any case, our results should be considered with caution, since the sample selection was affected by two important restraints. The first lies in the fact that only subjects of the male sex were assessed. According to other prior series, it would not be appropriate to assume absolutely that the results obtained for males can be fully extrapolated for females. The second restraint, more important still for the full validation of the study, is the limited reliability of the birth register in the country of origin of our supposed minors. According to UNICEF data, in 1998 Morocco was a country in which a birth certificate was officially required only for events such as school registration and marriage, but not for others such as health care during the paediatric period or for children's vaccination programmes. It is estimated that only 70–80% of all births are registered. On the other hand, the birth registration system in the Kingdom of Morocco is too slack in its application of registration rules, since a significant number of births occur outside a hospital environment. All this means that births occurring on Moroccan territory are not always registered immediately and, when they are recorded, the registered date of birth is not always reliable. This could introduce an incorrect classification bias into the truth criterion used (chronological age of the subjects) and consequently worse results in the validity indicators reported for the forensic age identification methods than would be expected with a strict truth criterion. We therefore have to consider the results of our study with some reserve, in spite of our attempt to use an appropriate method. In fact, a later review of the results, interviewing the supposed minor and his family, showed that the only false positive case (ethically unacceptable) obtained from the combined use of bone age and dental age could actually have corresponded to a false registration of chronological age, lower than the real age.

5. Conclusions

1. For the forensic estimation of age in supposed undocumented minors, we recommend the application of the diagnostic guidelines proposed by the Arbeitsgemeinschaft für Forensische Altersdiagnostik der Deutschen Gesellschaft für Rechtsmedizin.
2. For the forensic estimation of age in undocumented male subjects of Moroccan origin, the direct estimation of the results of applying the Greulich and Pyle method or the estimation of age based on Demirjian's technique applied to the third inferior molars may be acceptable. However,

the degree of error in the results obtained by both methods has to be considered by the physician as considerable, and may represent technically and ethically unacceptable errors.

3. For the forensic estimation of age in undocumented male subjects of around 18 years of age and Moroccan origin, when for reasons of accuracy it is necessary to rule out the appearance of false over-18 results (ethically unacceptable errors), it is recommendable to combine the results from applying the Greulich and Pyle method and Demirjian's technique in both third inferior molars. In these cases, the probability of false negative errors (subjects actually over 18 who are classified by the test as minors) increases significantly.

Acknowledgements

References

[1] A. Schmeling, A. Olze, W. Reisinger, et al., Age estimation of living people undergoing criminal proceedings, Lancet 358 (9276) (2001) 89–90.

[2] A. Schmeling, R. Schulz, W. Reisinger, et al., Studies on the time frame for ossification of the medial clavicular epiphyseal cartilage in conventional radiography, Int. J. Leg. Med. 118 (l) (2004) 5–8.

[3] W.W. Greulich, S.I. Pyle, Radiographic Atlas of Skeletal Development of the Hand and Wrist, Standford University Press, 1950–1959.

[4] A. Demirjian, L.H. Godstein, J.H. Tanner, A new system of dental age assessment, Hum. Biol. 42 (1973) 211–227.

[5] H.H. Mincer, E.F. Harris, H.E. Berryman, The ABFO study of third molar development and its use as an estimator of chronological age, J. Forensic Sci. 38 (2) (1993) 379–390.

[6] J.L. Fleiss, The Design and Analysis of Clinical Experiments, Wiley, New York, 1986.

[7] G. Dunn, Design and analysis of reliability studies, in: The Statistical Evaluation of Measurement Errors, Oxford University Press, London, 1989.

[8] K. Linnet, A review on the methodology for assessing diagnostic tests, Clin. Chem. 34 (1988) 1379–1386.

[9] J.C. Boyd, Mathematical tools for demonstrating the clinical usefulness of biochemical markers, Scand. J. Clin. Lab. Invest. 227 (Suppl.) (1997) 46–63.

[10] J.M. Tanner, Growth at Adolescence, Oxford Blackwell Scientific Publications, 1962.

[11] D.G. King, D.M. Steventon, M.P. O'Sullivan, et al., Reproducibility of bone ages when performed by radiology registrars: an audit of Tanner and Whitehouse II versus Greulich and Pyle methods, Brit. J. Radiol. 67 (1994) 848–851.

[12] L. Benso, S. Vanelli, L. Pastorin, et al., Variation of bone age progression in healthy children, Acta Pediatr. 423 (Suppl.) (1997) 109–112.

[13] G. Beunen, J. Lefevre, M. Ostyn, et al., Skeletal maturity in Belgian youths assessed by the Tanner–Whitehouse method (TW2), Ann. Hum. Biol. 17 (5) (1990) 355–376.

[14] R.R. Van Rijn, M.H. Lequin, S.G. Robbens, et al., Is the Greulich and Pyle atlas still valid for Dutch Caucasian children today? Pediatr. Radiol. 31 (10) (2001) 748–752.

[15] K. Koski, J. Haataja, M. Lappalainen, Skeletal development of hand and wrist in Finnish children, Am. J. Phys. Anthropol. 19 (1961) 379–382.

[16] E. Andersen, Comparison of TW2 and GP methods in a large scale Danish survey, Am. J. Phys. Anthropol. 35 (1971) 373–376.

[17] J. Taranger, J. Karlberg, B. Bruning, I. Engstrom, Standard deviation score charts of skeletal maturity and velocity in Swedish children assessed by the TW2 method (TW2-20), Ann. Hum. Biol. 14 (1987) 357–365.

[18] M. Hernández, E. Sánchez, B. Sobradillo, J.M. Rincon, Maduracion osea y prediction de la talla, Diaz de Santos, 1991.

[19] A. Wenzel, H. Droschl, B. Melsen, Skeletal maturity in Austrian children assessed by the GP and TW2 methods, Ann. Hum. Biol. 11 (1984) 173–177.

[20] M. Muramata, Population specific reference values for bone age, Acta Pediatr. 423 (Suppl.) (1997) 113–114.

[21] Y. Yi-Yan, W. Chuang Xing, C. Li-Zhi, Skeletal maturity of the hand and wrist in Chinese children in Changsha assessed by the TW2 method, Ann. Hum. Biol. 19 (4) (1992) 427–430.

[22] R.M. Rikhasor, A.M. Quershi, S.L. Rathi, N.A. Channa, Skeletal maturity in Pakistani children, J. Anat. 195 (2) (1999) 305–308.

[23] J. Tanner, D. Oshman, F. Bahhage, M. Healy, Tanner–Whitehouse bone age reference values for North American children, J. Pediatr. 131 (1997) 34–40.

[24] G.W. Gross, J.M. Boone, D.M. Bishop, Pediatric skeletal age: determination with neural networks, Radiology 195 (1995) 686–695.

[25] R. Lodler, D.T. Estle, K. Morrison, et al., Applicability of the Greulich and Pyle skeletal age standards to black and white children of today, AJDC 147 (1993) 1329–1333.

[26] F.K. Ontell, M. Ivanovic, D.S. Ablin, T.W. Barlow, Bone age in children of diverse ethnicity, AJR 167 (1996) 1395–1398.

[27] W.A. Marshall, M.T. Ashcroft, G. Bryan, Skeletal maturation of the hand and wrist in Jamaican children, Hum. Biol. 42 (1970) 419–435.

[28] V. Gilsanz, D.L. Skaggs, A. Kovanlikaya, et al., Differential effect of race on the axial appendicular skeletons of children, J. Clin. Endocrinol. Metab. 83 (1988) 1420–1427.

[29] A. Koc, M. Karaoglanoglu, M. Erdogan, et al., Assessment of bone ages: is the Greulich–Pyle method sufficient for Turkish boys, Pediatr. Int. 43 (6) (2001) 662–665.

[30] A. Schmeling, W. Reisinger, D. Loreck, et al., Effects of ethnicity on skeletal maturation: consequences for forensic age estimation, Int. J. Leg. Med. 113 (5) (2000) 253–258.

[31] A.E. Oestreich, Tanner–Whitehouse versus Greulich–Pyle in bone age determinations, J. Pediatr. 131 (1) (1997) 5–6.

[32] A.B. Jahari, J. Haas, M.A. Husaini, E. Pollitt, Effects of an energy and micronutrient supplement on skeletal maturation in undernourished children in Indonesia, Eur. J. Clin. Nutr. 54 (Suppl. 2) (2000) 74–79.

[33] K. Fleshman, Bone age determination in a pediatric population as an indicator of nutritional status, Trop. Doct. 30 (1) (2000) 16–18.

[34] B. Melsen, A. Wenzel, T. Miletic, et al., Dental and skeletal maturity in adoptive children: assessment at arrival and after one year in the admitting country, Ann. Hum. Biol. 13 (2) (1986) 153–159.

[35] L. Kulman, Accuracy of two dental and one skeletal age estimation method in Swedish adolescents, Forensic Sci. Int. 75 (1995) 225–236.

[36] P. Nambiar, H. Yaacob, R. Menon, Third molars in the establishment of adult status. Case report, J. Forensic Odonto-Stomatol. 14 (2) (1996) 30–33.

[37] A.C. Solari, K. Abramovitch, The accuracy and precision of third molar development as indicator of chronological age in hispanics, J. Forensic Sci. 47 (3) (2002) 531–535.

[38] P. Piscozzi, II ruolo dei terzi molari neH'accertamento della maggior eta, Minerva Med. Leg. 120 (2000) 199–205.