Available online at www.sciencedirect.com
SCIENCE DIRECT
Forensic Science International 159S (2006) S74–S77
Forensic Science International
www.elsevier.com/locate/forsciint

Reproducibility of radiographic stage assessment of third molars

K.S. Dhanjal, M.K. Bhardwaj, H.M. Liversidge *

Paediatric Dentistry, Dental Institute, Barts and The London School of Medicine and Dentistry, Turner Street, London E1 2AD, United Kingdom

Available online 10 March 2006

Abstract

The aim of this study was to determine intra- and inter-observer variability of the developing third molar from panoramic radiographs. Formation of third molars was assessed according to stages described by modified Demirjian et al.'s methods: Moorrees et al. [C.F.A. Moorrees, E.A. Fanning, E.E. Hunt, Age variation of formation stages for ten permanent teeth, J. Dent. Res. 42 (1963) 1490–1502] and Solari and Abramovitch [A.C. Solari, K. Abramovitch, The accuracy and precision of third molar development as an indicator of chronological age in Hispanics, J. Forensic Sci. 47 (2002) 531–535]; in addition, data were also analysed unmodified, i.e. Haavikko [K. Haavikko, The formation and alveolar and clinical eruption of the permanent teeth, an orthopantomograph study, Proc. Finn. Dent. Soc. 66 (1970) 104–170] and Demirjian et al. [A. Demirjian, H. Goldstein, J.M. Tanner, A new system of dental age assessment, Hum. Biol. 45 (1973) 211–227]. The sample was a random selection of 73 panoramic radiographs from patients aged 8–24 years. After training, the left maxillary and mandibular third molars were scored on two separate occasions without knowledge of previous scores. Cohen's Kappa and percentage agreement were calculated for each method, for maxillary, for mandibular third molars and combined. Percentage agreement for stages was also calculated.

Intra-observer agreement was greater for mandibular third molars compared to maxillary third molars, and better for methods with fewer stages. Kappa values indicated good agreement for most methods; the best was Demirjian et al.'s method for mandibular third molar with very good agreement ($K = 0.80$) for the first author, good agreement for the second author ($K = 0.75$) and good agreement between observers ($K = 0.75$). The stages with best agreement were Demirjian's stage E [A. Demirjian, H. Goldstein, J.M. Tanner, A new system of dental age assessment, Hum. Biol. 45 (1973) 211–227] and Moorrees et al.'s stage Cc and R1/4 [C.F.A. Moorrees, E.A. Fanning, E.E. Hunt, Age variation of formation stages for ten permanent teeth, J. Dent. Res. 42 (1963) 1490–1502].

Conclusions: Having clearly defined stages and fewer stages allowed better reproducibility of third molar formation.

Keywords: Third molar; Dental radiograph; Reproducibility

1. Introduction

Third molar formation is used to predict age during late adolescence and early adulthood. How accurately one can predict age is dependent on both the precision of stage assessment and the method of calculating age. The aim of this study was to assess the reproducibility of third molar formation stage assessment. The objectives were to determine intra- and inter-observer variability in radiographic stage assessment of developing third molars, to determine which method and which stage was most reproducible.

2. Materials and methods

The sample studied consisted of a random selection of dental panoramic tomograms taken in the course of diagnosis and treatment in the Dental Institute, Barts and The London School of Medicine and Dentistry. A total of 73 panoramic radiographs of patients (aged 8.97–23.79 years, mean 16.70 + 4.05 years) were examined. Left side third molars were assessed for crown or root stage. Scoring was first done using drawings from Moorrees et al. [1] and adapted Demirjian et al. which has two extra root growth stages [2]. Thereafter, data were collapsed into two methods [3–5] with fewer stages; Moorrees stage Coc collapsed into Haavikko C1/2, Ri into Cc, Cli into R1/4 and A1/2 into Rc. Solari stages F1 and F combined to Demirjian stage F and G1 and G into Demirjian G.

After training and calibration, radiographs were scored by the first two authors without knowledge of previous scores, and the results were compared. Cohen's Kappa [6] measuring strength of agreement was calculated for intra- and inter-observer agreement of each method. Results for maxillary teeth and mandibular teeth were calculated separately as well as combined. The percentage of identical readings for each

* Corresponding author. Tel.: +44 207 377 7058; fax: +44 207 377 7419.
E-mail address: h.m.liversidge@qmul.ac.uk (H.M. Liversidge).

0379-0738/$ – see front matter
doi:10.1016/j.forsciint.2006.02.020

Table 1
Intra-observer results showing Kappa values and percentage agreement

	Maxillary third molars, $N = 61$		Mandibular third molars, $N = 66$		All third molars, $N = 127$	
	Kappa	% equal	Kappa	% equal	Kappa	% equal
First author						
Moorrees et al.	0.65	69	0.63	67	0.65	67
Haavikko	0.74	79	0.79	74	0.73	74
Solari et al.	0.69	74	0.71	74	0.71	74
Demirjian et al.	0.74	79	0.80	83	0.78	81
Second author						
Moorrees et al.	0.56	61	0.65	68	0.61	64
Haavikko	0.68	74	0.71	74	0.70	74
Solari et al.	0.61	67	0.68	76	0.65	69
Demirjian et al.	0.62	69	0.75	79	0.70	74

Table 2
Inter-observer results showing Kappa values and percentage agreement

Method	Maxillary third molars, $N = 61$		Mandibular third molars, $N = 66$		All third molars, $N = 127$	
	Kappa	% equal	Kappa	% equal	Kappa	% equal
Moorrees et al.	0.43	49	0.59	62	0.52	56
Haavikko	0.63	67	0.65	70	0.64	69
Solari et al.	0.48	56	0.63	67	0.56	61
Demirjian et al.	0.61	67	0.75	79	0.68	73

developmental stage was calculated for each method for the first author for combined maxillary and mandibular third molars and compared where sample size ≥ 10.

3. Results

Intra-observer agreement measured using Kappa was highest for Demirjian et al.'s method for mandibular molars. Kappa values for mandibular third molars were greater than maxillary molars (Table 1). The two methods with fewer stages had higher Kappa values compared to other methods. Percentage agreement was highest for the mandibular molar using this Demirjian et al.'s method.

Inter-observer agreement showed moderate agreement for the methods of Moorrees and Solari, good agreement for Haavikko and Demirjian for the maxillary third molar. The highest Kappa value was seen for the mandibular third molar using Demirjian's method (Table 2). In all cases, mandibular third molars had higher Kappa values and higher percentage of identical readings than maxillary molars. In all cases, methods with fewer stages had better Kappa values and higher percentage of identical readings. Kappa values measuring strength of agreement for inter-observer variability show moderate agreement between observers (Table 3). Only the method of Demirjian et al. showed substantial/good agreement with considerably higher percentage agreement than other methods. Kappa value and percentage agreements were better for mandibular third molar compared to maxillary third molar. Demirjian method scored more than 79% agreement.

The stages with the highest identical readings were stages E for Demirjian's method and stage crown complete and root one quarter for Moorrees et al. with 90% or higher. Differences were one stage except for Solari stage G ($N = 6$), Moorrees C1/2 ($N = 2$) and Rc ($N = 3$) that were two stages different.

Table 3
Proportion of identical readings by stage (combined maxillary and mandibular molars)

Stage	Moorrees et al. [1]	Haavikko [3]
Ci		4/5
Cco		5/5
Coc	8/8	
C1/2	4/8	12/16
C3/4		3/5
Cc	15/16	18/21
Ri	2/5	
Cli	0/3	
R1/4	10/11	10/14
R1/2		5/7
R3/4		11/15
Rc		8/13
A1/2	5/15	
Ac	11/11	26/26

Stage	Solari and Abramovitch [2]	Demirjian et al. [4,5]
A		5/6
B		7/9
C		15/18
D		18/21
E		10/11
F	6/10	12/16
F1	5/6	
G	9/23	23/31
G1	6/8	
H		13/15

4. Discussion

Third molars differ from other molars in several respects. They are more variable in size, shape, timing of formation and eruption and agenesis [7,8]. The inclination of the developing third molar relative to the X-ray film may result in the crown appearing tilted on the radiograph, making crown stages difficult to assess. The roots of third molars are less divergent than other molars and are often fused making root stage assessment more difficult especially if stages include an estimate of root length rather than a clear view of the apical root cone angle. Most studies show that both intra- and inter-observer variability of third molars, using Demirjian's stages, is lower for this tooth compared to all other mandibular teeth. Mandibular first molars show better agreement (both intra- and inter-observer variability) than second molars [9–11] and this trend probably effects the entire molar field.

Calibration and training of stage assessment is necessary in order to reduce intra- and inter-observer variability [12]. Experience of the examiners, definition and number of stages also play a role. Reports in the literature measuring intra- and inter-observer variability varies and has been expressed as Kappa, percentage agreement although difference in dental age, median stage or correlation have also been used. Kappa values and percentage agreement are higher for intra-observer variation compared to that between observers, although this is not seen using weighted Kappa, that takes account of total as well as partial agreement [13]. Demirjian's method for all teeth has very good agreement (Kappa > 0.8) with more than 85% identical readings for intra-observer [9,14–17] and inter-observer variability [4,17]. Variability of assessment of the third molar using these stages is reported as 99% [18] for intra-observer agreement and 81.5% for maxillary and 85% for mandibular teeth between observers [19]. Kappa value of 0.88 is reported for inter-observer variability using these stages for mandibular third molars [20]. Results from the present study show lower values, although Kappa values have not previously been reported for third molars using these stages.

Stages first described by Gleiser and Hunt [21] and adaptions (including Moorrees et al.) show good agreement within one examiner [22], for roots of mandibular teeth [23], but reports vary for third molars; 0.65–0.79 Kappa for maxillary and 0.82–0.86 for mandibular intra-observer error for two examiners [24]. Inter-observer variability using these stages for maxillary and mandibular third molars crown stages was reported as 94% agreement and Kappa values of 0.87 and 0.88 for maxillary and 91–93% agreement and Kappa 0.86–0.89 for mandibular third molars [25]; for all stages, Kappa values of 0.84 and 0.87 for maxillary and mandibular third molars have also been reported [24]. Kullman described seven root stages for mandibular third molar formation and reports 57% agreement with breakdown by stage; the best agreement was observed for early root stages and worst for later stages particularly root complete apex closure not started [12]. Nolla's tooth stages show 97% agreement for mandibular third molars [26]; breakdown by stage for molars show that 'root cleft' stage had highest values followed by crown stages and the stage with least agreement was 'root complete'.

Reproducibility is higher for mandibular third molars than maxillary third molars. This is understandable as root formation of maxillary third molars can be superimposed with anatomical structures; including the hard palate, floor of the maxillary antrum, inferior border of the zygomatic arch, soft tissue and air shadows or artefacts.

In this study, we found that Demirjian's method showed the highest Kappa value (very good agreement) for both intra- and inter-observer readings for the mandibular third molar. The modified method with two extra stages [2] was lower in percentage agreement and Kappa values for both jaws and both within, and between examiners. Similarly, values for stages of Moorrees et al. [1] compared to Haavikko [3] were also lower. Increasing the number of tooth formation stages might improve accuracy, but too many reduces precision [27,28]. A comparison of reproducibility between stages of Demirjian and Haavikko found percentage agreement for Demirjian, similar to a previous study [14]. Clarity of definition was found to be more important than the number of traits in sex determination of skeletal material [29] and Demirjian's stages are the best described of any tooth assessment method. Moorrees et al. [1] provided 14 named stages and line drawings. Definition of crown and root stages was detailed in Demirjian's method [4,5] with several written criteria for each stage, a line drawing and a radiograph for mandibular anterior teeth, premolars and molars. In addition, instructions are helpful when not all criteria are met or a tooth presents between stages. One disadvantage of these eight stages is that only a few stages describe the latter half of root growth (stage F, G and H: mature apex). This is a limitation especially if a tooth has a long root and Thorson and Hägg [28] cite one study using 'late G'; others have added stages [2,30].

Stages with the highest agreement in this study were Moorrees et al.'s stage 'crown complete' and Demirjian's stage E; a result in accordance with previous findings that early molar root stages are more reliable than later stages [12,26].

Difficulties with the modified Demirjian stages were encountered with stages F1 and G1. This has been reported by De Salvia et al. [31] who mention the identification of stages F1 and G1 compared to F, G and H; they preferred the unmodified stages. These modified stages appear to have a similar accuracy in age prediction as the original eight stage system although neither age structure, variance or reproducibility are documented [31]. Another general difficulty in the present study was the fact that the radiographs were not selected for clarity or quality and many third molars were unclear and this may account for the slightly lower Kappa values than previously reported.

5. Conclusions

Demirjian's method of stage assessment of third molars showed very good agreement for both intra- and inter-examiner agreement. Having clearly defined stages and fewer intermediate stages allowed better reproducibility. We suggest that

Demirjian's method, rather than Solari et al., is to be preferred in assessing third molar development, as it is more reproducible. Further research is required to assess which method is most accurate in predicting age.

References

[1] C.F.A. Moorrees, E.A. Fanning, E.E. Hunt, Age variation of formation stages for ten permanent teeth, J. Dent. Res. 42 (1963) 1490–1502.

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

[3] K. Haavikko, The formation and alveolar and clinical eruption of the permanent teeth, an orthopantomograph study, Proc. Finn. Dent. Soc. 66 (1970) 104–170.

[4] A. Demirjian, H. Goldstein, J.M. Tanner, A new system of dental age assessment, Hum. Biol. 45 (1973) 211–227.

[5] A. Demirjian, H. Goldstein, New systems for dental maturity based on seven and four teeth, Ann. Hum. Biol. 3 (1976) 411–421.

[6] J.A. Cohen, A coefficient of agreement for nominal scales, Educ. Psychol. Meas. 20 (1960) 37–46.

[7] S.M. Garn, A.B. Lewis, B. Bonné, Third molar formation and its developmental course, Angle Orthod. 32 (1962) 270–279.

[8] G.Y. Levesque, A. Dermirjian, R. Tanguay, Sexual dimorphism in the development, emergence and agenesis of the mandibular third molar, J. Dent. Res. 60 (1981) 1735–1741.

[9] R. Nykänen, L. Espeland, S.I. Kvaal, O. Krogstad, Validity of the Demirjian method for dental age estimation when applied to Norwegian children, Acta Odont. Scand. 56 (1998) 238–244.

[10] G.Y. Levesque, A. Demirjian, The inter-examiner variation in rating dental formation from radiographs, J. Dent. Res. 59 (1980) 1123–1126.

[11] I.H. Leurs, E. Wattel, I.H.A. Aartman, E. Etty, B. Prahl-Andersen, Dental age in Dutch children, Eur. J. Orthod. 27 (2005) 309–314.

[12] L. Kullman, G. Tronje, A. Teivens, A. Lundholm, Methods of reducing observer variation in age estimation from panoramic radiograph, Dentomaxillofac. Radiol. 25 (1996) 173–178.

[13] M. Poyry, M. Nyström, R. Ranta, Comparison of two tooth formation rating methods, Proc. Finn. Dent. Soc. 82 (1986) 127–133.

[14] A. Olze, D. Bilang, S. Schmidt, K. Wernecke, G. Geserick, A. Schmeling, Validation of common classification systems for assessing the mineralizatin of third molars, Int. J. Legal Med. 119 (2005) 22–26.

[15] M. Nyström, J. Haataja, M. Kataja, M. Evalahti, L. Peck, E. Kleemola-Kujala, Dental maturity in Finnish children, estimated from the development of seven permanent mandibular teeth, Acta Odont. Scand. 44 (1986) 193–198.

[16] M. Nyström, R. Ranta, M. Kataja, H. Silvola, Comparisons of dental maturity between the rural community of Kuhmo in northeastern Finland and the city of Helsinki, Community Dent. Oral Epidemiol. 16 (1988) 215–217.

[17] C.S. Farah, D.R. Booth, S.C. Knott, Dental maturity of children in Perth, Western Australia, and its application in forensic age estimation, J. Clin. Forensic Med. 6 (1999) 14–18.

[18] J.L. Prieto, E. Barberia, R. Ortega, C. Magaña, Evaluation of chronological age based on third molar development in the Spanish population, Int. J. Legal Med. 119 (2005) 349–354.

[19] S. Arany, M. Iino, N. Yoshioka, Radiographic survey of third molar development in relation to chronological age among Japanese juveniles, J. Forensic Sci. 49 (2004) 534–538.

[20] P.M. Garamendi, M.I. Landa, J. Ballesteros, M.A. Solano, Reliability of the methods applied to assess age minority in living subjects around 18 years old. A survey on a Moroccan origin population, Forensic Sci. Int. 154 (2005) 3–12.

[21] I. Gleiser, E.E. Hunt, The permanent mandibular first molar: its calcification, eruption and decay, Am. J. Phys. Anthropol. 13 (1955) 253–283.

[22] G. Diaz, P. Maccioni, P. Zedda, F. Cabitza, I.M. Cortis, Dental development in Sardinian children, J. Craniofac. Genet. Dev. Biol. 13 (2003) 109–116.

[23] A.M. Grøn, Prediction of tooth emergence, J. Dent. Res. 41 (1962) 573–585.

[24] K. Mesotten, K. Gunst, A. Carbonez, G. Willems, Chronological age determination based on root development of a single third molar: a retrospective study based on 2513 OPGs, J. Forensic OdontoStomatol. 21 (2003) 31–35.

[25] M. Uzamiş, Ő. Kansu, T.U. Taner, R. Alpar, Radiographic evaluation of third-molar development in a group of Turkish children, J. Dent. Child. 136 (2000) 136–141.

[26] L.M.L. Lebret, Reproducibility of rating stages of tooth formation, Am. J. Orthod. 60 (1971) 90–91.

[27] E.A. Fanning, A longitudinal study of tooth formation and resorption, N. Z. Dent. J. 57 (1961) 202–217.

[28] J. Thorson, H. Hägg, The accuracy and precision of the third mandibular molar as an indicator of chronological age, Swed. Dent. J. 15 (1991) 15–22.

[29] D.E. Walrath, P. Turner, J. Bruzek, Reliability test of the visual assessment of cranial traits for sex determination, Am. J. Phys. Anthropol. 125 (2004) 132–137.

[30] R.L. Tompkins, Human population variability in relative dental development, Am. J. Phys. Anthropol. 99 (1996) 79–102.

[31] A. De Salvia, C. Calzetta, M. Orrico, D. De Leo, Third mandibular molar radiological development as an indicator of chronological age in a European population, Forensic Sci. Int. 146 (Suppl.) (2004) 9–12.