

# Bias and Accuracy of Age Estimation Using Developing Teeth in 946 Children

Helen M. Liversidge,<sup>1\*</sup> B. Holly Smith,<sup>2</sup> and Melissa Maber<sup>1</sup>

<sup>1</sup>*Institute of Dentistry, Bart's and The London School of Medicine and Dentistry, Queen Mary University of London, London E1 2AD, England, UK*

<sup>2</sup>*Museum of Anthropology, The University of Michigan, Ann Arbor, MI 48109*

**KEY WORDS** dental age; age determination; radiograph; crown; root

**ABSTRACT** Developing teeth are used to assess maturity and estimate age in several disciplines. The aim of the study was to determine which of the most well known dental age estimation methods was best at estimating age. The target sample of dental radiographs ( $N = 946$ , ages 3–16) was described by Maber et al. (*Forensic Sci Int* 159 (2006) S68–S73). Seven mandibular permanent teeth ( $I_1$ – $M_2$ ) were assessed, and dental age was calculated using four dental maturity scales and fifteen methods that use data for individual teeth. The mean difference between dental age and real age was calculated (bias) as well as several other measures of accuracy (mean/median absolute difference, percentage aged to within six months and to within 10% of real age). Most methods estimated age with significant bias and standard deviation of bias ranged from 0.86 to 1.03 years. Analysis by age group showed most methods over-aged younger children, and

considerably under-aged older children. The method that performed best was the dental maturity scale of Willems et al. (*J Forensic Sci* 46 (2001) 893–895) with bias of  $-0.14 \pm 0.86$  years ( $N = 827$ ), mean absolute difference of 0.66 years, 71% aged to 10% or less of age, and 49% aged to within six months. Two individual teeth,  $P_2$  and  $M_2$ , estimated age with bias not significantly different to zero for most formation stages using methods based on a large reference sample (L9a Demirjian stages) and a uniform age distribution (N25a Moorrees stages). Standard deviation of bias was least for early crown stages and most for late root stages. Methods that average ages for individual teeth improve if schedules for ‘mean age entering a stage’ are adjusted for prediction. Methods that directly calculate ‘mean age within stage’ can be improved by drawing from a uniform age distribution. *Am J Phys Anthropol* 143:545–554, 2010. ©2010 Wiley-Liss, Inc.

The developing dentition is used to assess maturity and estimate age in many disciplines including anthropology, archeology, forensic science, pediatric dentistry, and orthodontics. The development of each tooth can be divided into a series of maturity events—crown and roots stages. These biological age indicators are compared with a reference sample and from this we infer chronological age. During the last fifty years, numerous dental maturity studies have been reported and many are used to estimate maturity and age. Measures of performance and the terminology used to express accuracy of age estimation are varied and confusing. Some early studies report correlation between dental and chronological age, but this gives little information of the magnitude or direction of difference between dental age and real age. Accuracy refers to how close dental age is to chronological age. An age estimating method might consistently under- or over-estimate age and this is known as bias (Lovejoy et al., 1985). An accurate method has no bias, i.e. the mean difference between dental age and known age will be zero or close to zero. The standard deviation (SD) of the mean difference between dental age and real age, also known as the standard error of the estimate (Ritz-Timme et al., 2000), refers to the precision or reliability of estimated age. An age estimating method with high precision/reliability has a small SD, but could have substantial bias. A valid age estimating method is both accurate and precise, i.e. no bias and small SD. The terms precision and reliability are also used in the context of intra- and inter-observer reproducibility (see Ferranti and Cameriere, 2009). The difference between dental age and known age can be expressed in other ways

such as mean absolute difference (confusingly termed ‘accuracy’ by Lovejoy et al., 1985), median absolute difference, proportion aged to within an age interval, or to within a proportion of known age. Bayesian statistics are an alternate approach to measuring performance, but are not addressed in this study. The resolution of how age is measured is also of interest. Some studies report age up to two decimal points in 1, 3, or 6 monthly or year groups.

Numerous studies have investigated accuracy, precision, or reliability of various age estimating methods based on crown and root stages. Most report bias using a single method on small target samples of uneven age and different age ranges sometimes grouped into 5-year-olds, 6.5–9.5-year-olds, or children younger/older than 10. Several studies compare accuracy of two or more methods (Hägg and Matsson, 1985; Staaf et al., 1991; Saunders et al., 1993; Liversidge, 1994; Mørnstad et al., 1995; Rai and Anand, 2006; Rai, 2008). The findings from these studies are difficult to compare with conflicting results of

\*Correspondence to: Dr. Helen Liversidge, Institute of Dentistry, Bart's and The London School of Medicine and Dentistry, Queen Mary University of London, Turner Street, London E1 2AD, United Kingdom. E-mail: h.m.liversidge@qmul.ac.uk

Received 8 December 2009; accepted 20 April 2010

DOI 10.1002/ajpa.21349

Published online 8 July 2010 in Wiley Online Library (wileyonlinelibrary.com).

bias, but reliability of estimated age for an individual is low with SD of one year or less. Studies that compare age groups show that SD increases from younger to older age groups. More recently a Bayesian approach suggests that accuracy and reliability are not age related and geographic-specific methods do not improve the quality of age estimation (Braga et al., 2005).

This study set out to answer questions prompted by a request to the first author after the Tsunami in south-east Asia in 2004. Which method is best at estimating age and how is this best quantified? The aim of this study was to consider these questions by testing the most widely used dental maturity methods on one target sample and follows preliminary work by Maber et al. (2006).

## MATERIALS AND METHODS

The target sample consists of panoramic dental radiographs of 946 healthy children of known age attending a dental teaching hospital (Fig. 1). Subjects include 491 boys and 455 girls (mean age 9.80, standard deviation 4.05, range 3.00–16.99) with similar numbers of children from Bangladeshi and white ethnic origin for each year of age. The average age of permanent tooth formation is not significantly different between these ethnic groups (Liversidge, 2009). Tooth formation was assessed from panoramic radiographs taken with consent in the course of diagnosis and treatment. Radiographs were examined by the third author and seven mandibular teeth (excluding the third molar) on the left side were staged.

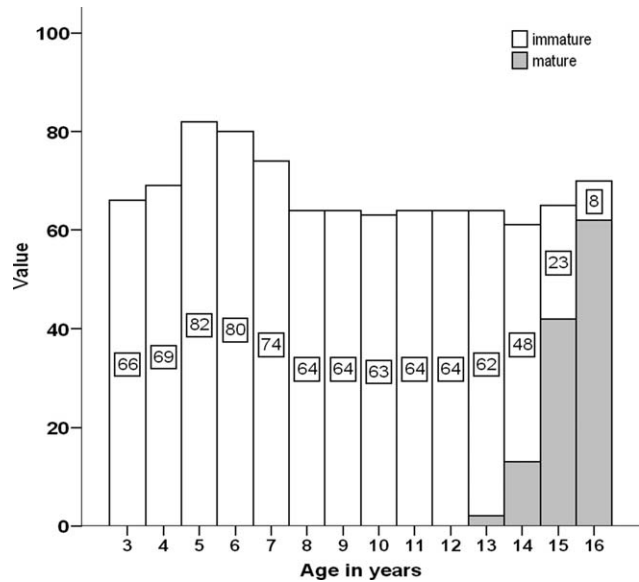


Fig. 1. Age distribution of the target sample. Shaded bars indicate number of individuals with all seven mandibular teeth mature.

Intraobserver error was assessed from duplicate scoring ten out of every hundred radiographs and showed good agreement (Maber et al., 2006).

Methods of estimating dental age are listed in Table 1. Dental age was calculated using four dental maturity scales and seven methods that use data for individual teeth. The dental maturity scales assessed were Nolla (1960), Demirjian (1994), Willems et al. (2001), and Chaillet et al. (2005). These methods rely on complete data from seven mandibular teeth and provide a maturity score that converts to a single dental age. The second group of methods includes those that directly calculate age of a developmental stage from their reference sample. These methods include Moorrees et al. (1963), Haavikko (1970), Anderson et al. (1976), Liversidge et al. (2006), Nyström et al. (2007), Liversidge (2010), and Liversidge (in prep). Many of these describe the age of 'entering' a stage of development and this can be adjusted for age prediction by adding half the interval to the onset of the next stage (Goldstein, 1979; Smith, 1991). Another modification is to select a uniform age distribution. Method L9 and L10 are Tables 9 and 10 in Liversidge et al. (2006). Method L10a (Liversidge, 2010) is selected to have a uniform age distribution from Liversidge et al. (2006). Methods Ny, Ny\_a, and Ny\_b are Tables II–V in Nyström et al. (2007). Methods N25a (adapted maturity data) and N25b (average age within a stage) are from Liversidge (in prep). Dental age was estimated for each individual in the target sample for this second group of methods by calculating the average of all developing teeth as well as by individual tooth type and stage. All dental ages for all methods were calculated from sex-appropriate tables.

Several other measures of accuracy were calculated including mean/median absolute difference between dental age and known age, the percentage of individuals aged to within 0.5 years and within 10% or less of age. Several very young children were too young to calculate dental age using Demirjian (1994) and maturity was calculated using tables from Demirjian and Goldstein

### Terminology of dental development and abbreviations

#### Tooth type:

I <sub>1</sub>	mandibular permanent central incisor
I <sub>2</sub>	mandibular permanent lateral incisor
C	mandibular permanent canine
P <sub>1</sub>	mandibular permanent first premolar
P <sub>2</sub>	mandibular permanent second premolar
M <sub>1</sub>	mandibular permanent first molar
M <sub>2</sub>	mandibular permanent second molar

#### Demirjian tooth stages:

A	initial cusp tips
B	fusion of cusp tips and outlined occlusal surface
C	occlusal enamel, dentine present, curved pulp roof
D	crown complete, initial root
E	bifurcation in molars, root one quarter, root length less than crown height
F	funnel shaped root ends, root length equal to crown height
G	root apical walls parallel
H	apex closed, normal periodontal ligament width

#### Moorrees et al. tooth stages:

Ci	initial cusp tips
Cco	coalescence of cusps tips
Coc	occlusal outline complete
C1/2	crown one half
C3/4	crown three quarters
Cc	crown complete
Ri	initial root
Rcl	root cleft
R1/4	root one quarter
R1/2	root half
R3/4	root three quarters
Rc	root complete
A1/2	apex half closed
Ac	apex closed.

TABLE 1. Methods of dental age assessed in this study: maturity scale, method that calculates one dental age from seven developing teeth; mean age entering, dental maturity data; midstage, mean age 'within a stage'

Method of analysis	Author reference	Abbreviation	Adaptation <sup>a</sup>	Adaptation <sup>b</sup>	Tooth stage description
Maturity scale	Nolla, 1960	N			Nolla, 1960
Maturity scale	Demirjian, 1994	D			Demirjian, 1994
Maturity scale	Willems et al., 2001	W			Demirjian, 1994
Maturity scale	Chaillet et al., 2005	Ch			Demirjian, 1994
Mean age entering	Moorrees et al., 1963	M	Ma <sup>c</sup>		Moorrees et al., 1963
Mean age entering	Haavikko, 1970	H	Ha		Moorrees et al., 1963 <sup>d</sup>
Mean age entering	Anderson et al., 1976	A	Aa		Moorrees et al., 1963
Mean age entering	Liversidge et al., 2006	L9	L9a		Demirjian, 1994
Mean age entering	Nyström et al., 2007	Ny	Ny_a		Demirjian, 1994
Mean age entering	Liversidge (in preparation)		N25a		Moorrees et al., 1963
Midstage	Liversidge et al., 2006	L10		L10a	Demirjian, 1994
Midstage	Nyström et al., 2007	Ny_b			Demirjian, 1994
Midstage	Liversidge (in preparation)			N25b	Moorrees et al., 1963

<sup>a</sup> Maturity data adapted for age prediction by adding half the interval to the next stage (Smith, 1991).

<sup>b</sup> Calculated from a uniform age distribution sample.

<sup>c</sup> Adapted by Smith, 1991.

<sup>d</sup> Haavikko, 1970 omits four Moorrees stages (Coc, Ri, Rcl, and A1/2).

TABLE 2. Bias (mean difference between dental and real age), mean/median absolute difference in years of dental age estimation methods

Method <sup>a</sup>	Type	N	Bias	SD	P	Bias rank	Mean absolute difference	Median absolute difference
Nolla stages								
N	Maturity scale	832	-1.04	0.95	**	16	1.11	0.88
Demirjian stages								
D	Maturity scale	827	0.25	0.86	**	9	0.71	0.60
W	Maturity scale	827	-0.14	0.86	**	=5	0.66	0.51
Ch	Maturity scale	827	-0.32	0.89	**	10	0.71	0.55
L9	Mean age entering	812	-1.13	0.95	**	17	1.22	0.98
L9a	Adapted	812	-0.21	0.98	**	7	0.75	0.56
L10	Midstage	827	-0.13	1.03	**	4	0.80	0.78
L10a	Midstage	827	-0.14	0.93	**	=5	0.75	0.57
Ny	Mean age entering	827	-1.34	1.03	**	19	1.38	1.18
Ny_a	Adapted	827	-0.35	0.93	**	11	0.74	0.54
Ny_b	Midstage	827	-0.23	1.02	**	8	0.79	0.59
Moorrees stages								
M	Mean age entering	833	-1.19	0.96	**	18	1.24	1.00
Ma	Adapted	833	-0.67	0.92	**	13	0.86	0.64
H	Mean age entering	832	-0.67	1.01	**	14	0.89	0.64
Ha	Adapted	832	0.04	0.96	Ns	=1	0.74	0.59
A	Mean age entering	833	-0.79	0.98	**	15	0.95	0.79
Aa	Adapted	833	-0.40	0.98	**	12	0.78	0.57
N25a	Adapted	833	-0.10	0.93	**	3	0.71	0.53
N25b	Midstage	833	-0.04	0.92	Ns	=1	0.70	0.55

<sup>a</sup> See Table 1 for abbreviations.

\*\*  $P < 0.01$ ; Ns bias not significant to zero.

N is the number of individuals with developing teeth; SD, standard deviation,

(1976) and Demirjian et al. (1973). Only developing teeth were used to estimate age and the proportion of individuals with all seven mandibular teeth mature for each year of age is shown as the shaded bars in Figure 1.

## RESULTS

Our results were analyzed firstly by a method combining all developing teeth and all stages for the entire age range, secondly by individual tooth type, where this was possible, and thirdly by separate stages of individual teeth. Results of the analysis combining all developing teeth and all stages for the entire age range are shown in Tables 2–4. Only two methods estimated age with bias not significant to zero, N25b and adapted Haavikko (1970). Other methods with little bias were adapted ma-

turity data (N25a), average age within stage from a large study (L10), and the maturity scale of Willems et al. (2001). Most methods under-estimated age, with two exceptions—Demirjian (1994) and adapted Haavikko (1970). Despite the large range of values for bias, the SD was between 0.86 and 1.03 years. The method with the smallest SD and smallest mean/median absolute difference was Willems dental maturity score. The mean absolute difference was similar for several other methods (N25b, N25a, Chaillet, Demirjian, and L10a). The methods that describe the age of 'entering' a stage of development (maturity data) all performed badly with the highest levels of bias and the worst values for other measures of accuracy (percentage aged to six months or to within 10% of age). Adjusting these methods for age prediction considerably improved performance. The method

TABLE 3. Bias and SD by year age cohort

Method <sup>a</sup>	Age cohorts <sup>b</sup>															
	3	4	5	6	7	8	9	10	11	12	13	14	15	16		
<b>Nolla stages</b>																
N	-0.41, 0.33	-0.48, 0.61	-0.68, 0.68	-0.73, 0.59	-0.77, 0.78	-0.87, 0.72	-1.00, 0.77	-1.24, 0.81	-1.18, 1.05	-1.33, 1.07	-1.60, 1.07	-1.83, 0.87	-2.35, 0.91	-3.33, 0.70		
<b>Demirjian stages</b>																
D	<b>0.12</b> , 0.75	0.65, 0.71	0.56, 0.61	0.53, 0.51	0.35, 0.72	0.19, 0.65	0.31, 0.91	0.46, 0.98	0.33, 0.89	0.30, 0.92	- <b>0.04</b> , 0.77	-0.42, 0.69	-1.03, 0.51	-1.95, 0.44		
W	0.22, 0.43	0.25, 0.50	- <b>0.04</b> , 0.63	<b>0.13</b> , 0.63	<b>0.16</b> , 0.85	<b>0.04</b> , 0.77	- <b>0.02</b> , 0.74	- <b>0.12</b> , 0.87	- <b>0.21</b> , 0.87	-0.30, 0.99	-0.52, 0.83	-0.95, 0.83	-1.48, 0.74	-2.61, 0.55		
Ch	<b>0.17</b> , 0.68	0.40, 0.51	<b>0.03</b> , 0.63	<b>0.05</b> , 0.52	-0.17, 0.72	-0.31, 0.71	-0.29, 0.84	-0.28, 0.92	-0.47, 0.77	-0.59, 0.86	-0.81, 0.93	-1.12, 0.83	-1.62, 0.74	-2.79, 0.63		
M	-0.57, 0.47	-0.38, 0.55	-0.75, 0.42	-0.87, 0.49	-0.98, 0.72	-1.12, 0.63	-1.04, 0.90	-1.07, 0.94	-1.26, 0.93	-1.47, 0.82	-1.90, 0.88	-2.38, 0.73	-3.09, 0.65	-3.87, 0.32		
Ma	-0.23, 0.48	- <b>0.03</b> , 0.56	-0.39, 0.43	-0.44, 0.56	-0.47, 0.78	-0.56, 0.64	-0.54, 0.86	-0.59, 0.94	-0.72, 0.97	-0.85, 0.94	-1.18, 1.02	-1.57, 0.85	-2.27, 0.75	-2.88, 0.35		
H	0.32, 0.41	<b>0.06</b> , 0.46	-0.38, 0.53	-0.33, 0.53	-0.48, 0.70	-0.65, 0.57	-0.59, 1.01	-0.41, 1.02	-0.57, 0.90	-0.92, 0.69	-1.44, 0.71	-2.06, 0.66	-2.77, 0.68	-3.77, 0.54		
Ha	0.96, 0.39	0.65, 0.47	0.23, 0.50	0.29, 0.57	0.23, 0.73	<b>0.11</b> , 0.57	<b>0.14</b> , 0.95	<b>0.23</b> , 0.98	<b>0.11</b> , 0.92	- <b>0.17</b> , 0.78	-0.57, 0.84	-1.10, 0.79	-1.79, 0.81	-2.69, 0.62		
A	0.27, 0.27	- <b>0.04</b> , 0.48	-0.48, 0.47	-0.47, 0.55	-0.53, 0.77	-0.67, 0.60	-0.75, 0.84	-0.80, 0.94	-0.93, 0.96	-1.12, 0.82	-1.49, 0.86	-1.96, 0.71	-2.68, 0.71	-3.65, 0.58		
Aa	0.45, 0.29	0.25, 0.53	-0.13, 0.49	- <b>0.11</b> , 0.54	- <b>0.14</b> , 0.83	-0.26, 0.66	-0.34, 0.88	-0.39, 1.00	-0.51, 1.00	-0.68, 0.92	-1.04, 0.98	-1.46, 0.84	-2.15, 0.79	-3.05, 0.58		
<b>Moorrees stages</b>																
L9	0.32, 0.44	-0.17, 0.39	-0.56, 0.49	-0.65, 0.50	-0.85, 0.64	-1.10, 0.60	-1.24, 0.74	-1.28, 0.94	-1.28, 1.00	-1.53, 0.87	-1.99, 0.64	-2.59, 0.48	-3.30, 0.42	-4.66, 0.29		
L9a	0.88, 0.37	0.54, 0.40	0.17, 0.50	0.14, 0.56	<b>0.06</b> , 0.72	- <b>0.09</b> , 0.64	-0.22, 0.74	-0.28, 0.92	-0.28, 1.00	-0.51, 0.91	-0.87, 0.75	-1.37, 0.59	-2.05, 0.51	-3.35, 0.42		
L10	0.89, 0.40	0.74, 0.44	0.40, 0.51	0.36, 0.57	0.27, 0.69	<b>0.07</b> , 0.60	- <b>0.15</b> , 0.69	-0.28, 0.85	-0.34, 0.94	-0.60, 0.85	-1.00, 0.68	-1.53, 0.51	-2.24, 0.43	-3.44, 0.34		
L10a	0.72, 0.37	0.50, 0.44	0.17, 0.52	0.19, 0.61	<b>0.15</b> , 0.73	<b>0.02</b> , 0.63	- <b>0.13</b> , 0.73	- <b>0.20</b> , 0.91	- <b>0.20</b> , 0.99	-0.40, 0.91	-0.76, 0.70	-1.25, 0.53	-1.96, 0.40	-3.05, 0.34		
Ny	-0.69, 0.63	-0.53, 0.59	-0.74, 0.60	-0.69, 0.55	-0.87, 0.62	-1.17, 0.55	-1.41, 0.66	-1.52, 0.88	-1.56, 1.00	-1.77, 0.90	-2.19, 0.72	-2.68, 0.54	-3.40, 0.44	-4.59, 0.39		
Ny_a	<b>0.10</b> , 0.60	0.29, 0.58	<b>0.08</b> , 0.57	<b>0.09</b> , 0.55	- <b>0.04</b> , 0.67	-0.26, 0.59	-0.43, 0.71	-0.49, 0.96	-0.42, 1.08	-0.57, 0.99	-0.89, 0.79	-1.33, 0.59	-2.03, 0.47	-3.19, 0.39		
Ny_b	0.60, 0.52	0.66, 0.50	0.37, 0.52	0.32, 0.52	0.16, 0.65	- <b>0.08</b> , 0.59	-0.28, 0.68	-0.43, 0.86	-0.47, 0.96	-0.71, 0.88	-1.07, 0.75	-1.54, 0.58	-2.23, 0.51	-3.45, 0.42		
N25a	0.37, 0.46	0.57, 0.66	0.33, 0.51	0.29, 0.49	<b>0.11</b> , 0.66	- <b>0.12</b> , 0.55	- <b>0.12</b> , 0.94	- <b>0.01</b> , 1.02	- <b>0.07</b> , 0.99	-0.25, 0.80	-0.67, 0.81	-1.17, 0.69	-1.94, 0.59	-2.57, 0.39		
N25b	0.45, 0.44	0.59, 0.61	0.29, 0.49	0.28, 0.53	0.21, 0.74	<b>0.07</b> , 0.58	<b>0.07</b> , 0.91	<b>0.11</b> , 0.97	- <b>0.01</b> , 0.92	-0.22, 0.80	-0.62, 0.80	-1.18, 0.68	-1.83, 0.50	-2.63, 0.31		
N range	51-66	69	82	80	74	64	64	63	64	62	60-62	48-50	23-25	8		

<sup>a</sup> See Table 1 for abbreviations.

<sup>b</sup> Cohort 3 includes individuals aged 3.00-3.99 etc.

Bold indicates bias not significantly different to zero.

TABLE 4. Percentage of individuals aged to within 0.5 years and  $\leq 10\%$  of age

Method <sup>a</sup>	Type	N	% < 0.5 years	% $\leq 10\%$ of age
Nolla stages				
N	Maturity scale	832	25	33
Demirjian stages				
D	Maturity scale	827	42	64
W	Maturity scale	827	49	71
Ch	Maturity scale	827	45	67
L9	Mean age entering	812	26	39
L9a	Adapted	812	46	63
L10	Midstage	827	40	51
L10a	Midstage flat	827	48	65
Ny	Mean age entering	827	19	30
Ny_a	Adapted	827	48	65
Ny_b	Midstage	827	43	62
Moorrees stages				
M	Mean age entering	833	20	33
Ma	Adapted	833	40	58
H	Mean age entering	832	40	59
Ha	Adapted	832	42	66
A	Mean age entering	833	37	52
Aa	Adapted	833	45	65
N25a	Adapted	833	47	68
N25b	Midstage flat	833	47	68

<sup>a</sup> See Table 1 for abbreviations.

with the highest percentage of individual aged to within six months of real age and the highest proportion aged to 10% or less of age was Willems at 49 and 71%, respectively. This seems straightforward until we separate tabulations by age group (Table 3, Fig. 2). Evidently, the method with the best overall bias is only truly best in the middle range ages, from 8 to 11 years. For the youngest ages, some of the apparent “worst” are “best,” at the oldest group of 13–16 where all methods suffer from under-estimation, Demirjian is clearly the best. Whatever the bias, all the methods show a trend in bias across age groups: young ages are over-aged, the middle is best, and older ages are under-aged. The wide range of bias is illustrated for N25b in Figure 3.

Results of bias and measures of accuracy for individual teeth (all stages combined) are shown in Tables 5 and 6. Several methods performed well with five individual teeth estimating age with bias not significant to zero (L9a, L10a, and Ny\_b). One method estimated age with no significant bias using I<sub>1</sub> and M<sub>1</sub>, whereas five methods estimated age with no bias using P<sub>2</sub>. Fewer tooth formation stages (Demirjian vs. Moorrees) resulted in better performance. The percentage of individuals aged to 0.5 years of known age for individual teeth varied from 54% (Anderson central incisor) to 14% (Moorrees canine). The percentage of individuals aged to within 10% or less of age for individual teeth was highest at 61% for M<sub>2</sub> (L10a, Ny\_a, and N25b); adapted Moorrees central incisor, N25a first premolar, and second molar scored 60%. The worst tooth was 22% using M<sub>1</sub> of Moorrees.

Results of individual tooth stages that estimated age with no significant bias are shown in Table 7 (Demirjian stages) and Table 8 (Moorrees stages). For many tooth stages, several methods performed well with similar values of standard deviation. The method with the most number (24) of Demirjian tooth stages was L10a (including all stages of M<sub>2</sub>) followed by L9a. Two methods performed well using individual Moorrees formation

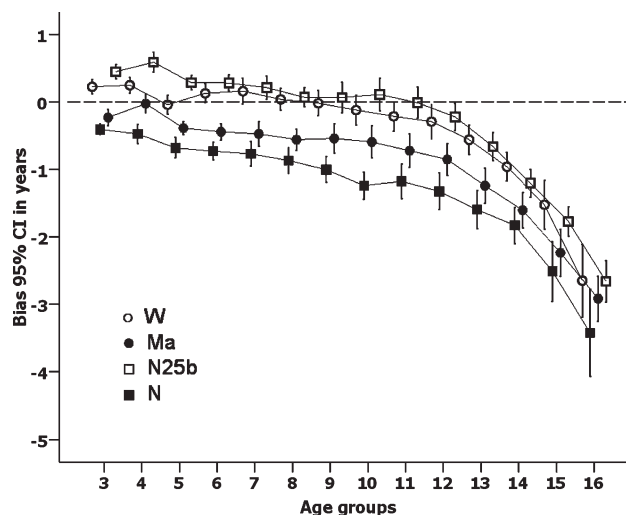


Fig. 2. Bias (95% confidence interval) of some methods in years by age cohort. Open circles, Willems; filled circles, adapted Moorrees; open squares, N25b; filled squares, Nolla. Dotted line is zero bias, when estimated age coincides with actual age.

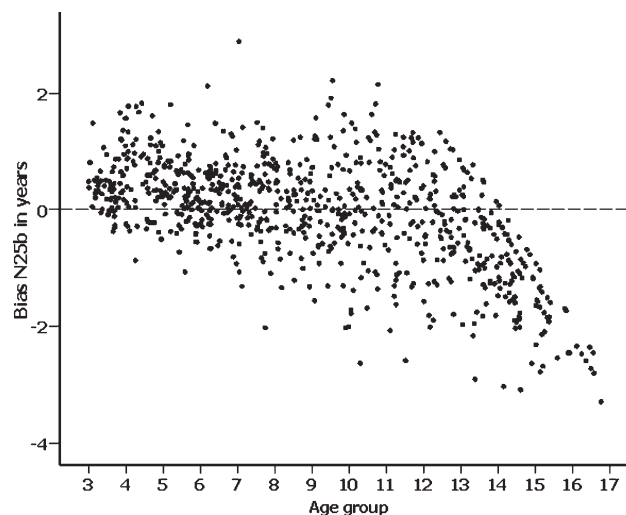


Fig. 3. Bias in years plotted against age for method N25b. Dotted line is zero bias, when estimated age coincides with actual age.

stages (30 stages using N25a and 26 using N25b), including many crown and root stages of P<sub>2</sub> and M<sub>2</sub>. Standard deviation of bias and mean absolute difference increased with development from early to late formation stage, from 0.37 to 1.32 years. Analysis by single tooth type or by age group is complicated by the fact that not all teeth continue development for the age range of the target sample. Incisors and first permanent molars complete maturation by around ten years of age whereas the formation of P<sub>2</sub> and M<sub>2</sub> is entirely encompassed within the age range of the target sample. It is therefore unsurprising that these two teeth perform best.

## DISCUSSION

Selecting the best method of estimating age is not as straightforward as it first appears. Criteria for choosing

TABLE 5. Bias and SD in years of individual teeth (all stages combined)

Method <sup>a</sup>	I <sub>1</sub>	I <sub>2</sub>	C	P <sub>1</sub>	P <sub>2</sub>	M <sub>1</sub>	M <sub>2</sub>
Demirjian stages							
L9	-0.65, 0.92	-0.77, 1.05	-1.07, 1.35	-0.91, 1.05	-0.92, 1.16	-1.30, 1.12	-1.08, 1.24
L9a	<b>0.00</b> , 0.92	<b>0.08</b> , 1.04	<b>-0.04</b> , 1.35	<b>0.00</b> , 1.05	<b>0.05</b> , 1.15	-0.26, 1.03	-0.14, 1.08
L10	0.24, 0.86	0.14, 0.97	0.14, 1.26	0.13, 1.05	0.11, 1.17	<b>0.05</b> , 0.99	<b>-0.04</b> , 1.11
L10a	0.10, 0.86	<b>0.01</b> , 0.96	<b>0.04</b> , 1.22	<b>0.04</b> , 1.01	<b>0.05</b> , 1.13	-0.14, 1.00	<b>-0.04</b> , 1.06
Ny	-0.93, 0.92	-0.97, 1.07	-1.35, 1.35	-1.14, 1.07	-0.98, 1.21	-1.26, 1.06	-1.11, 1.11
Ny <sub>a</sub>	-1.18, 0.88	-0.22, 1.04	-0.24, 1.28	-0.23, 1.05	-0.12, 1.14	-0.34, 0.99	-0.14, 1.06
Ny <sub>b</sub>	0.09, 0.87	<b>0.05</b> , 1.00	<b>-0.04</b> , 1.27	<b>-0.06</b> , 1.06	<b>0.04</b> , 1.19	-0.10, 0.99	<b>-0.01</b> , 1.10
Moorrees stages							
M	-0.56, 0.84	-0.63, 0.89	-1.47, 1.14	-1.09, 0.95	-0.90, 1.13	-1.20, 1.02	-1.04, 1.13
Ma	-0.29, 0.85	-0.29, 0.88	-0.88, 1.14	-0.57, 0.94	-0.39, 1.12	-0.73, 0.98	-0.58, 1.08
H	-0.71, 0.93	-0.74, 0.97	-0.95, 1.23	-0.33, 1.02	-0.46, 1.20	-0.50, 1.22	-0.38, 1.11
Ha	-0.16, 0.89	-0.09, 0.90	-0.15, 1.23	0.33, 1.02	0.15, 1.17	0.20, 1.03	0.35, 1.09
A	-0.30, 0.80	-0.49, 0.90	-0.64, 1.09	-0.84, 1.02	-0.82, 1.19	-0.35, 0.93	-0.74, 1.10
Aa	0.05, 0.80	-0.13, 0.92	-0.46, 1.15	-0.41, 1.01	-0.39, 1.17	0.10, 0.93	-0.32, 1.08
N25a	0.19, 0.90	-0.14, 0.97	-0.11, 1.18	0.18, 0.94	<b>-0.07</b> , 1.12	0.10, 0.93	0.09, 1.06
N25b	0.17, 0.89	<b>-0.06</b> , 0.88	<b>0.06</b> , 1.13	0.15, 0.95	<b>0.03</b> , 1.12	0.17, 0.93	0.14, 1.06
N range	431-490	799-808	507-724	647-695	699-760	431-490	799-808

<sup>a</sup> See Table 1 for abbreviations. Bold indicates bias not significantly different to zero.

TABLE 6. Percentage of individuals aged to 0.5 year and to ≤10% of age by individual tooth type

Method <sup>a</sup>	Percentage aged to 0.5 year of actual age							Percentage aged to ≤10% of actual age						
	I <sub>1</sub>	I <sub>2</sub>	C	P <sub>1</sub>	P <sub>2</sub>	M <sub>1</sub>	M <sub>2</sub>	I <sub>1</sub>	I <sub>2</sub>	C	P <sub>1</sub>	P <sub>2</sub>	M <sub>1</sub>	M <sub>2</sub>
Demirjian stages														
L9	33	31	23	31	28	23	29	40	38	37	44	44	25	42
L9a	43	39	32	37	36	33	39	52	51	53	56	55	45	59
L10	42	38	31	37	32	40	35	49	49	45	52	52	47	57
L10a	45	41	32	39	36	38	38	51	50	47	57	55	46	<b>61</b>
Ny	26	27	20	23	29	23	28	31	32	28	32	41	24	41
Ny <sub>a</sub>	39	40	30	39	38	39	<b>40</b>	47	48	44	53	<b>56</b>	46	<b>61</b>
Ny <sub>b</sub>	50	37	34	38	33	40	34	53	46	50	55	51	48	<b>56</b>
Moorrees stages														
M	45	38	14	24	30	19	30	42	43	42	57	55	48	59
Ma	49	45	26	37	37	32	<b>40</b>	56	56	52	55	55	47	57
H	36	33	27	41	38	44	37	54	48	23	33	46	22	44
Ha	48	<b>46</b>	34	37	35	37	36	<b>60</b>	<b>59</b>	40	51	55	38	56
A	<b>54</b>	44	33	29	32	47	36	57	52	48	41	45	54	53
Aa	52	44	34	36	37	<b>50</b>	38	59	51	48	51	53	57	59
N25a	45	43	36	42	38	48	37	53	50	52	<b>60</b>	<b>56</b>	<b>58</b>	60
N25b	47	<b>46</b>	<b>37</b>	<b>40</b>	<b>40</b>	46	39	53	54	<b>54</b>	58	55	53	<b>61</b>

<sup>a</sup> See Table 1 for abbreviations. Bold best per tooth.

a good method include low bias, low mean/median absolute difference, and high proportion of individuals aged to six months and to within 10% of age. Which method is best at estimating age? Overall, Willems stands out as performing better for all measures of accuracy despite a small significant under-estimation of age. If less than seven teeth are available and Moorrees stages are preferred, N25b is the method with least bias in estimating age. The best individual teeth using Demirjian stages was M<sub>2</sub> using L10a and for Moorrees stages it was P<sub>2</sub>, M<sub>1</sub>, and M<sub>2</sub> using N25a or N25b. The methods with bias of more than a year perform badly in all measures of accuracy and are not recommended to estimate age. These include Moorrees et al. (1963) and Nolla (1960).

What is the most useful way to quantify how good a method is at estimating age? The most important measure is the lack of bias. In our study, most methods with low bias also had low mean/median absolute difference as well as high percentage aged to within six months or within 10% of age and vice versa. For the full target sample, most methods had similar levels of reliability,

but for individual teeth, SD of early stages was much smaller than later stages.

Accuracy expressed as the proportion of the target group aged to within six months of known age is one way of assessing performance. Braga et al. (2005) report the proportion of individuals aged to ± three months of an age interval as 15%, with a range from 2 to 23% depending on geographic group or type of analysis. A level of 23–25% is reported for a small sample of 40 known-age-at-death skeletal remains (Heuzé and Cardoso, 2008). Our results (Willems 49% best, Nyström worst at 19%) are considerably better than this, but it is unclear if this is because of differences in the sample size, age range, age distribution of the target sample, or analysis by point estimate rather than age interval. Analysis by age interval is complicated when individuals close to the cut points are aged into the adjacent age category. For instance a child whose actual age is 53 months might have a dental age of 54 months. If the next age category begins at 54 months, accuracy for this child is ranked into the next age category, i.e. accuracy

TABLE 7. Bias and SD in years for individual tooth stages that estimate age with bias not significant to zero (Demirjian stages)

Tooth	Stage	N	Methods using Demirjian tooth descriptions <sup>a</sup>				
			L9a	L10	L10a	Ny_a	Ny_b
I <sub>1</sub>	E	159				0.03, 0.86	
	F	79	0.01, 0.79	0.16, 0.79	0.03, 0.79	0.01, 0.79	
	G	125		0.05, 0.97	-0.04, 0.98		-0.15, 0.96
I <sub>2</sub>	C	22		0.05, 0.47	0.13, 0.43		
	D	72			0.14, 0.78	-0.11, 0.78	
C	F	93		0.06, 0.88	0.00, 0.88		-0.04, 0.87
	C	68					-0.15, 0.65
	D	131			0.17, 0.99	0.13, 0.98	
P <sub>1</sub>	F	172	-0.07, 1.23	-0.09, 1.22	-0.01, 1.22		-0.17, 1.23
	B	45			0.11, 0.38		
	C	91	-0.05, 0.71		0.04, 0.71		0.01, 1.13
P <sub>2</sub>	E	145	0.10, 1.13		0.12, 1.13		
	F	153	-0.03, 1.14	-0.08, 1.14	-0.01, 1.14		
	B	55			-0.06, 0.63		
M <sub>1</sub>	C	111	0.04, 0.88		0.13, 0.88		
	D	130	0.05, 1.09			0.18, 1.08	
	E	125	-0.13, 1.24	-0.01, 1.25	0.01, 1.24	-0.15, 1.23	-0.04, 1.25
	F	140	-0.01, 1.28	-0.03, 1.27	0.12, 1.27		
	G	138	-0.13, 1.32		-0.21, 1.32		
	D	46					-0.02, 0.37
M <sub>2</sub>	E	143		0.02, 0.86			0.00, 0.86
	F	104		-0.02, 0.86			0.01, 0.86
	G	184		0.04, 1.22	-0.06, 1.22		
	A	35	-0.09, 0.45		0.09, 0.46		
	B	57	0.05, 0.61		0.12, 0.62	-0.06, 0.61	
	C	154	-0.03, 0.88		0.09, 0.88	0.06, 0.88	
M <sub>2</sub>	D	147	-0.07, 1.07	0.13, 1.07	-0.02, 1.07	-0.06, 1.07	0.13, 1.07
	E	119	-0.08, 1.10	-0.09, 1.10	-0.06, 1.10		-0.11, 1.11
	F	105			-0.15, 1.13		
	G	184			-0.10, 1.29	-0.18, 1.30	

<sup>a</sup> See Table 1 for abbreviations.

rank of more than six months but less than one year, rather than being ranked to the most accurate group of within  $\pm 3$  months.

What is the confidence interval of estimated age? The SD of bias from this study for all methods (all teeth combined and all stages combined) was around a year (Tables 2 and 5), making the 95% confidence interval of estimated age for an individual around  $\pm 2$  years. Further analysis by individual tooth stages (Tables 7 and 8) shows this to be age-related. Some early crown stages or stages that occur near the minimum age of the target sample (Demirjian stages C of I<sub>1</sub>, stage B of P<sub>1</sub>, M1 stage D, M2 stage A) have a SD of less than six months, whereas this value for some late root stages is over a year. The minimum and maximum SD for tooth stages with bias not significant to zero of methods L9a, L10, L10a, N25a, and N25b is 0.33 and 1.32 years, making the range of 95% confidence intervals from  $\pm 0.65$  to  $\pm 2.59$  years. This reflects the known increase in SD with tooth formation stage from 0.6 to 1.6 years (Anderson et al., 1976; Liversidge et al., 2006; Nyström et al., 2007; Liversidge, 2009).

The age distribution, structure, and sample size of both the reference and target samples are all important attributes (Konigsberg and Frankenberg, 1992, 2002; Hoppa and Vaupel, 2002; Steadman et al., 2006; Kimmerle et al., 2008; Prince and Konigsberg, 2008). Dental radiographic studies seldom include very young children and consequently many tooth stages are truncated at the minimum age and these stages will estimate age with bias. If the minimum age of a reference sample is three years, it is inappropriate to estimate age for individuals younger than this (Saunders et al., 1993). A child who is

only just three years of age and who is dentally delayed will not be aged accurately if the reference sample includes a small number of three-year-olds. The minimum age of our target sample (three years) explains why so few early stages of anterior teeth and the first permanent molar are included in Tables 6–8. The first permanent molar in three-year-olds of our target sample ranged from crown three quarters to quarter root. Stage C3/4 for M<sub>1</sub> is not represented by sufficient early, average, and late maturers, but only by a few delayed individuals. The maximum age of the target sample is also of importance. If we exclude the third molar, an individual will be dentally mature when the second molar distal apex closes, which in our target sample was earliest at 13 years (Fig. 1). Once this occurs, age cannot be estimated using either individual developing permanent teeth or a dental maturity scale, unless the third molar is assessed. As children reach dental maturity, they drop out of the target sample; these were not excluded in the initial analysis of Maber et al. (2006). This drop out results in fewer and fewer individuals in the older age groups whose dental age can be calculated. Those that remain are dentally delayed compared with an average child. In our sample, eight dentally delayed individuals represent the 16-year-old age group and all methods underestimate their age considerably.

A uniform age distribution with similar numbers for each year of age is desirable in both reference and target samples (Bocquet-Appel and Masset, 1982; Konigsberg and Frankenberg, 2002). Variance is inversely proportional to  $\sqrt{N}$  and in a normal distribution accuracy will be better at mean age and is poor at the age extremes where the sample size is small. This difficulty is overcome

TABLE 8. Bias and SD in years for individual tooth stages that estimate age with bias not significant to zero (Moorrees stages)

Tooth	Stage	N	Methods that use Moorrees tooth stages <sup>a</sup>						
			Ma	H	Ha	A	Aa	N25a	N25b
I <sub>1</sub>	Cc	67							
	R1/4	50 <sup>b</sup>	-0.03, 0.81		-0.11, 0.79	-0.14, 0.80		-0.05, 0.51	-0.04, 0.51
	R1/2	57		-0.07, 0.74		-0.14, 0.70			
	R3/4	52	0.02, 0.65		0.11, 0.64	-0.17, 0.68	0.16, 0.69		
	Rc	71					-0.18, 0.90	-0.15, 0.94	
I <sub>2</sub>	A1/2	68					-0.16, 0.99		
	C3/4	23				-0.11, 0.43	0.09, 0.43		
	R1/4	93	-0.12, 0.75			-0.09, 0.76			
	R1/2	60			-0.10, 0.70		0.13, 0.73		
	R3/4	50	-0.16, 0.73				-0.16, 0.73	-0.14, 0.76	-0.19, 0.74
C	A1/2	44	-0.31, 1.03				0.01, 1.06		
	C1/2	26					0.04, 0.51		
	Cc	125						-0.10, 0.91	0.05, 0.90
	R1/4	113			0.06, 1.01		-0.09, 1.04		
	R1/2	71	-0.23, 1.09		-0.11, 1.08			-0.13, 1.16	0.18, 1.07
P <sub>1</sub>	R3/4	101					-0.19, 1.31	0.12, 1.42	
	Rc	102						-0.14, 1.16	-0.16, 1.04
	Coc	13				0.02, 0.23			
	C1/2	58			0.11, 0.60		-0.12, 0.62		
	C3/4	82		0.08, 0.72					
	Cc	129							0.07, 0.88
	R1/4	110		0.01, 1.04			-0.07, 1.03		
	R1/2	70		0.11, 0.95				0.09, 0.97	
	R3/4	91						0.06, 1.04	0.20, 1.07
	Rc	56						-0.01, 1.08	-0.11, 1.09
P <sub>2</sub>	A1/2	55						-0.34, 1.06	
	Ci	33				0.09, 0.61	0.05, 0.56		-0.16, 0.56
	Cco	15		-0.13, 0.60		-0.30, 0.59	-0.12, 0.58	-0.14, 0.59	0.02, 0.61
	Coc	26	-0.12, 0.63				-0.10, 0.63	0.16, 0.63	
	C1/2	69	-0.13, 0.77	-0.06, 0.74			-0.10, 0.74		
	C3/4	77	-0.25, 0.83	0.01, 0.82					0.16, 0.85
	Cc	101						-0.16, 1.16	-0.07, 1.17
	R1/4	102	-0.20, 1.11	-0.11, 1.11			0.02, 1.13		
	R1/2	78	-0.24, 1.43		0.24, 1.21			-0.26, 1.37	0.04, 1.40
	R3/4	79			0.21, 1.10			0.02, 1.09	0.23, 1.11
M <sub>1</sub>	Rc	73						-0.28, 1.24	
	A1/2	80							-0.17, 1.28
	Cc	46 <sup>c</sup>		-0.10, 0.37			0.10, 0.32		0.00, 0.33
	Rcl	45					0.04, 0.67	0.04, 0.65	0.09, 0.66
	R1/4	66					-0.04, 0.67	0.16, 0.71	0.11, 0.68
	R3/4	76			0.04, 0.75				
	Rc	70					-0.23, 1.29	-0.26, 1.28	-0.05, 1.29
	A1/2	71					0.20, 1.13	0.03, 1.13	0.04, 1.12
	Ci	104	-0.09, 0.40	0.09, 0.38		-0.04, 0.38		-0.12, 0.40	
	Cco	16	-0.20, 0.93	-0.06, 0.71		-0.17, 0.91	0.08, 0.91	0.14, 0.93	0.15, 0.88
M <sub>2</sub>	Coc	39	0.06, 0.66			0.07, 0.67			
	C1/2	26					-0.03, 0.89		
	C3/4	25			0.07, 0.92			-0.04, 0.98	0.18, 0.98
	Cc	109							0.11, 1.17
	Ri	103						-0.24, 0.95	-0.38, 0.96
	Rcl	81						-0.23, 1.13	-0.37, 1.18
	R1/4	17	-0.20, 1.11	-0.23, 1.15			-0.13, 1.11	0.04, 1.11	0.16, 1.11
	R1/2	32		-0.05, 1.17			-0.09, 1.14	0.30, 1.13	0.23, 1.16
	R3/4	69		-0.03, 1.03			-0.03, 1.02		
	Rc	55			0.06, 1.37			0.11, 1.14	
A1/2	65						-0.20, 1.17		

<sup>a</sup> See Table 1 for abbreviations.

<sup>b</sup> N for A, Aa, N25a, and N25b I<sub>1</sub> stage R1/4 = 104.

<sup>c</sup> N for Aa first molar stage Cc = 27.

by selecting a uniform age distribution where variation between the extreme ages and mean age of the sample is similar. The two methods with a uniform age distribution, developed during this project, show this to be a useful approach.

The analytical method used to calculate the timing of age indicators of the reference sample is another impor-

tant characteristic, and estimating age in adults is more accurate if appropriate analytical methods are used (Konigsberg and Frankenberg, 2002; Kimmerle et al., 2008; Konigsberg et al., 2008; Prince and Konigsberg, 2008). Mean age of transition from one maturity event to the next, known as transition analysis (Milner et al., 2000; Boldsen et al., 2002; Konigsberg et al., 2008;



DiGangi et al., 2009), estimates age more accurately in adults than those based on reverse calibration. Calculating mean age of entry of a maturity event using probit or logistic regression and adapting this for prediction is similar in principle to transition analysis and in our study this adaptation improved measures of performance for the six methods we adapted.

Difficulties encountered when estimating age are 'attraction of the middle' and age mimicry (see Prince and Konigsberg, 2008). It is well documented that age is overestimated in younger individuals and underestimated in older individuals, whereas the middle age of the target sample shows little bias, a pattern also noted in our study. The age structure of the reference sample will influence age estimates, and age estimation mirrors the data upon which a method is based (Milner et al., 2000). This is known as age mimicry and occurs when inappropriate reference samples are used to estimate age (Bocquet-Appel and Masset, 1982). If the age distribution of the target sample differs from the reference sample (methods listed in Table 1), estimated age will be biased toward the reference sample. This may in part explain why the Moorrees method performs so badly in our study; it is one of the few radiographic studies from birth to age 25. A weakness in any study of age estimation is the target minimum age not being young enough. Yet, estimating age using Moorrees et al. (1963) on younger individuals from two known age-at-death collections also showed considerable bias (recalculated from Saunders et al., 1993; Liversidge, 1994). Mimicry cannot explain why the method of Moorrees performs badly for all age groups including older children in our target sample. Curiously, mean age for the first permanent molar from a selection of longitudinal radiographs from the same collection are consistently older (Gleiser and Hunt, 1955), and most fall within the 95% confidence interval of mean age of N25b (Liversidge, 2009).

The question of appropriate population-specific reference samples to estimate age in adults is important as regional differences in skeletal maturation have been demonstrated (Kimmerle et al., 2008; Konigsberg et al., 2008), but this has not been shown for mean age entering tooth stages. It is possible that subtle differences occur between our target sample and the reference methods. The average age entering permanent teeth stages were not significantly different between local white and Bangladeshi individuals aged 2–22 in London, United Kingdom (Liversidge, 2009)—a separate sample of radiographs to the target sample. This comparative study is part of a worldwide collaboration comparing maturation of permanent teeth from dental radiographs by the first author. Histological evidence shows only minor differences in molar crown duration in some world groups (Reid and Dean, 2006). Although little is documented for root formation, these differences are probably irrelevant to macroscopic crown and root stages.

There are several major challenges regarding the application of maturity data to estimate age. All of these relate to the inherent nature of maturation and the fact that it varies between individuals. A maturity event allows us to determine biological age and from this we infer chronological age. Herein lie two difficulties. The first is that a maturity event is a subjectively defined developmental stage in a continuum from cusp tip formation to apex maturation. In many cases, the formation stage is chosen as the one the tooth most closely resembles. Visual discrimination between formation stages is

improved by training and calibration but it remains subjective and a one stage difference can have a considerable impact on dental age. The second difficulty is that biological age differs to known age. Dental age is not the same for all children of a specific known age. For instance, a dental age of seven assumes the individual to be an average seven-year-old but that child could be a dentally advanced six-year-old or a dentally delayed eight-year-old and a confidence interval of estimated age is probably more appropriate.

## CONCLUSIONS

Which method can most accurately estimate age? If seven developing teeth are available, the dental maturity scores from Willems et al. (2001) is the method of choice. If less than seven teeth are available, tooth stages from Table 7 (Demirjian stages) and Table 8 (Moorrees stages) can be chosen to estimate age with little or no bias and similar levels of reliability. These methods include some stages from Haavikko (1970), Anderson et al. (1976), Smith (1991), Liversidge et al. (2006), Nyström et al. (2007), and Liversidge (2010, in prep). The best individual tooth was  $M_2$  using L10a (Liversidge, 2010) where Demirjian stages A to G estimated age with bias not significant to zero. What is the best measure of performance? Low bias is the most useful criterion to select the best method and methods with low bias performed well in other measures of accuracy.

Bias was not consistent across age groups but all methods considerably under-estimated age in 14 to 16-year-olds. Reliability/precision of age estimation was poor with a 95% confidence interval from  $\pm 0.65$  for early tooth stages to  $\pm 2.59$  years for late tooth stages. Methods for individuals of unknown sex are available from pooled sex data for Demirjian stages (Liversidge, 2009) and Moorrees stages (Willems et al., 2010). Methods that provide mean age entering tooth stages performed poorly and adjusting these methods for age prediction improved performance. Performance of methods based on mean age within stage was improved by using a uniform age distribution.

## LITERATURE CITED

- Anderson DL, Thompson GW Popovich F. 1976. Age attainment of mineralization stages of the dentition. *J Forensic Sci* 21:191–200.
- Bocquet-Appel JP, Masset C. 1982. Farewell to paleodemography. *J Hum Evol* 11:321–333.
- Boldsen JL, Milner GR, Konigsberg LW, Wood JW. 2002. Transition analysis: a new method for estimating age from skeletons. In: Hoppa RD, Vaupal JW, editors. *Paleodemography age distributions from skeletal samples*. Cambridge: Cambridge University Press. p 73–106.
- Braga J, Heuze Y, Chabadel O, Sonan NK, Gueramy A. 2005. Non-adult dental age assessment: correspondence analysis and linear regression versus Bayesian predictions. *Int J Legal Med* 119:260–274.
- Chaillet N, Nyström M, Demirjian A. 2005. Comparison of dental maturity in children of different ethnic origins: international maturity curves for clinicians. *J Forensic Sci* 50:1164–1174.
- Demirjian A. 1994. *Dental development*. CD-ROM, Silver Platter Education, University of Montreal, Montreal, 1993–1994.
- Demirjian A, Goldstein H. 1976. New systems for dental maturity based on seven and four teeth. *Ann Hum Biol* 3:411–427.

- Demirjian A, Goldstein H, Tanner JM. 1973. A new system of dental age assessment. *Hum Biol* 45:211–227.
- DiGangi EA, Bethard JD, Kimmerle EH, Konigsberg LW. A new method for estimating age-at-death from the first rib. *Am J Phys Anthropol* 138:164–176.
- Ferranti L, Cameriere R. 2009. Statistical methods to assess the reliability of measurements in the procedures for forensic age estimation. *Int J Legal Med* 123:277–283.
- Gleiser I, Hunt EE. 1955. The permanent mandibular first molar: its calcification, eruption and decay. *Am J Phys Anthropol* 13:253–283.
- Goldstein H. 1979. The design and analysis of longitudinal studies. London: Academic Press.
- Haavikko K. 1970. The formation and the alveolar and clinical eruption of the permanent teeth. An orthopantomographic study. *Proc Finn Dent Soc* 66:103–170.
- Hägg U, Matsson L. 1985. Dental maturity as an indicator of chronological age: the accuracy and precision of three methods. *Eur J Orthod* 7:25–34.
- Heuzé Y, Cardoso HF. 2008. Testing the quality of nonadult Bayesian dental age assessment methods to juvenile skeletal remains: the Lisbon collection children and secular trend effects. *Am J Phys Anthropol* 135:275–283.
- Hoppa RD, Vaupel JW. 2002. Paleodemography age distributions from skeletal samples. Cambridge: Cambridge University Press.
- Kimmerle EH, Konigsberg LW, Jantz RL, Baraybar JP. 2008. Analysis of age-at-death estimation through the use of pubic symphyseal data. *J Forensic Sci* 53:558–568.
- Konigsberg LW, Frankenberg SR. 1992. Estimation of age structure in anthropological demography. *Am J Phys Anthropol* 89:235–256.
- Konigsberg LW, Frankenberg SR. 2002. Deconstructing death in paleodemography. *Am J Phys Anthropol* 117:297–309.
- Konigsberg LW, Herrmann NP, Westcott DJ, Kimmerle EH. 2008. Estimation and evidence in forensic anthropology: age-at-death. *J Forensic Sci* 53:541–557.
- Liversidge HM. 1994. Accuracy of age estimation from developing teeth of a population of known age (0–5.4 years). *Int J Osteoarch* 4:37–45.
- Liversidge HM. 2009. Permanent tooth formation as a method of estimating age. *Front Oral Biol* 13:153–157.
- Liversidge HM. 2010. Demirjian stage tooth formation results from a large group of children. *Dent Anthropol* 23:16–23.
- Liversidge HM, Chaillet N, Mörnstad H, Nyström M, Rowlings K, Taylor J, Willems G. 2006. Timing of Demirjian tooth formation stages. *Ann Hum Biol* 33:454–470.
- Lovejoy CO, Meindl RS, Mensforth RP, Barton TJ. 1985. Multifactorial determination of skeletal age at death: a method and blind tests of its accuracy. *Am J Phys Anthropol* 68:1–14.
- Maber M, Liversidge HM, Hector MP. 2006. Accuracy of age estimation of radiographic methods using developing teeth. *Forensic Sci Int* 159:S68–S73.
- Milner GR, Wood JW, Boldsen JL. 2000. Paleodemography. In: Katzenberg AM, Saunders SR, editors. *Biological anthropology of the human skeleton*. New York: Wiley-Liss. p 467–497.
- Moorrees CFA, Fanning EA, Hunt EE. 1963. Age variation of formation stages for ten permanent teeth. *J Dent Res* 42:1490–1502.
- Mörnstad H, Reventlid M, Teivens A. 1995. The validity of four methods for age determination by teeth in Swedish children: a multicentre study. *Swed Dent J* 19:121–130.
- Nolla CM. 1960. The development of the permanent teeth. *J Dent Child* 27:254–266.
- Nyström M, Ranta HM, Peltola S, Kataja JM. 2007. Timing of developmental stages in permanent mandibular teeth of Finns from birth to age 25. *Acta Odont Scand* 65:36–43.
- Prince DA, Konigsberg LW. 2008. New formulae for estimating age-at-death in the Balkans utilizing Lamendin's dental technique and Bayesian analysis. *J Forensic Sci* 53:578–587.
- Rai B. 2008. The evaluation of two radiographic methods for age determination of children in an Indian population. *J Forensic Odontostomatol* 27:30–33.
- Rai B, Anand SC. 2006. Tooth developments: an accuracy of age estimation of radiographic methods. *World J Med Sci* 1:130–132.
- Reid DJ, Dean MC. 2006. Variation in modern human enamel formation times. *J Hum Evol* 50:329–346.
- Ritz-Timme S, Cattaneo D, Collins MJ, Waite ER, Schütz HW, Kaatsch HJ, Borrman HI. 2000. Age estimation: the state of the art in relation to the specific demands of forensic practice. *Int J Legal Med* 113:129–136.
- Saunders S, DeVito C, Herring A, Southern R, Hoppa R. 1993. Accuracy tests of tooth formation age estimations from human skeletal remains. *Am J Phys Anthropol* 92:173–188.
- Smith BH. 1991. Standards of human tooth formation and dental age assessment. In: Kelly MA, Larsen CS, editors. *Advances in dental anthropology*. New York: Wiley-Liss. p 143–168.
- Staaf V, Mörnstad H, Welander U. 1991. Age estimation based on tooth development: a test of reliability and validity. *Scand J Dent Res* 99:281–286.
- Steadman DW, Adams BJ, Konigsberg LW. 2006. Statistical bases for positive identification in forensic anthropology. *Am J Phys Anthropol* 131:27–32.
- Willems G, Thevissen PW, Belmans A, Liversidge HM. 2010. Willems II. Non-gender-specific dental maturity scores. *Forensic Sci Int* doi: 10.1016/j.forsciint.2010.04.033.
- Willems G, Van Olmen A, Spiessens B, Carels C. 2001. Dental age estimation in Belgian children: Demirjian's technique revisited. *J Forensic Sci* 46:893–895.