RÄTTSMEDICINALVERKET

Dnr. N22-90344
13 Okt 2022

Method description for the National Board of Forensic Medicine's medical age assessments; a probability model concerning the 18-year age limit in asylum cases

Background

The National Board of Forensic Medicine's (hereafter known as the Board) choice of method for medical age assessments was selected to meet requirements specified by the Swedish government in the assignment to the Board on 19 May 2016. The assignment states that the Board shall conduct medical age assessments based on current science and proven experience, and that this process must be legally safeguarded. The reasons given for the decision include that children have special rights and that resources intended for children should not benefit adult asylum seekers. Furthermore, the Board must take into account the requirements of scientific robustness, legal compliance and ethics, and ensure age assessments are carried out with respect for the privacy of the individual.

In support of choosing the Board's method ,the National Board of Health and Welfare's publications, *Metoder för radiologisk åldersbedömning [Methods for radiological age assessment] (1), and Åldersbedömning inom ramen för asylprocessen – En etisk analys [Age assessment in the context of the asylum process – An ethics analysis]* (2), as well as the publication *Åldersbedömning med icke-radiologiska metoder [Age assessment using non-radiological methods]* (3) issued by the Swedish Agency for Health Technology Assessment and Assessment of Social Services (SBU) were used. These publications reviewed parts of the scientific literature on medical age assessment. Information on how age assessment is carried out in other European countries was obtained through study visits and discussions, and is continuously monitored in publications by the European Asylum Support Office's (EASO)(4) and through articles in scientific journals. The method was chosen in dialogue with the experts[1] involved in the National Board of Health and Welfare's report *Metoder för radiologisk åldersbedömning*.

1 Carl-Erik Flodmark (Paediatrician, National Board of Health and Welfare), Andreas Cederlund (Dentist, National Board of Health and Welfare), Sandra Diaz Ruiz (Radiologist, Astrid Lindgren Children's Hospital, Karolinska Hospital), Sven Laurin (Radiologist)

Main points of the probability model

- No method currently used for medical age assessment can determine exact chronological age due to individual variations in biological development.
- The National Board of Forensic Medicine's method was developed to provide medical information for the Swedish Migration Agency's overall assessment in cases where the age stated by the asylum seeker is not likely.
- The described probability model represents an assessment of probability relative to the 18-year old threshold based on results from an MRI examination of the knee and plain X-ray examination of the wisdom teeth.
- The probability is expressed in a scale similar to other forensic reports, but is complemented with a margin of error that represents the proportion of the reference population that is misclassified according to the model.
- The assessments of tooth and knee development have two stages: "not reached final stage" and "reached final stage". The binary scale enables stage categorisation relating specifically to the 18-year old age limit rather than an assessment of the exact chronological age.
- This version of the probability model is based on 12 scientific studies and includes data from approximately 21,000 observations. The model will be updated as new data becomes available.
- All data used in the model are observed data of individuals from different studies. As the studies have grouped their data in diverse ways, individual data is not always presented. A mathematical model is therefore used to reproduce the data from these studies.
- Statistical models require assumptions. These affect the model's estimates.
- The uncertainty based on biological variation and the assumptions in the model are presented in this method description and in the statement.
- The impact of regional factors, such as ethnicity on the model is somewhat unclear. To account for variation, a large reference population is employed as it is assumed to be more appropriate for individuals with different backgrounds than a limited small population. The study populations on which the model is based spans nine countries spread over three different continents.

Introduction

The medical age assessments of the Board stems from the requirement for a medical opinion as part of the Swedish Migration Agency's decision-making process to determine whether a person is under or over the age of 18 years. The medical age assessments are based on the examination of two body examinations: wisdom teeth in the lower jaw by means of a plain X-ray (panoramic image) and the distal femur by means of magnetic resonance imaging (hereinafter referred to as MRI of the knee).

Following analysis of the wisdom tooth x-ray and knee MRI images by dentists and radiologists, respectively, the overall assessment is completed using a standardised matrix developed by the National Board of Forensic Medicine. The final assessment results in a medical age assessment of the examined person in relation to the 18-year age limit according to a predefined probability scale.

Dental maturity assessment

In order to minimise the amount of ionising radiation, the dental X-ray shall, as far as technically feasible, be limited to the wisdom teeth, where imaging may be repeated up to three times to achieve sufficient image quality.

Wisdom tooth maturity is assessed by X-ray analysis whereby two dentists assess whether any of the wisdom teeth in the lower jaw has reached stage H according to Demirjian's scale of tooth maturity (5). Stage H is defined as the fully mature stage exhibiting completely closed tooth roots (5). Stage categorisation is performed by two dentists, who both provide the Board an independent and blinded assessment. Each assessment is exported to the Board's case management system and marked with both the Swedish Migration Agency's reference number and the Board's case number. Independent assessment means that each dentist makes an assessment without having access to/knowledge of the other's assessment. Blinded assessment means that they are only have access to the X-rays and no further information. If categorisation is not possible due to either insufficient image quality (despite repeated imaging) or the absence of wisdom teeth, assessors may answer "Not assessable" and suggest the reason for this.

If the dental examination shows that both wisdom teeth in the lower jaw are absent, the assessment will be exclusively based on the degree of maturity of the lower growth zone of the femur. Furthermore, if one assessor indicate "not assessable", the image will be forwarded to a forensic odontologist at the Board, for a third assessment. The forensic odontologist's answer will be included in the age assessment matrix.If however, no assessment of the X-ray is possible, age assessment will again be exclusively based on the degree of maturity of the lower growth zone of the femur.

If wisdom teeth categorisation reveals findings that are deemed to require follow-up/examination within the medical/dental sector beyond routine care, the provider (the organisation charged with analysing the dental x-ray) shall inform the person examined and, where appropriate, the appointed guardian.

In order for the dental examination finding to be "reached final wisdom tooth stage", both dentists must make the assessment that the wisdom tooth has reached the final stage.

If one dentist has assessed that the wisdom tooth has not reached the final stage, the finding "final stage not reached" is sent to the Board's standardised matrix.

Bone maturity assessment

The lower part of the femur is examined by means of MRI, and the maturity of the growth zone is assessed according to the scale of bone maturity defined in Krämer et al. 2014 (6). The Board uses a binary scale ("reached final stage" and "not reached final stage") in its assessment of knee maturity. The Board defines the final stage as a growth zone that has reached at least stage IV, according to Krämer 2014. This stage means that the growth zone in the knee is completely ossified – with or without visible scarring. Consequently, a knee joint has not reached the final stage if it has not reached stage IV–V according to Krämer's definition.

Stage categorisation is conducted by two independent assessors (radiologists) that export their findings to the Board's case management system. Independent assessor means that the two radiologists must make their own assessment of each image, independently of one another. Blinded assessment means that they only have access to the images from the examination and the gender of the person to whom the image refers, but no additional information. If categorisation is not possible due to insufficient image quality, the assessors can indicate "Not assessable".

In order for the final assessment of skeletal maturity to be "final stage reached", both assessors must reach this conclusion independently of one another. If the findings are not in agreement, the assessment will be that the final stage has not been reached. Should the subject not be able to undergo the examination for any reason, such as the presence of metal fragments, the age assessment will be based on the wisdom teeth examination described previously.

If both assessors of the MRI images provide a "Not assessable" response, the dental maturity assessment will be the sole basis for the medical age assessment. If only one assessor indicates "Not assessable", the image will be examined by a third radiologist who is associated with the Board. If this assessment also results in "not assessable", the medical age assessment will be based exclusively on dental maturity; otherwise, the third radiologist's findings will be included in the age assessment matrix. If both the knee and dental examination result are "not assessable", the statement in the overall assessment will be concluded as "not assessable".

If skeletal maturity assessment of the distal femur reveals findings that require a follow-up/examination within the medical sector beyond routine care, the provider (the organisation charged with analysing the X-ray image) shall inform the person examined and, where appropriate, the appointed guardian of the condition.

Technical specification for MRI

The MRI sequence and slice plan used is identical to that used in the study by Krämer et. al. (6), i.e. a T1 weighted sequence of sagittal slices of the knee, continuous slices, 3 mm thick, covering the entire growth plate. The growth plate is assessed in all slices. If any aspect is fully or partially open on any slice, the growth zone is deemed not closed. If the growth zone is closed, with or without scarring (thin band-like changes) on all slices, the knee joint is designated as having reached the final stage.

The following MRI settings are to be used:

Sag T1 weighted.
TR/TE approx. 600-700/17 or similar.
3 mm slice thickness, gap 0.3.
Pixel size 0.5 x 0.5 mm or better.

A minimum field strength of 1.5 provides an acceptable resolution for assessing the growth zone. The disadvantage of higher field strength (3.0 Tesla) is that it is more sensitive to movements.

Factors that can affect skeletal and dental maturity

Factors such as socioeconomic status, stress and ethnicity have been suggested to affect skeletal maturity and tooth maturity. It is possible that socioeconomic factors may have an impact on skeletal maturity, as poorer living conditions cause bone to mature more slowly, resulting in underestimation of chronological age (7). Publications discussing medical age assessments using knee MRI include sample populations that are primarily of a good socioeconomic status, while those that will be subjected to the age assessment using MRI of the knee may have either the same or a lower socioeconomic affiliation. For this reason, it is reasonable to assume that applicants may be considered younger than their chronological age, but that it cannot be ruled out that these factors may have a different impact on certain populations or individuals.

One study suggests that increased stress may lead to earlier puberty for girls and possibly also for boys (8), but the impact on age assessment by means of skeletal and dental maturity was not investigated in this study (9).

Thevissen et al. 2010 (10) investigated how population-specific data from nine different countries can be used to reduce the misclassification rate in age assessments around the age of 18 when wisdom teeth are examined. This study concluded that the availability of population-specific data will only lead to a negligible reduction in the misclassification rate. It should therefore be possible to use a common set of tables for age assessment in different populations. The proportion of children misclassified as adults was shown to be lowest when a Belgian population was used as reference for other populations, but with a higher proportion of adults misclassified as children.

When populations of different ethnicities have been studied with respect to wisdom teeth maturity, differences have been noted in the earlier stages of maturity. However, no significant differences in the final fully mature stage of wisdom teeth have been observed between the studied populations (11,12). Thus, ethnicity appears to have less impact on age assessment when based on the fully mature stage of wisdom teeth.

According to a literature review by Schmeling et al. 2000 (13) on how ethnicity affects skeletal maturity and age assessment, the reference populations data can be used regardless of the subject's ethnicity. This is because ethnicity does not impact the age at which the bones mature in a manner significant to the assessment. On the other hand, a review article published in 2019 with calculations on meta-populations for age assessments from wrist x-rays, and another from 2021 on the use of dental x-rays warn of differences in skeletal or dental maturity in populations with different ethnicities (14,15). However, it can be noted that differences between populations attributed to ethnicity may be a consequence of the age distribution in the study population, as discussed in two systematic review articles by Dahlberg et al. and Rolseth et al. (16,17).

It is important to note that considerable variation exists between studies and more research along with careful analysis is needed to better examine the complex relationship between skeletal maturity, dental maturity, and factors such as ethnicity, socioeconomic status and stress. Nevertheless, the state of knowledge on how these factors influence the maturation process is considered sufficient for age assessment based on skeletal and dental maturity to be used as a part of an overall assessment.

It should also be mentioned that the distribution in dental and skeletal maturity within the studied populations is likely to include the above-mentioned influencing factors to some extent.

The medical age assessment model

When assessing age relative to the 18-year age threshold the overall results from the knee and dental examinations are used. The Board's model for medical age assessment is designed to lower the risk of incorrectly assessing children as adults, at the expense of a greater risk of failing to identify adults who are slightly, to a few years older than 18 as adults. Girls have an earlier skeletal maturity compared to boys, and the matrix is adapted to accommodate these gender differences.

In order to weigh results from different studies, a statistical model has been developed to assess whether a person is above or below 18 years of age at a given stage. The model is based on the scientific studies presented in Table 1. The studies included in the statistical model have been selected on the basis that they are consistent with the selected methodology for knee and tooth analysis. . The studies included must have used the same X-ray method (dental) and MRI sequence (knee), stage categorisation for both tooth/knee compatible with binary classification, present complete data (allowing the recreation of individual-based data and not previously published), and have a focus on age assessment related to the 18-year age threshold.

Table 1. Selected scientific studies

Indicator	Study	Males, number	Females, number	Country	Age range (years)	Method**
Knee	(18)	325	333	Germany	12–24	MRI
Knee	(6)	166	124	Germany	10–30	MRI
Knee	(19)*	220	180	Sweden	14–21	MRI
Knee	(20)	214	0	France	14–20	MRI
Knee	(21)	335	314	Turkey	10–29	MRI
Tooth	(22)	271	323	USA/Canada	14–24	RX
Tooth	(23)	1610	1691	South Korea	4–26	RX
Tooth	(24)	379	405	Turkey	8–23	RX
Tooth	(25)	1200	1900	China	4–26	RX
Tooth	(26)	540	540	Malaysia	11–25	RX
Tooth	(27)	258	347	Canada	11–29	RX
Tooth	(28)	1551	1661	China	5–25	RX

** The part of the study based on T1 sequence, called a bridge study*

***MRI= Magnetic resonance imaging, RX = radiological X-ray*

The model is based on observed data from individuals in the studies (from table 1). The studies however, have grouped the data in different ways, which means that data for individuals are not always presented. A mathematical model is therefore used to reproduce individual-based data where developmental stages are dichotomised according to the Board's binary classification, and related to chronological age. Reproduced individual-based data from all the studies are then compiled into metapopulations for both knee and tooth and male and female.

By developing the model in two steps, first starting from individual-based data and then performing a logistic regression analysis (stage given age), age mimicry (how the age distribution of those included in the study affects the results) is reduced, compared to directly using for example a reported mean/standard deviation of chronological age for the different stages (29–31). See the fitted logistical regression model with coefficient values in Appendix, Tables A.2 and A.3, and Figure A.1. The binary scale of final stage/not final stage is adapted so that the stage categorisation relates to the specific 18-year age threshold rather than an assessment of the exact chronological age.

A weighted estimate of probabilities and margins of error can be made by combining and transforming the calculated probabilities for knee and tooth using Bayes' theorem. This transformation involves reversing the conditional probability so that it is the odds of being over or under 18 years based on developmental stage that is provided (age given stage). The model presumes a uniform age distribution and an age interval between 15–21 years. This range includes the population close to 18 years of age, and at least 50% of the population has reached closed tooth or knee for the upper boundary (21 years), a limit used in similar statistical models (31). The actual age range of those undergoing medical age assessments is not known, however, the selected age range represents ages in which medical age assessment may be relevant. Since the assumption of age range affects the margin of error, it is important from a children's rights perspective to not set the upper limit of the age range too high. The higher the upper limit of the age range, the greater the risk is of misclassifying children. Conversely, the lower the upper limit, the greater the risk of misclassifying adults. The Board finds that an age range ending at 21 years is a reasonable trade-off, as it implies an equal probability that the individual is over or under 18 years of age. The choice of a uniform distribution was made based on its mathematical properties, in that probabilities and margins of error are then based on the study population with known tooth/knee age and stage of development. The developmental stages of the individual asylum seeker are thus compared with data from a population of individuals with known age between 15 and 21 years.

How knee and tooth development covary is not fully known and few studies have examined the covariation between different maturity indicators. A couple of studies investigating wrist and tooth covariation failed to demonstrate any correlation between their development (28, 29). Another study investigating wisdom tooth and knee development using MRI for both, found no significant covariation either (34). Thus, it can be assumed that bone maturation in the knee and wisdom teeth development proceeds independently of each other.

Further information detailing how different assumptions affect the model can be found under the heading "Uncertainties with the selected model and risk minimisation".

Assessment relative to the 18-year age limit and margins of error for boys/men

According to the literature, half the male population reach a final knee or wisdom tooth development stage after 18 years of age. We estimate this age to be 18.7 years for knees and 20.3 years for wisdom teeth based on the logistical regression analysis. In the model, male individuals are assessed as being above 18 years of age when one or both body parts have reached their final stage, but different degrees of certainty are used in the final report based on the probabilities of different combinations of development in the two body parts. The estimated probability of adults being identified as adults is 78% (sensitivity) and the corresponding figure for children being identified as children is 90% (specificity) for males/boys when the model based on the two body parts is used. This means that the model is better at identifying children than adults but also means that 80% (negative predictive value) of those identified as children are children and 89% (positive predictive value) of those identified as adults are adults; see Appendix Table A.1.

Calculations show that approximately 10% of all boys (aged 15.0–17.9) are expected to have reached the final stage in one or both body parts and are therefore at risk of being misclassified. Considering a population of males between 15 and 21 years of age, where one or both body parts have reached the final stage, the margin of error with the model corresponds to the probability that individuals with the examination finding in question are below 18 years of age, and this is estimated to be 11%. The statement is given with different degrees of certainty, based on the combination of development in the two body parts. These different degrees of certainty are reflected in the so-called probability scale used in forensic medical reports and other statements issued by the National Board of Forensic Medicine on a regular basis.

For male subjects for whom both wisdom tooth and knee reached the final stage, the certainty grade *"strongly indicates that the person examined is 18 years of age or older"* is used. This combination is uncommon in children aged 15.0–17.9 years, and is estimated to be less than 1%. In young adults up to 21 years of age (18.0–20.9), the combination occurs in approximately 28%. This means that the probability that individuals with the examination finding in question are under 18 years of age is approximately 2%, which corresponds to the margin of error. The margin of error for each one-year cohort is shown in Figure 1 (a).

The certainty grade *"indicates that the person examined is 18 years of age or older"* is used when the knee has reached the final stage, regardless of whether the wisdom tooth cannot be assessed or has not reached the final stage. This certainty grade is also used when the wisdom tooth has reached the final stage but the knee is not assessable. The probability that individuals are under 18 years of age with examination findings, wisdom tooth has reached the final stage but the knee is not assessable, is about 10%, which corresponds to the margin of error. For examination findings showing that the wisdom tooth has not reached the final stage, but that the knee has, the corresponding margin of error is approximately 12%. For the combination of the knee having reached the final stage, but the tooth not being assessable, the margin of error is estimated at approximately 8%. The margin of error for each one-year cohort is shown in Figure 1 (a).

The certainty grade *"possibly indicates that the person examined is 18 years of age or older*" is used for the combination where the wisdom tooth has reached the final stage but not the knee. This combination, which is uncommon in both the 15.0–17.9 age group (about 3%) and the 18.0–20.9 age group (about 6%), has a higher margin of error of about 35% and hence a lower scale step.

Boys in whom one body part has not reached the final stage and the other part cannot be assessed, as well as the boys in whom neither of the body parts have reached the final stage, are assessed with the model as being under 18 years of age. The certainty grade used is *"possibly indicates that the person examined is under 18 years of age"*. The probability that individuals are older than 18 years when examination findings show the knee to not have reached the final stage and the wisdom tooth as not assessable approximately 23%. The probability that neither the wisdom tooth nor the knee has reached the final stage is approximately 20%, which corresponds to the margins of error in the two groups. In cases where the findings indicate that wisdom tooth has not reached its final stage and the knee cannot be assessed, the probability that the individual is older than 18 years is approximately 41%, which corresponds to the margin of error. The margin of error for each one-year cohort is shown in Figure 1 (a).

Assessment relative to the 18-year age limit and margins of error for girls/women

Based on the scientific evidence, the age at which half of the female population reaches the final stage is before the age of 18 for the knee, but after the age of 18 for wisdom teeth. Half the population reached the final stage at 17.3 years of age for the knee and at 21.1 years of age for the wisdom teeth, according to the logistical regression analysis. This means that a knee that has not reached the end stage is informative for underage, but a knee that has reached the end stage is less informative.

If the model with both tooth and knee is used for women, the estimated probability of adults being identified as adults is 94% (sensitivity) and of children being identified as children is 68% (specificity). In contrast, the estimated probability of adults being identified as adults is 24% (sensitivity) and the corresponding probability of children being identified as children is 97% (specificity) for girls/women - when the model is based on tooth only. This means that the tooth only model is much better at correctly identifying children than adults. The assessment matrix for females has been therefore been adapted and is based on wisdom tooth development only. Knee development is used only in cases where the knee has not reached the final stage. This adaptation means that a relatively large proportion of young women will be incorrectly assessed as children; see Appendix Table A.1.

For women/girls different degrees of certainty are used in forensic statements of age based on the combination of development achieved in the body parts examined. These different degrees of certainty are reflected in the so-called probability scale used in forensic medical reports and other statements issued by the National Board of Forensic Medicine on a regular basis.

For female subjects for whom both wisdom tooth and knee have reached the final stage, or for whom the tooth has reached the final stage but no data for the knee is available, the certainty grade *"indicates that the person examined is 18 years of age or older"* is used. These combinations are uncommon in girls between 15.0–17.9 years of age, and are estimated to be approximately 1% and 3%, respectively. In young females up to 21 years of age (18.0–20.9), these combinations occur in approximately 23% and 24%, respectively, based on the scientific evidence. This means that the probability that individuals with the examination findings in question are under 18 years of age is approximately 6% and 12%, respectively, which corresponds to the margin of error. The margin of error for each one-year cohort is shown in Figure 1 (b).

As the knee has reached a final stage, and the wisdom tooth has either not reached the final stage or is not assessible, an age assessment relative to the 18-year age limit cannot be done.

The certainty grade "*indicates that the person examined is under 18 years of age*" is used for female applicants when both the tooth and knee joint have not reached the final stage, or when the knee has not reached the final stage and the wisdom tooth is not assessable. The probability that individuals with the examination findings in question are older than 18 years of age is approximately 8% and 9%, respectively, which corresponds to the margins of error. The margin of error for each one-year cohort is shown in Figure 1 (b).

The certainty grade *"possibly indicates that the person examined is under 18 years of age*" is used when the wisdom tooth has not reached the final stage but data for the knee joint is missing, or when the wisdom tooth has reached the final stage but the knee has not reached the final stage. The probability that individuals with the examination findings in question are over 18 years of age is approximately 44% and 34%, respectively, which corresponds to the margins of error. The margin of error for each one-year cohort is shown in Figure 1 (b).

Uncertainties of the selected model and risk minimisation

The starting point for medical age assessment is an individual of unknown chronological age. Due to individual variations in the biological maturation processes, it is not possible to translate biological maturation to an exact chronological age. The National Board of Forensic Medicine's probability model is based on scientific studies in which biological maturation based on defined developmental stages has been studied in relation to chronological age. The assessment of which developmental stage has been reached plays a crucial role in medical age assessment and therefore represents a risk of error that is important to limit. Statistical models require certain assumptions that introduce uncertainties that need to be considered.

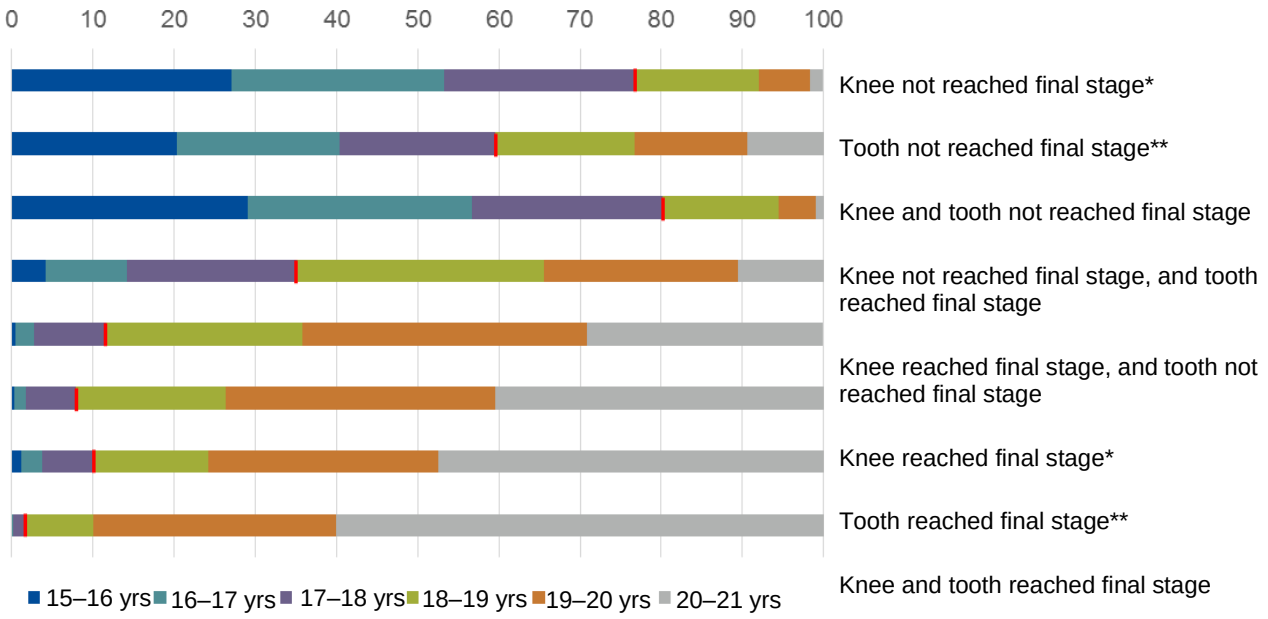
Uncertainties in the probability model

For the statistical probability model, three important assumptions are made. One of these is that knee and tooth development take place independently of each other, based on studies of bone development in the wrist or knee and tooth, as discussed in the section "The medical age assessment model".

The second assumption is to use a uniform age distribution. The third assumption is to select a relevant predetermined age range in the calculations, where the group of children is represented by the ages 15.0–17.9 and the group of young adults is represented by the ages 18.0–20.9. The use of a uniform age distribution means that no mathematical weighting is added when the Bayesian transformation of the conditional probability takes place. The results are instead determined by the logistical regression of the study population.

Selecting an age range and distribution is required in order to make a probability calculation based on the known developmental combination, and the reason for choosing the specific age range is given in the section "The medical age assessment model". The calculated margin of error can also be presented in full year intervals for the selected age range, which gives a picture of how the uncertainty is distributed in the age range, Figure 1.

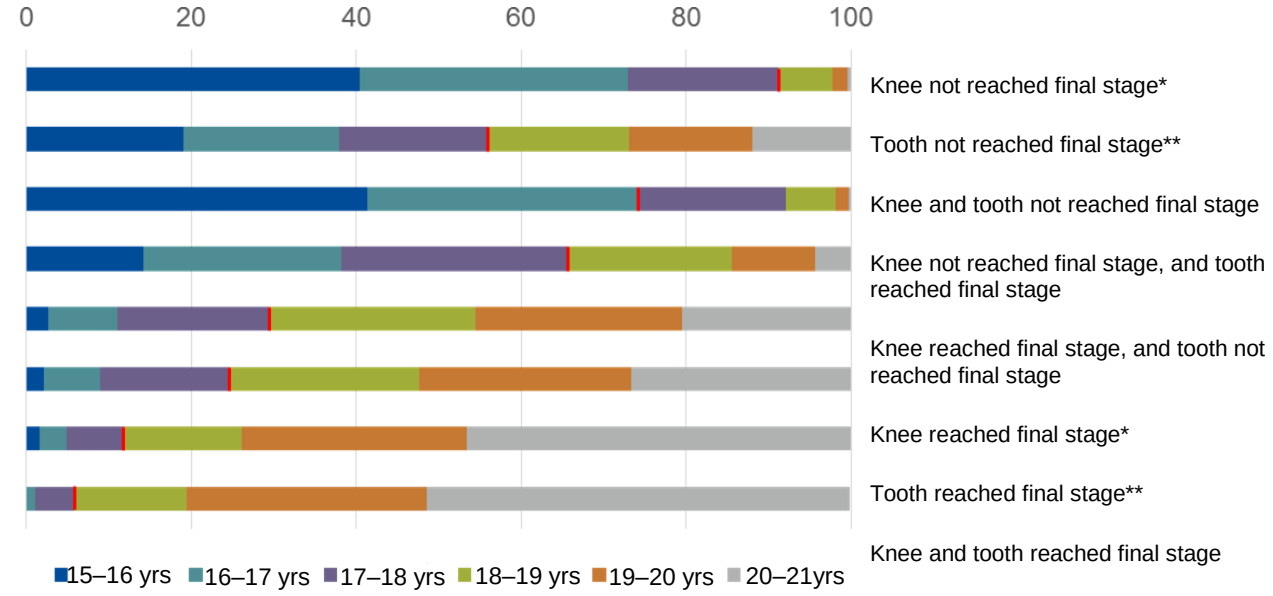
Figure 1 (a). Age by maturity combinations, males



** No information for tooth*

** ** No information for knee*

Figure 1 (b). Age by maturity combinations, females



** No information for tooth*

*** No information for knee*

Using a non-uniform age distribution affects the probabilities and margins of error in the model. On a theoretical level, the proportion of incorrectly assessed children and adults on a population basis could be reproduced more accurately if the true age distribution could be used instead of an even age distribution. However, if it is based on an estimated age distribution of those who applied for asylum in the past, it requires an assumption that the age distribution of those who apply for asylum in Sweden and are subject to medical age assessment is stable over time, which is not necessarily the case. Using such an age distribution would imply that the probability assessment for the individual asylum seeker was influenced by the population that had previously undergone medical age assessment, which is not reasonable. Table 2 (a and b) shows how the model is affected by other age distribution assumptions, where the probability of being under and over 18 years of age is reported for different combinations of examination findings. The three columns show the current assumption in the model with an *uniform age distribution,* i.e. that the comparative population is evenly distributed over the ages; and one with *a normal distribution centred around 18 years,* i.e. that the comparative population is dominated by individuals around 18 years of age and consists of an equal proportion of individuals who are slightly to a few years younger than 18 years of age; and one with a *mixed distribution,* where the age distribution starts as a normal distribution before 18 years of age and is then uniform, i.e. where the comparative population is assumed to consist of a uniformly distributed proportion of 18, 19 and 20 year olds, but fewer 17 year olds, and even fewer 16 year olds, etc.

In the final statement, the models estimated probabilities of being over or under 18 years of age for each developmental combination are expressed in words with a probability scale (certainty grades), along with the proportion in the reference population that is incorrectly assessed according to the model. The tables below show that although assumptions concerning the age distribution affect the probabilities, the probability would be described using the same certainty grades in the probability scale.

Table 2 (a) Males

	Uniform distribution Age range 15–21		Normal distribution centred around age 18 Age range 15–21		Mixed distribution* Age range 15–21		Outcome in the model with selected scale step
Combination	**<18 years old (%)**	**>18 years old (%)**	**<18 years old (%)**	**>18 (%)**	**<18 (%)**	**>18 (%)**	
Knee reached final stage (no info for tooth)	8	92	12	88	7	93	>18 Indicates
Tooth reached final stage (no info for knee)	10	90	15	85	8	91	>18 Indicates
Knee and tooth not reached final stage	80	20	75	25	72	28	<18 Possibly indicates
Knee and tooth reached final stage	1	99	3	97	1	99	>18 Strongly indicates
Knee reached final stage, and tooth not reached final stage	12	88	15	85	11	89	>18 Indicates
Knee not reached final stage, and tooth reached final stage	35	65	37	63	31	69	>18 Possibly indicates
Knee not reached final stage (no info for tooth)	77	23	71	29	68	32	<18 Possibly indicates
Tooth not reached final stage (no info for knee)	59	41	57	43	49	51	<18 Possibly indicates

* *Normal distribution before the age of 18 and thereafter even distribution*

Table 2 (b) Females

	Uniform distribution Age range 15–21		Normal distribution centred around age 18 Age range 15–21		Mixed distribution* Age range 15–21		Outcome in the model with scale step
Combination	**<18 years old (%)**	**>18 years old (%)**	**<18 years old (%)**	**> 18 years old (%)**	**<18 years old (%)**	**>18 years old (%)**	
Knee reached final stage (no info for tooth)	*25*	*75*	*30*	*70*	*28*	*72*	Does not allow assessment**
Tooth reached final stage (no info for knee)	12	88	17	83	13	87	>18 Indicates
Knee and tooth not reached final stage	92	8	88	12	90	10	<18 Indicates
Knee and tooth reached final stage	6	94	10	90	7	93	>18 Indicates
Knee reached final stage, and tooth not reached final stage	*29*	*71*	*34*	*66*	*32*	*67*	Does not allow assessment**
Knee not reached final stage, and tooth reached final stage	66	34	63	37	66	34	<18 Possibly indicates
Knee not reached final stage (no info for tooth)	91	9	87	13	89	11	<18 Indicates
Tooth not reached final stage (no info for knee)	56	44	54	46	54	46	<18 Indicates

** Normal distribution before age 18 and thereafter even distribution ** Based on knee reaching final stage is not informative for females as it occurs before age 18 in about half of all females.*

Stage categorisation and uncertainties in the assessment of MRI and X-ray images

The method for medical age assessments involves the review and assessment of MRI images and X-rays radiologists and dentists, respectively. All assessments are associated with a certain degree of uncertainty. In order to reduce uncertainties, the Board uses assessors with extensive experience in assessing MRI images and categorising wisdom teeth stages, respectively. The qualified assessors included in the agreement undergo training and calibration for the Board's method.

Demirjian's and Krämer's scales for the development of wisdom teeth and knees, respectively, originally comprise eight (A-H) and five (1-5) stages of development, respectively. In the National Board of Forensic Medicine's model, all stages of development, apart from the last in relation to the wisdom tooth (Stage H), and the last two in relation to the knee (4 and 5) are considered as "not reached the final stage". This choice simplifies the assessments, resulting in higher consistency and more reliable assessments. If the model were to include all stages of development instead, it would be possible to make narrower estimates at non-mature stages. However, a greater number of stages reduces the consistency between assessors and hence the reliability of the stage categorisation. Therefore, to minimise the risk of introducing errors early in the process, the assessment focuses solely on whether the wisdom tooth or knee has reached the final stage or not, referred to as a binary scale.

The binary scale means that individuals exhibiting developmental stages that are relatively strongly associated with being an adult (stage G for the wisdom tooth) receive the same assessment as an individual with a significantly less mature developmental stage. This results in a slightly higher proportion of young adults being incorrectly assessed as children, but also limits to some extent the ability of the model to make statements with a higher degree of certainty that the individual being examined is under 18 years of age – in cases with less advanced developmental stages.

The assessments are made independently and blinded as described in the sections "Dental maturity assessment" and "Skeletal maturity assessment". A procedure with a third assessor, if necessary, is also incorporated into the process, as described in the sections indicated above. The assessments are continuously followed up in consistency checks, where any discrepancies can be detected. The risk to the individual asylum seeker is also minimised by giving precedence to an assessment that the final stage has not been reached in cases where the assessors come to different conclusions regarding the stage of development. Overall, the process includes an increased level of certainty regarding image assessments compared to routine healthcare to minimise uncertainties and increase legal compliance.

In order to avoid similar cases (cases with the same examination findings) being assessed differently (being issued different statements), an assessment aid in the form of a matrix is used, where there is a predetermined outcome for each combination of developmental stages and gender. A review of the outcome is carried out in each case by a forensic pathologist trained in the method before the statement is issued.

Uncertainties with the model are reflected in the statements

To reflect the model's uncertainties in statements, a probability scale similar to that used in forensic medical reports and other statements issued by the National Board of Forensic Medicine is used. Statements indicate the probability that an individual in an evenly

distributed mixed population of both adults and children (15–21 years of age) with current examination findings is over or under 18 years of age, where the probability is reflected by the scale step.

What the calculated probabilities would look under other assumptions about the age distribution is described above in the section *Uncertainties in the probability model*. The uncertainty is also given in the form of a calculated margin of error, which refers to the group of children with chronological age close to 18 years, i.e. aged between 15.0–17.9, and young adults aged between 18.0-20.9 respectively, and corresponds to the proportion of the reference population that is misclassified by the model.

Probability scale for medical age assessments

The probability stages correspond to the certainty grades in forensic medical reports and other statements issued by the Board on a regular basis. In medical age assessments, the certainty grade is expressed in the individual statements together with the margin of error, based on the examination findings.

The assessment of the conducted examination **shows** that the examined individual is 18 years of age or older

The certainty grade is not used.

The assessment of the conducted examination **strongly indicates** that the examined individual is 18 years of age or older.

The certainty grade is only used for male applicants.

The assessment of the conducted examination **indicates** that the examined individual is 18 years or older

The certainty grade is used for both female and male applicants.

The assessment of the conducted examination **possibly indicates** that the examined individual is 18 years of age or older

The certainty grade is used for both female and male applicants.

The assessment of the conducted examination **possibly indicates** that the examined individual is under the age of 18

The certainty grade is used for both female and male applicants.

The findings of the conducted examination **indicate** that the individual examined is under the age of 18.

The certainty grade is only used for female applicants.

The findings of the conducted examination do not **allow assessment** of the age of the person examined in relation to the 18-year age limit.

The certainty grade is only used for female applicants.

References

1. The National Board of Health and Welfare. Metoder för radiologisk åldersbedömning: En systematisk översikt. 2016 p. 88.

2. The National Board of Health and Welfare. Åldersbedömning inom ramen för asylprocessen – En etisk analys. National Board of Health and Welfare, 2016.

3. SBU. Åldersbedömning med icke-radiologiska bedömningsmetoder. SBU; 2016.

4. European Asylum Support Office. EASO Practical Guide on Age Assessment: Second edition. 2018.

5. Demirjian A, Goldstein H, Tanner JM. A New System of Dental Age Assessment. Human Biology. May 1973;45(2):18.

6. Krämer JA, Schmidt S, Jürgens KU, Lentschig M, Schmeling A, Vieth V. Forensic age estimation in living individuals using 3.0T MRI of the distal femur. Int J Legal Med. May 2014;128(3):509–14.

7. Schmeling A, Schulz R, Danner B, Rösing FW. The impact of economic progress and modernization in medicine on the ossification of hand and wrist. Int J Legal Med. March 2006;120(2):121–6.

8. Axelsson I. Bone maturation cannot be used to estimate chronological age in asylum-seeking adolescents. Acta Paediatr. 07 December 2018;apa.14645.

9. Haglund M, Garamendi González PM. A problematic dismissal of forensic age estimation. Acta Paediatr. 09 May 2019;apa.14839.

10. 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. January 2010;124(1):35–42.

11. Olze A, Schmeling A, Taniguchi M, Maeda H, van Niekerk P, Wernecke KD, et al. Forensic age estimation in living subjects: the ethnic factor in wisdom tooth mineralization. Int J Legal Med. June 2004;118(3):170–3.

12. Cameriere R, De Luca S, Ferrante L. Study of the ethnicity's influence on the third molar maturity index (I3M) for estimating age of majority in living juveniles and young adults. Int J Legal Med. 23 May 2021;

13. Schmeling A, Reisinger W, Loreck D, Vendura K, Markus W, Geserick G. Effects of ethnicity on skeletal maturation: consequences for forensic age estimations. International Journal of Legal Medicine. 18 August 2000;113(5):253–8.

14. Alshamrani K, Messina F, Offiah AC. Is the Greulich and Pyle atlas applicable to all ethnicities? A systematic review and meta-analysis. Eur Radiol. June 2019;29(6):2910–23.

15. De Donno A, Angrisani C, Mele F, Introna F, Santoro V. Dental age estimation: Demirjian's versus the other methods in different populations. A literature review. Med Sci Law. January 2021;61(1_suppl):125–9.

16. Dahlberg PS, Mosdøl A, Ding Y, Bleka Ø, Rolseth V, Straumann GH, et al. A systematic review of the agreement between chronological age and skeletal age based on the Greulich and Pyle atlas. Eur Radiol. June 2019;29(6):2936–48.

17. Rolseth V, Mosdøl A, Dahlberg PS, Ding Y, Bleka Ø, Skjerven-Martinsen M, et al. Age assessment by Demirjian's development stages of the third molar: a systematic review. Eur Radiol. May 2019;29(5):2311–21.

18. Ottow C, Schulz R, Pfeiffer H, Heindel W, Schmeling A, Vieth V. Forensic age estimation by magnetic resonance imaging of the knee: the definite relevance in bony fusion of the distal femoral- and the proximal tibial epiphyses using closest-to-bone T1 TSE sequence. Eur Radiol. December 2017;27(12):5041–8.

19. The National Board of Health and Welfare. Om magnetkamera vid bedömning av ålder: En studie av validiteten i radiologisk undersökning. 2018 p. 75.

20. Saint-Martin P, Rérolle C, Pucheux J, Dedouit F, Telmon N. Contribution of distal femur MRI to the determination of the 18-year limit in forensic age estimation. Int J Legal Med. May 2015;129(3):619–20.

21. Ekizoglu O, Er A, Bozdag M, Basa CD, Kacmaz IE, Moghaddam N, et al. Forensic age estimation via magnetic resonance imaging of knee in the Turkish population: use of T1-TSE sequence. Int J Legal Med [Internet]. 24 August 2020 [cited 6 November 2020]; Available at: http://link.springer.com/10.1007/s00414-020-02402-0

22. Mincer HH, Harris EF, Berryman HE. The A.B.F.O. Study of Third Molar Development and Its Use as an Estimator of Chronological Age. J Forensic Sci. 01 March 1993;38(2):13418J.

23. Lee SH, Lee JY, Park HK, Kim YK. Development of third molars in Korean juveniles and adolescents. Forensic Science International. July 2009;188(1–3):107–11.

24. Karadayi B, Kaya A, Afsin H, Ozaslan A, Çetin G. The usage of third molars to determine legally relevant age thresholds in Turkey. Australian Journal of Forensic Sciences. 03 July 2015;47(3):275–82.

25. Zeng DL, Wu ZL, Cui MY. Chronological age estimation of third molar mineralization of Han in southern China. Int J Legal Med. March 2010;124(2):119–23.

26. Johan NA, Khamis MF, Jamal NSA, Ahmad B, Mahanani ES. The variability of lower third molar development in Northeast Malaysian population with application to age estimation. 2012;10.

27. Olze A, Pynn BR, Kraul V, Schulz R, Heinecke A, Pfeiffer H, et al. Studies on the chronology of third molar mineralization in First Nations people of Canada. Int J Legal Med. September 2010;124(5):433–7.

28. Guo YC, Lin XW, Zhang WT, Yan CX, Pan F, Yan TL, et al. Chronology of third molar mineralization in a northern Chinese population. Rechtsmedizin. February 2015;25(1):34–9.

29. Acharya AB, Bhowmik B, Naikmasur VG. Accuracy of Identifying Juvenile/Adult Status from Third Molar Development using Prediction Probabilities Derived from Logistic Regression Analysis. J Forensic Sci. May 2014;59(3):665–70.

30. Konigsberg LW. Multivariate cumulative probit for age estimation using ordinal categorical data. Annals of Human Biology. 04 July 2015;42(4):368–78.

31. Bleka Ø, Wisløff T, Dahlberg PS, Rolseth V, Egeland T. Advancing estimation of chronological age by utilizing available evidence based on two radiographical methods. Int J Legal Med. January 2019;133(1):217–29.

32. Varkkola O, Ranta H, Metsäniitty M, Sajantila A. Age assessment by the Greulich and Pyle method compared to other skeletal X-ray and dental methods in data from Finnish child victims of the Southeast Asian Tsunami. Forensic Sci Med Pathol. December 2011;7(4):311–6.

33. Gelbrich B, Frerking C, Weiß S, Schwerdt S, Stellzig-Eisenhauer A, Tausche E, et al. Combining wrist age and third molars in forensic age estimation: how to calculate the joint age estimate and its error rate in age diagnostics*. Annals of Human Biology. 04 July 2015;42(4):389–96.

34. Mostad P, Schmeling A, Tamsen F. Mathematically optimal decisions in forensic age assessment. Int J Legal Med. 15 December 2021;

Appendix

Table A.1: Estimated reliability of the model

Calculated measures	Male	Female
	Knee or tooth, or both reached final stage	Tooth reached final stage
Sensitivity	0.78	0.24
Specificity	0.90	0.97
PPV	0.89	0.88
NPV	0.80	0.56

Table A.1: Calculated sensitivity, specificity, positive predictive value (PPV= proportion of identified adults that are adults), negative predictive value (NPV= proportion of identified children that are children) for the model based on the combination of knee and tooth for males, and only tooth for females. The calculations are made assuming an uniform distribution between 15 and 21 years of age, with < 18 years being classified as a true negative and >18 years being classified as a true positive.

Table A.2(a): Logistical regression model's coefficient values and significance for knee and tooth, males.

	Coefficient value	Significance (p-value)
Knee intercept	- 29.3164	< 0.001
Knee age	1.5686	< 0.001
Tooth intercept	-18.6568	< 0.001
Tooth age	0.9189	< 0.001

Table A.2(b): Logistical regression model's coefficient values and significance for knee and tooth, females.

	Coefficient value	Significance (p-value)
Knee intercept	- 24.9819	< 0.001
Knee age	1.4476	< 0.001
Tooth intercept	-16.4067	< 0.001
Tooth age	0.7767	< 0.001

Table A.3(a): Logistic regression model's covariance matrix for knee and tooth, males.

	Intercept	**Age**
Knee intercept	3.8550	-0.2056
Knee age	-0.2056	0.0110
Tooth intercept	0.2472	-0.0122
Tooth age	-0.0122	0.0006

Table A.3(b): Logistical regression model's covariance matrix for knee and tooth, females.

	Intercept	**Age**
Knee intercept	4.0550	-0.2323
Knee age	-0.2323	0.0134
Tooth intercept	0.1338	-0.0064
Tooth age	-0.0064	0.0003

The logistical regression model *p(x)* is a function of the value *x* and is given as:

$$p(x) = \frac{1}{1 + e^{-(\alpha + \beta x)}}$$

where α and β are the coefficient values for the intercept and the independent variable age, respectively. The value *x* is a given age inserted into the function followed by the calculation of the function's value *p(x)*.This produces a probability between 0 and 1 that discloses the chance of the indicator at a given age being at the final stage.

Figure A.1 (a): Fitted logistical regression model and original data for males.

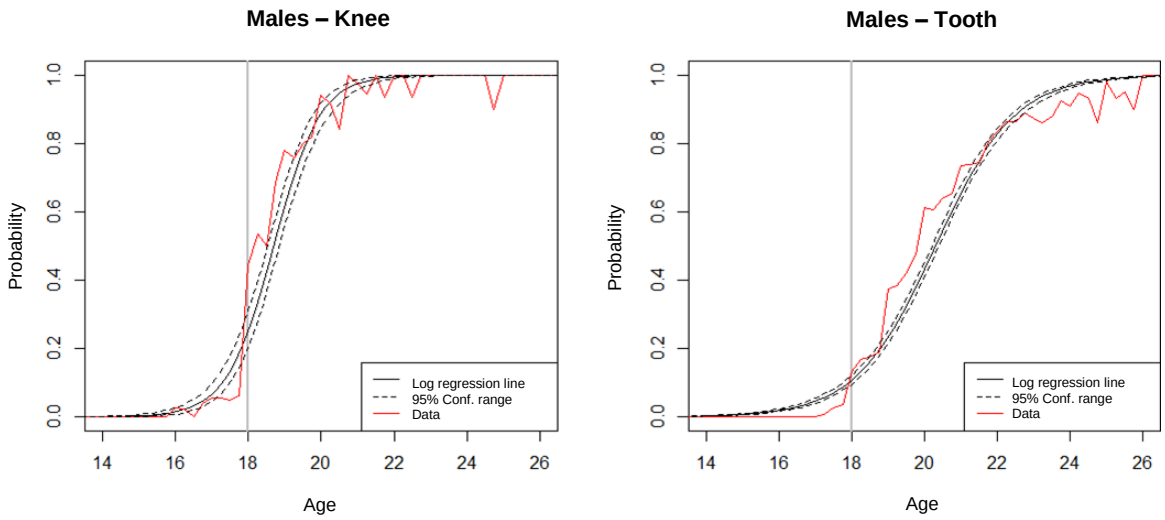


Figure A.1(a): Fitted logistical regression model for knee and tooth (black line) with 95% confidence interval (dashed line), males. The original data (red line) corresponds to the calculated proportion of data distributed quarterly.

Figure A.1 (b): Fitted logistic regression model and original data for females.

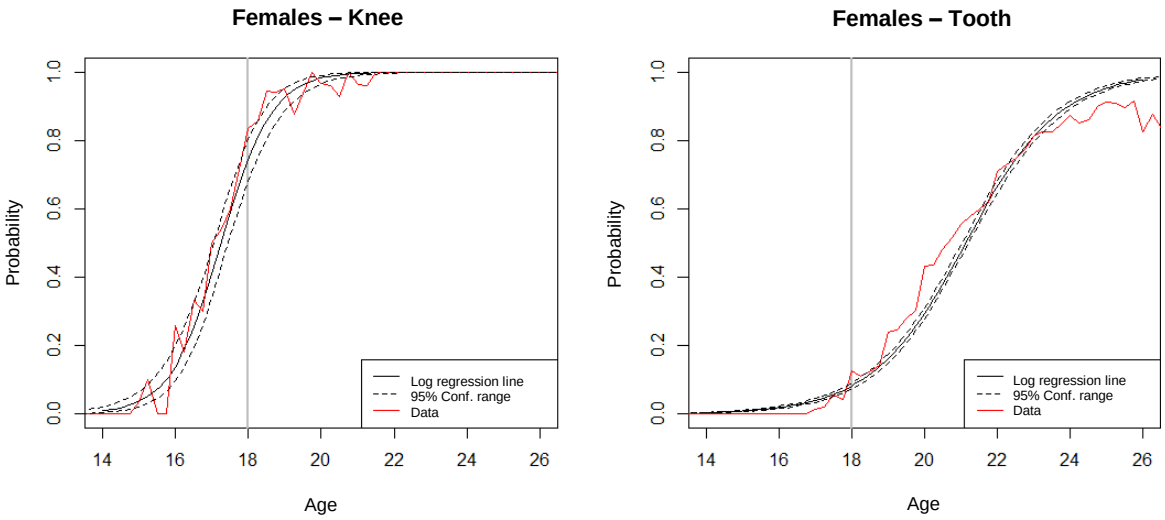


Figure A.1(b): Fitted logistic regression model for knee and tooth (black line) with 95% confidence range (dashed line), females. The original data (red line) corresponds to the calculated proportion of data distributed quarterly.