Accuracy of age estimation of radiographic methods using developing teeth

M. Maber, H.M. Liversidge *, M.P. Hector

Dental Institute, Bart’s and The London School of Medicine and Dentistry, Turner Street, London E1 2AD, UK

Available online 14 March 2006

Abstract

Developing teeth are used to assess maturity and estimate age in a number of disciplines, however the accuracy of different methods has not been systematically investigated. The aim of this study was to determine the accuracy of several methods. Tooth formation was assessed from radiographs of healthy children attending a dental teaching hospital. The sample was 946 children (491 boys, 455 girls, aged 3–16.99 years) with similar number of children from Bangladeshi and British Caucasian ethnic origin. Panoramic radiographs were examined and seven mandibular teeth staged according to Demirjian’s dental maturity scale [A. Demirjian, Dental development, CD-ROM, Silver Platter Education, University of Montreal, Montreal, 1993–1994; A. Demirjian, H. Goldstein, J.M. Tanner, A new system of dental age assessment, Hum. Biol. 45 (1973) 211–227; A. Demirjian, H. Goldstein, New systems for dental maturity based on seven and four teeth, Ann. Hum. Biol. 3 (1976) 411–421], Nolla [C.M. Nolla, The development of the permanent teeth, J. Dent. Child. 27 (1960) 254–266] and Haavikko [K. Haavikko, The formation and the alveolar and clinical eruption of the permanent teeth. An orthopantomographic study. Proc. Finn. Dent. Soc. 66 (1970) 103–170]. Dental age was calculated for each method, including an adaptation of Demirjian’s method with updated scoring [G. Willems, A. Van Olmen, B. Spiessens, C. Carels, Dental age estimation in Belgian children: Demirjian’s technique revisited, J. Forensic Sci. 46 (2001) 893–895]. The mean difference (±S.D. in years) between dental and real age was calculated for each method and in the case of Haavikko, each tooth type; and tested using t-test. Mean difference was also calculated for the age group 3–13.99 years for Haavikko (mean and individual teeth). Results show that the most accurate method was by Willems [G. Willems, A. Van Olmen, B. Spiessens, C. Carels, Dental age estimation in Belgian children: Demirjian’s technique revisited, J. Forensic Sci. 46 (2001) 893–895] (boys 0.05 ± 0.81, girls −0.20 ± 0.89, both −0.12 ± 0.85), Demirjian [A. Demirjian, Dental development, CD-ROM, Silver Platter Education, University of Montreal, Montreal, 1993–1994] overestimated age (boys 0.25 ± 0.84, girls 0.23 ± 0.84, both 0.24 ± 0.86), while Nolla [C.M. Nolla, The development of the permanent teeth, J. Dent. Child. 27 (1960) 254–266] and Haavikko’s [K. Haavikko, The formation and the alveolar and clinical eruption of the permanent teeth. An orthopantomographic study, Proc. Finn. Dent. Soc. 66 (1970) 103–170] methods underestimated age (boys −0.87 ± 0.87, girls −1.18 ± 0.96, both −1.02 ± 0.93; boys −0.56 ± 0.91, girls −0.79 ± 1.11, both −0.67 ± 1.01, respectively). For individual teeth using Haavikko’s method, first premolar and second molar were most accurate; and more accurate than the mean value of all developing teeth. The 95% confidence interval of the mean was least for mean of all developing teeth using Haavikko (age 3–13.99 years), followed by identical values for Demirjian and Willems (sexes combined).

Keywords: Accuracy; Developing teeth; Age determination

1. Introduction

Tooth formation is widely used to assess maturity and predict age. Within clinical dentistry this information aids in diagnosis and treatment planning. In archaeology and forensic odontology, age estimation methods can aid the identification of age at death of a deceased child and also give important information with regard to past populations [7]. Age estimation is also proving valuable when birth data is lacking or doubted in the management of immigration to help determine physiological age.

Previous research on accuracy of dental age estimation methods is complicated by different sample sizes, age structures, grouping and statistical analysis making comparisons difficult. The majority of studies have looked at a single method, others use several methods, some investigate living children [8–13], some report on skeletal remains [14,15]. The aim of this study was to determine the accuracy of four methods of age estimation using developing teeth from radiographs.
2. Materials and methods

2.1. Study design

The design of this study was a retrospective cross sectional study of radiographs. These were good quality panoramic radiographs taken in the course of diagnosis and treatment at the paediatric dental department of the Royal London Hospital.

2.2. Sample

The sample was 946 healthy children (491 males and 455 females) aged between 3.00 and 16.99 years of age shown in Fig. 1. The ethnic origin of the sample was Bangladeshi (258 boys and 219 girls) and British Caucasian (233 boys and 236 girls). Radiographs that were unclear or that showed hypodontia, gross pathology, and previous orthodontic treatment were excluded. The chronological age for each subject was calculated by subtracting the date of the radiograph from the date of birth after having converted both to a decimal age using Eveleth and Tanner [16].

2.3. Dental age methods tested

The radiographs were assessed to determine the developmental stages of the teeth according to the following methods: Demirjian [1], Nolla [4], and Haavikko [5]. The first method was Demirjian’s Dental Development CD-Rom [1] which calculates dental age. For a number of younger children, dental maturity was earlier than the programme could calculate and in these cases dental maturity score and dental age was calculated using the sex specific tables of Demirjian et al. [2,3]. Dental age was also calculated using Willems et al.’s [6] adjusted scores using tooth stages of Demirjian. The second method was Nolla [4]. Each tooth was assessed and assigned a stage of between 1 and 10. If the tooth was between stages an appropriate fraction (0.2, 0.5 or 0.7) was added as recommended by Nolla. The sum of scores was compared to the average sum for boys or girls and dental age calculated. The third method was using data from Haavikko [5]. Only stages up to and including root complete (parallel apical walls) were included. Dental age was calculated from the mean of all developing teeth, as well for each individual tooth type from Haavikko’s sex specific tables.

2.4. Calculating accuracy

Accuracy of dental age estimation was defined as how closely chronological age could be predicted, measured as the difference between chronological age and dental age. Dental age for each method was compared with chronological age for each subject. The chronological age was subtracted from the dental age and a positive result indicates an overestimation and a negative figure an underestimation. The significance of the difference between chronological and dental age was tested using Student’s t-test. This was done for each method of age estimation. In addition, the difference between ethnic groups was investigated using Student’s t-test using each method of age estimation. This was done for the entire age range for boys and girls separately. Accuracy was also determined for the whole age range and by age cohort and by tooth (using Haavikko’s method). Haavikko [17] suggested that her data were applicable only up to age 14 years, and accuracy was calculated for this method from 3 to 13.99 years; both mean age and by tooth type. All analyses were done for boys and girls separately and combined for use when sex is unknown.

2.5. Reproducibility

All assessment of tooth formation was done by the first author. A total of 120 radiographs were reassessed, after training and calibration prior to data collection. These were a random 10% sample of radiographs throughout the two years of this study. Cohen’s Kappa was calculated for each tooth formation method.

3. Results

No significant differences were observed between ethnic groups for any method for either sex and data from both ethnic groups were combined. Results of accuracy are shown in Tables 1 and 2. The 95% confidence interval of mean accuracy for each method and best tooth (boys and girls combined) is shown in Fig. 2. Figs. 3–6 show accuracy by method for each age cohort for boys and girls combined. The method of Willems was the most accurate, followed by Demirjian, Haavikko (age 3–13.99 years), Haavikko (all ages) and lastly Nolla. For individual teeth using Haavikko, first premolar and second molar were the most accurate. Using a single premolar or molar tooth to predict age was more accurate than using a mean of all developing teeth.

Demirjian’s method was found to overestimate age with a mean accuracy of 0.25 year for boys and 0.23 year for girls. The difference between chronological age and estimated dental age for both boys and girls was significant from zero ($P < 0.01$). Accuracy using Willems’ method was better.
especially for boys. This method yielded a mean under-estimation of −0.05 years for boys or −0.20 years for girls; accuracy between boys and girls was significantly different \((P < 0.01)\). For boys the estimated dental age using Willems was not statistically different from chronological age; for girls it was \((P < 0.01)\). The age group that was least accurately aged for both these methods was 16-year olds. Nolla’s method under-estimated age by −0.87 years for boys and −1.18 years for girls; also significantly different from chronological age \((P < 0.01)\). This underestimation occurred in all age groups and increased with age; for children ten years and older age the difference was more than one year.

Haavikko’s method (mean of all developing teeth) also under-estimated age by −0.56 years for boys and −0.79 for girls; both significantly different from chronological age \((P < 0.01)\). For children aged 13 years and older, the difference was greater than one year. This method differs from the maturity scales used in this study; in maturity scales, children with mature teeth (radiographic closed apices) contribute to dental age. Accuracy of Haavikko’s method was based on fewer children compared to other methods tested. This is because a mature tooth cannot be used to predict age, as the subject has past this stage by an unknown amount of time [18]. In a method where only developing teeth are used to predict age, older age groups are represented by only those children with delayed teeth. For instance, the least accurate age group using Haavikko’s average was 16 year olds, but this was made up of only six individuals who are clearly dentally delayed. Accuracy of individual teeth using Haavikko’s method show that the first premolar and second molar (for boys) were most accurate. Predicted age was significantly different \((P < 0.01)\) to chronological age for all analyses except the second molar for boys aged 3–13.99 years. All the teeth under-estimated age and the least accurate tooth was the canine followed by the incisors. Analysis of age group 3–13.99 years showed an improvement in mean accuracy of the second molar, both premolars and canine, but estimated age was still significantly different from chronological age \((P < 0.01)\). The accuracy of the second molar in boys was significantly improved, in fact using this tooth to predict age

\[
\begin{array}{c|c|c|c|c}
\text{Sex} & \text{Tooth} & \text{Age} 3.00–16.99 \text{years} & \text{Age} 3.00–13.99 \text{years} \\
\hline
\text{N} & \text{Mean} & \text{S.E.} & \text{S.D.} & \text{N} & \text{Mean} & \text{S.E.} & \text{S.D.} \\
\hline
\text{D} & \text{Boys} & 491 & 0.25 & 0.04 & 0.84 & 376 & −0.05 & 0.05 & 1.00 \\
 & Girls & 455 & 0.23 & 0.04 & 0.84 & & & & \\
 & Both & 946 & 0.24 & 0.03 & 0.86 & & & & \\
\text{W} & \text{Boys} & 491 & −0.05 & 0.04 & 0.81 & 353 & −0.19 & 0.05 & 1.04 \\
 & Girls & 455 & −0.20 & 0.04 & 0.89 & & & & \\
 & Both & 945 & −0.12 & 0.04 & 0.85 & & & & \\
\text{N} & \text{Boys} & 491 & −0.87 & 0.04 & 0.87 & 344 & −0.27 & 0.05 & 0.98 \\
 & Girls & 455 & −1.18 & 0.05 & 0.96 & & & & \\
 & Both & 946 & −1.02 & 0.03 & 0.93 & & & & \\
\text{H} & \text{Boys} & 437 & −0.56 & 0.04 & 0.91 & 350 & −0.74 & 0.06 & 1.15 \\
 & Girls & 395 & −0.79 & 0.06 & 1.11 & & & & \\
 & Both & 832 & −0.67 & 0.04 & 1.01 & & & & \\
\text{H} & \text{Boys} & 392 & −0.39 & 0.04 & 0.77 & & & & \\
 & Girls & 357 & −0.57 & 0.05 & 0.87 & & & & \\
 & Both & 749 & −0.47 & 0.03 & 0.82 & & & & \\
\end{array}
\]

\(D\): Demirjian; \(W\): Willems; \(N\): Nolla; \(H\): Haavikko; \(H < 14\): Haavikko age less than 14.00 years; \(S.E.\): standard error of mean; \(S.D.\): standard deviation.

Table 2
Mean accuracy (in years) for individual teeth using Haavikko’s method

\[
\begin{array}{c|c|c|c|c|c|c|c|c}
\text{Sex} & \text{Tooth} & \text{Age} 3.00–16.99 \text{years} & \text{Age} 3.00–13.99 \text{years} \\
\hline
\text{N} & \text{Mean} & \text{S.E.} & \text{S.D.} & \text{N} & \text{Mean} & \text{S.E.} & \text{S.D.} \\
\hline
\text{Boys} & \text{M2} & 420 & −0.20 & 0.05 & 1.07 & 376 & −0.05 & 0.05 & 1.00 \\
 & \text{M1} & 266 & −0.53 & 0.07 & 1.17 & 353 & −0.19 & 0.05 & 1.04 \\
 & \text{P2} & 380 & −0.36 & 0.06 & 1.18 & 344 & −0.27 & 0.05 & 0.98 \\
 & \text{P1} & 350 & −0.31 & 0.05 & 1.01 & 350 & −0.74 & 0.06 & 1.15 \\
 & \text{C} & 365 & −0.83 & 0.06 & 1.20 & & & & \\
 & \text{I2} & 238 & −0.73 & 0.06 & 0.96 & & & & \\
 & \text{I1} & 196 & −0.66 & 0.06 & 0.88 & & & & \\
\text{Girls} & \text{M2} & 376 & −0.60 & 0.06 & 1.20 & 340 & −0.40 & 0.06 & 1.04 \\
 & \text{M1} & 217 & −0.47 & 0.09 & 1.28 & 310 & −0.42 & 0.06 & 1.10 \\
 & \text{P2} & 326 & −0.56 & 0.07 & 1.26 & 300 & −0.32 & 0.06 & 1.00 \\
 & \text{P1} & 303 & −0.35 & 0.06 & 1.04 & 268 & −1.00 & 0.07 & 1.13 \\
 & \text{C} & 277 & −1.11 & 0.08 & 1.25 & & & & \\
 & \text{I2} & 152 & −0.77 & 0.08 & 1.02 & & & & \\
 & \text{I1} & 158 & −0.77 & 0.08 & 1.01 & & & & \\
\text{Both} & \text{M2} & 796 & −0.39 & 0.04 & 1.15 & 716 & −0.21 & 0.04 & 1.03 \\
 & \text{M1} & 483 & −0.50 & 0.06 & 1.22 & 663 & −0.30 & 0.04 & 1.07 \\
 & \text{P2} & 706 & −0.45 & 0.05 & 1.22 & 644 & −0.30 & 0.04 & 0.99 \\
 & \text{P1} & 653 & −0.33 & 0.04 & 1.02 & 618 & −0.85 & 0.05 & 1.14 \\
 & \text{C} & 642 & −0.95 & 0.05 & 1.23 & & & & \\
 & \text{I2} & 390 & −0.74 & 0.05 & 0.98 & & & & \\
 & \text{I1} & 354 & −0.71 & 0.05 & 0.94 & & & & \\
\end{array}
\]

M2: second molar; M1: first molar; P2: second premolar; P1: first premolar; C: canine; I2: lateral incisor; I1: central incisor; S.E.: standard error of mean; S.D.: standard deviation.
was more accurate than using all developing teeth (Haavikko) or any of the dental maturity scales tested, although 95% confidence interval of the mean was slightly better for Haavikko (3–13.99 years) compared to Demirjian and Willems. For combined sex, the tooth with the highest accuracy (Haavikko) was the second molar, with stages ‘Ci’, ‘R1/4’, ‘R1/2’, ‘R3/4’ and ‘Re’ having an accuracy of within +0.25 year (boys and girls combined, age 3–13.99 years). Of these, stage ‘Ci’ showed the lowest variance (−0.09 ± 0.38, N = 32) for combined sex.

Cohen’s Kappa measuring reproducibility for Demirjian’s stages was 0.86, Nolla’s method was 0.78 and for Haavikko’s stages 0.90. Disagreements were never more than one stage with similar number of over and under estimates for each method.

4. Discussion

Accuracy and precision are both important in assessing dental age. Accuracy refers to “the closeness of a computed value to its true value” [19] and was defined for this study as how close the difference between the estimated age and chronological age was to zero. Any differences found between the standard population and the sample population can be attributed to many variables including precision of the method, age structure of the sample, sample size, statistical approach and biological variation of individual children. Several previous studies have investigated accuracy of age estimation using developing teeth. Some use small sample size, or uneven age...
Few age estimation methods are designed for age estimation (discussed by Smith [18]). Demirjian’s maturity scale is designed to assess dental maturity of children of known age, and many studies have found that dental maturity is advanced compared to the original study. A number of difficulties were encountered with the methods tested in this study. Demirjian’s method involves calculated dental age by adding up a score for each stage of each tooth and this weighted scoring complicates analyses of accuracy. Towards the end of dental maturation fewer stages contribute more, so that a single stage change can lead to a large jump in dental age. For example a difference of one stage for a girl (M2 stage G, all other teeth mature, stage H) with a dental age 14.6 years, jumps to dental age 16 years for a girl with all seven teeth stage. Nolla’s study was one of the first to assess tooth formation longitudinally and quantify dental maturation. She presents a sum of tooth scores against each year of age and this is what we used to predict age into 1-year age groups. She also interpolates data back, providing an age at each stage for each tooth. Several researchers have used this double interpolated data of individual growth stages to predict age, but interpreting ordinal data as parametric is of questionable validity. The difficulty with Haavikko was deciding if the children who were considerably delayed in one tooth, were to be included in the full analyses in order to compare with other methods. We chose to present data for the entire age group as well as 3–13.99 years.

An important factor that influences reproducibility is the choice of tooth stage assessment. Those described by Demirjian are the most detailed using relative length of crown and root. When stages include fractions of crown and root growth, assessment is more difficult and subjective; one has to estimate what fraction of unknown future crown height or root length has formed. Such methods are thought to have a decreased precision, although in the present study precision for stages of Haavikko was good. The lower precision seen in Nolla’s method may be due to number of stages; the method allows the observer to choose between the 10 stages but also gives three inter stage options for each stage, increasing the possible stages to 40. Increasing the number of stages has been shown to decrease precision [29,30]. Choosing the tooth with the least variability is likely to prove most accurate in predicting age. For this, fast growing stages that change from being present in none of the children up to 100% in a short space of time might be ideal. Tooth formation stages are not equally spaced during growth and are not of equal duration. To date, our knowledge of exact dental formation times between crown/root fractions is meagre. The lack of difference between ethnic groups is important and suggests that suitable and accurate methods of age prediction are more widely suitable than previously thought. This is supported by recent findings of the nature of ethnic/racial differences in tooth formation. This appears to be a difference in timing of initial tooth formation and may involve only the third molar [31]; furthermore, histological evidence shows that duration of crown formation times between groups varies little [32].

Fig. 6. Accuracy of Haavikko’s method (mean of all developing teeth) showing mean age with standard error for each age groups (boys and girls combined). haavdiff difference between estimated age using Haavikko’s method and real age in years, agegr age group in 1 year cohorts.

distribution, or present results in a way that makes comparison difficult. A previous finding that age in younger children can be more accurately predicted than older children was not observed in the present study. A common finding is that Demirjian overestimated age [9–13,20], although an under-estimation has also been reported [21]. Hagg and Mattson [9] find this the most accurate method compared with two methods not tested in the present study [22,23]. Caro and Contreras [13] in a small study found it be least accurate compared to Nolla [4], Moorrees et al. [24] and their data adapted for prediction by Smith [18]. Two studies find the method of Haavikko to be more accurate. The first [12] compared accuracy to variations of Demirjian et al. [2,3] and adjusted data for Finnish children [25], although mean values are not reported; the second [11] compares accuracy to Demirjian et al. [2,3] and Liliequist and Lundberg [22], although the sample is split into two groups for Haavikko-children younger than 10 years and older than 10 years.

In this study the adapted scoring of Willems et al. [6] was the most accurate method followed by Demirjian [1–3]. This confirms findings by Willems et al. [6] on their test sample. The 95% confidence interval was marginally better for Haavikko (age 3–13.99 years) than similar value for Demirjian and Willems. The mean accuracy for Nolla’s maturity scale [4] showed a consistent underestimation, although a previous study found that real age in girls was similar to dental age, younger boys were dentally advanced [26]. Caro and Contreras [13] found Nolla to be more accurate than other methods they tested [1,18,24] although the sample size was small. None of these studies detail exactly how dental age was calculated using Nolla’s data [4]. Two studies investigate age of individual teeth using Nolla’s data but give few details of mean accuracy, making comparison with the present study difficult [27,28]. Using Haavikko to predict age gives an underestimation of age, similar to previous findings [11,12].
5. Conclusion

Willems adjusted data for Demirjian’s method was found to be the most accurate method for estimating age in this sample. The 95% confidence interval of the mean difference between estimated and real age was least using Haavikko (age 3–13.99 years), followed by identical values for Willems and Haavikko. Tooth formation described by Demirjian was the easiest and clearest method of stage assessment, although Kappa value was similar to Haavikko’s stages. All the methods, except Demirjian, were significantly more accurate at estimating age for boys compared to girls. The only two results where predicted and chronological age were not significantly different were Willems for boys and Haavikko second molar for boys (age 3–13.99 years). When sex is unknown, the most accurate method was Willems followed by Demirjian then Haavikko first premolar.

References