_ _

Dental age assessment on panoramic radiographs in a Swiss population: a validation study of two prediction models

_ _

Birchler, Flavia A; Kiliaridis, Stavros; Combescure, Christophe; Vazquez, Lydia

RESEARCH ARTICLE

Dental age assessment on panoramic radiographs in a Swiss population: a validation study of two prediction models

[1]Flavia A Birchler, [1]Stavros Kiliaridis, [2]Christophe Combescure and [3]Lydia Vazquez

[1]Department of Orthodontics, University Clinics of Dental Medicine, University of Geneva, Geneva, Switzerland; [2]Division of Clinical Epidemiology, Geneva University Hospital, Geneva, Switzerland; [3]Department of Orofacial Rehabilitation, Oral and Maxillofacial Radiology, University Clinics of Dental Medicine, University of Geneva, Geneva, Switzerland

Objectives: Dental age assessment methods are widely used for age estimation. This study aimed to analyse the accuracy of a meta-analysis method to estimate dental age in Swiss individuals and to detect potential limitations of the method. Precision of repeated tooth staging using Demirjian's classification on maxillary and mandibular teeth was also assessed.

Methods: Panoramic radiographs of 50 Swiss white healthy children were analysed. Developing teeth on the left maxilla and mandible and all third permanent molars were staged following Demirjian's classification. Dental age was calculated for each subject, using a random effects model and a fixed effect model, and compared with chronological age.

Results: The mean error of the dental age ranged between −3 and +1 months for both the calculation models. Dental age calculated with the fixed effect model overestimated the age of the subjects (average + 0.10 y, ranging from −1.95 y to +2.16 y) compared with their chronological age, whereas the random effects model underestimated the age (average −0.32 y, ranging from −2.24 y to +1.61 y).

Conclusions: Demirjian's method allowed a precise repeated staging of maxillary and mandibular developing teeth. For both calculation models, dental age correlated well, on average, with chronological age of Swiss subjects younger than 12 years. The random effects model showed a better accuracy for these subjects than the fixed effect model. However, both models underestimated the chronological age in subjects older than 12 years.

Dentomaxillofacial Radiology (2016) **45**, 20150137. doi: 10.1259/dmfr.20150137

Introduction

Methods of age assessment are used for individuals with unknown birth records in legal or forensic situations—such as schooling and adoption processes, immigration or identification of missing and deceased children[1]—but also in dental and orthodontic treatment planning.[2] Commonly used developmental indicators for age

estimations include skeletal maturity, body height and weight, sexual development, and tooth development and eruption.[3,4] Age estimation by means of tooth development has been used for several years, and it has been reported to be a reliable age estimation technique in growing subjects as it correlates better with chronological age than other techniques.[5,6] The method developed by Demirjian et al,[7] which classifies mandibular tooth development stages of the crown and the root, has been found to be accurate.[8] This technique, based on a sample of French–Canadian subjects, has been tested on various

population groups for its applicability when using only mandibular teeth.

Recently, the Dental Age Research London Information Group (DARLInG) established a dental age reference database of Caucasian subjects assessing left upper and lower teeth with Demirjian's method, as well as all the four wisdom teeth.[9,10] The authors used a meta-analysis approach to estimate a subject's dental age taking into account the developmental stage of each tooth[9] and were able to accurately estimate the age of young subjects around different age thresholds.[10–12]

The aim of this study was to analyse the accuracy of dental age assessment using the DARLInG meta-analysis method in a Swiss population with developing teeth and to detect potential limitations of the method. The second aim was to assess the precision of repeated tooth staging using Demirjian's classification on maxillary and mandibular teeth.

Methods and materials

The panoramic radiographs for this retrospective study were randomly selected from pre-treatment records of patients of the Department of Orthodontics (University of Geneva, Geneva, Switzerland). The sample consisted of 50 randomly chosen Swiss white healthy children (25 males and 25 females) born between 31 January 1990 and 27 September 2002 with developing teeth. Radiographs of children with conditions that could influence the rate of dental development and those of poor quality were excluded.

Panoramic radiographs were performed with a conventional panoramic unit (Cranex® 3+; Soredex® Orion Corporation, Tuusula, Finland) and digitized on a transmitted-light scanner (Epson® Expression® 1600 Pro; Seiko Epson Corporation, Suwa, Japan) set to 300 dpi and stored as TIFF images. OsiriX® 1.2 software program (Pixmeo, Geneva, Switzerland) installed in an iMac OS X v. 10.6.8 workstation (iMac 27-inch Quad Core 3.4 GHz Intel® Core™ i7; Apple Inc., Cupertino, CA) was used to view and analyse the digitized images. Image manipulation (zoom, sharpening, brightness, contrast and gamma adjustments) to facilitate image analysis was performed when necessary.

The research protocol was approved by the Ethics committee of the University Hospital of Geneva, Geneva, Switzerland (#12–261).

Two examiners (FB and LV) underwent training and calibration in panoramic radiograph assessment. After this calibration phase, the examiners analysed the panoramic radiographs for the study together. All developing teeth on the left hand side (quadrants two and three) and all four third permanent molars were assessed[9] using the Demirjian et al[7] classification, which divides the process of tooth development into eight stages. All radiographs were re-examined after an interval of 2 weeks. The examiners were blinded to the age and sex of each subject when assessing tooth development stages. When

tooth anatomy was difficult to see or when a tooth was missing, assessment of the contralateral tooth was performed. If the latter was not interpretable, the tooth was not included in the analysis.

Chronological age of each subject was calculated by subtracting the date of birth from the date of the radiograph. Dental age for every patient was calculated using the DARLInG meta-analysis method[9] based on sex-specific data obtained from Caucasian subjects. Every tooth was hereby assigned a dental age according to its root development. An overall dental age was computed using information on all developing teeth of a given subject.

The calculation of a dental age was undertaken using two different types of meta-analysis, both based on the same information from the DARLInG database. Meta-analysis is a quantitative procedure that is used in statistical methodology to combine and summarise the results of several studies that address a particular research hypothesis. In the context of dental age assessment, the meta-analysis approach provides an estimate of the expected dental age of a subject by calculating a weighted mean of the mean ages of the subject's tooth development stages.[9] In general, a fixed effect model assumes that all studies are estimating the same effect size, thus weight is assigned based on the amount of information apprehended by a study. On the contrary, a random effects model assumes that each study is estimating a different effect size.[13] Applied to dental age assessment, the fixed effect model gives more weight to the tooth development stages with larger number of observations in the DARLInG database, whereas the random effects model balances the weight between the different tooth development stages more equally.

Finally, the estimated dental age was compared with the chronological age of every patient.

On a sample of 50 patients, Roberts et al[9] found an average difference between dental age and chronological age of +3.5 months with a standard deviation of 10 months. We based our sample size on this report, calculating that 50 patients would allow a precision (half-length of the 95% confidence interval) of ±3 months around the mean error, which we considered to be accurate.

Dental age was calculated using the random effects model, and the fixed effect model was proposed by Roberts et al.[9] The chronological and dental ages were described by mean, standard deviation, median, minimum and maximum. The reproducibility of the evaluation of development stages between the two sessions was assessed using Kappa coefficients. For the dental ages, Bland and Altman analyses[14] were conducted, and intraclass correlation coefficients were assessed. The limits of agreement, indicating the interval covering 95% of the individual difference between sessions, were reported. The prediction error, defined as the dental age minus the chronological age, was graphically represented against the chronological age and described by the mean for both fixed effect and random effects

Table 1 Intersession agreement of tooth development stages

Tooth	N^a	κ (95% CI)	Tooth	N^a	κ (95% CI)
UL1	N = 50	1.000	LL1	N = 50	1.000
UL2	N = 49	1.000	LL2	N = 49	1.000
UL3	N = 49	0.969 (0.910; 1.000)	LL3	N = 50	0.969 (0.909; 1.000)
UL4	N = 49	0.936 (0.851; 1.000)	LL4	N = 50	0.969 (0.908; 1.000)
UL5	N = 47	0.942 (0.864; 1.000)	LL5	N = 49	1.000
UL6	N = 50	1.000	LL6	N = 50	1.000
UL7	N = 49	0.920 (0.831; 1.000)	LL7	N = 50	0.972 (0.917; 1.000)
UL8	N = 36	0.964 (0.894; 1.000)	LL8	N = 39	0.961 (0.886; 1.000)
UR8	N = 36	1.000	LR8	N = 37	1.000

CI, confidence Interval; LL1, lower left central incisor; UL1, upper left central incisor.
*Number of evaluated teeth per group of teeth. Missing information due to missing teeth or unreadable structures. The missing information was congruent for both evaluations.

models. Limits of agreement were reported. The mean error was modelled by the chronological age using a piecewise linear regression model. Moreover, the individual prediction error varies around the mean, and these variations are given by the residuals of the regression model. Therefore, we reported the interval covering 95% of individual prediction error. Assuming that the individual prediction errors are normally distributed, this interval was assessed by the mean ± 1.96 times the standard deviation of the residuals. The assumption of the normality of the distribution was graphically inspected. Prediction errors were compared between fixed effect and random effects models using a Wilcoxon test for paired data.

Statistical analysis was performed using S-plus® 8.0 for Windows® (Insightful Corp., Seattle, WA). The statistical tests were two-sided with a significance threshold of 5%.

Results

A total of 851 teeth were evaluated per session. The development stages could be described for all teeth in 28 (56%) subjects. For 12 subjects, 1 or 2 stages could not be defined because of unreadable pictures (7 teeth) or missing teeth (49 teeth). Ten subjects presented with more than two indefinable stages (up to five stages). The mean chronological age of those subjects with all teeth present was 12.38 years (SD 1.68 years), whereas in the group with at least one missing tooth to evaluate, the mean chronological age was 10.86 years (SD 1.74 years). The difference between the mean chronological age of the two groups was statistically significant ($p = 0.004$) (Figures A1 and A2). Most of the missing stages (52/60, 86.7%) were located amongst the wisdom teeth. This was either because of agenesis or later development of these teeth. In evaluable teeth, the evaluation of development stages was reproducible: Kappa coefficients for evaluations between sessions were above 0.92 for all teeth (Table 1).

The chronological age distribution of the studied subjects is presented in Figure 1. Chronological age of the subjects ranged from 8 to 15 years with a mean of 11.71 years and was similar in females and males. The mean dental age was 11.82 years with the fixed effect model and 11.40 years with the random effects model. The calculated dental age was very similar in Sessions 1 and 2 for both fixed effect and random effects models (Table 2).
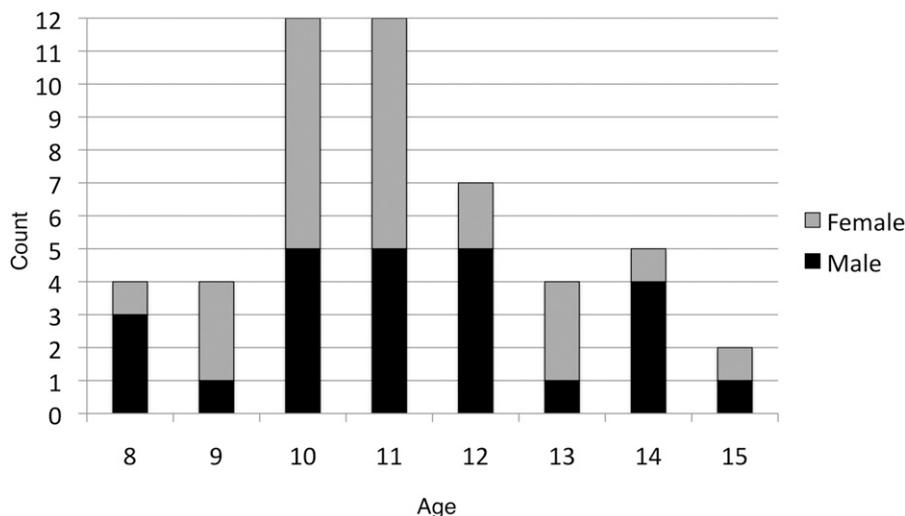

Figure 1 Chronological age distribution of sample. Age 8 indicates age group from 8.00 to 8.99 years, etc.

Table 2 Descriptive statistics of the chronological age and the dental age predicted by fixed and random effects methods at both sessions. Mean (standard deviation), minimum and maximum are reported. Age is expressed in years

| Evaluated subjects | Chronological age | Age predicted by meta-analysis method | | | |
| | | Fixed effect model | | Random effects model | |
		Session 1	Session 2	Session 1	Session 2
All subjects (*n* = 50)	11.71 (1.85)	11.82 (1.93)	11.82 (1.94)	11.40 (1.36)	11.40 (1.38)
	[7.99; 15.85]	[8.01; 14.41]	[7.87; 14.41]	[7.96; 13.57]	[7.81; 13.57]
Female (*n* = 25)	11.48 (1.70)	11.60 (1.88)	11.60 (1.88)	11.31 (1.28)	11.31 (1.29)
	[7.99; 15.23]	[8.17; 14.23]	[8.17; 14.23]	[8.13; 13.12]	[8.13; 13.12]
Male (*n* = 25)	11.95 (2.00)	12.03 (2.00)	12.03 (2.02)	11.48 (1.46)	11.48 (1.49)
	[8.59; 15.85]	[8.01; 14.41]	[7.87; 14.41]	[7.96; 13.57]	[7.81; 13.57]

The intraclass correlation coefficients were >0.99, and the individual differences between Sessions 1 and 2 did not exceed +0.25 years except for one subject (Figure 2). Because of high intraclass correlation coefficients, the following statements are true for both sessions and only results for Session 1 will be presented subsequently.

The error committed by the two studied models in the prediction of age was defined as the dental age minus the chronological age. This prediction error was close to zero on average (+0.10 year for the fixed effect method and −0.32 years for the random effects method). The prediction error was similar between females and males (fixed effect method: $p = 0.91$, random effects method: $p = 0.16$). However, the prediction error was associated with the chronological age.

With the fixed effect method, the prediction error was −0.59 years for a chronological age of 8 years and increased significantly ($p = 0.01$) until +0.77 years for a chronological age of 12 years (Table 3, Figure 3). Beyond 12 years of age, the mean error decreased significantly ($p < 0.0001$) and reached −1.66 years for a chronological age of 16 years.

With the random effects method, the prediction error was stable ($p = 0.68$) and close to zero until a chronological age of 12 years. After 12 years, the mean prediction error decreased significantly ($p < 0.0001$) and reached −2.83 years for a chronological age of 16 years. Above 12 years of age, the underestimation increased with every year of chronological age on average by 0.61 years (fixed effect) and 0.73 years (random effects).

Beside the observed bias in dental age, the limits of agreement (Table 3, Figure 3) captured the scattering of the individual prediction errors around the mean: the limits of agreement represent the interval containing approximately 95% of the individual prediction errors. With the fixed effect method, the limits of agreements were +1.72 years around the mean prediction error. With the random effects method, the limits of agreements were narrower: +1.15 years around the mean.

Dental ages predicted by fixed and random effects methods were compared (Figure 4). In subjects with a chronological age of 12 years or younger, the dental age predicted by the fixed effect model was slightly, but significantly, higher than the dental age predicted by random effects method (median difference: 0.05 year, $p = 0.04$). The difference had the same direction in subjects with a chronological age of 13 years or more and was more pronounced (median difference: 0.91 year, $p = 0.001$).
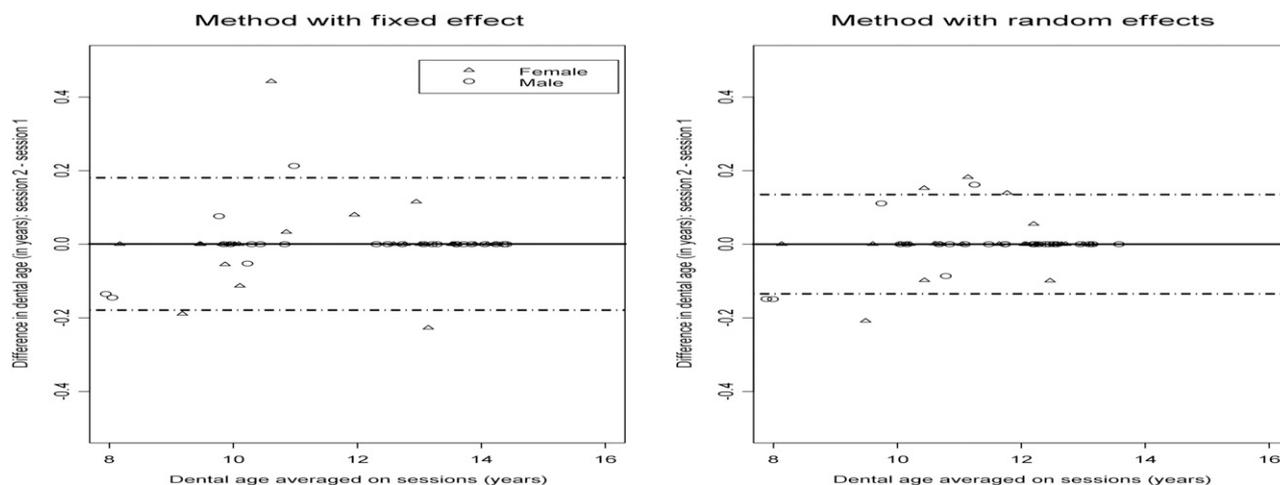

Figure 2 Bland and Altman's plot for the intrarater agreement. The solid line represents the mean difference of dental age between sessions. The dashed lines represent the limits of agreement: for approximately 95% of subjects, the difference between sessions is in this interval.

Table 3 Prediction error (predicted age minus the chronological age) for fixed and random effects methods at Session 1. The mean prediction error was assessed in relation to the chronological age (linear regression). The limits of agreement show the interval containing approximately 95% of the individual prediction error

	Prediction error (years)					
	Fixed effect model			Random effects model		
		Limits of agreement			Limits of agreement	
Chronological age (years)	Mean error	Lower	Upper	Mean error	Lower	Upper
8	−0.59	−2.31	1.14	0.23	−0.92	1.39
9	−0.25	−1.97	1.48	0.20	−0.96	1.35
10	0.09	−1.63	1.82	0.16	−0.99	1.32
11	0.43	−1.29	2.16	0.13	−1.03	1.28
12	0.77	−0.95	2.50	0.09	−1.06	1.24
13	0.16	−1.56	1.89	−0.64	−1.79	0.51
14	−0.45	−2.17	1.28	−1.37	−2.52	−0.22
15	−1.05	−2.78	0.67	−2.10	−3.26	−0.95
16	−1.66	−3.39	0.06	−2.83	−3.99	−1.68
All	0.10	−1.95	2.16	−0.32	−2.24	1.61

Discussion

Tooth development, a genetically controlled process,[15] is mostly unrelated to skeletal and sexual development[3,16] as well as exogenic factors such as malnutrition or disease.[17] Radiological evaluation of dental maturity is considered a simple, noninvasive and reliable method to determine the chronological age in children.[16] Tooth development evaluation on extraoral and intraoral radiographs allows assessment of tooth mineralization from the moment radiopaque spots become visible before tooth calcification until the tooth apex is closed.[18] Demirjian et al[7] described a widely used dental maturity method,[19] only applicable for subjects up to 16 years of age, as it is restricted to the development of all teeth excluding third molars.[20] The present study used Demirjian's anatomical descriptions of tooth development stages. For dental age calculation, the DARLInG method[9]

added upper teeth and third molars evaluation to improve the overall accuracy of dental age estimation. Our tooth analysis on the left side was based on findings of similar chronological development of homologous teeth in each arch.[1] We classified all third molars with Demirjian's stages; validity of five classification systems on third molars reported Demirjian's classification to be the most accurate.[21] We evaluated third molars in all four quadrants, even though mandibular third molars show no significant side differences in mineralization.[22] This may be advantageous in cases of tooth agenesis or extraction. We used sex-specific data, because it has been shown that overall, teeth develop earlier in girls,[1,23] with the exception of third molars that develop faster in boys.[17,22,24]

Demirjian's classification was developed to stage mandibular teeth from incisors to second molars only, but following the DARLInG guidelines, we also used it on upper teeth and all third molars. Our results with high
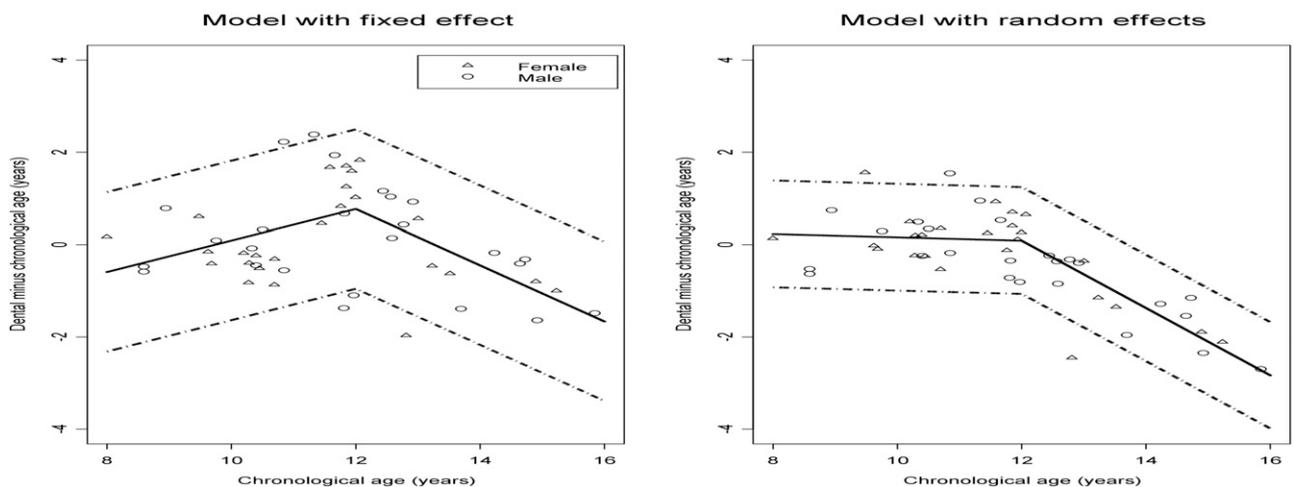
Session 1

Figure 3 Prediction error (difference between dental and chronological age) for fixed effect model and random effects model in the first evaluation session. Values are reported in years. The solid curve marks the mean prediction error, whereas the area between the two dotted lines (limits of agreement) contains approximately 95% of the data.
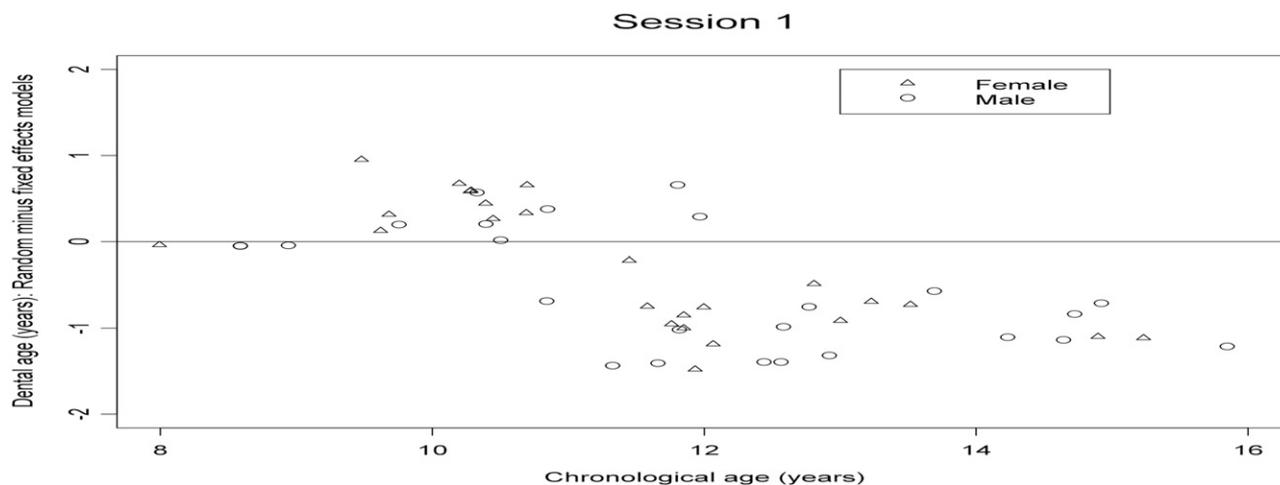
Figure 4 Agreement between the values calculated by random effects (RE) model and fixed effect (FE) model (Difference RE − FE) in relation to chronological age (True age). Values are given in years. Y-axis: the solid line marks a null difference between RE and FE models.

Kappa values—ranging from 0.81 to 1.00—indicate an "almost perfect" agreement[25] between the two sessions and suggest high precision of Demirjian's classification in this study. These results can be explained by the selection of good quality radiographs and an intense calibration phase prior to the study; calibration of both examiners led to good interobserver and intraobserver agreements.[26] Furthermore, during panoramic radiographic examination, the two calibrated examiners jointly agreed upon a stage for each analysed tooth. Consensus on stage development between the examiners might have increased the accuracy of the assessment. The impact of joint evaluation on tooth staging has not been reported, and further studies should be performed to clarify this issue. Our findings, showing slightly higher reproducibility for mandibular teeth evaluation compared with maxillary teeth, were in agreement with previous reports.[27] Anatomic factors, such as superimpositions of the hard and soft tissues on upper teeth, can lead to challenging radiographic analysis.

When compared with other studies dealing with dental age assessment, the size of our sample was rather small and some authors consider this a misleading factor.[17] Flood et al[28] on the other hand indicated that smaller samples might be used for age determination, even though larger samples should be favoured. Even though our sample size was limited, it allowed us to detect potential limitations of dental age assessment with two different types of meta-analysis methods, and therefore provides a satisfactory starting point for future clinical researches. Whether our findings are true for bigger samples will have to be confirmed by larger studies designed for these specific purposes.

In 22 subjects (44%), at least one tooth could not be evaluated. The patients in the group with all teeth were older (statistically significant) than in the group with some missing teeth. It is interesting to note that the group with missing teeth had a significant lower mean age than did the group with all teeth present for evaluation. The overall precision of dental age assessment was better in younger patients. Even with some teeth lacking for evaluation, the calculation of a dental age is closer to chronological age in younger patients. This is probably due to the fact that more roots are being calcified in younger patients than in older patients. To give a conclusive answer, further clinical research on patients with missing teeth or teeth that cannot be evaluated for dental age assessment is necessary.

In the present study, the technique of meta-analysis proposed by Mitchell et al[10] was used for dental age calculation with two different models (fixed effect and random effects). The mean error of −0.32 to +0.1 years was relatively low for both the calculation models, and the limits of agreement ranged from −2.24 up to +2.16 years. These values corroborate previous reports; Chaillet et al[23] estimated age of children between 2 and 18 years with an average of ±1.95 years. Liversidge et al[29] and Kullman et al[30] found a precision of ±2 years. Thorson and Hägg[24] found differences between estimated and true chronological age (±4.5 years in girls and ±2.8 years in boys), which were larger than those found in our findings. A mean age difference between dental age and chronological age of ±1 year was considered accurate by some authors,[23] whereas Peiris et al[31] stated that a difference greater than 6 months was not suitable for age determination.

The method of calculation is a major source of age assessment discrepancy.[32] Cruz-Landeira et al[33] also used two models for dental age estimation and demonstrated that the choice of equation had a big impact on the average age score. In our study, the mean difference between the calculated age values was within the limits of agreement of less than 3 months, even though in 2 out of 50 patients, 2 teeth were rated differently between the two sessions. This was true for both calculation models, although less effect was found with the

random effects model. In case of disagreement between the two sessions, the attributed scores were neighbouring developmental stages, which is in accordance with Urzel and Bruzek.[34]

Factors other than the calculation method might have impacted our results. One is the choice of the studied population; many researchers found differences in dentition development among different population groups;[1,31,33] Azrak et al[16] stated that genetic and nutritional differences have to be taken into account when using reference data for age determination. Our findings corroborate this report; age estimation of Swiss subjects was not as good as for the subjects of British origin using the same reference data.[10–12] On the contrary, Willershausen et al[35] failed to find differences in growth patterns based on national and/or ethnic background; and Liversidge considered it unnecessary to adapt scores for different population groups[36] for dental age evaluation.

Another factor impacting our results might be the chronological age of the subjects. For subjects younger than 13, the random effects model allowed a better age prediction. Both models underestimated age for subjects older than 13 years but the age predicted by the fixed effect model was lower than with the random effects model. For every year increase in chronological age, the average predicted age declined by 0.61 years (fixed effect) and 0.73 years (random effects). The calculation of our sample size was not adapted to detect such a finding, and it would need confirmation with a study powered for this specific purpose. Nevertheless, other reports also showed a better accuracy of age estimation in younger patients.[8,29] More developing teeth are available for staging and there is a greater variability in tooth formation in older subjects.[30,32] Because only developing teeth serve as a basis for dental age calculation, the result is more precise when more teeth are in a Stage A–G.

Urzel and Bruzek[34] found an age limit of 14 years to be critical for age estimation, which does not concur with our observations: at 14 years of age, the mean error is more than 1 year with the random effects model. The fixed effect model shows a more acceptable tendency with a mean error of −0.61 years. This difficulty in predicting age in older patients was also reported by Gunst et al.[17] Our results showed that, on average, both prediction models showed a good correlation between predicted dental age and chronological age in Swiss subjects younger than 12 years. To determine which prediction model should be used for dental age assessment in legal and forensic circumstances, future investigations should focus on precise age prediction in older subjects.

Conclusion

Demirjian's method allowed a precise repeated staging of maxillary and mandibular developing teeth. The two DARLInG-based calculation models showed that the dental age correlated well, on average, with the chronological age of Swiss subjects younger than 12 years. The random effects model showed a better accuracy for these children than did the fixed effect model. However, both the models underestimated the chronological age in subjects older than 12 years.

Acknowledgments

References

1. Feijoo G, Barberia E, De Nova J, Prieto JL. Permanent teeth development in a Spanish sample. Application to dental age estimation. *Forensic Sci Int* 2012; **213**: e1–3.
2. Maret D, Peters OA, Dedouit F, Telmon N, Sixou M. Cone-beam computed tomography: a useful tool for dental age estimation? *Med Hypotheses* 2011; **76**: 700–2. doi: 10.1016/j.mehy.2011.01.039
3. Schmeling A, Olze A, Reisinger W, Geserick G. Age estimation of living people undergoing criminal proceedings. *Lancet* 2001; **358**: 89–90. doi: 10.1016/S0140-6736(01)05379-X
4. Demirjian A, Buschang PH, Tanguay R, Kingnorth Patterson D. Interrelationships among measures of somatic, skeletal, dental, and sexual maturity. *Am J Orthod* 1985; **88**: 433–8. doi: 10.1016/0002-9416(85)90070-3
5. Garn SM, Lewis AB, Kerewsky RS. Genetic, nutritional, and maturational correlates of dental development. *J Dent Res* 1965; **44**: 228–42. doi: 10.1177/00220345650440011901
6. Willems G. A review of the most commonly used dental age estimation techniques. *J Forensic Odontostomatol* 2001; **19**: 9–17.
7. Demirjian A, Goldstein H, Tanner JM. A new system of dental age assessment. *Hum Biol* 1973; **45**: 211–27.
8. Hägg U, Mattson L. Dental maturity as an indicator of chronological age: the accuracy and precision of three methods. *Eur J Orthod* 1985; **7**: 25–34.
9. Roberts GJ, Parekh S, Petrie A, Lucas VS. Dental age assessment (DAA): a simple method for children and emerging adults. *Br Dent J* 2008; **204**: 192–3. doi: 10.1038/bdj.2008.21
10. Mitchell JC, Roberts GJ, Donaldson AN, Lucas VS. Dental age assessment (DAA): reference data for British Caucasians at the 16 year threshold. *Forensic Sci Int* 2009; **189**: 19–23. doi: 10.1016/j.forsciint.2009.04.002
11. Yadava M, Roberts GJ, Lucas VS. Dental age assessment (DAA): reference data for British children at the 10-year-old threshold. *Int J Legal Med* 2011; **125**: 651–7. doi: 10.1007/s00414-010-0491-z
12. Chudasama PN, Roberts GJ, Lucas VS. Dental age assessment (DAA): a study of a Caucasian population at the 13 year threshold. *J Forensic Leg Med* 2012; **19**: 22–8. doi: 10.1016/j.jflm.2011.09.008
13. Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. Fixed-effect *versus* random-effects models. In: Borenstein M (ed). *Introduction to meta-analysis*. Chichester, UK: John Wiley & Sons Ltd; 2009. pp 77–86.
14. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986; **1**: 307–10.
15. Pelsmaekers B, Loos R, Carels C, Derom C, Vlietinck R. The genetic contribution to dental maturation. *J Dent Res* 1997; **76**: 1337–40. doi: 10.1177/00220345970760070201

16. Azrak B, Victor A, Willershausen B, Pistorius A, Hörr C, Gleissner C. Usefulness of combining clinical and radiological dental findings for a more accurate noninvasive age estimation. *J Forensic Sci* 2007; **52**: 146–50. doi: 10.1111/j.1556-4029.2006.00300.x

17. Gunst K, Mesotten K, Carbonez A, Willems G. Third molar root development in relation to chronological age: a large sample sized retrospective study. *Forensic Sci Int* 2003; **136**: 52–7. doi: 10.1016/S0379-0738(03)00263-9

18. Rózyło-Kalinowska I, Kiworkowa-Rączkowska E, Kalinowski P. Dental age in Central Poland. *Forensic Sci Int* 2008; **174**: 207–16.

19. Liversidge HM, Chaillet N, Mörnstad H, Nyström M, Rowlings K, Taylor J, et al. Timing of Demirjian's tooth formation stages. *Ann Hum Biol* 2006; **33**: 454–70. doi: 10.1080/03014460600802387

20. Jayaraman J, Wong HM, King NM, Roberts GJ. The French-Canadian data set of Demirjian for dental age estimation: a systematic review and meta-analysis. *J Forensic Leg Med* 2013; **20**: 373–81. doi: 10.1016/j.jflm.2013.03.015

21. Olze A, Bilang D, Schmidt S, Wernecke KD, Geserick G, Schmeling A. Validation of common classification systems for assessing the mineralization of third molars. *Int J Legal Med* 2005; **119**: 22–6. doi: 10.1007/s00414-004-0489-5

22. Meinl A, Tangl S, Huber C, Maurer B, Watzek G. The chronology of third molar mineralization in the Austrian population—a contribution to forensic age estimation. *Forensic Sci Int* 2007; **169**: 161–7. doi: 10.1016/j.forsciint.2006.08.014

23. Chaillet N, Nyström M, Kataja M, Demirjian A. Dental maturity curves in Finnish children: Demirjian's method revisited and polynomial functions for age estimation. *J Forensic Sci* 2004; **49**: 1324–31.

24. Thorson J, Hägg U. The accuracy and precision of the third mandibular molar as an indicator of chronological age. *Swed Dent J* 1991; **15**: 15–22.

25. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; **33**: 159–74. doi: 10.2307/2529310

26. Kullman L, Tronje G, Teivens A, Lundholm A. Methods of reducing observer variation in age estimation from panoramic radiographs. *Dentomaxillofac Radiol* 1996; **25**: 173–8. doi: 10.1038/dmfr.25.4.9084269

27. Dhanjal KS, Bhrdwaj MK, Liversidge HM. Reproducibility of radiographic stage assessment of third molars. *Forensic Sci Int* 2006; **159**: 74–7. doi: 10.1016/j.forsciint.2006.02.020

28. Flood SJ, Mitchell WJ, Oxnard CE, Turlach BA, McGeachie J. To evaluate the utility of smaller sample sizes when assessing dental maturity curves for forensic age estimation. *J Forensic Sci* 2011; **56**: 1604–9. doi: 10.1111/j.1556-4029.2011.01884.x

29. Liversidge HM, Lyons F, Hector MP. The accuracy of three methods of age estimation using radiographic measurements of developing teeth. *Forensic Sci Int* 2003; **131**: 22–9. doi: 10.1016/S0379-0738(02)00373-0

30. Kullman L, Johanson G, Akesson L. Root development of the lower third molar and its relation to chronological age. *Swed Dent J* 1992; **16**: 161–7.

31. Peiris TS, Roberts GJ, Prabhu N. Dental age assessment: a comparison of 4- to 24-year-olds in the United Kingdom and an Australian population. *Int J Paediatr Dent* 2009; **19**: 367–76. doi: 10.1111/j.1365-263X.2009.00984.x

32. Braga J, Heuze Y, Chabadel O, Sonan NK, Gueramy A. Non-adult dental age assessment: correspondence analysis and linear regression versus Bayesian predictions. *Int J Legal Med* 2005; **119**: 260–74. doi: 10.1007/s00414-004-0494-8

33. Cruz-Landeira A, Linares-Argote J, Martínez-Rodríguez M, Rodríguez-Calvo MS, Otero XL, Concheiro L. Dental age estimation in Spanish and Venezuelan children. Comparison of Demirjian and Chaillet's scores. *Int J Legal Med* 2010; **124**: 105–12. doi: 10.1007/s00414-009-0380-5

34. 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. doi: 10.1111/1556-4029.12221

35. Willershausen B, Löffler N, Schulze R. Analysis of 1202 orthopantograms to evaluate the potential of forensic age determination based on third molar developmental stages. *Eur J Med Res* 2001; **6**: 377–84.

36. Liversidge HM. Interpreting group differences using Demirjian's dental maturity method. *Forensic Sci Int* 2010; **201**: 95–101. doi: 10.1016/j.forsciint.2010.02.032
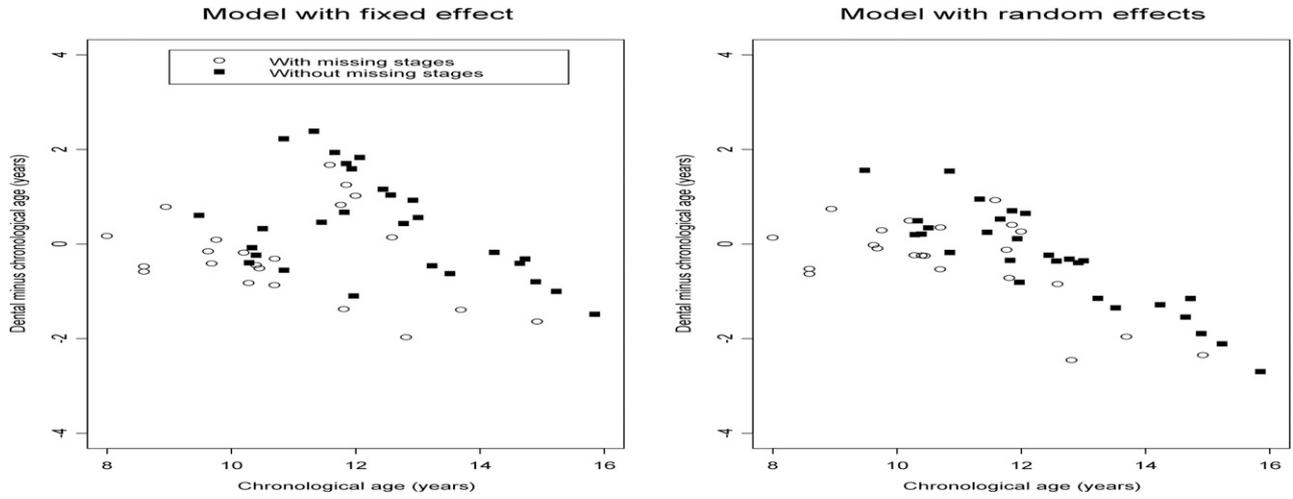
Appendix

Figure A1 Distribution of subjects with and without missing teeth for evaluation for fixed effect model and random effects model. The results were the same in both evaluation sessions.
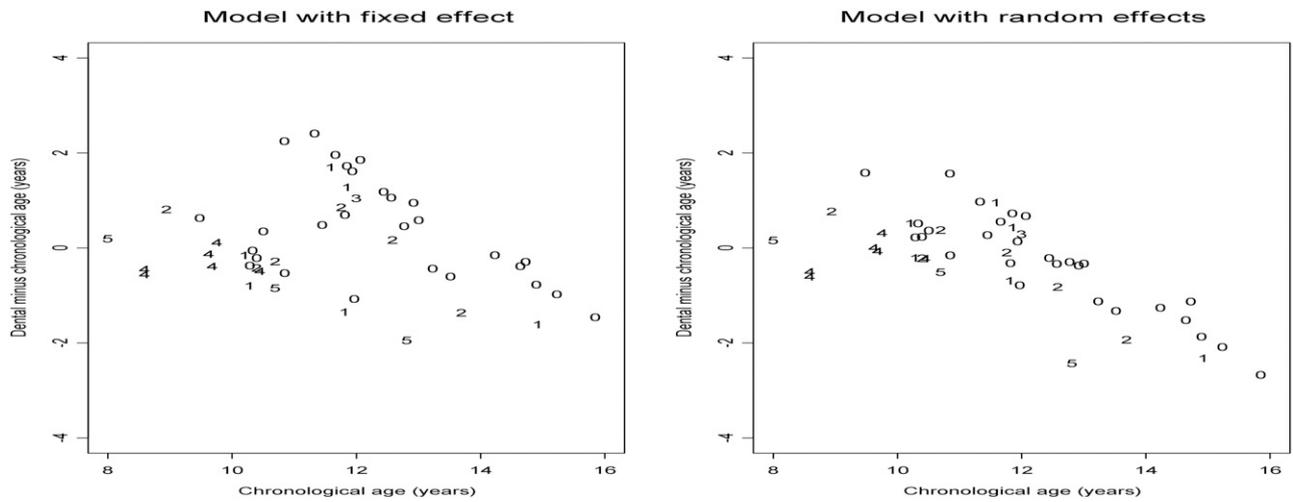

Figure A2 Distribution of number of undefinable stages per subjects for fixed effect model and random effects model. The results were the same in both evaluation sessions.