TITLE: Combining dental and skeletal evidence in age classification: pilot study in a sample of Italian sub-adults.

AUTHORS: Vilma Pinchi*, Federica De Luca* , Martina Focardi*, Francesco Pradella*, Giulia Vitale*, Federico Ricciardi,Gian-Aristide Norelli***

*Department of Health Sciences, Sectionof Medical Forensic Sciences, University of Florence, Largo Brambilla 3, 50134, Florence, Italy

**Department of Statistics, Computer Science, Application (DISIA) University of Florence, Viale Morgagni 59, 50134, Florence, Italy.

Corresponding Author :VilmaPinchi, pinchi@unifi.it

Department of Health Sciences, Forensic Medicine Sciences Section, University of Florence, Largo Brambilla 3, 50134, Florence, Italy

Abstract

Background: Dental and skeletal maturation have proved to be reliable evidence for estimating age of children and prior studies and internationally accredited guidelines recommend to evaluate both evidence in the same subject to reduce error in age prediction. Nevertheless the ethical and legal justification of procedures that imply a double exposition of children stands as a relevant issue. This study aims to evaluate the accuracy of age estimation provided by a combination of skeletal and dental methods applied in the same sample of children.

Materials and methods: The sample consisted of 274 ortopantomographies and left hand-wrist X-rays of Italian children, (aged between 6 and 17 years) taken on the same day. Greulich and Pyle's (GP), Tanner-Whitehouse's version 3 (TW3) and Willems' (W) and the Demirjian's (D) methods were respectively applied for estimating skeletal and dental age. A combination of skeletal and dental age estimates through Linear Discriminant Analysis (LDA) is proposed to obtain a classifier respect to an age threshold.

Results: The combination of D and TW3 obtained an improvement of accuracy in classifying female subjects respect to the 12 years threshold respect to the original methods (from about 77% using either original methods to 83.3% combining TW3+D) as well as a consistent reduction of false positives rate (from 17-21% for original methods to 5.6% with TW3+D). For males the LDA classifier (based on TW3 and W) enable a small improvement in accuracy, whilst the decreasing of false positives was as noticeable as for females (from 17.6- 14.1% for original methods to 6.2% combining TW3+W).

Conclusions. Although the study is influenced by the limited size and the uneven age distribution of the sample, the present findings support the conclusion that age assessment procedures based on both dental and skeletal age estimation can improve the accuracy and reduce the occurrence of false positives.

Combining dental and skeletal evidence in age classification: pilot study in a sample of Italian sub-adults.

INTRODUCTION

In forensic practice, age estimation assessments in living children and adolescents have increased due to irregular immigration, asylum seeker proceedings, criminality in adolescence and adoption procedures [1]. Legal requirements depend on whether age estimation is requested for criminal proceedings or for civil or administrative purposes. Moreover, legal age thresholds of criminal responsibility vary considerably worldwide [2]. Dental calcification and wrist-hand bone maturation have proved to be reliable tools for estimating chronological age or classifying individuals with respect to an age threshold. Many studies [3–16] have concluded that an age assessment for legal or forensic purposes should rely on estimation of both dental and skeletal age. Nevertheless very few researchers had the opportunity to compare the estimation of dental age with bone age in the same sample of sub-adults because of the difficulties in collecting samples of known age, who underwent wrist X-rays and ortopantomography in the same period of time [9–11].

In the past dentists requested dental pantomographies and wrist-hand radiographies before orthodontic treatments of children. Nowadays orthodontists can obtain the requested auxological stage from cervical vertebrae development on lateral X-rays of the skull used for cephalograms, thus avoiding the additional exposition due to wrist radiography [12–14]. This difficulty in collecting appropriate sample of dental and wrist radiographies explains the low number of papers dealing with these observations. The aim of this study is to evaluate the accuracy of age estimation provided by different skeletal and dental methods applied in orthopantomography and wrist X-rays taken from the same children on the same day. Additionally, a procedural combination of skeletal and dental age estimates by means of Linear Discriminant Analysis (LDA) is experimented in order to verify whether the accuracy of final age assessment improves when a dental method is combined with a skeletal one.

MATERIALS AND METHODS

A sample of the left wrist-hand radiographies and ortopantomographies from 274 individuals (139 males and 135 females) 6-17 years of age (table 1) was selected using the following inclusion criteria:

- dental pantomography and left hand and wrist radiography taken on the same day;

- Unremarkable medical history;

- Known gender, date of birth and date of X-ray examination

A specialist in forensic medicine examined the left hand and wrist x-rays and estimated skeletal age according to the atlas method of Greulich and Pyle (GP) [17] and the score method of Tanner and Whitehouse, RUS, version 3 (TW3) [18]. GP and TW3 were selected since GP is largely used in Italy and TW3 proved to be reliable for bone age estimation in Italian children [7], whilst other methods (Gisland and Ratib digital atlas, e.g.) are not used in Italy [19,20]. A forensic odontologist provided estimations of dental age based on teeth development applying Willems' (W) [21] and Demirjian's (D) methods [22,23]. For the present study, the age of 12 is selected as the age threshold of interest, since it is relevant in some countries and because it is the most suitable to be analysed according to the age distribution of the sample (Table 1).

The reproducibility of the methods was evaluated using the Intra-class Correlation Coefficient (ICC). Inter-rater variability was investigated by submitting 27 (about 10% of the sample) randomly selected X-rays to two different experts provided by similar scientific background and experience of the original operators.

Finally we experimented the use of LDA to combine dental and bone age estimates. LDA is a statistical methodology used to find a linear combination of variables that is adequate to separate two (or more) classes of objects or events. In this context it aims to express a dependent variable (the classification of a subject with respect to the age threshold) as a function of a set of explicatory variables, which are the different estimates of the age provided by the different applied methods (TW3, GP, W and D).

The best dental and skeletal method were selected by comparing the mean error (both *overall* and *per gender*), accuracy, false positive and false negative rates, when classifying subjects with respect to the age threshold of 12 years. LDA was then applied to the selected methods and the performance of the LDA classifier [24] was compared to each original method. Since the estimated ages were assumed as independent and continuous variables, and in LDA independent variables must be normal, Shapiro-Wilk normality tests were performed and a graphical overview of normality is given by means of Q-Q plots.

In order to validate the approach based on LDA, the sample was randomly divided using the R statistical computing environment and individuals were assigned to a **training sample,** which was composed of 60% of the original sample, and to a **test sample.** The predictive performance of the

LDA model was then evaluated using test sample. This is a quite common technique for assessing how the results of a statistical analysis will generalize to another dataset.

Table 1. about here

For a preliminary comparison among methods, we considered the difference, express through Mean Error, between estimated age and chronological age. The mean error gives a measure of differences between chronological and estimated ages that is not affected by the sign (positive or negative) of the difference, thus providing a more easily interpretable measure of the distance between "reality" and "guess".

The Correct Classification rate (also referred to as "Accuracy" throughout the paper) of each method indicates the proportion of individuals that are correctly classified with respect to the age threshold of interest.

The false positive rate (children younger than 12 misclassified as older) and the false negative rate (individuals older than 12 misclassified as younger) have been assessed.

RESULTS

The ICC values resulted to be 0.93 for GP, 0.90 for TW3, 0.85 for W and 0.82 for D. Hence all applied methods demonstrated to be highly reproducible since all these values are higher than the minimum satisfactory level, traditionally fixed at 0.80 [25].

Table 2 shows the mean error of estimations for each method calculated for the whole sample and separately for each gender. Skeletal methods yielded slightly lower mean error than dental methods in both genders. Furthermore TW3 and GP produced similar mean error (0.922 GP and 0.916 TW3) while, among dental methods, D was associated with lower overall errors, even if W performed better for male subjects.

Table 2. about here

The performance of the methods in the classification of individuals with respect to the age threshold of 12 is displayed in Table 3.

For accuracy, it emerged that TW3 performs better for male individuals (83.5%), whilst other methods (GP, D, W) obtained similar values for both males and females (ranging from 75% to 79%). The Willems' and the Demirjian's methods were consistently much more prone to age underestimation for both genders than skeletal methods, leading to higher occurrence of false negatives. Regarding the false positive rate, our data revealed that W was the safer method and that GP was affected by an impressively high percentage of overestimation, being 27% of males and 31% of females misclassified over the age threshold.

Table 3 about here

The comparison of mean error and accuracy values showed that TW3 performed better than GP, thus it was chosen for the LDA framework. The results were less definitive for dental methods since D was associated with lower overall mean error whilst W was the more accurate in males (79.1% W vs 75.5% D). Since LDA requires that independent variables are normal for each level of the grouping variable, the normality tests were performed for TW3 and dental methods using the Shapiro-Wilk test for normality according its approximated version proposed by Royston [26]. Figure 1 and Table 4 report the normality test results: D slightly failed in males <12 years (p-value 0.045) and W in females <12 years (p-value 0.044). Hence results of the normality tests, the accuracy and mean error values lead us to choose D for females and W for males to be combined with TW3 as skeletal method in the LDA framework.

Table 4 . about here

Figure 1 : about here

The test and training sub-samples are assumed as independent due to their random composition and the LDA classifier obtained from the training sample was then applied to the test sample. The Correct Classification rate, false negative rate and false positive rate were calculated for LDA classifier separately for males and females. Table 3 (last line) reports the values of the Correct Classification rate, the false positive rate and false negative rate yielded by LDA classifiers for both genders (TW3+D for females and TW3 + W for males). An improvement of the performance emerged when the LDA classifiers were used to classify individuals' age at the threshold of 12 years instead of single methods (TW3, D, W).

For female individuals the LDA based on TW3 and D improved the accuracy to 83.3 % and reduced consistently the false positive rate to 5.6%, whilst the false positive rate of the original methods were respectively 21.9% (TW3) and 17.7% (D). In male individuals the LDA enables an accuracy improvement (80.4%) with respect to D method (75.5%), but a little worsening in comparison to TW3 (83.5%). As for females the most remarkable improvement emerged from the false positive rate that decreases to 6.2% for LDA whilst for the original methods the values were respectively of 17.6% for TW3 and 14.1% for W.

DISCUSSION

In many countries there is a growing demand for forensic age estimation of living children and adolescents related to criminal or civil proceedings or asylum seeker procedures and age thresholds for criminal liability vary ranging from 8-10 to 18-21 years in different legal systems [27]. Skeletal and dental evidence have both proved effective and reliable for age estimation, but the occurrence of prediction errors continues to stimulate researches looking for ever more accurate methods and procedures as well as suitable statistical approaches [3–7, 9–11, 28-41]. Garamendi et al. [10] highlighted the difference between technically unacceptable errors (misclassifications under the threshold) and ethically unacceptable errors (misclassification over the threshold). Ethically unacceptable errors are the most detrimental to the child's best interest, particularly in criminal proceedings and the false positive rate of the applied methods should be accurately investigated and reported. Since age estimation is mainly based on radiological methods for the analysis of teeth and bone maturation, the X-rays exposition of growing subjects for forensic purposes, without clinical indications, stands as a relevant ethical issue [42-44]. The combined dental and skeletal age estimation in the same subject to improve accuracy and reduce misclassifications of the age implies double expositions for children, the involvement of two forensic experts, more expensive and time consuming procedures. As stated by Maber et al. [29] the scientific evidence can be built by comparing similar studies with consistent conclusions, but the comparison of this study with very few prior researches is limited due to different samples, methods and statistical approaches used by Authors.

Among dental methods, W and D were chosen since were reported to be reliable for estimating age of Italian subadults [23] and D produced a general underestimation of age in both genders. These results are consistent with results of Cameriere et al. [34], who reported that D underestimates the ages of Italian, Spanish and Croatian female subadults and Maber et al. [29], whose research revealed a mean underestimation of -0.05 years for boys and -0.20 years for girls. On the contrary, many studies [19, 37, 41, 45-47] reported that the Demirjian's method tended to overestimate age. Similarly to other studies [19, 29, 41, 48,49], we found W to be more accurate than D.

Regarding the bone age assessment this study cannot address properly the possible influence of ethnic, environmental or socio-economic status on dental or bone maturation since the sample was ethnically homogenous and no information about life conditions were available [19]. Consistently with previous researches [10, 50-53], GP showed the highest false positive rate among all methods for both genders (27% for males and 31% for females). In agreement with Serinelli et al. [52], TW3 resulted to be more accurate than GP yielding a false positive rate of 17.6% for male and 21.9% for female individuals. Among all methods applied here TW3 resulted the most accurate for male individuals (83.5%), whilst all methods produced very similar accuracy (~77%) for females. As previously reported in literature wrist-hand bone maturation and dental calcification resulted to be affected by gender [15, 23,54-59].

The combination of dental and skeletal age estimation methods has consistently been regarded as a procedure endowed with higher accuracy than age assessments performed using only either wrist or dental maturation evidence [54-58]. Despite limitations due to the sample size and the uneven age distribution, the combination of dental method with a skeletal method through LDA improves accuracy in classifying subjects over/under the age threshold of 12 years and reduces false positive rate for both genders with respect to original methods. Hence this approach reduces the risk of misclassifications over the threshold of age, thereby supporting the legal and ethical justification [24] of procedures that require double expositions (dental and wrist radiographies) of young individuals undergoing an age assessment. The results of this study are consistent with very few similar studies retrievable in literature. Santoro et al. [9] compared GP and D in an Italian sample and noticed that mean errors were greater in older children for both GP and D. The study is limited to a comparison of estimated ages (dental and skeletal), whilst a real combination of the two methods was not tested. Cameriere et al. [56] reported an improvement of accuracy in all age cohorts by combining estimates based on dental and skeletal evidence, but the remarkable differences about the applied methods for age assessment and the age range of the sample jeopardize the possibility of further comparison with our results. Garamendi et Al. [10] found that a combination of lower third molar calcification and wrist development enables consistent improvements in accuracy and reduction of false positive cases (specificity) in age estimation of Moroccan boys. As rigorously reported by the Authors themselves, two relevant biases affected this study: only male subjects were considered and the chronological ages of the sample was uncertain due to the limited reliability of the Moroccan birth registration system.

CONCLUSIONS

Robust scientific literature demonstrated that hand and wrist maturation and dental calcification are both reliable pieces of evidence that can be used to predict the age of children.

The results of this research show that skeletal methods based on wrist/hand bones (GP and TW3) resulted slightly more accurate than dental methods (D and W), but the latter produce lower rates of false positives, which are the worst ethical and legal errors.

The LDA classifiers, which combine TW3 with W for females and D for males, gained an improvement of accuracy and a consistent reduction of false positive rate respect to original methods separately applied.

The present findings support the conclusion that age assessment procedures based on both dental and skeletal age estimations is a valid approach that can improve the accuracy of the final prediction and reduce the occurrence of overestimations. Future developments of this study should consider larger samples, with more homogenous age distributions so that 14 or 16 age thresholds could be properly investigated.

REFERENCES

1. Schmeling A, Olze A, Reisinger W, Rösing FW, Geserick G, Forensic age diagnostics of living individuals in criminal proceedings, Homo 2003;54: 162-9
2. Janes L, Criminal liability of minors and severity of penalties: European trends and Developments, The Howard League for penal reform, England and Wales (2008). http://www.europeanrights.eu/public/commenti/LauraJanes_en.pdf. Accessed 28Th August 2015
3. Ritz-Timme S, Cattaneo C, Collins MJ, Waite ER, Schütz HW, Kaatsch HJ, Borrman HI, Age estimation: the state of the art in relation to the specific demands of forensic practice, Int. J. Legal Me 2000; 113: 129-36.
4. Cunha E, Baccino E, Martrille L, Ramsthaler F , Prieto J, Schuliar Y, Lynnerup N, C. Cattaneo, The problem of aging human remains and living individuals: a review, Forensic Sci. Int. 193(1-3) (2009) 1-13.

5. Schmeling A, Geserick G, Reisinger W, Olze A, Age estimation, Forensic Sci Int 2007; 165:178-81.
6. Schmeling A, Olze A, Reisinger W, Geserick G, Age estimation of living people undergoing criminal proceedings, Lancet 2001; 358: 89-90.
7. Pinchi V, De Luca F, Ricciardi F, Focardi M, Piredda V, Mazzeo E, Norelli GA Skeletal age estimation for forensic purposes: A comparison of GP, TW2 and TW3 methods on an Italian sample Forensic Sci Int. 2014;238:83-90.
8. Solheim T, Vonen A, Dental age estimation, quality assurance and age estimation of asylum seekers in Norway, Forensic Sci 2006;159 , Suppl 1:S56-60.
9. Santoro V, Roca R, De Donno A, Fiandaca C, Pinto G, Tafuri G, Introna F, Applicability of Greulich and Pyle and Demirjian aging methods to a sample of Italian population, Forensic Sci Int. 2012; 153.e1-5.
10. Garamendi PM, Landa MI, Ballesteros J, Solano MA, 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 2005; 154: 3-12.
11. Kullman L, Accuracy of two dental and one skeletal age estimation method in Swedish adolescents, Forensic Sci Int 1995; 75: 225-36.
12. Baccetti T, Franchi L, McNamara JA, An improved version of the cervical vertebral maturation (CVM) method for the assessment of mandibular growth, Angle Orthod. 2002; 72: 316-23.
13. Santiago TC, de Miranda Costa LF, Vitral RW, Fraga MR, Bolognese AM, Maia LC, Cervical vertebral maturation as a biologic indicator of skeletal maturity, Angle Orthod. 2012; 82: 1123-31.
14. Baccetti T, Franchi L, De Toffol L, Ghiozzi B, Cozza P, The diagnostic performance of chronologic age in the assessment of skeletal maturity, Prog Orthod 2006; 7: 176-88.
15. Schmeling A, Olze A, Reisinger W, König M, Geserick G, Statistical analysis and verification of forensic age estimation of living persons in the Institute of Legal Medicine of the Berlin University Hospital Charité, Leg Med (Tokyo), 2003; 5: S367–S371
16. Schmeling A, Garamendi PM, Prieto JL, Landa MI, Forensic Age Estimation in Unaccompanied Minors and Young Living Adults, http://cdn.intechopen.com/pdfs-wm/19163.pdf, accessed 30th August 2015
17. Greulich W, Pyle SI, Radiographic Atlas of Skeletal Development of the Hand and Wrist, second ed., Stafford University Press, Stafford, 1959.
18. Tanner JM, Healy MJR, Goldstein H, Cameron N, Assessment of Skeletal Maturity and Prediction of Adult Height (TW3 method), Saunders, London, 2001.
19. Gisland V, Ratib O, Hand bone age. A digital atlasfor skeletal matutity, Springer-Verlag Berlin Heidelberg 2005
20. Schmidt S1, Nitz I, Schulz R, Tsokos M, Schmeling A. The digital atlas of skeletal maturity by Gilsanz and Ratib: a suitable alternative for age estimation of living individuals in criminal proceedings? Int J Legal Med. 2009 Nov;123(6):489-94.
21. Willems G,Van Olmen A, Spiessens B, Carels C, Dental age estimation in Belgian children: Demirjian's technique revisited, J Forensic Sci 2001; 46: 893–895.
22. Demirjian A, Goldstein H, Tanner JM, A new system of dental age assessment, Hum Biol 1973;45: 221–227.

23. Pinchi V, Norelli GA, Pradella F,Vitale G, Rugo D, Nieri M, Comparison of the applicability of four odontological methods for age estimation of the 14 years legal threshold in a sample of Italian adolescents, J Forensic Odontostomatol 2012; 30:17-25.
24. Härdle W, Simar L, Applied Multivariate Statistical Analysis, 2nd extended Ed, Springer-Verlag, Berlin, Heidelberg, 2007.
25. Ferrante L, Cameriere R, Statistical methods to assess the reliability of measurements in the procedures for forensic age estimation, Int J Leg Med, 2009; 123: 277-283.
26. Royston P, A pocket-calculator algoritm for the Shapiro Francia Test for non-normality: an application to medicine. Statistics in medicine, 1993; 12: 181-184
27. Children's Rights and the Minimum Age of Criminal Responsibility. A global perspective, Ashgate Publishing, Farnham, Surrey UK, 2009
28. Corradi F, Pinchi V, Barsanti I, Garatti S.Probabilistic classification of age by third molar development: the use of soft evidence. J Forensic Sci. 2013 Jan;58(1):51-9.
29. Maber M, Liversidge H.M., Hector M.P., Accuracy of age estimation of radiographic methods using developing teeth, Forensic Sci. Int159 Suppl 1(2006) S68-73
30. Thevissen PW, Fieuws S, Willems G, Human third molars development: Comparison of 9 country specific populations. Forensic Sci Int 2010;201:102-5.
31. Liversidge HM, Similarity in dental maturation in two ethnic groups of London children. Ann Hum Biol2011; 38: 702-15.
32. Corradi F, Pinchi V, Barsanti I , Manca R , Garatti R, Optimal age classification of young individuals based on dental evidence in civil and criminal proceedings, Int J Legal Med 2013; 127: 1157-64.
33. Akkaya N, Yilanci HO, Göksülük D, Applicability of Demirjian's four methods and Willems method for age estimation in a sample of Turkish children, Leg Med (Tokyo), 2015 in press
34. Cameriere R, Ferrante L, Liversidge HM, Prieto JL, Brkic H, Accuracy of age estimation in children using radiograph of developing teeth, Forensic Sci Int 2008;176: 173–7
35. Trevino-Tijerinaa mc, Valenzuela-Garach A, Elizondo-Pereo RA, Cerda-Flores RM, Vargas-Villarreal J, González-Salazarc F, Age estimation of teenagers from Monterrey (Mexico) by the evaluation of dental mineralization after multi-slice helical computed tomography, Aus J Forensic Sci, 2015; 47: in press
36. Sironi E, Gallidabino M, Weyermann C, Taroni F,Probabilistic graphical models to deal with age estimation of living persons. Int J Legal Med. 2015 Mar 21, in press
37. Deitos AR, Costa C, Michel-Crosato E, Galić I, Cameriere R, Biazevic MG, Age estimation among Brazilians: Younger or older than 18? J Forensic Leg Med. 2015 Jul;33:111-5
38. Lewis JM, Senn DR. Dental age estimation utilizing third molar development: A review of principles, methods, and population studies used in the United States. Forensic Sci Int. 2010 Sep 10;201(1-3):79-83.
39. Pinchi V, Pradella F, Vitale G, Rugo D, Nieri M, Norelli GA.Comparison of the diagnostic accuracy, sensitivity and specificity of four odontological methods for age evaluation in Italian children at the age threshold of 14 years using ROC curves. Med Sci Law. 2016 Jan;56(1):13-8.

40. Ambarkova V, Galić I, Vodanović M, Biočina-Lukenda D , Brkić H, Dental age estimation using Demirjian and Willems methods: cross sectional study on children from the Former Yugoslav Republic of Macedonia, Forensic Sci Int. 2014; 234: 187.e1-7
41. Thevissen P, S.I. Kvaal, K. Dierickx, G. Willems, Ethics in Age Estimation of Unaccompanied Minors, J. Forensic Odontostomatol. 30(2012) 85-102.
42. Aynsley Green A, Unethical Age assessment, Br Dent J 2009; 206: 337.
43. Focardi M, Pinchi V, De Luca F, Norelli GA Age estimation for forensic purposes in Italy: ethical issues Int J Legal Med. 2014;128(3):515-22.
44. Urzel V, Bruzek J, Dental age assessment in children: a comparison of four methods in a recent French population, J Forensic Sci 2013; 58: 1341-7.
45. Foti B, Lalys L, Adalian P, Giustiniani J , Maczel M, Signoli M, Dutour O, Leonetti G, New forensic approach to age determination in children based on tooth eruption, Forensic Sci Int 2003; 132: 49-56.
46. El-Bakary AA, M.H. Shaza, M. Fatma, Dental age estimation in Egyptian children, comparison between two methods, J. Forensic Leg. Med. 17(2010)363-367
47. Schmeling A, Olze A, Reisinger W, Geserick G, Forensic age estimation and ethnicity, Leg. Med. (Tokyo) , 2005; 7: 134-7
48. Mani S. Naing L, John J, Samsudin AR, Comparison of two methods of dental age estimation in 7–15-year-old Malays, Int J Paediatr Dent 2008; 18: 380–388.
49. Grover S, Marya CM, Avinash J, Pruthi N, Estimation of dental age and its comparison with chronological age: accuracy of two radiographic methods, Med Sci Law 2012;52:32-5.
50. Schmidt S, Koch B, Schulz R, Reisinger W, Schmeling A, Studies in use of the Greulich–Pyle skeletal age method to assess criminal liability, Leg. Med. (Tokyo) 10(2008) 190-5.
51. Molinari L, Gasser T, Largo RH. TW3 bone age: RUS/CB and gender differences of percentiles for score and score increments. Ann Hum Biol 2004; 31: 421-35.
52. Serinelli S, Panetta V, Pasqualetti P, Marchetti D, Accuracy of three age determination X-ray methods on the left hand-wrist: a systematic review and meta-analysis, Leg. Med. (Tokyo). 2011;13: 120-33.
53. Büken B, Safak AA, Yazici B, Büken E, Mayda AS, Is the assessment of bone age by the Greulich-Pyle method reliable at forensic age estimation for Turkish children? Forensic Sci Int 2007; 173: 146-53.
54. Thevissen PW, Kaur J, Willems G, Human age estimation combining third molar and skeletal development. Int J Legal Med 2012; 126: 285-92.
55. Willems G, Thevissen P, The triple test: age estimation protocol for unaccompanied fugitives developed at the KU Leuven, Belgium, J Forensic Odontostomatol.2013; 31:77-8.
56. Cameriere R , Ferrante L, 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 2008;174: 60-3.
57. Suma G, Rao BB, Annigeri RG, Rao DJ, Goel S, Radiographic correlation of dental and skeletal age: Third molar, an age indicator, J Forensic Dent Sci 2011;3:14-8.
58. Cameriere R, De Luca S , Biagi R, Cingolani M, Farronato G, Ferrante L, Accuracy of three age estimation methods in children by measurements of developing teeth and carpals and epiphyses of the ulna and radius, J Forensic Sci 2012; 57: 1263-70.

59. Pruvost MO, Boraud C, Chariot P. Skeletal age determination in adolescents involved in judicial procedures: from evidence-based principles to medical practice. J Med Ethics. 2010 Feb;36(2):71-4.

Figure

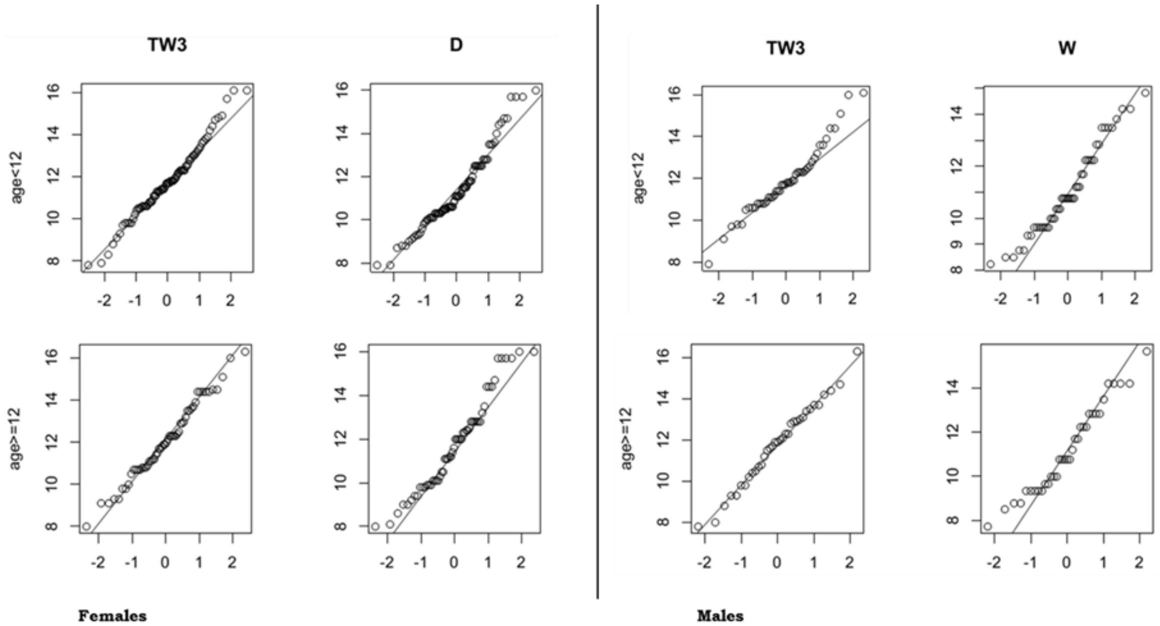

Figure 1 : Q-Q plots for TW3 and D for females and for TW3 and W for males

Table

Age	Females	Males	Total
6	1	0	1
7	1	3	4
8	4	4	8
9	35	8	43
10	31	38	69
11	24	32	56
12	17	27	44
13	15	12	27
14	3	10	13
15	3	1	4
16	0	3	3
17	1	1	2
Total	**135**	**139**	**274**

Table 1. Composition of the sample

	Mean Error		
Method	**Overall**	**Females**	**Males**
GP	0.922	0.945	0.901
TW3	0.916	0.950	0.882
W	1.187	1.267	1.110
D	1.108	1.081	1.134

Table 2. Overall (left), Females (middle) and Males (right) Mean Errors (ME) related to the four methods considered.

	Females				Males		
Method	**CC**	**FPr**	**FNr**		**CC**	**FPr**	**FNr**
TW3	77.8%	21.9%	23.1%		83.5%	17.6%	14.8%
GP	77.0%	31.2%	2.6%		78.4%	27.1%	13.0%
D	77.0%	17.7%	35.9%		75.5%	20.0%	31.5%
W	77.8%	9.4%	53.8%		79.1%	14.1 %	31.5%
LDA TW3 + D	83.3%	5.6%	38.9%	**LDA TW3 + W**	80.4%	6.2%	37.5%

Table 3. Accuracy (CC), False Positive (FP) and False Negative (FN) rates for the four methods considered, by gender. CC, FPr and FNr obtained by linear discriminant analysis (LDA) classifier for females (TW3+D) and for males (TW3+W).

	Females		Males	
Methods	**<12 years**	**>12 years**	**<12 years**	**> 12 years**
TW3	0.24	0.12	0.09	0.95
D	0.48	0.41	0.04	0.57
W	0.04	0.12	0.09	0.23

Table 4. The table shows the results (p-values) of the Shapiro-Wilk tests for normality applied to the methods on the sample of female and male individuals subdivided by the age threshold of 12 years