Forensic age estimation based on development of third molars: a staging technique for magnetic resonance imaging

Jannick De Tobel[1,2,3,4]
Inès Phlypo[5]
Steffen Fieuws[6]
Constantinus Politis[3]
Koenraad L. Verstraete[1]
Patrick W. Thevissen[2]

[1]Department of Radiology and Nuclear Medicine, Ghent University, Belgium
[2]Department of Oral Health Sciences – Forensic Dentistry, KU Leuven and Department of Dentistry – University Hospitals Leuven, Belgium
[3]Department of Oral and Maxillofacial Surgery, Leuven University Hospital, Belgium
[4]Department of Head, Neck and Maxillofacial Surgery, Ghent University Hospital, Belgium
[5]Department of Dentistry – Special Care in Dentistry, PaeCoMeDiS, Ghent University, Ghent, Belgium
[6]Leuven Biostatistics and Statistical Bioinformatics Centre (L-BioStat), Leuven, Belgium

Corresponding author:
jannick.detobel@ugent.be

ABSTRACT

Background: The development of third molars can be evaluated with medical imaging to estimate age in subadults. The appearance of third molars on magnetic resonance imaging (MRI) differs greatly from that on radiographs. Therefore a specific staging technique is necessary to classify third molar development on MRI and to apply it for age estimation.
Aim: To develop a specific staging technique to register third molar development on MRI and to evaluate its performance for age estimation in subadults.
Materials and methods: Using 3T MRI in three planes, all third molars were evaluated in 309 healthy Caucasian participants from 14 to 26 years old. According to the appearance of the developing third molars on MRI, descriptive criteria and schematic representations were established to define a specific staging technique. Two observers, with different levels of experience, staged all third molars independently with the developed technique. Intra- and inter-observer agreement were calculated. The data were imported in a Bayesian model for age estimation as described by Fieuws et al. (2016). This approach adequately handles correlation between age indicators and missing age indicators. It was used to calculate a point estimate and a prediction interval of the estimated age. Observed age minus predicted age was calculated, reflecting the error of the estimate.
Results: One-hundred and sixty-six third molars were agenetic. Five percent (51/1096) of upper third molars and 7% (70/1044) of lower third molars were not assessable. Kappa for inter-observer agreement ranged from 0.76 to 0.80. For intra-observer agreement kappa ranged from 0.80 to 0.89. However, two stage differences between observers or between staging sessions occurred in up to 2.2% (20/899) of assessments, probably due to a learning effect. Using the Bayesian model for age estimation, a mean absolute error of 2.0 years in females and 1.7 years in males was obtained. Root mean squared error equalled 2.38 years and 2.06 years respectively. The performance to discern minors from adults was better for males than for females, with specificities of 96% and 73% respectively.
Conclusion: Age estimations based on the proposed staging method for third molars on MRI showed comparable reproducibility and performance as the established methods based on radiographs.
.

INTRODUCTION
Staging third molars' development
The development of third molars can be evaluated with medical imaging to estimate age in subadults by allocating developmental stages. A staging method should cover the entire maturation sequence of the structure that is evaluated.[1, 2] Thevissen et al. (2013) pointed out that the choice of staging method should depend on the number of stages in the period of interest.[2] Moreover, stages should be defined unambiguously with clear threshold between them.[2] In forensic context, the most important question to be answered is whether or not the individual is a minor or an adult. Therefore, in most countries, an adequate staging method should encompass stages defined by changes that occur around the 18th birthday. A balance should be sought between a comprehensive method with a sufficient number of stages and a performant method with sufficient reproducibility and accuracy.[3] Stage characteristics should be straightforward and simple, in order to facilitate the learning process for observers and to exclude stage overlap caused by different interpretations.

Numerous staging techniques have been described, all of them based on radiological appearance of developing teeth.[4, 5] In all published papers on dental age estimation based on MRI of third molars, the radiological staging methods were extrapolated without any MR specific validity testing.[6-10] All of them used the Demirjian staging technique, whereas De Tobel et al. (2017) used both the Demirjian and the Köhler technique.[11, 12] Demirjian stages are defined by objective criteria, while Köhler stages are based on predictions of crown and root lengths. De Tobel et al. reported considerations to take into account when transferring Demirjian and Köhler stages to MRI. Still, major concerns remain regarding the different appearance on MRI compared with radiographs, which cannot be overcome using the existing staging techniques. Therefore, a specific staging technique is necessary to classify third molar development on MRI and to apply it for age estimation.

Statistical approach to age estimation
It has been stated that a Bayesian approach renders the most appropriate age estimation using developmental stages.[13, 14] Although the prediction outcome does not strongly outperform the classical regression result, it circumvents some assumptions that are not true in age estimation: (1) a linear relationship between age and stages, (2) a normal distribution of the variation of age around the mean with a constant variance and (3) uncorrelated development of the different anatomical structures[13, 15]. The major drawback of a Bayesian approach is its computational burden when combining multiple dependent predictors. However, this can be circumvented. Fieuws et al. (2016) reported a practical approach using Bayes' rule combining multiple age indicators based on Boldsen et al. (2002).[16, 17] The ad-hoc procedure allows to construct an approximate confidence interval without the need to model the multivariate correlation structure between the indicators.[16]

Aims
The aims of the current study were (1) to develop an MRI specific staging technique for the development of third molars and (2) to evaluate the age estimation performance of a Bayesian approach using this MRI specific staging.

MATERIALS AND METHODS
Study population
The local ethics committee approved the study and written informed consent was obtained from every participant. In case the participant was a minor, the parents' consent was also obtained. A study sample of 309 healthy Belgian and Dutch Caucasian volunteers (163 females, 146 males) were prospectively included. Table 1 shows the age distribution of the participants per sex. Additionally, four younger children were scanned (two girls of age 7 and 11; two boys of age 9 and 13). Their images were used to illustrate certain stages, but they were not included for analyses. Part of the study

population was included in previous papers.[9,10] None of the participants were relatives up to the third degree. Neither had any of them had any removal of a third molar. Socio-economic background was documented. Teeth were named according to the International Standards Organisation Designation System.

Table 1: Number of participants per age per sex.

Age (y)	Frequency		
	Female	**Male**	**Total**
14	11	11	22
15	11	10	21
16	10	10	20
17	11	10	21
18	13	10	23
19	15	14	29
20	20	10	30
21	14	11	25
22	12	12	24
23	12	10	22
24	11	11	22
25	13	12	25
26	10	15	25
Total	**163**	**146**	**309**

Image acquisition
Between March 2012 and May 2017, 3T MRI was conducted according to the protocol described in De Tobel et al. (2017) with a Siemens scanner (Magnetom Trio Tim, Siemens, Erlangen, Germany).[9] Fast spin echo (FSE) T2 images were available in three planes. Sagittal images were made along the long axis of the teeth per side. Axial images were made parallel to the occlusal plane, whereas coronal images were made perpendicular to the occlusal plane (Figure 1). A bilateral flexible four-channel surface head

coil (Model NMP-001D-ST-4, Nova Medical Inc., Wilmington, NC, USA) and an individualised bite bar were used. In ten cases (3.2%), the scan had to be done over because of motion artefacts (9 cases) or wrong coil positioning (1 case). In three cases the head positioning was too extended at the neck, causing motion artefacts because the participant could not keep the lower teeth still in the bite plate. This was resolved in the second scan by making a new bite plate, allowing for a more neutral neck position.

Image analysis
MRI specific staging technique
The considerations recommended by De Tobel et al. (2017)[10], together with other concerns made by the authors of the current study, were included to develop an MRI specific staging technique for third molar development.
The allocation of stages based on MRI should be conducted scrolling through the whole stack of slices depicting the considered tooth. When a fluid containing structure is seen in the jaw where the third molar is expected, stage 0 can be allocated. In fact, one can only be sure that a third molar is present when calcification appears (stage 1). After all, the crypt may be a cyst in which no tooth will develop (this situation is similar when evaluating radiographs). Therefore stage 0 should not be included in any analysis for age estimation. Since the youngest participants in the study sample were 14 years of age, it could be decided that the third molar was agenetic when no possible crypt or calcified tooth part was seen at the third molar region.[1,18]
Since on regular MR-images no distinction can be made between enamel, dentin and cementum, criteria based on these materials were omitted. For instance, the cemento-enamel junction cannot be identified on MR-images. As a consequence, the MR crown height was defined as being the distance between the tips of the cusps and the pulp horns (Figures 2 and 3). Corresponding to Demirjian's rules, when the different cusps are not at the same level, the midpoint between them is considered the highest reference point. Similarly, the lowest reference point is the midpoint between the pulp horns. Lines to define MR crown height should be perpendicular to the tooth axis. New criteria based on this MR crown height were formulated.
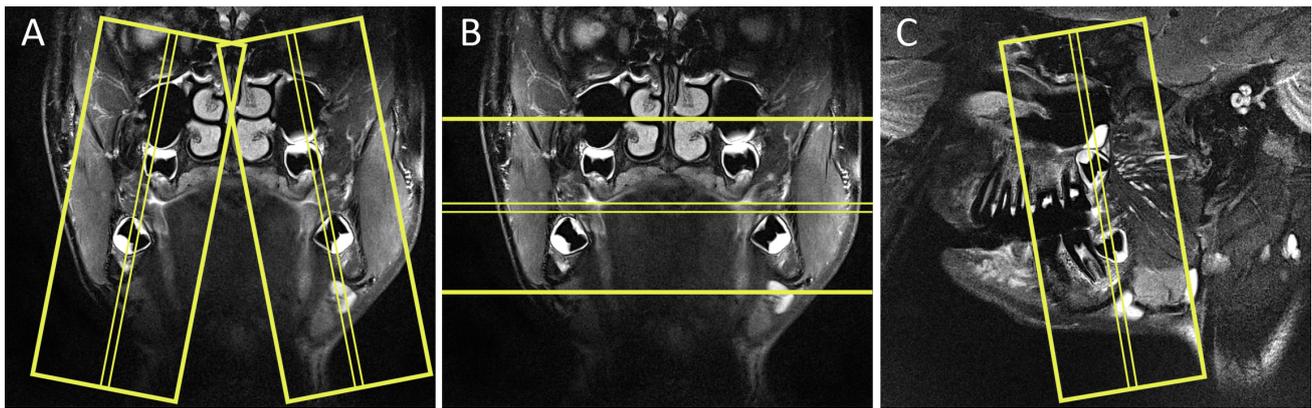
Figure 1: Yellow boxes depict the stack of scanned MRI slices

Figure 2: Definition of MR crown height and MR root length on MRI. Lines are perpendicular to the tooth axis. Distances (arrows) are evaluated along the tooth axis.
Line D is at the distal cusp tip, while M is at the mesial cusp tip. Line C represents the midpoint between the distal and mesial cusp tips.
Line H is at the pulp horns, which are both at the same level.
Line A is at the most apical point of the roots, which are all at the same line.
The distance between lines C and H is the MR crown height. The distance between lines H and A is the MR root length. In this case MR root length is more than one and a half MR crown height, so the tooth is in stage 5
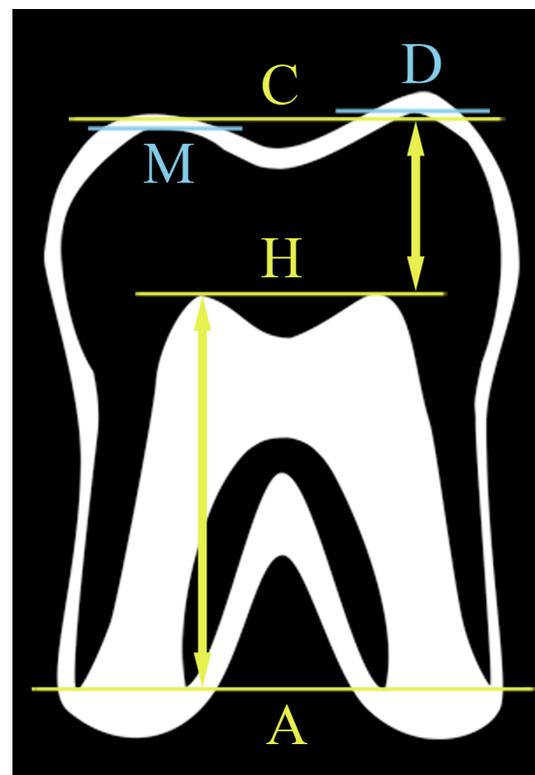
The MR crown height clearly differs from the clinical crown height, which – in contrast to radiographs or CT – cannot be determined on MRI. When evaluating relative lengths, a pair of dividers can be used to compare MR root lengths with MR crown height. In borderline cases, the measure tool of the viewing software can be used to compare absolute measures. MR root length is measured from the nearest pulp horn to the most apical point of the root, at the least developed root wall (Figures 2 and 3). The least developed root should be considered in case different roots are in different developmental stages.

Stages 3, 4 and 5 depend on MR root length compared with MR crown height. In case doubt prevails even after using the measure tool, the youngest stage should be allocated. Because the measure tool can only be used on one slice, it is impossible to measure tooth proportions when the tooth is depicted over multiple images. Still, when a certain MR root length is nearly reached on one slice and

Figure 3: (A, B) Lower right mandibular third molar depicted on two consecutive MRI slices. Slice (A) is situated more buccally than slice (B). The pulp chamber has a trapezoidal shape, corresponding to stage 3. To exclude stage 4, MR crown height and MR root length have to be evaluated as illustrated in images (C, D).

(C) Copy of image (A), with marked landmarks and distances to consider in order to allocate a stage. Lines are perpendicular to the tooth axis. Distances (arrows) are evaluated along the tooth axis. Line Dc is at the distal cusp tip, while Mc is at the mesial cusp tip. Line C represents the midpoint between the distal and mesial cusp tips. Line Dh is at the distal pulp horn while line Mh is at the mesial pulp horn. Line H represents the midpoint between the distal and mesial pulp horns. Line Da is at the most apical point of the distal root. Line Ma is at the most apical point of the mesial root.

The distance between lines C and H is the MR crown height based on this slice. The distance between lines Dh and Da is the distal MR root length, while the distance between lines Mh and Ma is the mesial MR root length.

(D) Copy of image (B). Both cusp tips are at the same level on this slice, represented by line C. MR crown height is larger than on the previous image, whereas the distal MR root length is smaller.

To allocate a stage, MR crown height on image (D) is the most appropriate, while the distal MR root length on image (C) is the most appropriate. Because the third molar is tilted bucco-lingually, the observer has to scroll through consecutive slices to decide on the most appropriate measures to consider. In slice (C) the crown is transsected more buccally, so part of the crown is not depicted. By contrast, the distal root apex is situated more buccally than the crown, so it is best depicted in slice (D). Because the distal root is shorter than MR crown height, this third molar is in stage 3.
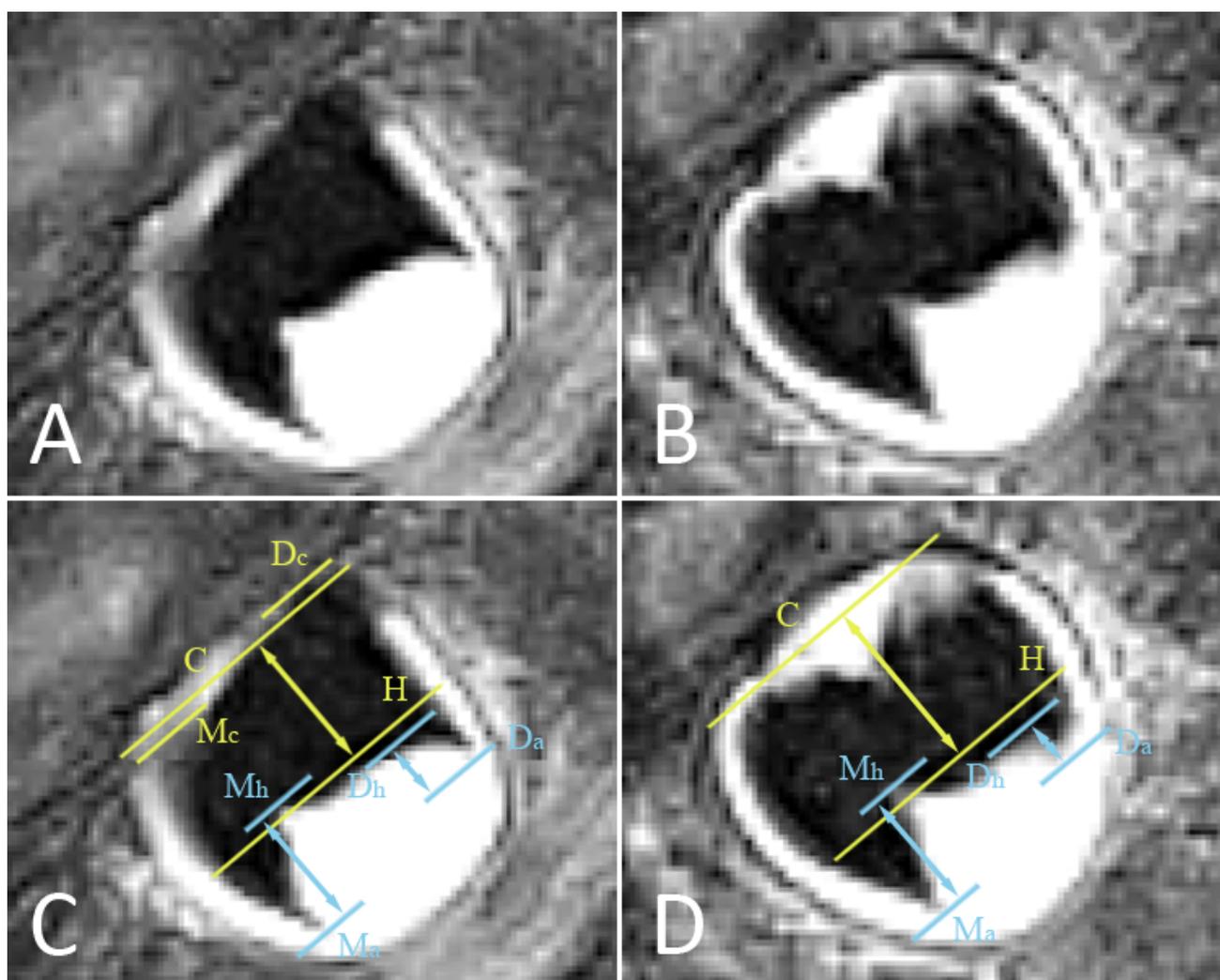
Figure 4: Lower right mandibular third molar depicted on consecutive MRI slices from buccally to lingually. MR crown height is most appropriately measured on slice D. When MR root length would only be based on slice C, stage 3 would be allocated. In fact the tooth is in stage 4, since MR root length on slice B is slightly longer than MR crown height and the root is depicted over several slices (keeping in mind that slice thickness is 2 mm).
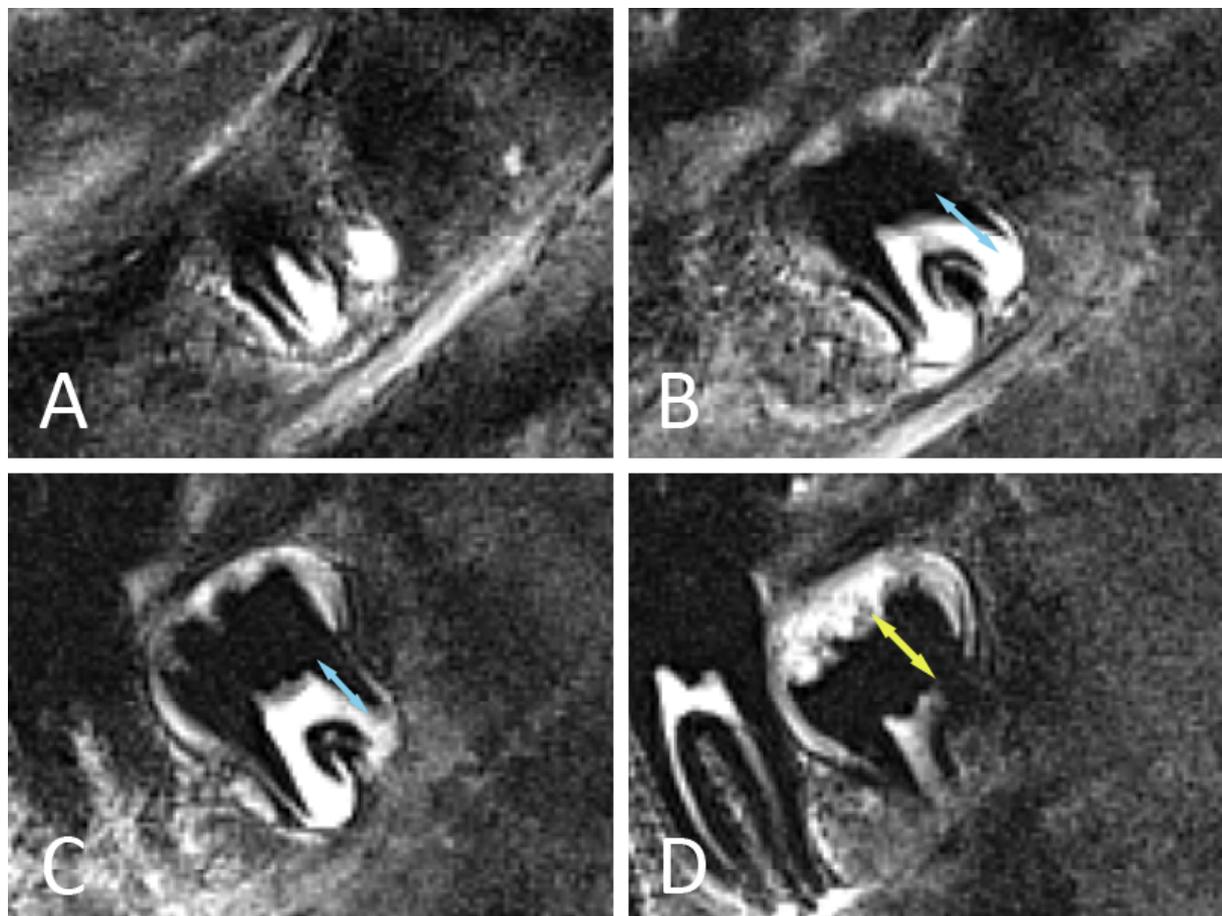
the root is spread over multiple consecutive sagittal slices, the higher stage can be allocated (Figure 4). After all, one has to take into account that the MR-sequence has a slice thickness of 2 mm. In the final stages, the root dentin at the apex changes from thin and parallel (stage 6), over thicker and converging (stage 7), to thick and closed (stage 8).All of these considerations resulted in the stages defined in Table 2 and Figure 5. To have reached a certain stage, the appearance of the root has to comply with the given criteria. When two criteria are given, the molar has reached the stage if the first criterion applies. When three criteria are given, the first two have to apply to allocate the concerning stage. Both schematic representations of uniradicular third molars and multiradicular third molars are given. Examples of the appearance of the different stages on MRI are shown in Figures 6 to 8.

Observers and media
Images were anonymised and evaluated by two observers independently. Observers were blinded to the age of the participant. Five participants per age per gender between 14 and 26 years old were assessed in a first session. After four months, both observers evaluated all 309 participants in a second session. The first observer (J.D.T.) was a resident at maxillofacial surgery studying forensic dentistry. He had been involved in research on age estimation for 8 years, including 6 years of dental age estimation. The second observer (I.P.) was a dentist in the first year after graduation. She had been involved in dental age estimation research for 1 year.
Images were assessed using a Barco MFGD monitor (3280 x 2048 pixels, Barco, Kortrijk, Belgium).

Table 2: Descriptive criteria for developmental stages of third molars on MRI.

Stage	Description
Stage 0	The crypt of the third molar is suspected without any calcification.
Stage 1	A beginning of calcification is seen at the superior level of the crypt in the form of an inverted cone or cones. There is no fusion of these calcified points.
Stage 2	a) Fusion of the calcified points forms one or several cusps which unite to give a regularly outlined occlusal surface. b) The outline of the pulp chamber has a flat or curved shape at the occlusal border. c) Initial formation of the radicular bifurcation is seen in the form of a hypo-intense calcified point.
Stage 3	a) The pulp chamber has a trapezoidal shape. The outline of the pulp horns is pointy and shaped like an umbrella top. b) Further downshaping of the crown and/or beginning of root formation is seen in the form of a spicule. The spicule is shorter than MR crown height. c) The calcified region of the bifurcation has developed further into a hypo-intense semi-lunar shape.
Stage 4	a) MR root length reaches at least one MR crown height. b) The calcified region of the bifurcation still has a semi-lunar shape or has developed further down.
Stage 5	a) MR root length reaches at least one and a half MR crown height. b) The calcified region of the bifurcation has developed further down from its semi-lunar shape to give the roots a more definite and distinct outline with funnel shaped endings. The funnel shape persists for some millimetres (i.e. it is not limited to a few pixels on the image).
Stage 6	a) The walls of the distal root canal are parallel and its apical end is still partially open. b) The walls at the apex of the root canal show relatively thin dentin. c) Remnants of the dental follicle are seen in the form of a hyper-intense area surrounding the apex.
Stage 7	a) The walls of the distal root canal are convergent and its apical end is still partially open. b) The walls at the apex of the root canal show relatively thin dentin. c) Remnants of the dental follicle are seen in the form of a hyper-intense area surrounding the apex.
Stage 8	a) The apical end of the distal root canal is completely closed. b) The walls at the apex of the root canal show relatively thick dentin.

Figure 5: Schematic representation of developmental stages of third molars on MRI. Mineralized tissues appear black on MRI. By contrast, the dental follicle, pulpal tissues, the periodontal space and saliva appear white. The upper panels show stages for multiradicular molars, while in the lower panels stages for monoradicular molars are shown (this also corresponds with the appearance of the palatal root in upper molars).
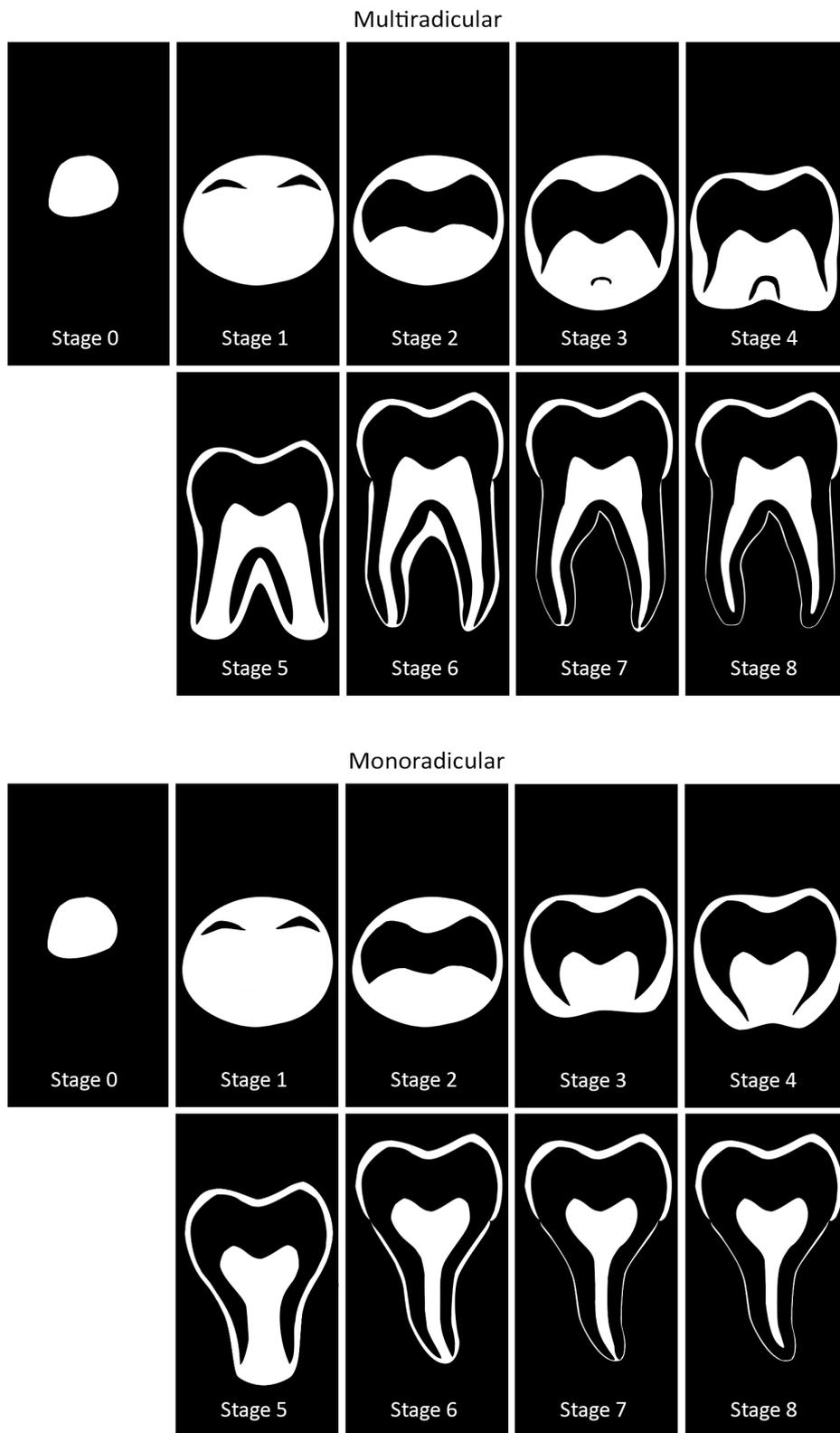
Figure 6: Representative examples of third molars in developmental stages 0 to 3, in the upper (A-E) and lower jaw (F-J). For some stages different appearances are illustrated.

(A) Stage 0. The crypt of the third molar shows no calcification. It is seen as a clearly delineated white area.

(F) Stage 1. Cusp tips are seen as separate black areas within the crypt.

(B, G) Early stage 2. Cusps are fused. The roof of the pulp chamber is quite flat.

(C, H) Late stage 2. The roof of the pulp chamber is more curved than in (B, G). Notice that in (H), the mesial side of the pulp chamber is more mature than the distal side. Thus, for staging the distal side should be considered. The distinction between early and late stage 2 was considered too subjective to consider them as separate stages.

(D, I) Stage 3. Notice the pointy appearance of the pulp horns. No furcation was present.

(E, J) Stage 3. Notice the furcation. In (J) the distal pulp horn appears curved on this sagittal slice. However, scrolling through the slices and including the coronal slices in the assessment, it was clear that both pulp horns were pointy, like an umbrella top.

Figure 7: Representative examples of third molars in developmental stages 4 and 5, in the upper (A-C) and lower jaw (D-F). In (A-C) palatal roots are depicted. For stage 5 different appearances are illustrated.

(A, D) Stage 4. Notice that the distal root in (D) is less developed than the mesial root.

(B, E) Early stage 5. Root walls are clearly funnel shaped at the root apex.

(C, F) Late stage 5. The funnel shape of the root walls at the apex is more subtle than in (B, E). The distinction between early and late stage 5 was considered too subjective to consider them as separate stages. Moreover, variability in root length would hinder a subclassification of stage 5.
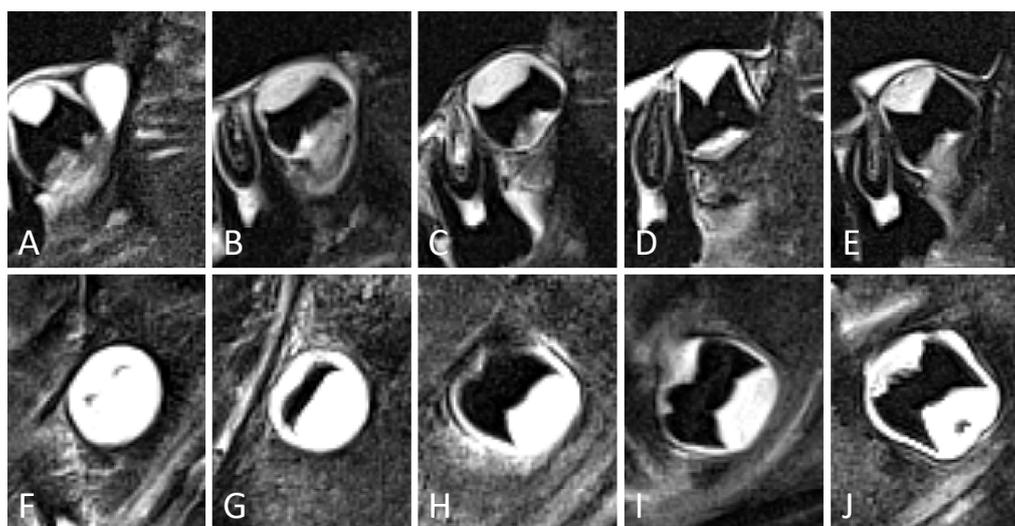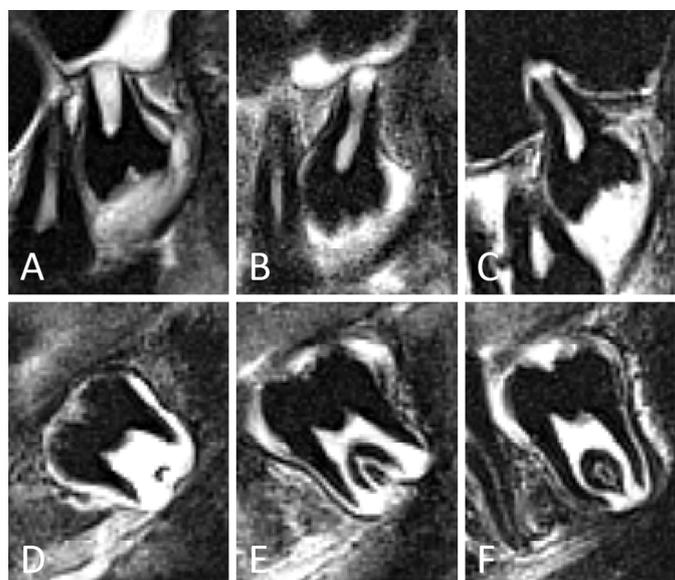
Figure 8: Representative examples of third molars in developmental stages 6 to 8, in the upper (A-D) and lower jaw (E-H). In (A-D) palatal roots are depicted. For stage 8 different appearances are illustrated.
(A, E) Stage 6. The width of the root canal differs depending on tooth anatomy. Still, parallel root walls are clear. Notice that the thin dentin at the apex in (A) might give the impression of a small funnel shape. However, it is stated in the criteria for stage 5 that the funnel shape should be more extensive than is seen in this example. Therefore, stage 6 is appropriate.
(B, F) Stage 7. The apices have clearly started closing. In (F) remnants of the dental follicle can be seen as white areas surrounding root apices.
(C, G) Stage 8. The apical dentin is relatively thin, but clearly continuous.
(D, H) Stage 8. Not only is the apical dentin continuous, but in these examples it is also relatively
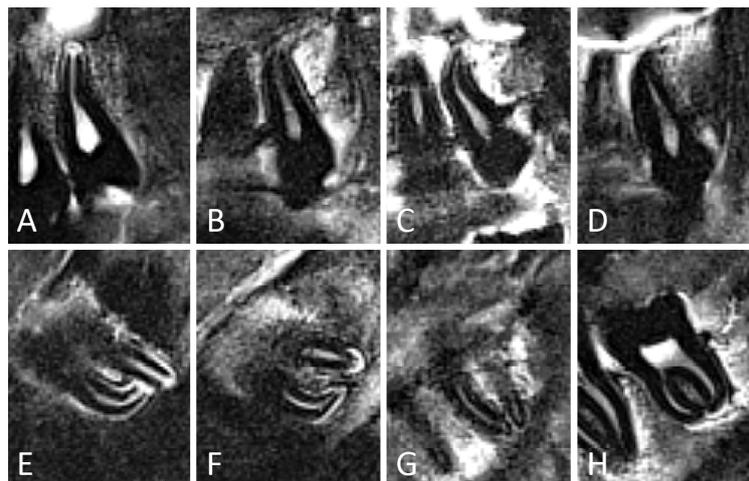
Studied variables

Using Microsoft Access forms, both observers assessed the images gathering data on four variables. First, a developmental stage was allocated or it was decided that the tooth could not be evaluated. Reasons for the latter were included in Table 3.[10]

Second, it was documented which root was considered to decide on the stage. Third, assessability of the roots was noted (Table 4). Fourth, observers indicated which planes they used to allocate a stage, allowing for combinations to be ticked off.

Statistical analysis

All data were transferred from Microsoft Access to SPSS Statistics 24.0 (IBM SPSS Statistics for Windows, Armonk, NY, USA) and SAS 9.4 (SAS Institute, Cary NC, USA). Descriptive statistics were calculated. Results of the second session by both observers were combined to report on the root used to stage, assessability of the roots and essential planes.

A paired Wilcoxon test was used to compare development between upper and lower third molars on the same side and to compare left and right third molars in the same jaw. Inter- and intra-observer agreement regarding stage allocation were quantified using proportion agreement and weighted kappa statistics. Cross tabulation of the observations allowed to check for systematic differences. The difference in marginal score distribution (between two observers or between two measurement occasions) was verified with Bowker's test of symmetry.

The data from the first observer were implemented into an ad-hoc procedure to obtain a point estimate of age and appropriate prediction intervals.[16] Participants with all available third molars in stage 8 were not included in the analysis (N = 28). Hence, the prediction pertains to subjects with not all available third molars fully developed. Motivation for this approach was that the point prediction (and thus the error) for participants with fully developed third molars is heavily influenced by the age range of included participants.

The ad-hoc procedure was based on application of Bayes' rule, using continuation-ratio models assuming conditional independence. The model takes third molar position into account, so that

Table 3: Reasons for the MRI being not assessable, with their frequencies.

Reason for being not assessable	Upper third molars		Lower third molars	
Insufficient contrast between apex tip and surrounding bone	2.6%	(28/1096)	2.1%	(22/1044)
Apex tip falls in between slices	0.5%	(6/1096)	1.8%	(19/1044)
Poor coil positioning	0.0%	(0/1096)	0.0%	(0/1044)
Poor image quality (e.g. poor signal to noise ratio)	0.3%	(3/1096)	0.4%	(4/1044)
Artefacts due to motion of the participant	0.9%	(10/1096)	1.5%	(16/1044)
Other artefacts (e.g. susceptibility due to metal)	0.4%	(4/1096)	0.6%	(6/1044)
Other, please specify	0.0%	(0/1096)	0.3%	(3/1044)

Table 4: Relative assessability of the different roots

Elements	Assessablity	Frequency	
Upper third molars	Only the staged root is assessable	4%	(32/914)
	Other roots are also assessable	81%	(739/914)
	Only the staged root is present	16%	(143/914)
Lower third molars	Only the staged root is assessable	2%	(16/844)
	Other roots are also assessable	97%	(817/844)
	Only the staged root is present	3%	(27/844)

even when one or more third molars are agenetic, the other third molars contribute to the model. The influence of agenetic third molar(s) to the posterior density distribution was illustrated in Figure 9. Non-proportional odds were allowed in the continuation-ratio model. Linearity was assumed for the relation between age and the logits. Note that for the same reason, the more simplistic model assuming proportional odds would lead to a more stable solution. Due to the low number of scores equal to 2, these were combined with scores 3 into a single level. Evaluation of the performance was based on 10-fold cross-validation and the approach was performed separately for males and females. The creation of the folds was stratified on age category (1 year interval).

Observed age minus predicted age was calculated, reflecting the error of the estimate. The maximum likelihood (ML) estimate was given as point prediction (this equals the modus of the posterior distribution of age), as well as the mean and 95% trimmed mean of the posterior distribution. The interval estimate (prediction interval) for age corresponded to the 95% age values of highest probability density. The difference between the posterior density assuming conditional independence and the correction by the ad-hoc procedure was illustrated in Figure 10.The proportion of cases, whose chronological age fell inside the 95%

confidence interval (CI), known as coverage, was calculated. Since the ML estimate is known not to minimize the root mean squared error (RMSE), the mean of the posterior distribution, as well as the mean of the posterior distribution in the prediction interval (= trimmed mean) were also reported.Spearman correlation between chronological age and staging was calculated, as well as Pearson correlation between chronological age and error of the estimated age. The first reflecting the degree of change in development explained by a change in age. The latter reflecting the degree of bias of the age estimate. Accuracy, sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) were used to evaluate the minor-adult distinction. Accuracy represented the proportion of correctly classified subjects. Sensitivity indicated the proportion of correctly classified adults, while specificity indicated the proportion of correctly classified minors. PPV equalled the proportion of adults within estimated adults. NPV was the proportion of minors within estimated minors. The area under the receiver operator characteristic (ROC) curve (AUC) reflected the percentage of times that a randomly selected individual from the older age category would have a more advanced root compared to a randomly chosen individual from the younger age category. Finally, the AUC probability to be older than 18 years was calculated.

Statistical tests were performed two-sided and evaluated at the 0.05 significance level.

Figure 9: Influence of agenetic third molars on the posterior density. The posterior density becomes smaller as a function of increasing information (plot obtained from males).

Figure 10: Influence of agenetic third molars on the posterior density. The posterior density becomes smaller as a function of increasing information (plot obtained from males).
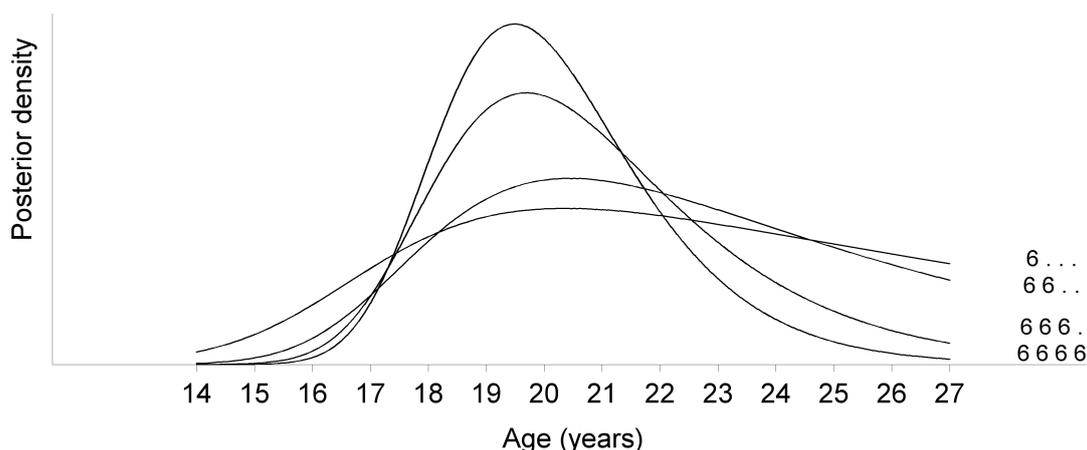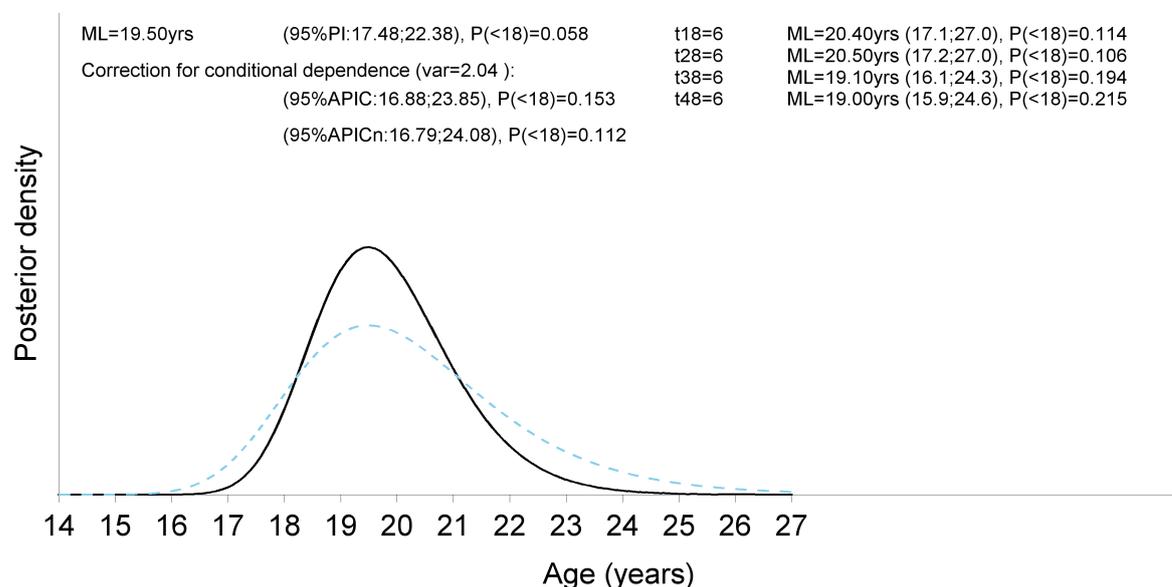
RESULTS
Tooth development and assessability

In the study sample, agenesis of one or more third molars was frequently seen: teeth 18, 28, 38 and 48 were agenetic in 40, 30, 50 and 46 out of 309 cases respectively (Table 5).

Tables 4 and 6 summarise results on the relative assessability of the different roots and root used to stage per third molar. The root used to stage was always the least developed assessable root.

These results included all assessable third molars from stage 4 on. No stage could be allocated in 5% (51/1096) of upper and 7% (70/1044) of lower third molars (Table 3). In a few cases monoradicular third molars were encountered (Table 4). Nineteen percent of the upper third molars had unassessable (mostly buccal) roots, compared to only 3% of lower third molars (mesial or distal were approximately equally distributed).

Table 5: Patterns of agenesis of one or more third molars with their frequencies.

Agenetic elements	Frequency
18	6
28	3
38	7
48	6
18, 28	7
38, 48	14
18, 38	2
18, 48	2
28, 38	1
18, 28, 38	1
18, 38, 48	3
28, 38, 48	2
18, 28, 38, 48	10
Total	**64**

Table 6: Frequency of root used to stage (least developed).

Elements	Root used to stage	Frequency	
Upper third molars	Palatal	88 %	(741/838)
	Mesiobuccal	9 %	(75/838)
	Distobuccal	3 %	(22/838)
Lower third molars	Mesial	11 %	(91/845)
	Distal	89 %	(754/845)

Essential plane

Sagittal images were essential to allocate a stage in 95% (1040/1096) of assessable upper third molars. Only in a few upper third molars did the coronal (5% = 55/1096) plane contribute to staging, while the axial images were never useful. In lower third molars frequencies were 92% (966/1044), 4% (41/1044) and 1% (9/1044) for sagittal, coronal and axial images respectively. Coronal and/or axial images proved to be useful when the tooth was extremely tilted, when the apex seemed to fall in between sagittal slices or to differentiate stage 2 from stage 3.

Staging and age estimation

In fourteen participants all third molars were not allocated a stage, because they were either agenetic (n = 10) or unassessable due to motion artefacts (n = 2) or susceptibility to metal (n = 2). A systematic difference in development between upper and lower third molars was statistically confirmed (P = 0.001 right, P < 0.001 left), with lower third molars overall being in the same or more advanced stages than upper ones. Left and right third molars in the same jaw did not differ significantly in development (P = 0.283 upper, P = 0.085 lower).

Reproducibility of staging

Table 7 shows inter- and intra-observer agreement for stage allocation. Table 8 shows the cross tabulation of frequencies of allocated stages by both observers at both staging sessions. A one stage difference was frequently seen. Two stage differences were also encountered between staging sessions in 1.3% (5/379) and 1.4% (5/350) of assessments, and between observers in 1.4% (5/364) and 2.2% (20/899) of assessments. Moreover, a systematic difference in allocated stages was noticed for both observers. In the second session more frequently a higher stage was allocated. Bowker's test of symmetry indicated no statistically significant asymmetry between both staging sessions for observer 1 (P = 0.21), while it was significant for observer 2 (P < 0.001). This also resulted in a significant asymmetry between both observers in the second session (P < 0.001), while in the first session asymmetry was not significant (P = 0.51).

Table 7: Reproducibility of staging third molar development. The proportion agreement as well as two versions of the weighted kappa are reported for all third molars combined, as well for each third molar separately. The kappa with the linear weights is the weighted kappa, typically reported in most agreement studies. The kappa with quadratic weights is similar to the intra-class correlation (ICC). Note that results based aggregated data from the four third molars do not take into account the correlation between the four third molars.

Element	Intra-observer agreement			Inter-observer agreement		
	Agreement (SE)	Weighted kappa		Agreement (SE)	Weighted kappa	
		Linear (SE)	Quadratic (SE)		Linear (SE)	Quadratic (SE)
Observer 1				Session 1		
All third molars	(N = 379) 0.760 (0.022)	0.873 (0.013)	0.954 (0.006)	(N = 364) 0.808 (0.021)	0.893 (0.012)	0.959 (0.006)
18	0.722 (0.046)	0.857 (0.025)	0.953 (0.009)	0.868 (0.036)	0.932 (0.019)	0.977 (0.007)
28	0.760 (0.043)	0.848 (0.030)	0.933 (0.017)	0.781 (0.042)	0.866 (0.028)	0.944 (0.015)
38	0.769 (0.044)	0.890 (0.023)	0.966 (0.008)	0.773 (0.045)	0.868 (0.029)	0.946 (0.015)
48	0.791 (0.043)	0.889 (0.025)	0.959 (0.012)	0.809 (0.042)	0.903 (0.022)	0.968 (0.008)
Observer 2				Session 2		
All third molars	(N = 350) 0.760 (0.022)	0.834 (0.015)	0.954 (0.006)	(N = 899) 0.620 (0.016)	0.790 (0.010)	0.922 (0.005)
18	0.713 (0.049)	0.846 (0.028)	0.945 (0.012)	0.648 (0.032)	0.803 (0.021)	0.926 (0.010)
28	0.674 (0.049)	0.804 (0.032)	0.919 (0.017)	0.617 (0.032)	0.784 (0.021)	0.920 (0.010)
38	0.774 (0.046)	0.877 (0.027)	0.958 (0.010)	0.588 (0.034)	0.764 (0.023)	0.907 (0.012)
48	0.655 (0.051)	0.809 (0.030)	0.931 (0.013)	0.624 (0.033)	0.804 (0.019)	0.933 (0.008)

Table 8: Cross tabulation of frequencies of allocated scores by both observers in both staging sessions.

Intra-observer agreement

Observer 1		Stage session 1							
	Frequency	2	3	4	5	6	7	8	Total
Stage session 2	2	4	0	0	0	0	0	0	4
	3	0	52	3	0	0	0	0	55
	4	0	6	58	6	0	0	0	70
	5	0	0	11	50	1	1	0	63
	6	0	0	1	13	37	9	0	60
	7	0	0	0	1	11	30	5	47
	8	0	0	0	0	2	21	57	80
	Total	4	58	73	70	51	61	62	379

Inter-observer agreement

Session 1		Stage observer 2							
	Frequency	2	3	4	5	6	7	8	Total
Stage observer 1	2	4	0	0	0	0	0	0	4
	3	1	51	6	0	0	0	0	58
	4	0	5	47	21	0	0	0	73
	5	0	0	5	64	0	1	0	70
	6	0	0	0	5	35	4	1	45
	7	0	0	0	1	5	48	3	57
	8	0	0	0	0	2	10	45	57
	Total	5	56	58	91	42	63	49	364

Observer 2		Stage session 1							
	Frequency	2	3	4	5	6	7	8	Total
Stage session 2	2	4	2	0	0	0	0	0	6
	3	1	40	1	0	0	0	0	42
	4	0	14	33	0	0	0	0	47
	5	0	0	22	76	0	0	0	98
	6	0	0	2	15	27	5	1	50
	7	0	0	0	0	6	33	11	50
	8	0	0	0	0	2	22	33	57
	Total	5	56	58	91	35	60	45	350

Session 2		Stage observer 2							
	Frequency	2	3	4	5	6	7	8	Total
Stage observer 1	2	20	7	0	0	0	0	0	27
	3	5	77	30	1	0	0	0	113
	4	0	4	64	74	3	0	0	145
	5	0	0	2	112	29	0	0	143
	6	0	0	1	33	102	30	3	169
	7	0	0	0	0	33	75	14	122
	8	0	0	0	0	12	61	107	180
	Total	25	88	97	220	179	166	124	899

Age estimation

An overview of chronological age and estimated age for the study sample is presented in Figure 11. Figure 12 presents posterior distributions of the Bayesian approach. Table 9 shows examples of point predictions with prediction intervals and probabilities to be adult for different patterns of allocated stages per sex.

Applying the Bayesian model for age estimation, using the mean of the posterior distribution as point prediction rendered better results than using the trimmed mean or ML estimate. The mean absolute error was 2.0 years in females (median (Me) = 1.7, interquartile range (IQR) 0.8–2.7) and 1.7 years in males (Me = 1.6, IQR 0.6–2.5) based on the mean of the posterior distribution. The mean error was 0.1 years in females (Me = 0.0, IQR -1.7–1.8) and males (Me = 0.0, IQR -1.4–1.9). Root mean squared error equalled 2.38 (95% confidence interval (CI) 2.11–2.65) for females and 2.06 (95% CI 1.79–2.33) for males. Coverage of the 95% prediction interval was 94.7% (142/150) for females and 91.4% (107/117) for males.

Moreover, the error of the age estimate clearly depended on age. The dependency was lowest using the mean of the posterior distribution as

point prediction with Spearman correlation (r) for mean error in females equal to 0.51 (95% CI 0.38–0.62, P < 0.001) and in males equal to 0.50 (95% CI 0.35–0.62, P < 0.001). For mean absolute error r = 0.10 in females (95% CI -0.06–0.25, P = 0.22) and r = 0.29 in males (95% CI 0.11–0.44, P = 0.0017).

Performance of the Bayesian procedure to discriminate between minors and adults is summarized in Table 10. In forensic age estimation in the living, one should strive for an approach with high specificity and NPV (specificity being the major concern). Estimating age based on the ML estimate rendered the highest specificity in females and males. The highest NPV was obtained using the mean of the posterior distribution as point prediction of age. The AUC was very similar for all three point predictions. The AUC probability to be older than 18 years was 0.869 (95% CI 0.811–0.926) for females and 0.948 (95% CI 0.908–0.988) for males.

Figure 11: Graphs comparing chronological age (dots) with the point (triangle) and interval (line) prediction in females (A) and males (B). The point prediction is the ML estimate.
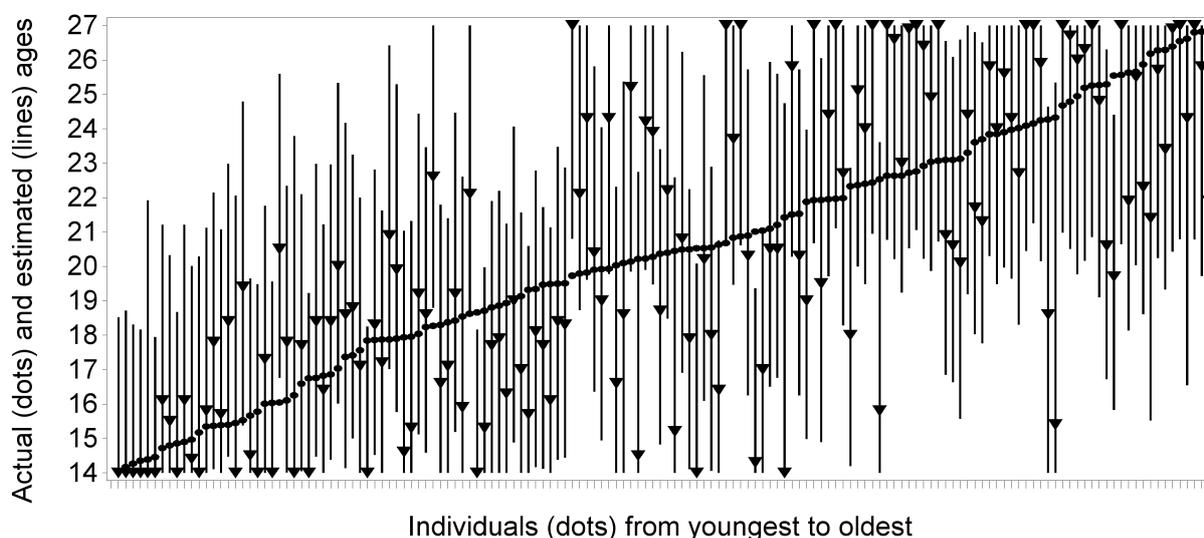
A   Females
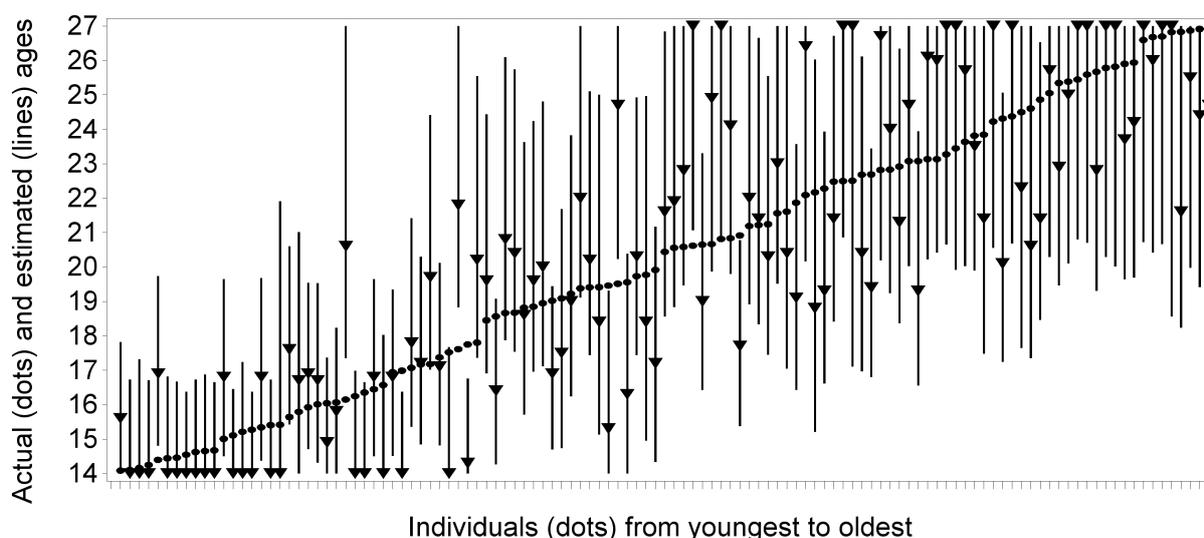
B   Males

Figure 12: Posterior density for all possible homogenous stage patterns (same stage for all third molars) in males. When all third molars are in stages equal to or lower than three (3333), the distribution of age is right-skewed. This smoothly evolves to a left-skewed age distribution when all third molars are fully mature (8888). Around the age of 18 years, most individuals have third molars in stage 5. Per situation the probability to be adult is represented by the area under the posterior density curve to the right of the 18 years threshold (blue vertical line).
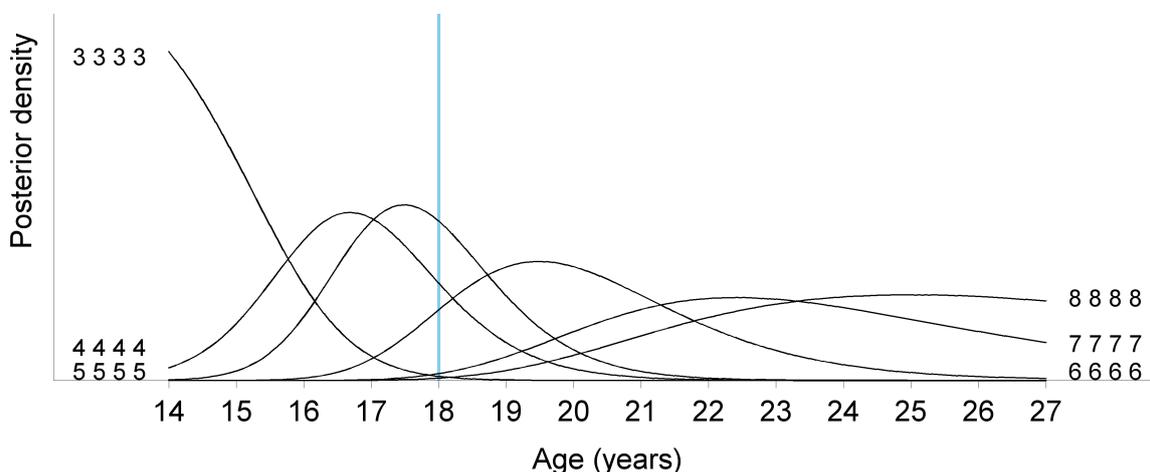

Table 9: Examples of point predictions with prediction intervals and probabilities to be adult for different patterns of allocated stages per sex. Point predictions of age are based on the mean of the posterior distribution. Notice that the lower limit of the prediction interval in the lowest stage reflects the minimum age in the study sample. Similarly, the upper limit in the highest stages equals the maximum age in the study sample.

Element				Point prediction	95% Prediction interval	Probability to be adult
18	28	38	48			
			Females			
3	3	3	3	15.53	(14.00–18.33)	0.0600
4	4	4	4	16.89	(14.00–20.78)	0.2496
5	5	5	5	18.59	(14.35–22.74)	0.5760
6	6	6	6	21.88	(17.76–26.73)	0.9509
7	7	7	7	23.96	(20.15–27.00)	0.9956
8	8	8	8	24.68	(21.17–27.00)	0.9989
			Males			
3	3	3	3	15.03	(14.00–16.68)	0.0037
4	4	4	4	16.90	(14.35–19.45)	0.1765
5	5	5	5	17.75	(15.35–20.33)	0.3746
6	6	6	6	20.17	(16.79–24.08)	0.8883
7	7	7	7	22.85	(19.25–27.00)	0.9926
8	8	8	8	23.76	(20.04–27.00)	0.9976

Table 10: Performance to discriminate between minors and adults.

	Predicted age based on	ML estimate	Mean	Trimmed mean
		Rate (95% CI)		
Females	Accuracy	77.3 (69.8–83.8)	79.3 (72.0–85.5)	78.0 (70.5–84.4)
	Sensitivity	78.9 (70.0–86.1)	84.4 (76.2–90.6)	82.6 (74.1–89.2)
	Specificity	73.2 (57.0–85.8)	65.8 (49.4–79.9)	65.8 (49.4–79.9)
	PPV	88.7 (80.6–94.2)	86.8 (78.8–92.6)	86.5 (78.4–92.4)
	NPV	56.6 (42.3–70.2)	61.4 (45.5–75.6)	58.7 (43.2–73.0)
	AUC	0.865 (0.809–0.922)	0.873 (0.817–0.928)	0.874 (0.818–0.929)
Males	Accuracy	90.6 (83.8–95.2)	90.6 (83.8–95.2)	89.7 (82.8–94.6)
	Sensitivity	91.0 (82.4–96.3)	92.3 (84.0–97.1)	91.0 (82.4–96.3)
	Specificity	89.7 (75.8–97.1)	87.2 (72.6–95.7)	87.2 (72.6–95.7)
	PPV	94.7 (86.9–98.5)	93.5 (85.5–97.9)	93.4 (85.3–97.8)
	NPV	83.3 (68.6–93.0)	85.0 (70.2–94.3)	82.9 (67.9–92.8)
	AUC	0.950 (0.912–0.988)	0.949 (0.909–0.988)	0.949 (0.909–0.988)

Accuracy = proportion of correctly classified subjects; Sensitivity = proportion of correctly classified adults; Specificity = proportion of correctly classified minors; PPV = proportion of adults within estimated adults; NPV = proportion of minors within estimated minors; AUC = area under the ROC curve.

DISCUSSION
Staging third molars' development.

Considerations for developing a staging technique

Staging development of third molars has been developed based on radiographs. However, with the increasing demand of imaging for age estimation that doesn't use ionizing radiation[19], MRI is being studied for dental age estimation by several research groups.[6-10] Since the appearance of teeth on MRI differs greatly from that on radiographs, a mere extrapolation of radiographical stages seems inappropriate. After all, criteria for staging based on crown height and root length cannot be applied when the cemento-enamel junction cannot be identified unambiguously. Dedicated MR sequences, in which it is possible to differentiate between the hard dental tissues, have been developed.[20-22] However, their use is not common practice yet, rendering them unavailable for forensic purposes.

Therefore, a universally applicable MRI specific staging technique for third molar development was proposed in this paper.
Some authors stated that predictions of crown height and root length should be avoided because, especially in third molars, dimensions are highly variable and unpredictable.[18] Predictions always imply subjectivity, while objective criteria for stages should be pursued. It has been reported that precision of the staging method might be reduced – i.e. compromise the feasibility to register all of the stages correctly – if thresholds between stages are based on predictions of lengths of tooth parts.[2, 23] Moreover, the fully developed crown height cannot be used to predict the future mature root length.[24] In the current study, it was considered inappropriate to include a stage in which MR root length would be at least twice MR crown height, because it was noticed that the roots of some third molars never reached this length, even when fully matured. In literature, it was also

stated that because of variability between second and third molars, predictions of third molar lengths should not be based on or compared with the dimensions of neighbouring teeth.[18] Therefore, only objective criteria were used in the proposed staging technique.

To differentiate between stage 5 and stage 6, one could check the tooth's eruption. In stage 5 it is still in infra-occlusion, while in stage 6 it has reached the occlusal plane. However, third molars are often impacted or they over-erupt (when they don't have an antagonising tooth), which makes these characteristics inapplicable. Therefore, eruption was not included in the criteria for MR specific staging.

Tooth development and assessability

It can be considered a limitation of the current study that both participants with and without third molar impaction, agenesis or extraction of other teeth were included to generate the model for age estimation. However, in the general population several patterns of agenesis/extraction/impaction are present and it is not feasible to take all different patterns into account for age estimation. It has been stated that agenesis and impaction might delay third molar development[11, 25-31], while extractions might accelerate it.[32, 33] It is our intention in future research to study this on MRI in the current study population.

In about 90% of upper third molars the palatal root was the least developed one, meaning it was either less developed or equally developed as the buccal roots. Also in about 90% of lower third molars the distal root was the least developed. These numbers are lower than reported based on a subset of the current study population (98% and 95% respectively).[10] In some cases not all roots could be assessed on MRI. This was more frequent in the upper jaw, mainly due to the small dimensions of the buccal roots, as previously reported by De Tobel at al. (2017).[10] Few studies reported on the relative development of different roots within the same third molar.[29] In any case, the least developed root should be considered, to grant the benefit of the doubt.

Baumann et al. (2015) mentioned that 5% (15/307) of molars could not be assessed, due to technical reasons (e.g. motion artefacts), equally distributed among upper and lower jaw. However, this also included first and second molars. In the study by Guo et al. (2015) 2% (13/530) of lower third molars could not be assessed due to due to insufficient image quality. In the current study slightly higher numbers of unassessable teeth were encountered with 5% (51/1096) of upper and 7% (70/1044) of lower third molars. Whether this should be attributed to the MRI scanner and/or MRI sequence used, could be subject of future studies. However, the used sequence proved to be the most suitable after a selection process in De Tobel et al. (2017).[9]

Essential plane

Regarding the plane in which slices are deemed suitable for stage allocation, De Tobel et al. (2017) reported that sagittal slices were essential in 99% of cases.[10] They contributed this to the anatomy of third molars. This corresponds with the current number of 94%. Coronal and axial slices were less frequently useful in the current study (4% and 0% respectively) than in the previous study (11% and 8% respectively).[10]

Statistical approach to age estimation

Reproducibility of staging

It has been stated that reproducibility depends on the staging technique.[34] Inter- and intra-observer agreement in the current study were similar to or lower than those reported in previous studies on third molar development, as seen on 3D imaging modalities (Table 11).[6, 7, 10, 35, 36] It appears that studies including a larger sample of staged molars had relatively lower reproducibility values. Possibly, more easy to stage cases ended up in the small subsamples, used for reproducibility calculations. Furthermore, one might expect that staging based on computed tomography (CT) would be more reproducible than based on MRI, since MRI is more prone to artefacts and is more influenced by surrounding tissue and motion. From Table 11 however, it is clear that staging on MRI shows similar reproducibility as on CT. Although the presented staging technique did not outperform the established techniques (developed on radiographs), one could question the suitability of the Demirjian technique, since it is based on criteria that cannot be visualised with MRI, and the Köhler technique, since it is based on predictions of root lengths which are highly variable.[18]

Although intra- and inter-observer agreement was high, a substantial proportion of

disagreement remained (Tables 7 and 8). Two striking observations can be made: two stage differences occurred and systematically a higher stage was allocated during the second staging session. Both observations might be explained by a learning curve. In the first session, more often when observers doubted about the appropriate stage, the lower stage was allocated. Possibly they were more confident during the second session, with the first observer being more confident than the other. However, since some degree of uncertainty remained, more divergent results were obtained, compared to the first session. The learning effect was present in both observers, although observer 1 was more experienced than observer 2. An explanation might be that although observer 1 had seen more teeth on MRI than observer 2 and he had staged some series of third molars on MRI and panoramic radiograph for previous research, he had never staged a series this large. As was stated by other researchers, more experienced observers generate more consistent results.[37-41] An alternative explanation for two stage differences was when the considered third molar was depicted over several slices. In those cases one observer might have been more conservative and allocate the lower stage because in most slices that seemed appropriate. By contrast, the other observer might have reasoned that the root was in a higher stage, incorporating the slice thickness. When age estimation is done in practice, it is wise to assess the images with at least two observers and allocate stages in consensus.

Finally, also calibration of the observers influences reproducibility. Both observers were from the same research group and were trained in a similar way. It would be useful to see which results would be obtained by an independent observer, e.g. someone from another research group. Anyway, future research is necessary to verify the reproducibility of the proposed staging technique.

The only way to eliminate inter- and intra-observer variability is to conduct automated age estimation. Urschler et al. (2015) reported promising results on automated age estimation based on hand and wrist MRI.[42] Whether or not this approach can be extrapolated to other anatomical structures, such as third molars, is subject of further research.[43]

Table 11: Reproducibility of staging third molar development based on 3D imaging modalities.

Reference	Imaging modality	Elements	Staging technique	Intra-observer agreement		Inter-observer agreement	
				Statistic	N	Statistic	N
Baumann (2015)[6]	MRI	All molars	Demirjian	-	-	Cohen's kappa	0.51	312
Guo (2015)[7]	MRI	Lower left third molar	Demirjian	Kappa	0.89	60	Kappa	0.83	60
De Tobel (2017)[10]	MRI	All third molars	Demirjian	ICC	0.94-0.97	48-50	ICC	0.85-0.94	44-47
			Köhler	ICC	0.96-0.97	48-50	ICC	0.86-0.95	44-47
Current	MRI	All third molars		Weighted kappa	0.80-0.89	350-379	Weighted kappa	0.76-0.80	899
				Agreement	0.66-0.79	350-379	Agreement	0.77-0.87	899
Bassed (2011)[35]	CT	Lower third molars	Demirjian	Kappa	0.949	25	Kappa	0.842	25
				Agreement	0.96	25	Agreement	0.88	25
Cantekin (2013)[36]	CBCT	Lower third molars	Demirjian	Kappa	0.896	70	Kappa	0.692	70

(CB)CT = (cone beam) computed tomography; ICC = intra-class correlation coefficient; MRI = magnetic resonance imaging; N = number of assessed molars.

Age estimation

Baumann et al. (2015) demonstrated that compared with staging on MRI, slightly lower stages were allocated to the same third molar on panoramic radiographs.[6] Guo et al. (2015) found that the minimum age for a fully mature 38 on MRI was 19.57 years for females and 17.77 years for males.[7] Consequently, it seems that a mature third molar 38 on MRI could act as a sign of adulthood in females. In the current sample minimum ages were not used to discern minors from adults. Instead probabilities were calculated using the Bayesian model. When homogenous stage pattern 8 is seen, it is highly probable that an individual is over 18 years old (99.89% in females, 99.76% in males). However, around the age of 18, most individuals will have third molars in stage 5 (Figure 12).

With a mean absolute error of 2.0 years in females and 1.7 years in males, age estimation based on MRI of third molars is less accurate than a similar approach based on radiographs which had an overall mean absolute error of 1.13 years (Me = 0.89, IQR 0.44–1.62).[14] Third molar stages of 2513 individuals were included in their Bayesian model. It can be expected that age estimation performance based on MRI would increase when the reference sample would be larger. However, third molars are not routinely scanned with MRI, so retrospective data collection is impossible. Since several research groups are gathering third molar MRI data prospectively, joining forces could generate a more robust age estimation model.

The performance to discern minors from adults was better for males than for females, with specificities of 96% and 73% respectively. The AUC equalled 0.873 for females and 0.949 for males. Based on lower left third molar staging on radiographs, Liversidge and Marsden (2010) reported a specificity of 96% (females and males combined).[44] However, they reported separate statistics for the different stages, since they did not apply statistical modelling to estimate age. In their study AUC was 0.904 (95% CI 0.889–0.919).[44] Based on staging all third molars on radiographs, Thevissen et al. (2010) reported that specificity ranged between 33% and 87%, without obvious better results for either sex, using

country-specific data in a Bayesian model.[45] An AUC of 0.853 was reported in another paper by the same research group.[14]

Because of the inherent inter-individual variability of development, several anatomical structures should be incorporated into the ad-hoc procedure. It has been demonstrated that combining the information of several developing structures increases accuracy of age estimation. [46-52] However, when combinations are used for age estimation, appropriate statistical methods should be used. Simple regression will generate unrealistically narrow prediction intervals. Instead, a Bayesian approach has been demonstrated to be the most suitable statistical method.[16, 53, 54] In view of adding information of other anatomical structures to the ad-hoc procedure used in the current study, the upper age limit of the study population (26 years of age) was higher than in other studies about third molar development (25 years of age[44], 24 years of age[7], 23 years of age[6], 22 years of age[14]).

CONCLUSION

A mere extrapolation of staging techniques for third molar development based on radiographs to MRI was considered inappropriate. Therefore, an MRI specific staging technique was proposed. Reproducibility was similar to other staging techniques. Although embedding this technique into a Bayesian model for age estimation did not outperform established age estimation methods based on radiographs, it opens the perspective of combining developmental MRI information for age estimation. Other anatomical structures can be added to the used third molars model.

REFERENCES

1. Liversidge HM. Timing of human mandibular third molar formation. Ann Hum Biol. 2008;35(3): 294-321.

2. Thevissen PW, Fieuws S, Willems G. Third Molar Development: Evaluation of Nine Tooth Development Registration Techniques for Age Estimations. J Forensic Sci. 2013.

3. Serin J, Rerolle C, Pucheux J, Dedouit F, Telmon N, Savall F, Saint-Martin P. Contribution of magnetic resonance imaging of the wrist and hand to forensic age assessment. Int J Legal Med. 2016;130(4):1121-8.

4. Olze A, Bilang D, Schmidt S, Wernecke KD, Geserick G, Schmeling A. Validation of common classification systems for assessing the mineralization of third molars. Int J Legal Med. 2005;119(1):22-6.

5. Thevissen PW, Fieuws S, Willems G. Third molar development: evaluation of nine tooth development registration techniques for age estimations. J Forensic Sci. 2013;58(2):393-7.

6. Baumann P, Widek T, Merkens H, Boldt J, Petrovic A, Urschler M, Kirnbauer B, Jakse N, Scheurer E. Dental age estimation of living persons: Comparison of MRI with OPG. Forensic Sci Int. 2015;253(0):76-80.

7. Guo Y, Olze A, Ottow C, Schmidt S, Schulz R, Heindel W, Pfeiffer H, Vieth V, Schmeling A. Dental age estimation in living individuals using 3.0 T MRI of lower third molars. Int J Legal Med. 2015;129(6):1265-70.

8. Ottow C, Krämer JA, Olze A, Schmidt S, Schulz R, Wittschieber D, Heindel W, Pfeiffer H, Ribbecke S, Vieth V, Schmeling A. Magnetresonanztomographiestudie zur Altersschätzung von unbegleiteten minderjährigen Flüchtlingen. Rechtsmedizin. 2014;25:12-20.

9. De Tobel J, Hillewig E, Bogaert S, Deblaere K, Verstraete K. Magnetic resonance imaging of third molars: developing a protocol suitable for forensic age estimation. Ann Hum Biol. 2017;44(2):130-9.

10. De Tobel J, Hillewig E, Verstraete K. Forensic age estimation based on magnetic resonance imaging of third molars: converting 2D staging into 3D staging. Ann Hum Biol. 2017;44(2):121-9.

11. Köhler S, Schmelzle R, Loitz C, Puschel K. [Development of wisdom teeth as a criterion of age determination]. Ann Anat. 1994;176(4):339-45.

12. Demirjian A, Goldstein H, Tanner JM. A new system of dental age assessment. Hum Biol. 1973;45(2):211-27.

13. Hillewig E, Degroote J, Van der Paelt T, Visscher A, Vandemaele P, Lutin B, D'Hooghe L, Vandriessche V, Piette M, Verstraete K. Magnetic resonance imaging of the sternal extremity of the clavicle in forensic age estimation: towards more sound age estimates. Int J Legal Med. 2013;127(3):677-89.

14. Thevissen PW, Fieuws S, Willems G. Human dental age estimation using third molar developmental stages: does a Bayesian approach outperform regression models to discriminate between juveniles and adults? Int J Legal Med. 2010;124(1): 35-42.

15. Thevissen PW. Dental age estimation: striving for an optimal approach [Doctoral thesis]. Leuven: Leuven University Press, 2013.

16. Fieuws S, Willems G, Larsen-Tangmose S, Lynnerup N, Boldsen J, Thevissen P. Obtaining appropriate interval estimates for age when multiple indicators are used: evaluation of an ad-hoc procedure. Int J Legal Med. 2016;130(2):489-99.

17. Boldsen JL, Milner GR, Konigsberg LW, Wood JW. Transition analysis: a new method for estimating age from skeletons. In: Hoppa RD, Vaupel JW, eds. Paleodemography: Age Distributions from Skeletal Samples. 1st ed. Cambridge: Cambridge University Press, 2002; p. 73-106.

18. Thevissen PW, Fieuws S, Willems G. Third molar development: measurements versus scores as age predictor. Arch Oral Biol. 2011;56(10):1035-40.

19. European Asylum Support Office (EASO). Practical Guide on Age Estimation, 2nd edition. Malta: in progress, unpublished, 2016.

20. Bracher AK, Hofmann C, Bornstedt A, Boujraf S, Hell E, Ulrici J, Spahr A, Haller B, Rasche V. Feasibility of ultra-short echo time (UTE) magnetic resonance imaging for identification of carious lesions. Magn Reson Med. 2011;66(2):538-45.

21. Manoliu A, Ho M, Nanz D, Dappa E, Boss A, Grodzki DM, Liu W, Chhabra A, Andreisek G, Kuhn FP. MR neurographic orthopantomogram: Ultrashort echo-time imaging of mandibular bone and teeth complemented with high-resolution morphological and functional MR neurography. Journal of Magnetic Resonance Imaging. 2016;44(2):393-400.

22. Idiyatullin D, Corum C, Moeller S, Prasad HS, Garwood M, Nixdorf DR. Dental magnetic resonance imaging: making the invisible visible. J Endod. 2011;37(6):745-52.

23. Corradi F, Pinchi V, Barsanti I, Manca R, Garatti S. Optimal age classification of young individuals based on dental evidence in civil and criminal proceedings. Int J Legal Med. 2013;127(6):1157-64.

24. Altalie S, Thevissen P, Willems G. Classifying stages of third molar development: crown length as a predictor for the mature root length. Int J Legal Med. 2014.

25. Olze A, van Niekerk P, Schulz R, Ribbecke S, Schmeling A. The influence of impaction on the rate of third molar mineralisation in male black Africans. Int J Legal Med. 2012;126(6):869-74.

26. Olze A, Otto A, Tsokos M. Einfluss der Retention auf die Mineralisationsgeschwindigkeit dritter Molaren. Rechtsmedizin. 2012;22(2):110-4.

27. Knell B, Schmeling A. Einfluss der Retention auf die Weisheitszahnmineralisation. Rechtsmedizin. 2010;20(6):469-74.

28. Guo YC, Yan CX, Lin XW, Zhang WT, Zhou H, Pan F, Wei L, Tang Z, Liang F, Chen T. The influence of impaction to the third molar mineralization in northwestern Chinese population. Int J Legal Med. 2014;128(4):659-65.

29. Friedrich RE, Ulbricht C, Ljuba ABvM. The influence of wisdom tooth impaction on root formation. Ann Anat. 2003;185(5):481-92.

30. Gelbrich B, Hirsch A, Dannhauer KH, Gelbrich G. Agenesis of second premolars and delayed dental maturation. J Orofac Orthop. 2015;76(4):338-50.

31. Lebbe A, Cadenas de Llano-Perula M, Thevissen P, Verdonck A, Fieuws S, Willems G. Dental development in patients with agenesis. Int J Legal Med. 2017;131(2):537-46.

32. Halicioglu K, Toptas O, Akkas I, Celikoglu M. Permanent first molar extraction in adolescents and young adults and its effect on the development of third molar. Clin Oral Investig. 2014;18(5): 1489-94.

33. Yavuz I, Baydas B, Ikbal A, Dagsuyu IM, Ceylan I. Effects of early loss of permanent first molars on the development of third molars. Am J Orthod Dentofacial Orthop. 2006;130(5):634-8.

34. Kimmerle EH, Prince DA, Berg GE. Inter-observer variation in methodologies involving the pubic symphysis, sternal ribs, and teeth. J Forensic Sci. 2008;53(3):594-600.

35. Bassed RB, Briggs C, Drummer OH. Age Estimation and the Developing Third Molar Tooth: An Analysis of an Australian Population Using Computed Tomography. J Forensic Sci. 2011;56(5): 1185-91.

36. Cantekin K, Sekerci AE, Buyuk SK. Dental computed tomographic imaging as age estimation: morphological analysis of the third molar of a group of Turkish population. Am J Forensic Med Pathol. 2013;34(4):357-62.

37. Pinchi V, Norelli GA, Caputi F, Fassina G, Pradella F, Vincenti C. Dental identification by comparison of antemortem and postmortem dental radiographs: influence of operator qualifications and cognitive bias. Forensic Sci Int. 2012;222(1-3): 252-5.

38. Hogge JP, Messmer JM, Doan QN. Radiographic identification of unknown human remains and interpreter experience level. J Forensic Sci. 1994;39(2):373-7.

39. Koot MG, Sauer NJ, Fenton TW. Radiographic human identification using bones of the hand: a validation study. J Forensic Sci. 2005;50(2):263-8.

40. Lynnerup N, Belard E, Buch-Olsen K, Sejrsen B, Damgaard-Pedersen K. Intra- and interobserver error of the Greulich-Pyle method as used on a Danish forensic sample. Forensic Sci Int. 2008;179(2-3):242 e1-6.

41. Wittschieber D, Schulz R, Vieth V, Kuppers M, Bajanowski T, Ramsthaler F, Puschel K, Pfeiffer H, Schmidt S, Schmeling A. Influence of the examiner's qualification and sources of error during stage determination of the medial clavicular epiphysis by means of computed tomography. Int J Legal Med. 2014;128(1):183-91.

42. Urschler M, Grassegger S, Stern D. What automated age estimation of hand and wrist MRI data tells us about skeletal maturation in male adolescents. Ann Hum Biol. 2015;42(4):358-67.

43. Stern D, Kainz P, Payer C, Urschler M. Multi-Factorial Age Estimation from Skeletal and Dental MRI Volumes. Med Image Comput Comput Assist Interv. 2017;1-8. Accepted.

44. Liversidge HM, Marsden PH. Estimating age and the likelihood of having attained 18 years of age using mandibular third molars. Br Dent J. 2010;209(8):E13.

45. Thevissen PW, Alqerban A, Asaumi J, Kahveci F, Kaur J, Kim YK, Pittayapat P, Van VM, Zhang Y, Fieuws S, Willems G. Human dental age estimation using third molar developmental stages: Accuracy of age predictions not using country specific information. Forensic Sci Int. 2010;201(1-3):106-11.

46. Bassed RB, Briggs C, Drummer OH. Age estimation using CT imaging of the third molar tooth, the medial clavicular epiphysis, and the spheno-occipital synchondrosis: a multifactorial approach. Forensic Sci Int. 2011;212(1-3):273.e1-5.

47. Garamendi PM, Landa MI, Ballesteros J, Solano MA. 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. 2005;154(1):3-12.

48. Thevissen PW, Kaur J, Willems G. Human age estimation combining third molar and skeletal development. Int J Legal Med. 2012;126(2):285-92.

49. Pinchi V, De Luca F, Focardi M, Pradella F, Vitale G, Ricciardi F, Norelli GA. Combining dental and skeletal evidence in age classification: Pilot study in a sample of Italian sub-adults. Legal Medicine. 2016;20:75-9.

50. Cameriere R, Ferrante L. Age estimation in children by measurement of carpals and epiphyses of radius and ulna and open apices in teeth: A pilot study. Forensic Sci Int. 2008;174(1):60-3.

51. Cameriere R, De Luca S, Biagi R, Cingolani M, Farronato G, Ferrante L. Accuracy of three age estimation methods in children by measurements of developing teeth and carpals and epiphyses of the ulna and radius. J Forensic Sci. 2012;57(5): 1263-70.

52. Baccino E, Ubelaker DH, Hayek LA, Zerilli A. Evaluation of seven methods of estimating age at death from mature human skeletal remains. J Forensic Sci. 1999;44(5):931-6.
53. Prince DA, Kimmerle EH, Konigsberg LW. A Bayesian approach to estimate skeletal age-at-death utilizing dental wear. J Forensic Sci. 2008;53(3): 588-93.
54. Prince DA, Konigsberg LW. New formulae for estimating age-at-death in the Balkans utilizing Lamendin's dental technique and Bayesian analysis. J Forensic Sci. 2008;53(3):578-87.