

Received:
30 October 2016
Revised:
25 November 2016
Accepted:
9 December 2016

Heliyon 2 (2016) e00216



Northern Chinese dental ages estimated from southern Chinese reference datasets closely correlate with chronological age

Hai Ming Wong^{a,*}, Yi Feng Wen^a, Jayakumar Jayaraman^b, Jing Li^c, Ling Sun^a, Nigel Martyn King^d, Graham J. Roberts^e

^a Paediatric Dentistry, Faculty of Dentistry, The University of Hong Kong, Prince Philip Dental Hospital, 34 Hospital Road, Sai Ying Pun, Hong Kong

^b School of Dentistry, International Medical University, Kuala Lumpur, Malaysia

^c Department of Pediatric Dentistry, Peking University School and Hospital of Stomatology, Beijing, China

^d Paediatric Dentistry, School of Dentistry, University of Western Australia, Oral Health Centre of WA, Australia

^e Paediatric Dentistry, King's College London Dental Institute, London, United Kingdom

* Corresponding author at: 2/F, Prince Philip Dental Hospital, 34 Hospital Road, Hong Kong.

E-mail addresses: wonghmg@hku.hk (H.M. Wong), pyfwen@hku.hk (Y.F. Wen), jayakumar83@hotmail.com (J. Jayaraman), lijing975@sina.com (J. Li), theuser98@163.com (L. Sun), nigel.king@uwa.edu.au (N.M. King), graham.j.roberts@kcl.ac.uk (G.J. Roberts).

Abstract

While northern and southern Chinese are genetically correlated, there exists notable environmental differences in their living conditions. This study aimed to evaluate validity of the southern Chinese reference dataset for dental age estimation applied to northern Chinese. Dental panoramic tomographs of 437 northern Chinese aged 3 to 21 years were analysed. All the left maxillary and mandibular permanent teeth plus the 2 third molars on the right side were scored based on Demirjian's classification of tooth development stages. Mean and standard error of dental age were obtained for each tooth development stage, followed by random effect meta-analysis for mean dental age estimation. Validity of the method was examined through measures of agreement (95% limits of agreement, standard error of measurement, and Lin's concordance correlation

coefficient) and measure of reliability (Intraclass correlation coefficient). On average, the estimated dental age overestimated chronological age by only around 1 month in both females and males. The Intraclass correlation coefficient values were 0.99 for both sexes, suggesting excellent reliability of the method. Reference dataset for dental age estimation developed on the basis of southern Chinese was applicable for use among the northern Chinese.

Keywords: Dentistry, Anatomy, Evidence-based medicine, Medical imaging, Pediatrics

1. Introduction

Many physical indicators can be used for age evidence if there is no documentation to confirm a person's chronological age. Such indicators include skeletal maturation, height/weight measurement, and emergence of secondary sex characteristics, tooth wear, and dental development [1, 2, 3, 4]. Comparisons have been made among these methods and the conclusions revealed that dental development correlates more closely with chronological age for children, adolescents and emerging adults as it is less affected by environmental factors [5, 6, 7]. Among the several methods of dental age estimation, the radiographic approach is simple, non-invasive, and reproducible [8]. The reason why Tooth Development Stages (TDSs) can be used for age estimation is because the timing of all stages of tooth development follows a sequential and organised pattern, from the initial calcification of enamel until complete root formation. TDSs have been extensively studied since 1960s and several methods of dental age estimation such as the Demirjian's method [9, 10], Nolla's method [11], and Cameriere's method [12, 13] have been developed. In particular, the Demirjian's method [9] proved to be effective and reliable [14, 15]. Following Demirjian's method, different datasets for dental age estimation have been set up for different ethnic groups, for example, the French-Canadian dataset [9], the UK dataset of Caucasians, and the Afro-Trinidadian dataset [16]. To facilitate the development of these reference datasets, Dental Age Research London Information Group (DARLInG) has established a universal database for dental age estimation. In conjunction with DARLInG, a whole reference dataset for southern Chinese population has been built up and validated by Jayaraman and his colleagues [17]. Considering the wide applicability of the Demirjian's method and the availability of existing reference dataset for southern Chinese, the Demirjian's method was used for dental age estimation in the present study.

China is a unified multi-ethnic country located in eastern Asia with the world's third largest territory and the largest population. Based on "Out-of-Africa" hypothesis, the first entry of modern humans into the southern part of East Asia occurred about 18,000–60,000 years ago from Africa, followed by a great

northward migration extended into northern China and Siberia. Therefore, southern populations in eastern Asia are much more polymorphic than their northern counterparts [18, 19]. In China, an apparent genetic difference can be observed between northern and southern residents [20]. The differences in distinctive patterns of Y chromosome [21], mitochondrial DNA (mtDNA) polymorphism [22], 14-bp insertion/deletion frequencies in HLA-G, allele and haplotype frequencies of HLA-G and HLA-A loci [23], 13 contiguous single nucleotide polymorphisms (SNPs), zinc finger protein (ZNP) 510 and ZNP782 [24] have already been detected by researchers.

Notable environmental differences exist between southern and northern China which can be geographically divided into the northern and southern area by the line of the Huai River-Qin Mountains. Southern China has a warmer and more humid subtropical climate than northern China's temperate climate. The dietary pattern also varies significantly between the southern and the northern Chinese that closely adheres to the climatic conditions. The northern Chinese consume more wheat based diet compared to southern whose staple diet is mainly rice. Such genetic and environmental differences have led to distinctive physical and dentofacial [25] characteristics in southern and northern Chinese. According to general phenotypes, northern Chinese may have taller, broader body types, heavier body weight, smaller eyes, fairer skin, and longer faces; while southern Chinese may have shorter, narrower body types, lighter body weight, a darker complexion, rounder faces, bigger eyes with double eyelids, and shorter necks and limbs. It has been claimed that due to a warmer temperature, southern Chinese might have higher metabolism rates, and thus an earlier but shorter pubertal growth period [26]. According to a national survey, the average male height of northern China was 175.31 ± 6.09 cm in Shandong Province and 174.97 ± 6.76 cm in Beijing, while the average male height may range between 167.51 ± 5.84 (Guizhou province) and 172.42 ± 5.19 cm (Fujian province) in southern China [27]. Variations in skeletal development between the northern and southern Chinese also initiate a question of such existence in dental development.

Several differences in phenotypic expressions between northern and southern Chinese have been established; however, the pattern of dental development between these populations has not been explored. Hence, it is uncertain whether the dataset for dental age estimation of the southern Chinese is applicable to the northern Chinese population. Dental age estimation is important because it is frequently used for anthropological and archaeological studies, and for civil, legal, criminal and forensic practices. The method could be possibly employed to estimate children of northern Chinese origin whose births had not been registered. Using retrospectively collected panoramic radiographs, the purpose of the present study was to evaluate validity of the southern Chinese reference dataset for

radiological dental age estimation according to Demirjian et al. [9] applied to a northern Chinese population.

2. Materials and methods

2.1. Sampling

The protocol for this study was approved by the Institutional Review Board of the University of Hong Kong – West cluster Hospital authority (IRB No. UW 12-280). The sample of northern Chinese was selected from the most recent patient list in the archives of the School and Hospital of Stomatology, Peking University using a systematic random sampling method. Dental panoramic radiographs taken previously for diagnostic purposes were re-used in this study. The inclusion criteria were:

- chronological age of participants aged 3 to 21 years
- both parents of northern Chinese origin of Han ethnicity
- both date of birth (DOB) and date of radiograph (DOR) available

The exclusion criteria were participants with:

- systemic diseases or syndromes that would affect skeletal and dental growth
- localised oral pathology, anomalies or impacted teeth that would affect dental growth
- severe malocclusion
- present or history of orthodontic treatment
- DPT of poor quality in which one or more targeted teeth cannot be scored

The sample was grouped into 18 age levels by a yearly interval. Similar to a previous study [13], we required least 10 participants for each age by sex group.

2.2. Data collection

Participants' panoramic radiographs and their personal information related to chronological age (Date of birth, DOB and Date of radiograph, DOR) were collected from the existing records. Each panoramic radiograph was assigned a code, scanned at a resolution of 300 dpi in grey-scale format, and stored as a JPEG image with dimensions of 2440 × 1280 pixels (Epson scanner 1000XL, Epson Inc, USA). Chronological ages of the participants were calculated by subtracting DOB from DOR and were stored in unit of years with two decimal places. The digitised panoramic radiographs and the participants' chronological ages collected from Beijing were brought to the Hong Kong centre for scoring and analysis.

2.3. Scoring of the radiographs

All of the scanned panoramic radiographs were scored independently and randomly (using the random numbers generated electronically) by two trained and calibrated examiners who were blinded to the chronological age (CA) and sex of each participant. The digitised panoramic radiograph was viewed on a widescreen monitor (Philips 201E, Philips Electronics, Taiwan). Microsoft Office Picture Manager (Microsoft Corp, USA) was used and when required the panoramic radiographs were magnified up to 2x times for identification of the dental development stages. All the left maxillary and mandibular permanent teeth plus the 2 third molars on the right side (total 18 teeth per participant) were scored based on Demirjian's classification of tooth developmental stages [9]. The classification system has eight stages starting from the initial calcification (Stage A) up to root completion (Stage H), see Fig. 1. When a tooth on the left side was missing or difficult to read, the contralateral tooth was assessed. A Microsoft Access (Microsoft Corp, USA) database was used for data entry. Ten percent of the panoramic radiographs were randomly selected using the random numbers generated electronically, and the tooth developmental stages were re-evaluated two weeks later to test the inter-examiner and intra-examiner reliabilities.

2.4. Dental age estimation

For each panoramic radiograph, mean dental age (DA) and the corresponding standard deviation and standard error were obtained from the established southern Chinese dataset [17]. A tooth at stage H, which denotes a closed apex, was excluded from data analysis, as it was impossible to identify the time of root completion from a single snapshot radiograph [28]. Using the meta-analysis routine of STATA statistical package (StataCorp, Version 13.1, Texas, USA), the dental age of each participant was calculated using the standard error as the weighting factor for the random effects model [29]. The final output of the meta-analysis computation was generated as forest plots (see Table 1 and Fig. 2 for an example of one subject in this study). The estimated mean dental age, together with its corresponding upper and lower 95% confidence limits, was used as DA in this study.

Developmental stage of each tooth is indexed by four characters, where the first distinguishes upper (U) from lower (L) dentition, the second distinguishes left (L) from right (R) side, in the third place is a number indicating location of a tooth in each quadrant (1 to 8 represent tooth from central incisor to third molar), which is followed by the scored developmental stage of the tooth according to the Demirjian's method.

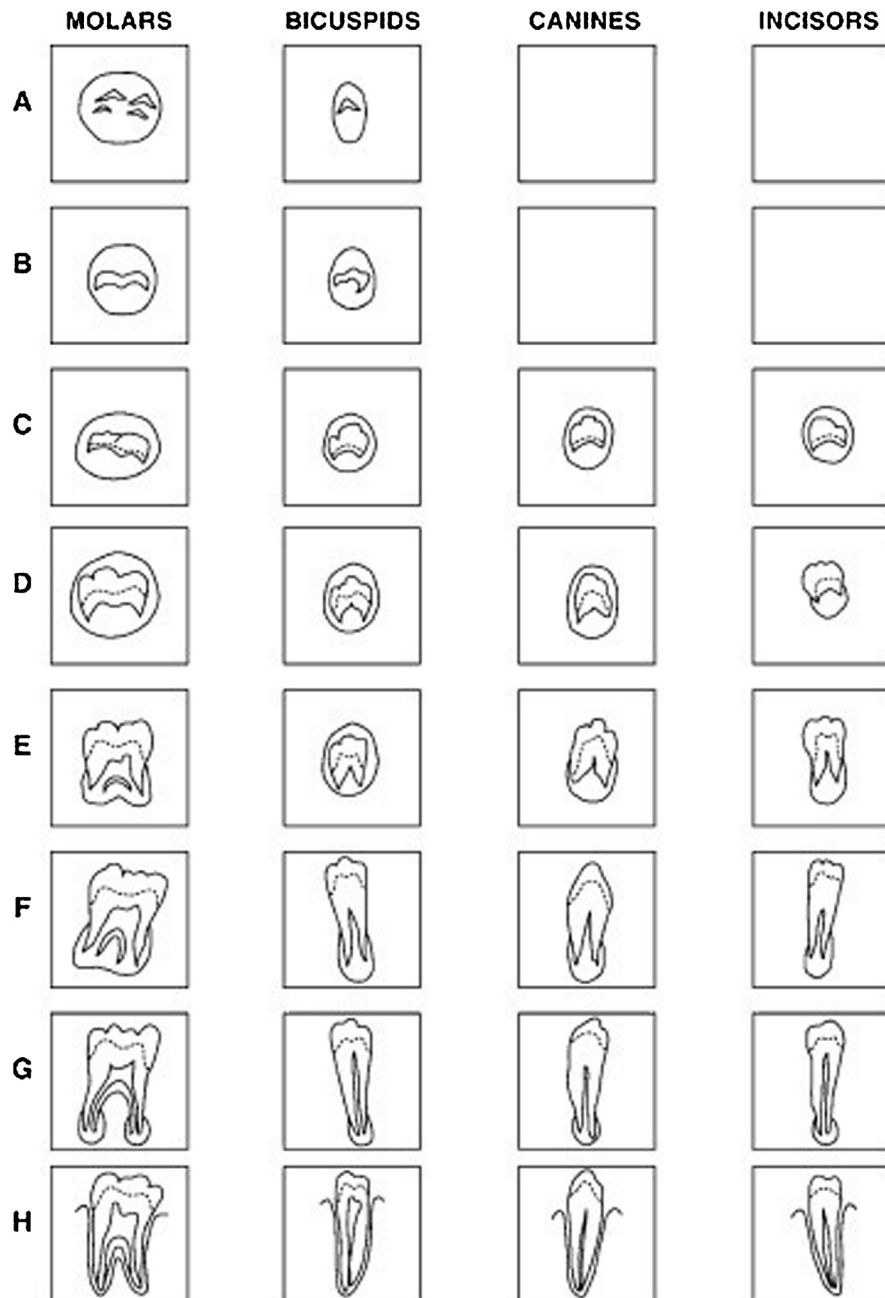


Fig. 1. The tooth development stages described by Demirjian et al. [9].

2.5. Statistical analysis

Inter- and intra-examiner agreement in assessment of tooth developmental stages were estimated using Cohen's Kappa and the coefficients were interpreted according to the classification scheme proposed by Landis and Koch [30].

Table 1. An example of the data entered into STATA software for a female subject (CA = 10.23 years). The corresponding forest plots of meta-analysis of dental age estimation is shown in Fig. 2.

TDS	DA for TDS	se
UL1H	–	–
UL2G	10.12	0.179
UL3F	9.09	0.116
UL4F	10.18	0.142
UL5F	10.56	0.135
UL6H	–	–
UL7F	11.69	0.143
LL1H	–	–
LL2H	–	–
LL3F	8.87	0.096
LL4F	10.21	0.136
LL5F	10.97	0.119
LL6H	–	–
LL7F	11.65	0.128
LL8B	11.28	0.213
LR8A	9.97	0.163

TDS, tooth development stage; DA for TDS: mean dental age for each TDS assigned by the southern Chinese dataset; se, standard error.

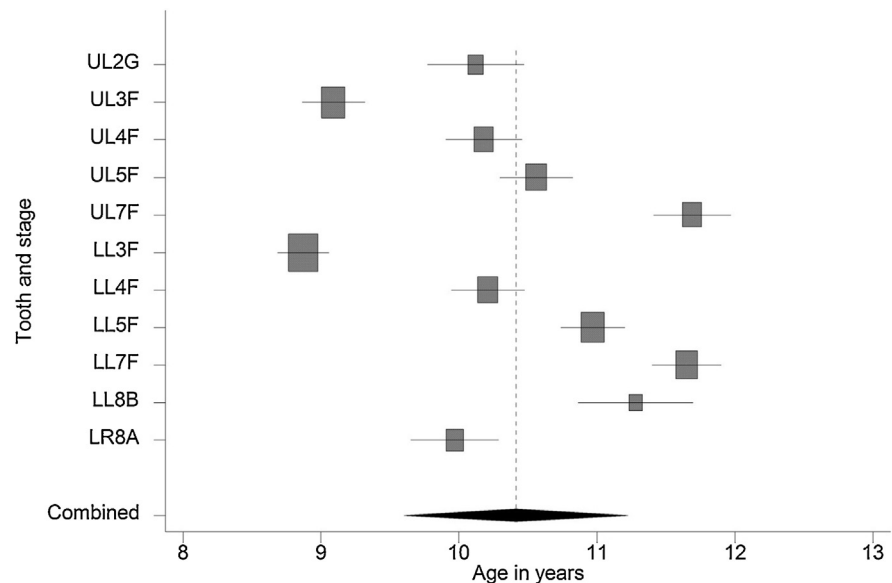


Fig. 2. Forest plot generated from meta-analysis weighed by standard error of each TDS for a female subject (the same subject for Table 1, CA = 10.23 years). The DA calculated from this analysis was 10.41 years.

Descriptive statistics for chronological age (CA) and dental age (DA) were given as means and standard deviations for each age by sex group as well as for all age groups pooled by sex. For the purpose of validation of southern Chinese reference dataset, we distinguished measures of agreement from measures of reliability. A measure of agreement quantifies the absolute magnitude of difference between CA and DA in unit of years [31]. It reflects the ability of the two age assessment methods to yield the same value. Measures of reliability, on the other hand, are relative measures of variability among participants in relation to measurement error [31]. They provide insight into the ability of the proposed method to differentiate participants of different ages. Parameters of agreement and reliability are both estimated in this study in that it has been noted that dependence solely on one set of parameters might lead to wrong conclusions [32].

2.5.1. Level of Agreement

The 95% limits of agreement were calculated as mean difference between chronological age (CA) and dental age (DA) $\pm 1.96 X$ standard deviation of the differences as proposed by Bland and Altman [33]. This was carried out separately for each age by sex as well as for all age groups pooled by sex. The 95% limits of agreement represent the range of values that 95% of the difference between CA and DA would fall into. Interpretation of the 95% limits of agreement should be focused on the lower or upper bound of this interval, whichever is larger in absolute magnitude, to determine if difference up to that magnitude is clinically acceptable. Bland-Altman plots [33] were obtained for both sexes by plotting the difference between CA and DA against their average value for each participant.

Standard error of measurement (SEM) has been advocated as a suitable measure of agreement [32]. This measure was calculated as the square root of error variance and was also reported in this study. Lin's concordance correlation coefficient, a bias corrected version of Pearson correlation coefficient that determines the degree of deviation from the 45° line of concordance, was also calculated [34]. In addition, CA was plotted against DA for each participant separately for females and males.

2.5.2. Reliability testing

Intra-class correlation coefficient (ICC) was used as a measure of reliability. Two-way random effects model for single measurement, corresponding to ICC (2,1) of Shrout and Fleiss' six types of ICC [35], was applied to data for each sex. Values of ICC represent the proportion of variability in participants' estimated ages attributable to real chronological age differences. All measures of agreement and reliability were calculated in R version 3.3.0 [36]. Estimation of

Cohen's Kappa and ICC were performed using the psych package version 1.6.4 [37] in R.

3. Results

3.1. Participants

A total of 437 participants were recruited, among which 240 were females and 197 were males. Chronological age of the participants ranged from 3 to 21 years for females and 3 to 20 years for males (Table 2 and Table 3). Inter- and intra-examiner agreements were tested on 50 panoramic radiographs with 774 Tooth Developmental Stages (TDSs). The kappa values were 0.866 (almost perfect) and 0.854 (almost perfect) for intra-examiner agreement and were 0.802 for inter-examiner agreement (almost perfect). All discrepancies were within one TDS.

Table 2. Descriptive statistics (mean values and corresponding standard deviations) and measures of agreement by age level for females between the southern and northern Chinese.

Age range	n	CA	DA	CA-DA	LoA	SEM
3.00–3.99	20	3.63 (0.31)	3.94 (0.29)	−0.31	(−0.76, 0.15)	0.27
4.00–4.99	23	4.32 (0.27)	4.37 (0.44)	−0.05	(−0.80, 0.71)	0.27
5.00–5.99	10	5.30 (0.22)	5.08 (0.36)	0.22	(−0.24, 0.69)	0.22
6.00–6.99	16	6.51 (0.36)	6.65 (0.61)	−0.14	(−1.15, 0.86)	0.36
7.00–7.99	11	7.37 (0.31)	7.39 (0.37)	−0.02	(−0.80, 0.76)	0.27
8.00–8.99	12	8.44 (0.23)	8.94 (0.94)	−0.50	(−2.17, 1.17)	0.68
9.00–9.99	12	9.45 (0.33)	9.66 (1.06)	−0.21	(−2.01, 1.59)	0.64
10.00–10.99	12	10.44 (0.29)	11.12 (1.10)	−0.68	(−2.39, 1.04)	0.76
11.00–11.99	15	11.55 (0.27)	12.23 (0.63)	−0.78	(−1.86, 0.50)	0.63
12.00–12.99	18	12.46 (0.28)	12.51 (0.40)	−0.05	(−1.01, 0.91)	0.34
13.00–13.99	10	13.40 (0.29)	13.29 (0.55)	0.11	(−0.79, 1.01)	0.32
14.00–14.99	15	14.54 (0.29)	14.43 (0.66)	0.11	(−1.10, 1.32)	0.43
15.00–15.99	11	15.46 (0.30)	15.37 (1.12)	0.09	(−1.83, 2.02)	0.66
16.00–16.99	11	16.47 (0.34)	16.70 (1.26)	−0.23	(−2.12, 1.65)	0.67
17.00–17.99	11	17.32 (0.28)	17.08 (1.03)	0.24	(−1.54, 2.02)	0.64
18.00–18.99	10	18.33 (0.26)	18.36 (1.34)	−0.03	(−2.42, 2.37)	0.82
19.00–19.99	10	19.15 (0.11)	18.36 (1.24)	0.79	(−1.65, 3.23)	1.01
20.00–20.99	13	20.34 (0.30)	19.61 (0.50)	0.73	(−0.51, 1.99)	0.68
Overall	240	11.17 (5.27)	11.23 (5.14)	−0.06	(−1.58, 1.47)	0.55

n, number of participants in the corresponding age group; CA, chronological age in years; DA, dental age in years; LoA, limits of agreement; SEM, standard error of mean.

Table 3. Descriptive statistics (mean values and corresponding standard deviations) and measures of agreement by age level for males between the southern and northern Chinese.

Age range	n	CA	DA	CA-DA	LoA	SEM
3.00–3.99	14	3.72 (0.21)	3.91 (0.28)	−0.19	(−0.74, 0.36)	0.24
4.00–4.99	15	4.41 (0.30)	4.38 (0.29)	0.03	(−0.77, 0.83)	0.28
5.00–5.99	12	5.47 (0.36)	5.57 (0.55)	−0.10	(−0.90, 0.71)	0.29
6.00–6.99	10	6.41 (0.22)	6.48 (0.63)	−0.07	(−1.23, 1.09)	0.40
7.00–7.99	13	7.49 (0.34)	7.47 (0.68)	0.02	(−1.09, 1.14)	0.39
8.00–8.99	10	8.41 (0.29)	8.54 (0.55)	−0.13	(−1.35, 1.10)	0.43
9.00–9.99	14	9.43 (0.33)	9.53 (0.72)	−0.10	(−1.46, 1.26)	0.48
10.00–10.99	10	10.35 (0.22)	10.82 (0.42)	−0.47	(−1.38, 0.43)	0.46
11.00–11.99	10	11.38 (0.19)	11.34 (0.66)	0.04	(−1.34, 1.42)	0.47
12.00–12.99	16	12.45 (0.30)	12.52 (0.67)	−0.07	(−1.43, 1.29)	0.48
13.00–13.99	10	13.30 (0.20)	13.27 (0.34)	0.03	(−0.56, 0.62)	0.20
14.00–14.99	13	14.42 (0.30)	14.32 (0.87)	0.10	(−1.24, 1.45)	0.47
15.00–15.99	10	15.35 (0.29)	15.83 (1.41)	−0.48	(−2.96, 1.99)	0.91
16.00–16.99	10	16.33 (0.27)	16.88 (1.02)	−0.55	(−2.49, 1.39)	0.77
17.00–17.99	10	17.54 (0.31)	17.31 (0.87)	0.23	(−1.51, 1.97)	0.62
18.00–18.99	10	18.39 (0.31)	18.62 (1.28)	−0.23	(−2.84, 2.39)	0.91
19.00–19.99	10	19.45 (0.28)	19.41 (0.40)	0.04	(−0.61, 0.70)	0.23
Overall	197	11.01 (4.87)	11.11 (4.93)	−0.10	(−1.48, 1.28)	0.50

n, number of participants in the corresponding age group; CA, chronological age in years; DA, dental age in years; LoA, limits of agreement; SEM, standard error of mean.

3.2. Level of agreement

Table 2 and Table 3 summarises descriptive statistics for chronological age (CA) and dental age (DA) and measures of agreement for females and males, respectively. Mean CA were 11.17 years for females and were 11.01 years for males. Lower and upper bound of the 95% limits of agreement for the age groups were mostly within 2 years. Using data pooled across age groups, Bland-Altman plot (Fig. 3a and b) showed that the differences between CA and DA were randomly distributed around the mean difference of -0.06 years (SD 0.78 years) in females and -0.10 years (SD 0.70 years) in males. However, a trend was noted that points representing participants of older ages appeared to be more scattered than those representing younger participants. A similar trend was also observed when CA was plotted against DA (Fig. 4a and b). The 95% limits of agreement were also illustrated on the plot. Approximately 95% of the

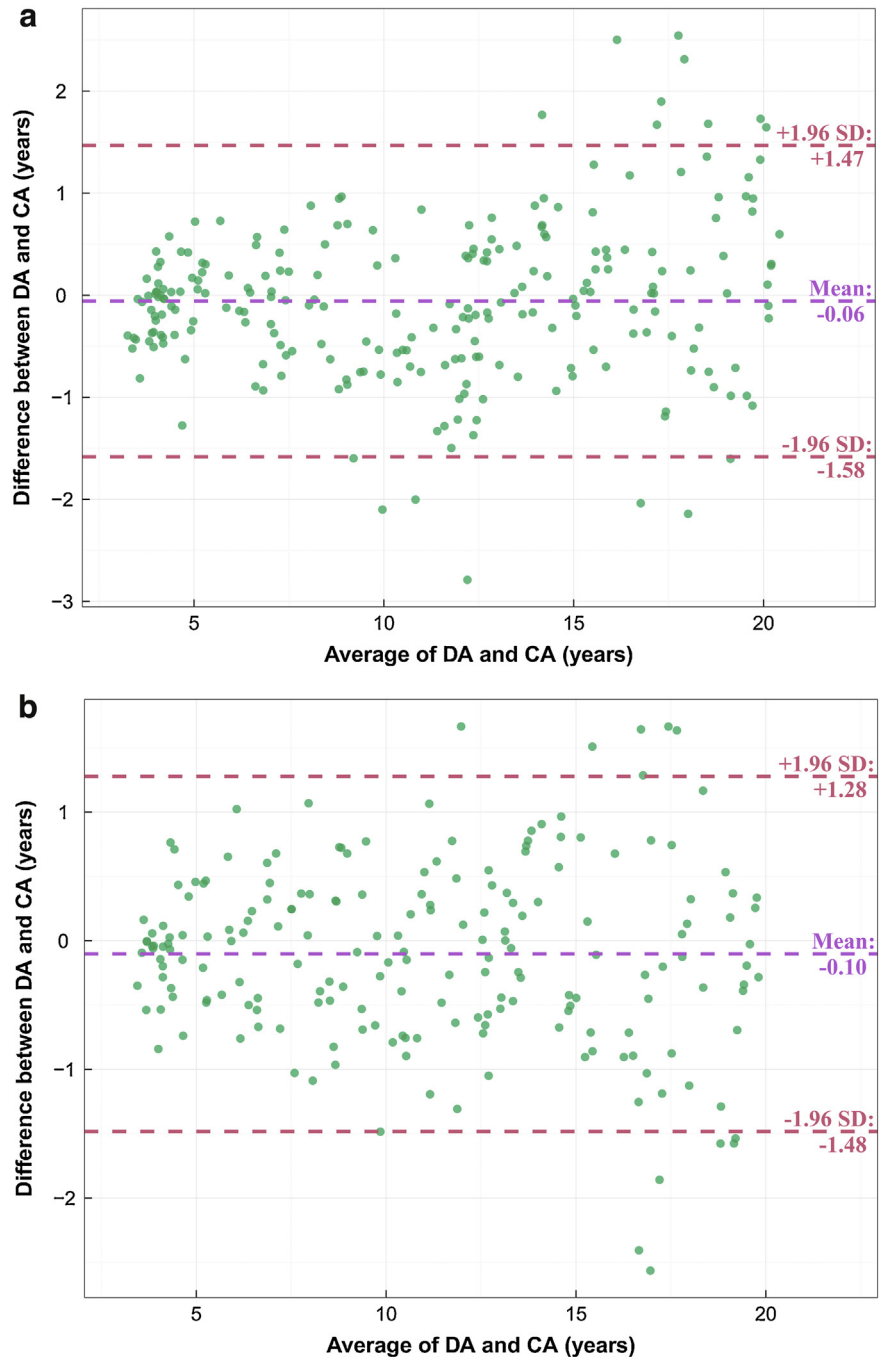


Fig. 3. (a) Bland-Altman plot for Chronological Age (CA) and Dental Age (DA) estimated from the southern Chinese Reference Data Set in females. (b) Bland-Altman plot for Chronological Age (CA) and Dental Age (DA) estimated from the southern Chinese Reference Data Set in males.

discrepancies between CA and DA were expected to lie within -1.58 years to 1.47 years for females and -1.48 years to 1.28 years for males. For both sexes, most of the discrepancies were within 1 year and more than half were within 0.5 years.

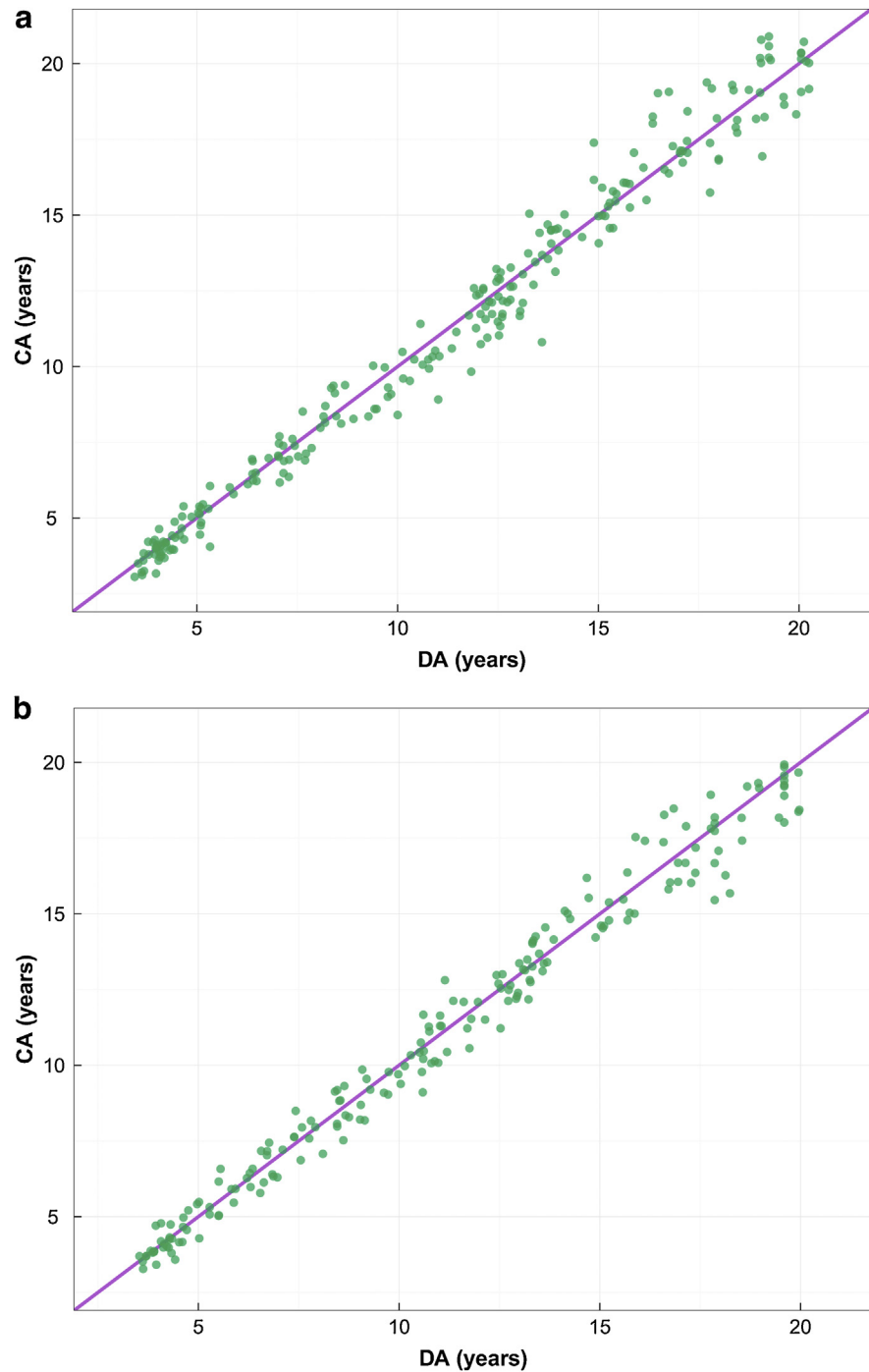


Fig. 4. (a) Plot of Chronological Age (CA) against Dental Age (DA) estimated from southern Chinese Reference Data Set in females. (b) Plot of Chronological Age (CA) against Dental Age (DA) estimated from southern Chinese Reference Data Set in males.

Except for the female 19.00–19.99 age group that showed a standard error of measurement (SEM) of 1.01 years, SEM for all other female and male age groups were smaller than 1 year. Lin's concordance correlation coefficient was 0.989 for females (95% CI 0.986–0.991) and 0.989 for males (95% CI 0.986–0.992), both of which were very close to the upper limit of 1.0 when there is perfect concordance.

3.3. Reliability scores

The intraclass correlation coefficient (ICC) values were as high as 0.99 for both sexes, suggesting excellent reliability according to Cicchetti et al.'s [38] guideline of cut-offs for ICC. Thus almost all variability in the assessed dental age reflected genuine chronological age differences among the participants, while measurement error accounted for only 1% of the total variability.

4. Discussion

China had a population of 1.37 billion by the end of 2014, which constitutes about one-fifth of the world's population [39]. The issue of unregistered children has been noted with its huge number and characters, such as descendants of floating population, unplanned births, female children, or adopted children [40]. Age assessment is necessary in many situations for people without birth registrations, such as children applying for social welfare and benefits, getting access to health care, education, or even employment, marriage registration, young criminals claiming to be underage, and so on. China is undergoing huge economic reformation, and establishments of special economic zones in southern part of China had seen high influx of economic migrants in this region. It is mandatory for the job seekers to provide documentary evidence of their age before employment and when absent, age estimation is conducted on these individuals. Because dental age estimation is a reliable method for the age estimation of children and emerging adults [29], it is recommended in most circumstances. It is noteworthy that different ethnic groups may have different dental development characteristics [16, 41]; therefore, the establishment of ethnic-specified reference datasets is necessary.

Noting that the UK and French-Canadian reference dataset for dental age estimation were inapplicable to the southern Chinese children [42, 43, 44], specific dataset for the southern Chinese was established [17]. This dataset comprised a sample of 1,123 females and 1,183 males aged from 2 to 25 years. The established data were validated on southern Chinese using 484 participants (229 males and 255 females) from 2 to 25 years of age with an average of 10 participants in each age/sex group [17]. Genetic and environmental differences between southern and northern China has led to many differences between the southern and northern population, including body stature, facial characteristics, and body fat ratio [45]. On the other hand, the economic reform in China in recent decades has also

fostered population transportation between northern and southern China. Both interbreeding and migration have actively narrowed the differences between northern and southern residents. Thus, it remains uncertain as to whether the southern Chinese dataset is applicable to northern Chinese. Such uncertainties make it necessary to validate the dataset developed for southern Chinese to northern Chinese before it can be used on northern Chinese.

The teeth on the left and right sides usually develop symmetrically; therefore, only the teeth on the left side were included in the dental age estimation, except for the third molars. Both left and right third molars were included in the assessment because third molar is the only tooth type still developing in young adults and their prevalence of hypodontia is higher. Teeth in stage H were not assigned any dental age because it is hard to determine the actual age of apex closure. This is also the reason why dental age estimation can only be performed on developing dentitions.

Close examination of the Bland-Altman plots revealed that points representing participants of older ages appeared more scattered around the line of mean difference than points representing younger participants. In line with this, Birchler et al. [46] found that dental age was estimated with higher level of precision for younger participants. Tomás et al. [47] reported that the predictive capacity of Demirjian decreased with age and the method was unable to predict age in adults. Similarly, Urzel and Bruzek [48] cautioned age estimation for individuals aged over 14 years. Possible explanation for this may be that there are more developing teeth in younger individuals available for scoring, which allows age of younger participants to be better estimated. In addition, higher degree of variability in tooth development has been noted among older individuals, which may also account for the larger discrepancy between their chronological age and estimated age [49, 50]. To improve accuracy of age estimation for older individuals, we included both upper and lower left permanent teeth as well as 2 contralateral third molars for age estimation.

The discrepancy between chronological age (CA) and dental age (DA) for older participants in this study may stem from the inherent inaccuracy in estimating age of older individuals using the Demirjian method. It may also be due to differences in temporal arrangement of tooth development between southern and northern Chinese. These two sources of discrepancy could not be disentangled in the present study. Further research comparing validity of the southern Chinese reference dataset applied to northern Chinese participants with teeth that cannot be evaluated and those with all teeth available for evaluation will help determine the relative contributions of the two sources of discrepancy. Although the two sources of discrepancy could not be distinguished in this study, the degrees of discrepancy between CA and DA for all age groups were within an acceptable range as shown by the small magnitude of absolute difference between CA and DA.

Our findings suggest that in spite of the differences in physical and dentofacial characteristics between northern and southern Chinese [25], the two populations shared similar chronological patterns of dental development. This lends further support to the relative stability of using tooth developmental stage as a means of dental age estimation. However, several limitations of the present study bear noting. First, although the panoramic radiographs for this study were randomly selected, the sampling frame was patients attending a particular hospital rather than the total population. As a result, we were unable to completely rule out the possibility of selection bias. Caution should therefore be exercised when generalizing the present findings. Second, sample size for the age by sex categories was generally small. This resulted in relatively large boundary values for the 95% limits of agreement and SEM despite the small magnitude of absolute difference between CA and DA. Third, due to availability of the reference dataset, only the Demirjian's method of dental age estimation was used in the present study. Future studies establishing reference dataset for northern/southern Chinese based on other methods such as the Nolla's method [11], modified Demirjian's method [10], and the Cameriere's method [12, 13] are warranted.

5. Conclusion

Dental age for northern Chinese estimated from reference dataset developed on the basis of southern Chinese demonstrated high level of agreement and reliability compared with Chronological age. Although lower agreement was noted for participant in the older age groups, the discrepancies were within an acceptable range. Such findings suggest that reference dataset for southern Chinese could be employed for dental age estimation in the northern Chinese.

Declarations

Author contribution statement

Hai Ming Wong: Conceived and designed the experiments; Performed the experiments; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Yi Feng Wen, Ling Sun: Performed the experiments; Analyzed and interpreted the data; Wrote the paper.

Jayakumar Jayaraman: Conceived and designed the experiments; Performed the experiments; Wrote the paper.

Nigel Martyn King, Graham J. Roberts: Conceived and designed the experiments; Performed the experiments; Wrote the paper.

Competing interest statement

The authors declare no conflict of interest.

Funding statement

This work was supported by the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. 17122914).

Additional information

No additional information is available for this paper.

Acknowledgements

The work described in this paper was fully supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. 17122914).

References

- [1] R.F. Sognaes, Forensic stomatology (second of three parts), *N. Engl. J. Med.* 296 (1977) 79–85.
- [2] J. Tanner, M. Healy, H. Goldstein, N. Cameron, *Assessment of Skeletal Maturity and Prediction of Adult Height: TW3 Method*, Saunders, Philadelphia, 2001.
- [3] V. Maled, B. Manjunatha, K. Patil, B.M. Balaraj, The chronology of third molar root mineralization in south Indian population, *Med. Sci. Law.* 54 (2014) 28–34.
- [4] V. Maled, S.B. Vishwanath, The chronology of third molar mineralization by digital orthopantomography, *J. Forensic Leg. Med.* 43 (2016) 70–75.
- [5] A. Demirjian, P.H. Buschang, R. Tanguay, D.K. Patterson, Interrelationships among measures of somatic skeletal, dental, and sexual maturity, *Am. J. Orthod.* 88 (1985) 433–438.
- [6] S.M. Garn, A.B. Lewis, R.S. Kerewsky, Genetic nutritional, and maturational correlates of dental development, *J. Dent. Res.* 44 (1965) 228–242.
- [7] A.B. Lewis, S.M. Garn, The relationship between tooth formation and other maturational factors, *Angle Orthod.* 30 (1960) 70–77.
- [8] A.S. Panchbhai, Dental radiographic indicators, a key to age estimation, *Dentomaxillofac. Radiol.* 40 (2011) 199–212.

- [9] A. Demirjian, H. Goldstein, J. Tanner, A new system of dental age assessment, *Hum. Biol.* (1973) 211–227.
- [10] Y. Zhai, H. Park, J. Han, H. Wang, F. Ji, J. Tao, Dental age assessment in a northern Chinese population, *J. Forensic Leg. Med.* 38 (2016) 43–49.
- [11] C.M. Nolla, The development of permanent teeth, *J. Dent. Child.* 27 (1960) 254–266.
- [12] R. Cameriere, L. Ferrante, M. Cingolani, Age estimation in children by measurement of open apices in teeth, *Int. J. Leg. Med.* 120 (2006) 49–52.
- [13] Y.C. Guo, C.X. Yan, X.W. Lin, H. Zhou, J.P. Li, F. Pan, et al., Age estimation in northern Chinese children by measurement of open apices in tooth roots, *Int. J. Leg. Med.* 129 (2015) 179–186.
- [14] M. Maber, H. Liversidge, M. Hector, Accuracy of age estimation of radiographic methods using developing teeth, *Forensic Sci. Int.* 159 (2006) S68–S73.
- [15] A. Olze, D. Bilang, S. Schmidt, K.-D. Wernecke, G. Geserick, A. Schmeling, Validation of common classification systems for assessing the mineralization of third molars, *Int. J. Leg. Med.* 119 (2005) 22–26.
- [16] K. Moze, G. Roberts, Dental age assessment (DAA) of Afro-Trinidadian children and adolescents. Development of a reference dataset (RDS) and comparison with Caucasians resident in London. UK, *J. Forensic Leg. Med.* 19 (2012) 272–279.
- [17] J. Jayaraman, H.M. Wong, N.M. King, G.J. Roberts, Development of a Reference Data Set (RDS) for dental age estimation (DAE) and testing of this with a separate Validation Set (VS) in a southern Chinese population, *J. Forensic Leg. Med.* 43 (2016) 26–33.
- [18] L. Jin, B. Su, Natives or immigrants: modern human origin in East Asia, *Nat. Rev. Genet.* 1 (2000) 126–133.
- [19] B. Su, et al., Y-chromosome evidence for a northward migration of modern humans into Eastern Asia during the last Ice Age, *Am. J. Hum. Genet.* 65 (1999) 1718–1724.
- [20] J. Chu, et al., Genetic relationship of populations in China, *Proc. Natl. Acad. Sci. U. S. A.* 95 (1998) 11763–11768.
- [21] Y. Ke, et al., Y-chromosome haplotype distribution in Han Chinese populations and modern human origin in East Asians, *Sci. China C Life Sci.* 44 (2001) 225–232.

- [22] Y.-G. Yao, Q.-P. Kong, H.-J. Bandelt, T. Kivisild, Y.-P. Zhang, Phylogeographic differentiation of mitochondrial DNA in Han Chinese, *Am. J. Hum. Genet.* 70 (2002) 635–651.
- [23] W. Tian, J. Cai, F. Wang, L. Li, Y. Cao, HLA-G* 0105 N and HLA-G 14 bp dimorphisms in exon 8 in four distinct populations in mainland China, *Tissue Antigens* 75 (2010) 227–234.
- [24] S.-F. Lei, et al., Genome-wide association scan for stature in Chinese: evidence for ethnic specific loci, *Hum. Genet.* 125 (2009) 1–9.
- [25] Y. Gu, U. Hagg, J. Wu, S. Yeung, Differences in dentofacial characteristics between southern versus northern Chinese adolescents, *Aust. Orthod. J.* 27 (2011) 155–161.
- [26] W. Lin, C. Hu, A study of environment difference on Chinese youth growth, *Acta Anthropologica Sinicae* 9 (1990) 152–159.
- [27] Z. Ying-Xiu, W. Shu-Rong, Geographic variation of stature in Chinese youth of age 18+, *Anthropologist* 13 (2011) 103–106.
- [28] T. Boonpitaksathit, N. Hunt, G.J. Roberts, A. Petrie, V.S. Lucas, Dental age assessment of adolescents and emerging adults in United Kingdom Caucasians using censored data for stage H of third molar roots, *Eur. J. Orthod.* 33 (2011) 503–508.
- [29] G. Roberts, S. Parekh, A. Petrie, V. Lucas, Dental age assessment (DAA): a simple method for children and emerging adults, *Br. Dent. J.* 204 (2008) E7.
- [30] J.R. Landis, G.G. Koch, The measurement of observer agreement for categorical data, *Biometrics* 33 (1977) 159–174.
- [31] A.F. De Winter, et al., Inter-observer reproducibility of measurements of range of motion in patients with shoulder pain using a digital inclinometer, *BMC Musculoskelet. Disord.* 5 (18) (2004).
- [32] H.C. de Vet, C.B. Terwee, D.L. Knol, L.M. Bouter, When to use agreement versus reliability measures, *J. Clin. Epidemiol.* 59 (2006) 1033–1039.
- [33] J.M. Bland, D. Altman, Statistical methods for assessing agreement between two methods of clinical measurement, *Lancet* 327 (1986) 307–310.
- [34] L.I.-K. Lin, A concordance correlation coefficient to evaluate reproducibility, *Biometrics* (1989) 255–268.
- [35] P.E. Shrout, J.L. Fleiss, Intraclass correlations: uses in assessing rater reliability, *Psychol. Bull.* 86 (1979) 420.

- [36] R Development Core Team, A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria, 2016.
- [37] W. Revelle, psych: Procedures for personality and psychological research, Northwestern University, Evanston Illinois, USA, 2016. Version = 1.6.4 <http://CRAN.R-project.org/package=psych>.
- [38] D.V. Cicchetti, S.A. Sparrow, Developing criteria for establishing interrater reliability of specific items: applications to assessment of adaptive behavior, *Am. J. Ment. Defic.* (1981) 127–137.
- [39] China Statistical Yearbook 2015, The State Bureau of Statistics, Beijing, 2015. Accessed at September 28th, 2016 <http://www.stats.gov.cn/tjsj/ndsj/2015/indexeh.htm>.
- [40] Y. Zhou, Uncovering children in marginalization: Explaining unregistered children in China, Paper presented at the IUSSP General Conference, Tours, France, 2005.
- [41] T.S. Peiris, G.J. Roberts, N. Prabhu, Dental Age Assessment: a comparison of 4- to 24-year-olds in the United Kingdom and an Australian population, *Int. J. Paediatr. Dent.* 19 (2009) 367–376.
- [42] J. Jayaraman, N.M. King, G.J. Roberts, H.M. Wong, Dental age assessment: are Demirjian's standards appropriate for southern Chinese children? *J Forensic Odontostomatol.* 29 (2011) 22–28.
- [43] J. Jayaraman, G.J. Roberts, N.M. King, H.M. Wong, Dental age assessment of southern Chinese using the United Kingdom Caucasian reference dataset, *Forensic Sci. Int.* 216 (2012) 68–72.
- [44] J. Jayaraman, H.M. Wong, N.M. King, G.J. Roberts, The French–Canadian data set of Demirjian for dental age estimation: a systematic review and meta-analysis, *J. Forensic Leg. Med.* 20 (2013) 373–381.
- [45] Z. Huang, et al., A north-south comparison of blood pressure and factors related to blood pressure in the People's Republic of China: a report from the PRC-USA Collaborative Study of Cardiovascular Epidemiology, *J. Hypertens.* 12 (1994) 1103–1112.
- [46] F.A. Birchler, S. Kiliaridis, C. Combescure, L. Vazquez, Dental age assessment on panoramic radiographs in a Swiss population: a validation study of two prediction models, *Dentomaxillofac. Radiol.* 45 (2015) 20150137.
- [47] L.F. Tomás, L.S. Mónico, I. Tomás, P. Varela-Patiño, B. Martín-Biedma, The accuracy of estimating chronological age from Demirjian and Nolla methods in a Portuguese and Spanish sample, *BMC Oral Health* 14 (2014) 160.

- [48] V. Urzel, J. Bruzek, Dental age assessment in children: a comparison of four methods in a recent French population, *J. Forensic Sci.* 58 (2013) 1341–1347.
- [49] L. Kullman, G. Johanson, L. Akesson, Root development of the lower third molar and its relation to chronological age, *Swed. Dent. J.* 16 (1991) 161–167.
- [50] J. Braga, Y. Heuze, O. Chabadel, N. Sonan, A. Gueramy, Non-adult dental age assessment: correspondence analysis and linear regression versus Bayesian predictions, *Int. J. Leg. Med.* 119 (2005) 260–274.