International Journal of Legal Medicine
https://doi.org/10.1007/s00414-024-03324-x

ORIGINAL ARTICLE

# A probability model for estimating age in young individuals relative to key legal thresholds: 15, 18 or 21-year

**Nina Heldring[1,2] · Ali-Reza Rezaie[1] · André Larsson[3] · Rebecca Gahn[1] · Brita Zilg[1,2] · Simon Camilleri[4] · Antoine Saade[5] · Philipp Wesp[6,7] · Elias Palm[1] · Ola Kvist[8,9]**

Received: 23 May 2024 / Accepted: 29 August 2024


## Abstract

Age estimations are relevant for pre-trial detention, sentencing in criminal cases and as part of the evaluation in asylum processes to protect the rights and privileges of minors. No current method can determine an exact chronological age due to individual variations in biological development. This study seeks to develop a validated statistical model for estimating an age relative to key legal thresholds (15, 18, and 21 years) based on a skeletal (CT-clavicle, radiography-hand/wrist or MR-knee) and tooth (radiography-third molar) developmental stages. The whole model is based on 34 scientific studies, divided into examinations of the hand/wrist (15 studies), clavicle (5 studies), distal femur (4 studies), and third molars (10 studies). In total, data from approximately 27,000 individuals have been incorporated and the model has subsequently been validated with data from 5,000 individuals. The core framework of the model is built upon transition analysis and is further developed by a combination of a type of parametric bootstrapping and Bayesian theory. Validation of the model includes testing the models on independent datasets of individuals with known ages and shows a high precision with separate populations aligning closely with the model's predictions. The practical use of the complex statistical model requires a user-friendly tool to provide probabilities together with the margin of error. The assessment based on the model forms the medical component for the overall evaluation of an individual's age.

**Keywords** Age distribution · Bayesian theorem · Biological variation · Population · Forensic anthropology · Validation study

✉ Nina Heldring
nina.heldring@rmv.se

1 Department of Forensic Medicine, Swedish National Board of Forensic Medicine, Retzius Väg 5, 171 65 Stockholm, Sweden

2 Department of Oncology-Pathology, Karolinska Institutet, Retzius V. 3, 171 77 Stockholm, Sweden

3 Paindrainer, Medicon Village, 223 81 Lund, Sweden

4 Faculty of Dentistry, Oral and Craniofacial Sciences, Tower Wing, Guys' Hospital St Thomas Street, London, England

5 Department of Orthodontics, Faculty of Dental Medicine, Lebanese University, Beirut, Lebanon

6 Department of Radiology, LMU University Hospital, LMU Munich, Marchioninistraße 15, 81377 Munich, Germany

7 Munich Center for Machine Learning (MCML), Geschwister-Scholl-Platz 1, 80539 Munich, Germany

8 Pediatric Radiology Department, Karolinska University Hospital, Stockholm, Sweden

9 Department of Women's and Children's Health, Karolinska Institute, Stockholm, Sweden

## Introduction

There are many shortcomings in all medical age assessments that are being applied in different countries. No current method can determine an exact chronological age (CA) due to the individual variations in biological development. Still, there are practical needs to assess age in various legal contexts with minimal error rates. Age estimation is relevant for pre-trial detention and sentencing in criminal cases as well as part of the evaluation in asylum processes to protect the rights and privileges of minors. The European Asylum Support Office (EASO) recommends using the least intrusive examination for medical age assessments methods in their practical guide [1] with radiation free procedures argued to be preferable in children and young adults. The lack of validated or standardized methods has rendered countries within or outside the EU to choose various methods of medical age assessment [1, 2]. In addition, the mission differs slightly between countries in terms of the questions that are expected

Published online: 18 September 2024

to be answered as well as which party carries out the task. In many nations, adopting a minimum age concept is a prevalent strategy aimed at minimizing the risk of misclassifying minors. However, this strategy overlooks the potential drawbacks of erroneously classifying adults as minors. Such consequences include misallocation of resources intended for minors to adults and hindrance to the proper administration of justice, as adults may escape prosecution in criminal cases. Probability methods provide a most likely age distribution based on a large reference population rather than an indeterminable CA. The overall approach to provide a probability of an individual being below or above a certain age includes, as a first step, to examine the developmental stages of a selected skeletal component together with the wisdom tooth, and then comparing this to the age distribution of the reference population of the same sex and developmental stages. The probabilities are supplemented with the margin of error, represented by the minor portion of the reference population distribution in relation to the chosen age threshold. The order of magnitude of the margin of error reflects the certainty level of the assessment. Notably, there is a knowledge gap of how one can objectively use multiple anatomical locations and statistical models to estimate the age of an individual more accurately. Having validated models ensures fairness and accuracy as far as possible in legal proceedings. This study seeks to develop and present a validated statistical model for estimating an age relative to key legal thresholds (15, 18, and 21 years) based on skeleton (CT-clavicle, radiography-hand/wrist, MR-knee) and teeth radiography-third molar) developmental stages.

## Methods

### Data included in the model

A literature search was conducted to identify scientific studies investigating hand/wrist, third molar, distal femur or clavicle maturity in relation to age. After removal of duplicate articles and categorization based on title and abstract, full text articles were read and the following exclusion criteria were applied:

1) Imaging method other than radiography (hand/wrist, third molar), MRI (distal femur), CT (clavicle). 2) Incomplete data: the study does not present all the data needed to recreate individual-based data. 3) Different staging than Greulich & Pyle (hand/wrist), Demirjian (third molar) Krämer (Distal femur), Schmeling (Clavicle). 4) The study population does not include ages on both sides of the 15- and 18-year boundaries (Distal femur only). 5) Other anatomical structure than selected indicators. 6) Previously published results, e.g. analysis or review of previous data. 7) Post-mortem study population. 8) Full text not available in English, Swedish, Danish or Norwegian. 9) Study based on data that is not available. 10) Study population includes individuals with a disease that may affect skeletal maturity. 11) Study population has uneven age distribution according to Chi-square test (type 3 data only).

All the hand/wrist studies investigated skeletal age based on radiographs where the developmental stages are classified according to Greulich & Pyle [3]. Studies were identified through targeted searches on PubMed using the strategy (skeletal matur* OR ossifi* OR age estimat* OR forensic age OR age asses* OR age determin*) AND (radiography OR radiograph* OR x-ray OR ionizing) AND (Greulich OR Pyle) and Embase, which generated 727 studies. The data included in the model were obtained from 15 hand/wrist studies that met the criteria (Table 1).

All the dental studies related the development of the third molar in the lower jaw, imaged with plain radiographs and classified by Demirjian, to CA in the study populations. Dental studies were identified from the summaries previously made in BioAlder 1.3 [4–7]. A total of 58 articles were identified, all of which were read in full text and 10 studies met the criteria and were included in the model (Table 1).

The distal femur studies related the development of the upper knee joint (distal femur), examined by magnetic resonance imaging (MRI) with field strength of at least 1.5T and T1 weighting, to CA after classification according to Krämer 2014 [8]. Studies were identified from Heldring et al. 2022 [7], supplemented with articles from an internal literature monitoring procedure on distal femur studies. A total of 27 studies were identified and read in full text and 4 of these met the criteria and were selected for inclusion (Table 1.)

Original clavicle studies where the development of clavicles according to Schmeling's staging (1–5) [9] and CA was studied, were identified. This was done by a literature search in PubMed using the string ((skeletal matur* OR ossifi* OR age estimat* OR forensic age OR age asses* OR age determin*) AND (clavicle OR medial epiphysis OR medial end OR medial clavicular epiphysis OR sternal epiphysis OR sternal end) AND (CT scan OR computed tomography OR CT OR scanner OR Schmeling's method OR "chest radiographs" OR "forensic radiology") which generated 296 articles and 5 clavicle studies met the criteria for inclusion (Table 1).

### Data extraction and simulating population age distributions

The method of data extraction is adapted to how the data is presented in each study. In order to fit the probabilistic model to the datasets, all data must include a list with known CA and corresponding developmental stage for each individual. The format of type 1 data provides CA presented together with the development stage for each individual either in a

**Table 1** Studies included in the probability model

| Indicator | Study | Men | Women | Country | Age span (men) | Age span (women) | Method | Type of data |
|---|---|---|---|---|---|---|---|---|
| Hand | Alcina 2018 [52] | 590 | 560 | Spain | 0–19 | 0–18 | x-ray | Type 1b |
| Hand | Bala 2010 [53] | 80 | 80 | India | 8–14 | 8–14 | x-ray | Type 4 |
| Hand | Büken 2007 [54] | 251 | 241 | Turkey | 11–20 | 11–19 | x-ray | Type 4 |
| Hand | Cantekin 2012 [55] | 342 | 425 | Turkey | 7–18 | 7–18 | x-ray | Type 1b |
| Hand*~ | Chaumoitre 2017 [56] | 1423 | 1191 | France | 1–21 | 1–20 | x-ray | Type 3 |
| Hand | Dembetembe 2012 [57] | 131 | - | South Africa | 13–22 | 13–22 | x-ray | Type 1b |
| Hand | Elamin 2017 [58] | 487 | 627 | Sudan | 1–28 | 2–37 | x-ray | Type 4 |
| Hand | Hackman 2013 [59] | 249 | 157 | Scotland | 1–20 | 1–20 | x-ray | Type 1b |
| Hand | Koc 2001 [60] | 225 | - | Turkey | 7–17 | - | x-ray | Type 4 |
| Hand | Mora 2001 [61] | 265 | 269 | USA | 0–19 | 0–19 | x-ray | Type 4 |
| Hand | Paxton 2013 [62] | 276 | 130 | Australia | 1–18 | 1–18 | x-ray | Type 1b |
| Hand | Soudack 2012 [63] | 375 | 304 | Israel | 1–18 | 1–18 | x-ray | Type 1b |
| Hand* | Tisé 2011 [64] | 359 | 125 | Italy | 11–19 | 11–19 | x-ray | Type 3 |
| Hand | van Rijn 2001 [65] | 278 | 294 | Netherlands | 5–20 | 5–20 | x-ray | Type 1b |
| Hand | Zabet 2015 [66] | 100 | 90 | France | 10–18 | 10–19 | x-ray | Type 1b |
| Distal femur* | Ekizoglu 2020 [67] | 335 | 314 | Turkey | 12–30 | 12–30 | MRI | Type 3 |
| Distal femur | Krämer 2014 [68] | 166 | 124 | Germany | 10–31 | 10–31 | MRI | Type 3 |
| Distal femur | Ottow 2017 [69] | 326 | 335 | Germany | 12–25 | 12–26 | MRI | Type 3 |
| Distal femur | Saint-Martin 2014 [70] | 214 | - | France | 14–20 | - | MRI | Type 2 |
| Clavicle* | Ekizoglu 2015 [71] | 362 | 141 | Turkey | 10–35 | 10–35 | CT | Type 3 |
| Clavicle | Franklin 2015 [72] | 185 | 148 | Australia | 10–35 | 10–35 | CT | Type 2 |
| Clavicle* | Pattamapaspong 2015 [32] | 249 | 160 | Thailand | 11–29 | 11–29 | CT | Type 3 |
| Clavicle | Uysal 2017 [73] | 399 | 202 | Turkey | 10–35 | 10–35 | CT | Type 2 |
| Clavicle | Zhang 2015 [74] | 370 | 382 | China | 15–25 | 15–25 | CT | Type 3 |
| Third molar | Duangto 2017 [75] | 877 | 990 | Thailand | 8–23 | 8–23 | x-ray | Type 2 |
| Third molar | Hassan 2021 [76] | 170 | 180 | Egypt | 14–24 | 14–24 | x-ray | Type 3 |
| Third molar | Hegde 2016 [77] | 664 | 475 | India | 5–16 | 5–16 | x-ray | Type 2 |
| Third molar | Johan 2012 [78] | 540 | 540 | Malaysia | 14–25 | 14–25 | x-ray | Type 2 |
| Third molar | Kasper 2009 [79] | 804 | 1019 | USA (Latino) | 12–22 | 12–22 | x-ray | Type 2 |
| Third molar | Lee 2009 [80] | 786 | 964 | South Korea | 7–22 | 7–24 | x-ray | Type 2 |
| Third molar | Li 2012 [81] | 989 | 1089 | China | 5–23 | 5–23 | x-ray | Type 2 |
| Third molar | Liu 2018 [82] | 1012 | 1196 | China | 8–23 | 8–23 | x-ray | Type 2 |
| Third molar | Lopez 2013 [83] | 236 | 315 | Brazil | 15–23 | 15–23 | x-ray | Type 3 |
| Third molar | Quispe 2017 [84] | 102 | 106 | Peru | 14–22 | 14–22 | x-ray | Type 3 |

Studies included in the probability model. MRI = Magnetic resonance imaging, x-ray = Radiological X-ray, CT = Computed tomography. * Only data for women were used as the age distribution was not evenly distributed for men. ~ Some stages were excluded as these did not fulfil the normal distribution assumption according to the research article

table by the authors (type 1a) or extracted from a figure with PlotDigitizer (type 1b) [10], hence can be included without recreation.

However, datasets where both CA and corresponding developmental stage are not reported for each individual require recreation of individual-based datasets. Type 2 data are reported as the frequency of different stages within integer age intervals, either as counts or as fractions together with the total number of individuals for the different intervals. Individual-based data is recreated by calculating the number of individuals with a specific stage in each of the age-cohorts and CAs are assigned randomly within each age interval assuming a uniform distribution. If minimum and maximum of CA for a given developmental stage is provided in addition to the frequency data, the simulated uniform values are further limited to this specified interval.

Type 3 data present the number of individuals at each stage, alongside essential statistical measures such as the min, max and lower, median and upper quartile of the CA within each stage (type 3b), or the mean and standard deviation for each stage (type 3a). In the case of type 3a data, a normal distribution is used to generate the individual ages,

however, if an age range [*a, b*] is additionally specified for each specific stage by the study, a truncated normal distribution is fitted to the reported values. The *truncorm* package (version 1.0–9) in R was used [11] to perform this. For type 3b data, which reports the quantiles of the measured age distributions for each stage, a normal distribution of CA is assumed, for every stage *s*. A truncated normal is fitted through a numerical optimization process that minimizes the errors between the quantiles of the simulated truncated normal distribution and quantiles reported in the study. In the full dataset, CA from type 3a and 3b datasets are therefore simulated with either a normal or truncated normal distribution using the estimated parameters as described above. Further details on this approach and the truncated normal can be found in Supplementary Appendix.

Type 4 data reports mean age, standard deviation, and Pearson's correlation for an age-cohort of both the CA and skeletal age. To simulate populations, the process includes a two-step approach, as described in Bleka et al. [5]. In short, the additional information provided by the Pearson's correlation coefficient is incorporated by fitting a multivariate normal distribution to the data, including the conditional dependence between CA and stage. The resulting bivariate normal distribution is then used to re-creat333333333 and stage *s* for each individual in the study. All resulting statistics in this report are derived from 10,000 simulated populations, unless stated otherwise.

## The probability model

The first step in generating the probabilities is to obtain an estimate of chronological stage *s* through finding the probability of stage given age, $P(S = s \mid A)$, by fitting ordinal/logistic regression models to the datasets of each individual developmental indicator. In the second step, these results are used in equation 1,

$$P(A|S = s) = \frac{P(S = s|A)\,P(A)}{\int_b^a P(S = s|y)\,P(y)dy} \quad (1)$$

to obtain the inverse probability of age given stage, $P(A \mid S = s)$ for each indicator. As this equation only depends on $P(S = s \mid A)$, assuming a uniform prior, we can find the normalizing factor in the denominator by requiring the total area of the probability density function (PDF) to be one. Finally, we end up with a probability density function $P(A \mid S = s)$ for each stage/combination of stages *s,* which can be integrated to find the relevant statistics, such as the probability of stage *s* for being below or above a certain age threshold. This two-step approach also using re-created population data was taken to minimize the influence of age mimicry [12]. The probability of being below 15, 18 or 21-year thresholds is calculated based on all 10,000 simulations with bootstrapping for each stage, and the 50th percentile is selected as the estimate. From the bootstrap sample, we also determine a 95% confidence interval for the calculated statistics based on the 2.5th and 97.5th percentile. In addition, the probability of the one-year age-cohorts within the assumed age distribution is computed by applying the 50th percentile value from all simulated 10,000 populations.

## Prior age distribution

The selected uniform prior ensures that all information is derived from the data in the posterior distribution as the purpose is to generate the conditional PDF without any subjective influence. This approach with a non-informative prior requires a defined lower and upper limit of the uniform distribution being determined by the assumed age range within the model. Based on the endpoint of the second-to-last stage for hand/wrist, 20 years of age for females and 21 for males was chosen as the upper bound (Roberts et al. (2015) [13]. In order to avoid an increased risk of type 1 errors (identifying children as adults) in the third molar model, the upper limit is set in accordance with Knell et al. (2009) [14] and Olze et al. (2010) [15], at the age when 50% of the population reaches stage H (21 years for both genders) due to the wide distribution of the second-to-last stage G. The lower bound for both the hand/wrist as well as the third molar model is set to 7 years for both sexes. Data from clavicle studies typically span ages 10–35, and it is noted that stage 4 of the clavicle can still be detected among 35-year-olds for both genders. Similar to the third molar model, the upper limit for the clavicle model is set at the age when 50% of the population reaches the last development stage (stage 5). Hence, the assumed age range was considered between 10–30 years for females and 10–32 years for men, for the clavicle model. For distal femur, we adopted an age range of 15–21, as proposed in Heldring et al. (2022) [7].

## Additional assumptions when combining two indicators

In order to obtain an estimate of CA when the stages of several different developmental indicators are combined, we assume that stages are conditionally independent from each other. Previous probability models similar to this one assume a conditional independence between skeleton development and third molar development [5, 7, 16] based on studies investigating hand/wrist and third molar development [17–19]. The study that is comparing models that included or excluded a co-dependence between indicators on a combination dataset concluded that there was no statistically significant improvement in the accuracy of age estimation when including a conditional dependence between indicators [5]. However, this assumption does not apply between skeletal

indicators, rendering the calculation of probabilities in those combinations inaccurate.

The probability of one skeleton indicator being in stage $s_s$ and the third molar indicator being in stage $s_t$ for a given age, can be expressed as.

$$P(S_s = s_s, S_t = s_t|A) = P(S_s = s_s|A) \cdot P(S_t = s_t|A) \quad (2)$$

assuming conditional independence between the indicators. To obtain the reverse conditional probability, probability of age given stage $s$ (Eq. 2) is applied analogous to the calculations in Eq. 1.

For the combined clavicle and third molar model, the upper limit is set to 26.0 years, as the data is truncated at this age for the third molar model. The upper limit is set to 21 years for both females and males for the third molar and hand/wrist combination, as well as the third molar and distal femur model. In addition, the dichotomous distal femur model in combination with third molar is based on the age range 15–21 years and includes the relevant Demirjian stages D-H.

## Model selection

Two candidate ordinal regression models, cumulative and continuous-ratio (CR), with either logit or probit for the linking functions and using either parallel or non-parallel odds-ratios were considered (Supplementary Appendix). This is similar to models previously described in the BioAlder tool [5].

The best model was selected based on a goodness-of-fit of the data for each indicator and gender combination. For each 10,000 populations, the Akaike information criterion (AIC) [20] was computed for every model combination and the final model was selected based on the lowest median AIC value. The choice of AIC was motivated by its ability to penalize the addition of extra parameters estimated in the ordinal model, thereby balancing model complexity. This process was carried out individually for each indicator and gender, yielding a total of 8 distinct models. Both the cumulative and the CR model will be equivalent to a simple logistic regression model for indicators with only two separate stages as in the distal femur model.

The model was written in R (Version 4.3.1) [21]. The ordinal/logistic regression models were fitted by applying the *vglm* function in the VGAM (Version 1.1–9) package [22]. The different conditional PDFs were created by extracting the corresponding parameters from the ordinal/logistic models followed by applying Bayes' theorem. To calculate the area under the curve of the conditional PDF for a given threshold or one-year cohorts, the *integrate* function was applied. The method for estimating the prediction intervals (PI) of the CA is described in the Supplementary Appendix.

## Collection of validation populations

The access to independent datasets is mainly dependent on other researchers. In our initial search for studies to be included when building the model, we identified studies where data is presented in a format that was not suitable or had a high risk of age mimicry. We invited some of the authors of these studies and additional studies found in later searches to share their primary data (CA, development stage and gender) to be used as independent validation populations (Table 2). In addition, an independent study of clavicles with CT was performed. The study was retrospective in its design with all cases extracted from Karolinska University Hospital, Stockholm, and approved by the Swedish Ethical Review Authority (Dnr 2024–00531-01). Individuals aged 17.0 to 25.0 years examined during routine clinical practice and with known CA and sex were selected. Scans with poor image quality and individuals with an injury or a skeletal disease that could

**Table 2** List of validation datasets

| Indicator | Study | Men | Women | Country | Age span (men) | Age span (women) | Method | Type of data |
|---|---|---|---|---|---|---|---|---|
| Hand | Maggio 2018 [85] | 104 | 108 | Australia | 0–25 | 0–25 | x-ray | Type 1a |
| Hand | Saade 2017 [28] | 115 | 108 | Lebanon | 8–16 | 8–15 | x-ray | Type 1a |
| Hand | Zafar 2010 [86] | 167 | 84 | Pakistan | 0–18 | 0–18 | x-ray | Type 1a |
| Distal femur* | Socialstyrelsen 2018 [87] | 217 | 178 | Sweden | 14–22 | 14–22 | MRI | Type 1a |
| Third molar | Jayaraman 2022 [88] | 189 | 186 | USA (Latino) | 8–17 | 7–17 | x-ray | Type 1a |
| Third molar | Knell 2009 [14] | 551 | 622 | Switzerland | 15–22 | 15–22 | x-ray | Type 1a |
| Third molar | Malta collection | 553 | 651 | Malta | 8–26 | 7–25 | x-ray | Type 1a |
| Third molar | Saade 2017 [28] | 1113 | 119 | Lebanon | 8–16 | 8–15 | x-ray | Type 1a |
| Clavicle | Swedish collection | 199 | 201 | Sweden | 17–24 | 17–25 | CT | Type 1a |
| Clavicle | Wesp 2024 [50] | 28 | 22 | Germany | 15–29 | 15–29 | CT | Type 1a |

List of validation datasets including number of individuals, age span and country. *Observations with MRI T1 TSE sequence (bridging study)

affect clavicle development were excluded. Selected scans were subsequently assessed with regard to development stage in agreement with the Schmeling staging system [9, 23] on the most developed side by one radiologist with 14 years of musculoskeletal (MSK) radiology experience and 8 years with focus on pediatric MSK radiology experience.

### Validation of the statistical model with independent datasets

We used the true development stages of the independent individual observations for the classification of whether they fall below or above the 15-, 18- or 21-year age threshold limits. This classification process involves selecting a cutoff point of the given probability where probabilities below the cutoff will classify the individual as above the threshold while probabilities above the cutoff will generate a classification of the individual as below the age threshold. While a common method involves ROC curve analysis to determine an optimal cutoff point to maximize sensitivity and specificity, the chosen cutoff point of 0.35 was based on being an acceptable error of the mean for a final evaluation. This strategy consequently leads to minimizing type 1 errors (classifying underage as overage) and as a consequence will classify more individuals being over the age threshold as under than the opposite if applied. The individuals and proportions being correctly or incorrectly classified are visualized and presented in distribution-plots, point-plots, bar graphs and line-graphs (Fig. 3, 4, 5, and 6 and Supplementary Fig. 12). The distribution of the collected validation populations is visualized as interpolated kernel density estimator (KDE) of the different study distributions and all the studies combined (Supplementary Fig. 1 (a-b)). The KDE is fitted with the geom_density function in the ggplot2 package [24].

In order to calculate the minimum sample size required to estimate the precision of the models, the *pmsampsize* function from the pmsampsize package [25] was used in R. To calculate the minimal sample size needed for external validation of prediction models with a binary outcome (correct or incorrect classification) [26, 27] included a conservative outlook with a c-statistics of 0.85 and a prevalence of 0.15, meaning 15% misclassification of events are expected. This resulted in 195 individuals for a validation sample size for males and females, respectively.

## Results

### Data included in the model

Observations from approximately 27,000 individuals from 6 geographic regions are included in the model (Table 1 and Supplementary Table 1).

### Selected model

We found that the continuation-ratio model with logit link function and a non-parallel slope coefficient provided the best fit for the clavicle and third molar model (both sexes). A continuation-ratio model with probit link function and a non-parallel slope coefficient fitted the data best for the hand/wrist model in both sexes. For distal femur, where only two stages are used (not closed/closed), logistic regression with a logit link function for both sexes was the best fit and used in the final model. A graphic representation of how the fitted parametric regression model relates to the calculated semi-annually proportion of underlying data (non-parametric), calculated as the fraction of individuals with a specific stage in the simulated datasets, is presented in Supplementary Fig. 2–8.

We refrained from log-transforming the CA variable to avoid potentially increasing complexity within the model, as the non-parallel fit gives the posterior distributions more flexibility as they were being estimated and because of the assumption of normal distributions among stages. This is in contrast to previous models where a parallel slope coefficient for all models and log-transformation was applied [16]. We demonstrate that certain third molar stages, fitted with the KDE from one of the randomly generated populations compared with its fitted PDF, appear to be approximately normal distributed (Supplementary Fig. 1 (c-n)) when the influence of age mimicry is low, i.e. where the chronological age of the data is approximately uniformly distributed (Supplementary Fig. 1 (a-b)).

### Age prediction model

The estimated 75% and 95% PI's of CA for the hand/wrist and third molar stages of development are shown in Supplementary Fig. 9, separately (a) and in combination (b), as the median from 10,000 simulated populations. The age distributions are wider when using a single indicator compared to combining the third molar with hand/wrist, indicating that multifactorial age estimations are more accurate compared to using a single anatomical site. This is also seen for the combination with the distal femur (Supplementary Fig. 10) or clavicle (Supplementary Fig. 11). The PDF's for hand/wrist, third molar, distal femur, and clavicle assuming normally distributed ages for each indicator and stage are shown for males (a-d) and females (e–h) in Fig. 1. The distributions display one randomly selected distribution from the 10,000 generated populations for each stage.

### Combining indicators

From the known probability of being in a stage given age, we derived the conditional PDF for age within this stage by using Bayes' theorem (Eq. 1). The assumption of conditional independence does not apply between skeletal indicators, rendering the three skeletal indicators inappropriate to combine. Hence, the current combinations are third molar with either one of the skeletal indicators. Age distributions for selected combinations are shown in Fig. 2 for males (a and b) and females (c and d). The probability of age in relation to a certain threshold is represented by the part of a specific combination's distribution being on either side of the age limit. The distribution as well as probabilities are affected by the chosen upper age limit for each indicator. A sensitivity analysis was performed with several upper age limits (Table 3, hand/wrist and third molar, Supplementary Table 2 clavicle and Supplementary Table 3 clavicle and third molar). We observe that the probabilities of being under 18 years of age is only minimally affected if the upper age limit is increased for the combination of hand/wrist and third molar (Table 3). We also noted that the probabilities of being under the 21-year threshold for stage 4 or 5 in the clavicle model do not vary significantly when changing the upper boundary between 30 and 35 years (Supplementary Table 2). This demonstrates that the chosen distribution predicts reliable probabilities.

### Validation with independent test populations

To assess how well the model performs on independent data, a number of datasets for populations of known age have been collected and used for validation (Table 2). Aside from the Swedish collection of a clavicle dataset that was collected specifically for the purpose of the validation of the model, the datasets are from published studies or collections, kindly provided by authors and researchers upon contact. Each indicator was validated separately, except the combination of third molar and hand/wrist where examination and developmental stage were studied in the same individual for one of the datasets [28].

### Validation of the third molar model

The validation set for third molar included in total 1406 males (Fig. 3(a)) and 1578 females (Fig. 3(b)), spanning an age interval between 7–26 years (Table 2) and originates from 4 separate datasets (Fig. 3). In total, 93% of the male and 87% of the female populations were correctly classified regarding the 18-year threshold, corresponding to the separate model's total accuracy (Table 4 and Fig. 3 (c-f)). In addition, the model accuracy with regard to the 15-year threshold is 90% for males and 87% for females (Table 4 (a)). The sensitivity (adults identified as adults) of the male third molar model is 90% and specificity (children identified as children) is 95% for the 18-year threshold, while the positive predictive value (identified as adults that are adults) is 91% and the negative predictive value (identified as children that are children) is 94% (Table 4 (a)). The corresponding sensitivity in the female third molar model is 75% and the specificity 94% (Table 4 (a)). Not surprisingly, very early stages cause few errors in the assessments of both the 15-and the 18-year threshold (Fig. 3 (c-f)). Most of the incorrectly classified individuals are in the development stages C-F for the 15-year threshold and D-H for the 18-year threshold in both males (c and e) and females (d and f). These individuals are fewer compared to correctly classified individuals (Fig. 3 (g-h)), and represent both individuals with an age close to the limit and individuals with either early or late third molar development (Fig. 3 (c-f)). The proportion of the independent population being under 15 (orange full line) or 18 (blue full line) years overlaps almost completely with the predicted probabilities (dashed lines) for the model (Fig. 3 (g-h)), for both males (g) and females (h). This demonstrates a high reliability of the probability model.

### Validation of the hand/wrist model

In total, 386 males (Fig. 4 (a)) and 301 females (Fig. 4 (b)), spanning an age interval between 7–26 years and originating from 3 separate datasets (Fig. 4 (a-b)) are included in the independent validation set for hand/wrist. What distinguishes the hand/wrist model from the dental model is that it is suitable for assessing the 15-year threshold but is of limited use for the 18-year threshold as the last developmental stage begins before the age of 18 to a large extent (Fig. 1 (a, e)). In total, 88% of the male and 91% of the female populations were correctly classified regarding the 15-year threshold (Table 4 (b)). Similar to the third molar model, incorrectly classified individuals are not found in the early development stages but have reached skeletal age (SA) 13 up to 18 (Fig. 4 c-f) in both males (c) and females (d). The incorrectly classified individuals are fewer compared to correctly classified (Fig. 4) in both males (g) and females (h) except for SA 16 and 17 in females with regard to the 15-year threshold where it is equal (h). With regard to the 18-year threshold, the model has an acceptable precision when it comes to below 18 (Fig. 4(e–f)), while the development stages of hand/wrist do not seem to allow for accurate age estimations with regard to above18 years of age. The proportion of individuals being under 15 (orange full line) or 18 (blue full line) in the independent validation population of the hand/wrist model basically follows the probabilities of being under 15 (orange dashed line) or 18 (blue dashed line) according to the model for males and females (Fig. 4 (g-h). However, the non-smoothness of the curves reflects

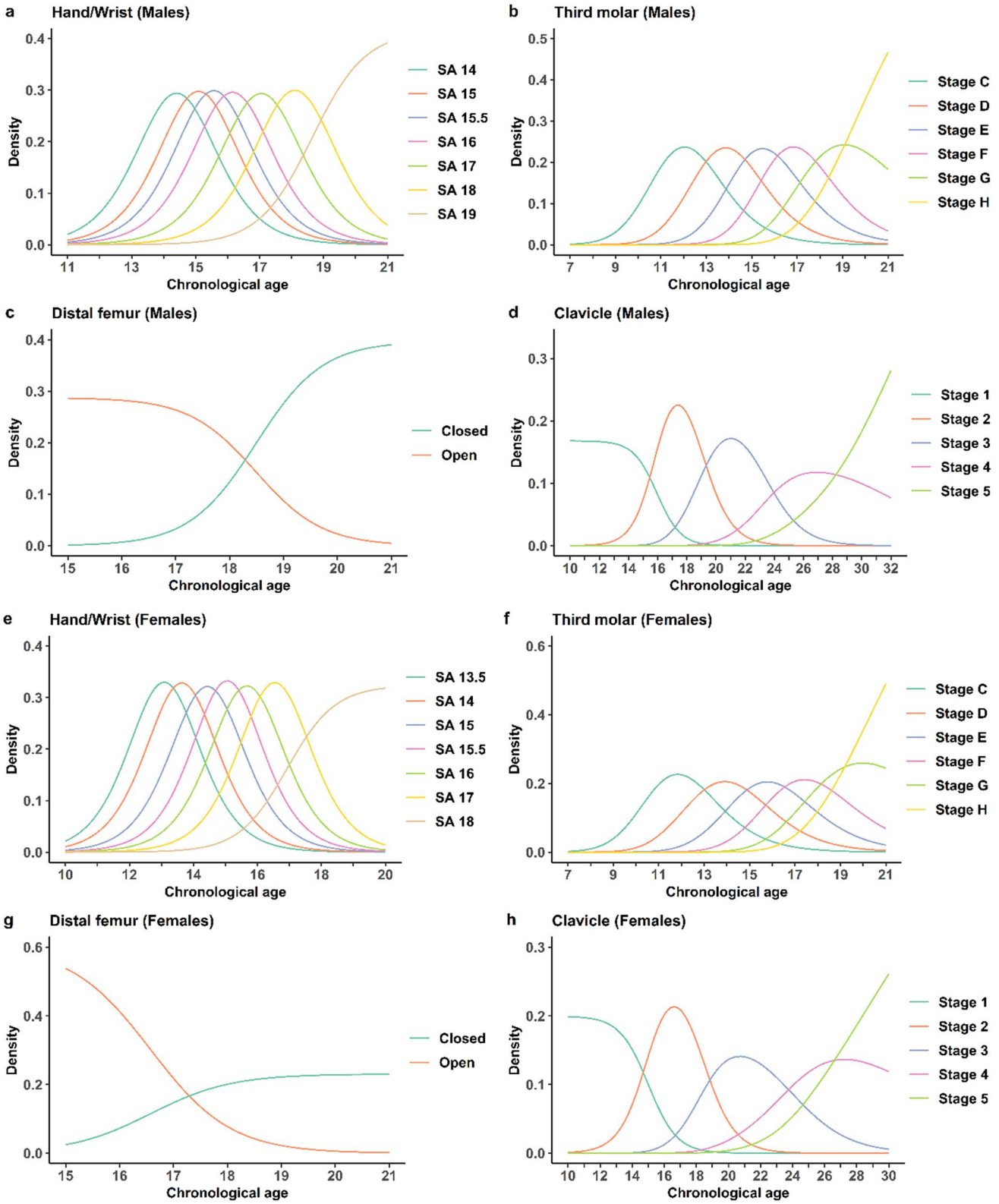



**Fig. 1** Probability density functions. Age distributions for hand/wrist, third molar, distal femur and clavicle stages for male (**a**-**d**) and female (**e**–**h**) individuals in terms of density of developmental stage hand/wrist skeletal age 14–19 and 13.5–18 respectively (Greulich & Pyle) (**a** and **e**), third molar stage C-H (Demirjian) (**b** and **f**), distal femur reached final stage or not (Krämer) (**c** and **g**) and clavicle stage 1–5 (Schmeling) (**d** and **h**)

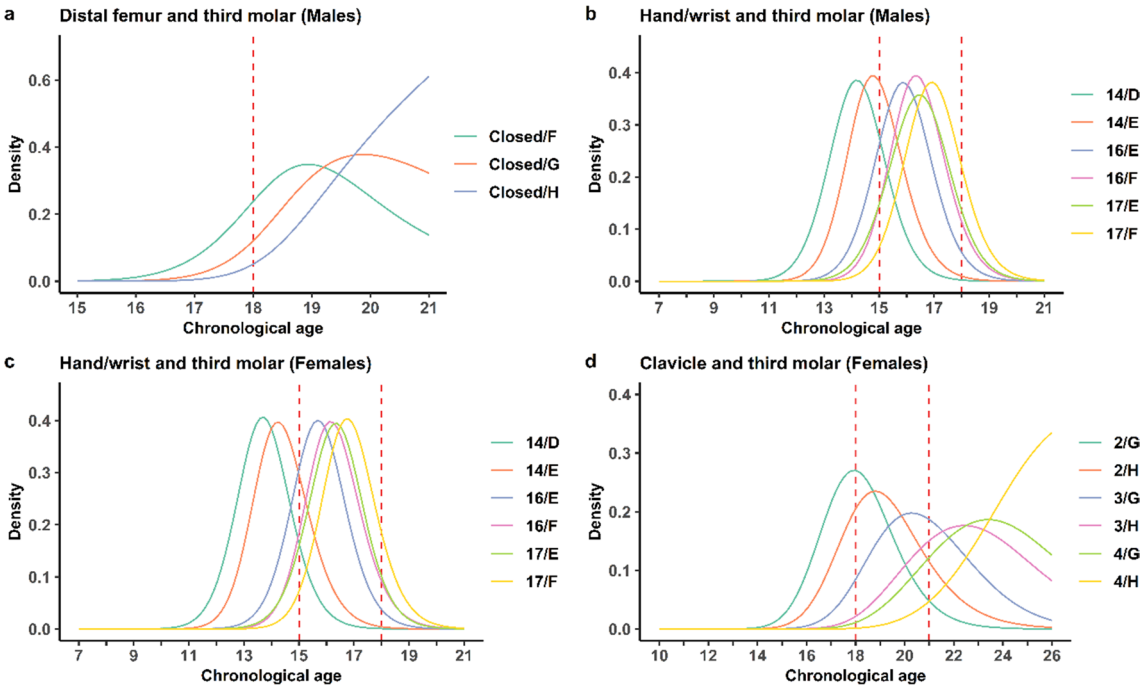



**Fig. 2** Probability density functions for combinations. Age distributions for selected combinations in terms of density of developmental stages for distal femur in combination with third molar (males) (**a**), hand/wrist in combination with third molar (males) (**b**), hand/wrist in combination with third molar (females) (**c**), and clavicle in combination with third molar (females) (**d**). Red dotted line represents age thresholds of interest

the limited number of individuals being in some of the SA development stages in the validation population. The sensitivity (aged over 15 identified as aged over 15) of the male hand/wrist model is 81% and specificity (under 15 identified as under 15) is 92% for the 15-year threshold (Table 4 (b)). The corresponding sensitivity of the female hand/wrist model is 89% and specificity is 91% for the 15-year threshold (Table 4 (b)). Keeping in mind that the proportion of individuals above 18-years of age in the independent population is limited (Fig. 4 (c-f)), the total accuracy with regard to the 18-year threshold for the male model is 93% and for the female model, 90% (Table 4 (b)).

## Validation of the distal femur model

The validation set of the distal femur model included a population of total 217 males (Fig. 5 (a)) and 217 females (Fig. 5 (b)), spanning an age interval between 12–23 years and originates from one dataset (Fig. 5 (a-b) and Table 2). The distal femur model is based on dichotomous development where the Krämer stages 1–3 are defined as open and 4–5 are defined as closed [7, 8], rendering the model useful exclusively for the 18-year threshold. In total 88% of the independent male and 84% of the female population were correctly classified with regard to the 18-year threshold (Table 4 (c)) corresponding to the accuracy. The incorrectly classified individuals are in minority compared to correctly classified (Fig. 5) in both males (e) and females (f). In regard to the 18-year threshold, the model has an acceptable precision when it comes to men (Fig. 5 (c) and (e)), while a closed distal femur in women generates a lower precision (Fig. 5 (d) and (f)). The proportion of individuals being under 18-years of age (blue full line) in the independent population used for validation of the distal femur model basically follows the probabilities of being under 18-years of age (blue dashed line) according to the model (Fig. 5) for males (e) and females (f). The sensitivity (adults identified as adults) of the male distal femur model is 82% and specificity (children identified as children) 96% for the 18-year threshold (Table 4 (c)). The corresponding sensitivity in the female third molar model is 89% and specificity 80% (Table 4 (c)).

## Validation of the clavicle model

The validation set of the clavicle model included a population of total 227 males (Fig. 6 (a)) and 223 females (Fig. 6 (b)), spanning an age interval between 14–30 years and originates from two datasets (Fig. 6 (a-b) and Table 2). Being a skeletal indicator that still develops after 18-years of age renders the clavicle model particularly useful for the 21-year threshold. The validation has been performed for both the 18- and the 21-year threshold. In total 77% of the

**Table 3** Sensitivity analysis of upper age limits for hand/wrist and third molar stages

| Gender | G&P | Demirjian | 19 yr | 20 yr | 21 yr | 22 yr | 23 yr | 24 yr | 25 yr |
|---|---|---|---|---|---|---|---|---|---|
| Male | 18 | - | 0.65 | 0.49 | 0.41 | 0.36 | 0.33 | 0.32 | 0.31 |
| Male | 19 | - | 0.50 | 0.29 | 0.20 | 0.14 | 0.11 | 0.09 | 0.07 |
| Female | 17 | - | 0.85 | 0.78 | 0.76 | 0.74 | 0.74 | 0.74 | 0.74 |
| Female | 18 | - | 0.62 | 0.44 | 0.33 | 0.27 | 0.22 | 0.19 | 0.17 |
| Male | - | F | 0.82 | 0.74 | 0.71 | 0.69 | 0.68 | 0.68 | 0.68 |
| Male | - | G | 0.59 | 0.41 | 0.33 | 0.28 | 0.26 | 0.24 | 0.23 |
| Male | - | H | 0.41 | 0.21 | 0.12 | 0.08 | 0.06 | 0.04 | 0.04 |
| Female | - | F | 0.76 | 0.65 | 0.59 | 0.56 | 0.54 | 0.53 | 0.53 |
| Female | - | G | 0.56 | 0.37 | 0.27 | 0.22 | 0.19 | 0.18 | 0.16 |
| Female | - | H | 0.41 | 0.21 | 0.12 | 0.08 | 0.05 | 0.04 | 0.03 |
| Male | 18 | F | 0.71 | 0.62 | 0.58 | 0.57 | 0.56 | 0.56 | 0.56 |
| Male | 19 | F | 0.58 | 0.43 | 0.37 | 0.34 | 0.33 | 0.32 | 0.32 |
| Male | 18 | G | 0.51 | 0.35 | 0.28 | 0.25 | 0.24 | 0.24 | 0.23 |
| Male | 19 | G | 0.39 | 0.21 | 0.14 | 0.11 | 0.09 | 0.08 | 0.08 |
| Male | 18 | H | 0.36 | 0.18 | 0.12 | 0.09 | 0.07 | 0.06 | 0.06 |
| Male | 19 | H | 0.26 | 0.10 | 0.05 | 0.03 | 0.02 | 0.01 | 0.01 |
| Female | 17 | F | 0.82 | 0.76 | 0.75 | 0.74 | 0.74 | 0.74 | 0.74 |
| Female | 18 | F | 0.62 | 0.47 | 0.40 | 0.37 | 0.35 | 0.34 | 0.34 |
| Female | 17 | G | 0.66 | 0.54 | 0.49 | 0.48 | 0.47 | 0.47 | 0.47 |
| Female | 18 | G | 0.44 | 0.26 | 0.18 | 0.14 | 0.12 | 0.10 | 0.10 |
| Female | 17 | H | 0.52 | 0.36 | 0.29 | 0.26 | 0.25 | 0.25 | 0.25 |
| Female | 18 | H | 0.33 | 0.14 | 0.08 | 0.05 | 0.03 | 0.02 | 0.02 |

Sensitivity analysis of upper age limits. Probabilities of being under 18 when the upper age limit is varied on the uniform distribution for hand/wrist or third molar stages and for the combination of these indicators. The green color indicates the upper limit applied in the statistical model

male and 85% of the female validation population were correctly classified with regard to the 18-year threshold and 75% of the males and 78% of the females to the 21-year threshold (Table 4 (d)) corresponding to the accuracy. The sensitivity (above 21 identified as above) of the male clavicle model is 59% and the specificity (below 21 identified as below 21) is 96% for the 21-year threshold (Table 4 (d)). The corresponding sensitivity in the female clavicle model is 64% and specificity 95% (Table 4 (d)). The incorrectly classified individuals, with regard to the 21-year threshold is mainly individuals in development stage 3 (Fig. 6) for both males (e and g) and females (f and h). For the 18-year threshold, the incorrectly classified individuals are mainly in development stage 2. The proportion of individuals being under 21-years of age (orange full line) in the independent population used for validation of the clavicle model basically follows the probabilities of being under 21-years of age (orange dashed line) according to the model (Fig. 6) for males (g) and females (h), indicating a high reliability of the prediction model. In regard to the 18-year threshold, the validation (blue full line) deviates more from the probabilities according to the prediction model (dashed blue lines) indicating a lower precision compared to the 21-year threshold (orange) (Fig. 6 (g and h).

## Validating the model on a test set with both third molar and hand/wrist

The precision of the age estimation increases when the result from multiple developmental indicators are combined, which

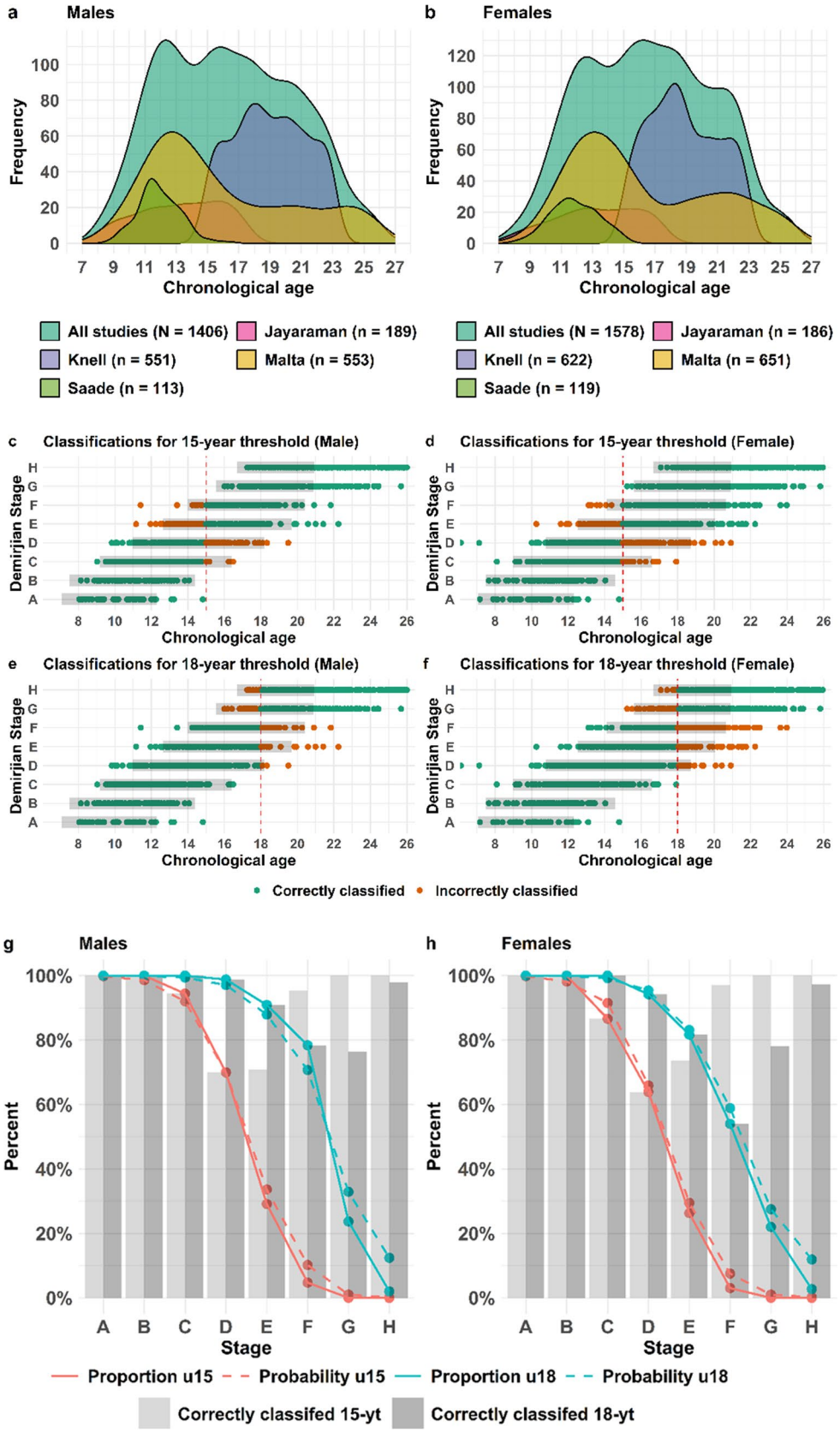



**Fig. 3** Validation of the third molar model Validation of the third molar model. Distribution of the full validation dataset and the separate studies are shown for males (**a**) and females (**b**). Point plots displaying the chronological age and corresponding Demirjian development stage of the third molar together with classification with regard to the 15-or 18- year threshold for males ((**c**) and (**e**)) and females ((**d**) and (**f**)). Grey bars in (**c-f**) represents the 95% PI for each development stage. The proportion in the validation set (full line) being under 15(orange) or 18 (blue) for each development stage together with the predicted probability according to the statistical model (dashed lines) for males (**g**) and females (**h**). The proportion of the validation set being correctly classified (**g-h**) with regard to the 15-year threshold (light grey bar) and the 18-year threshold (dark grey bar) is displayed for each development stage for males (**g**) and females (**h**)

**Table 4** Quantitative reliability of the models

| (a) | | | | |
|---|---|---|---|---|
| Measure | Third molar model | | | |
| | Males | | Females | |
| | 15-year threshold | 18-year threshold | 15-year threshold | 18-year threshold |
| Sensitivity | 0.90 | 0.90 | 0.86 | 0.75 |
| Specificity | 0.89 | 0.95 | 0.89 | 0.94 |
| PPV | 0.93 | 0.91 | 0.93 | 0.89 |
| NPV | 0.85 | 0.94 | 0.78 | 0.86 |
| Accuracy | 0.90 | 0.93 | 0.87 | 0.87 |
| (b) | | | | |
| Measure | Hand/wrist model | | | |
| | Males | | Females | |
| | 15-year threshold | 18-year threshold | 15-year threshold | 18-year threshold |
| Sensitivity | 0.81 | 0.33 | 0.89 | 0.00 |
| Specificity | 0.92 | 0.99 | 0.91 | 1.00 |
| PPV | 0.81 | 0.67 | 0.73 | - |
| NPV | 0.91 | 0.96 | 0.97 | 0.94 |
| Accuracy | 0.88 | 0.95 | 0.91 | 0.94 |
| (c) | | | | |
| Measure | Hand/wrist model | | | |
| | Males | | Females | |
| | 15-year threshold | 18-year threshold | 15-year threshold | 18-year threshold |
| Sensitivity | 0.81 | 0.33 | 0.89 | 0.00 |
| Specificity | 0.92 | 0.99 | 0.91 | 1.00 |
| PPV | 0.81 | 0.67 | 0.73 | - |
| NPV | 0.91 | 0.96 | 0.97 | 0.94 |
| Accuracy | 0.88 | 0.95 | 0.91 | 0.94 |
| (d) | | | | |
| Measure | Distal femur model | | | |
| | Males | | Females | |
| | 15-year threshold | 18-year threshold | 15-year threshold | 18-year threshold |
| Sensitivity | - | 0.82 | - | 0.91 |
| Specificity | - | 0.96 | - | 0.78 |
| PPV | - | 0.96 | - | 0.77 |
| NPV | - | 0.82 | - | 0.92 |
| Accuracy | - | 0.88 | - | 0.84 |
| (e) | | | | |
| Measure | Clavicle model | | | |
| | Males | | Females | |
| | 18-year threshold | 21-year threshold | 18-year threshold | 21-year threshold |
| Sensitivity | 0.76 | 0.59 | 0.86 | 0.64 |
| Specificity | 0.82 | 0.96 | 0.82 | 0.95 |
| PPV | 0.97 | 0.95 | 0.97 | 0.94 |
| NPV | 0.32 | 0.65 | 0.45 | 0.69 |
| Accuracy | 0.77 | 0.75 | 0.85 | 0.78 |

Quantitative measures of the model's reliability for the third molar model (a) hand/wrist model (b) distal femur model (c) clavicle model (d) and the combination of hand/wrist with third molar (d). Quantifications based on the independent dataset validations. Sensitivity (CA above age limit and identified as above the age limit), specificity (CA under age limit identified as under the age limit), positive predictive value ((PPV) individuals identified as above the age limit with CA above the age limit), negative predictive value ((NPV) individuals identified as below the age limit with CA below the age limit) and total accuracy presented for the 15-year and/or 18-year and/or 21-year threshold for the male and female models

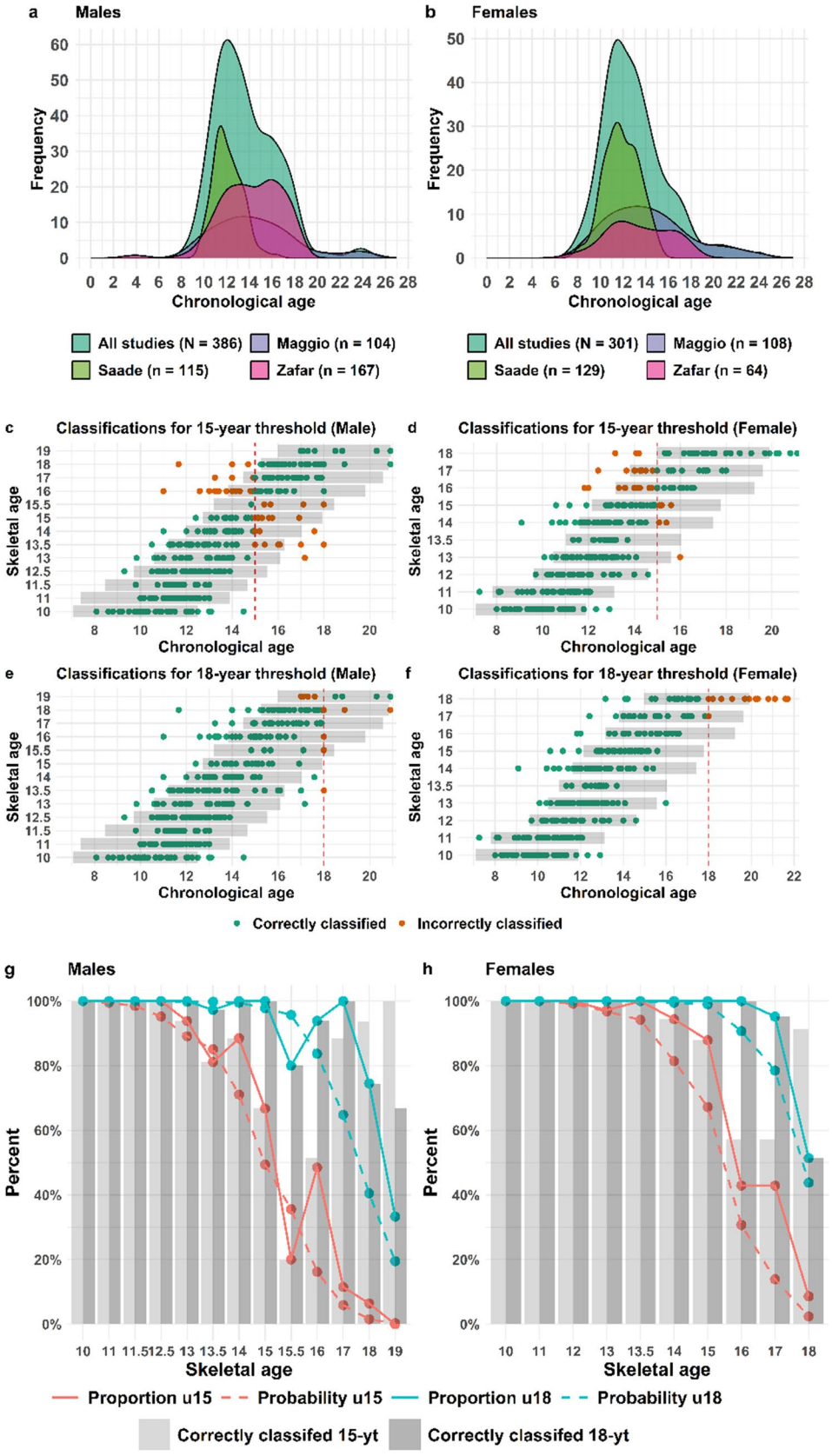



**Fig. 4** Validation of the hand/wrist model. Distribution of the full validation dataset and the separate studies for males(**a**) and females(**b**). Point plots displaying the chronological age and corresponding G&P development stage of hand/wrist together with classification with regard to the 15- or 18- year threshold for males (**c**) and (**e**) and females (**d**) and (**f**). Grey bars in (**c**-**f**) represents the 95% PI for each development stage in the model. The proportion in the validation set (full line) being under 15 (orange) or 18 (blue) for each development stage together with the predicted probability according to the statistical model (dashed lines) for males (**g**) and females (**h**). The proportion of the validation set being correctly classified with regard to the 15-year threshold (light grey bar) and the 18-year threshold (dark grey bar) displayed for each development stage in the model for males (**g**) and females (**h**)

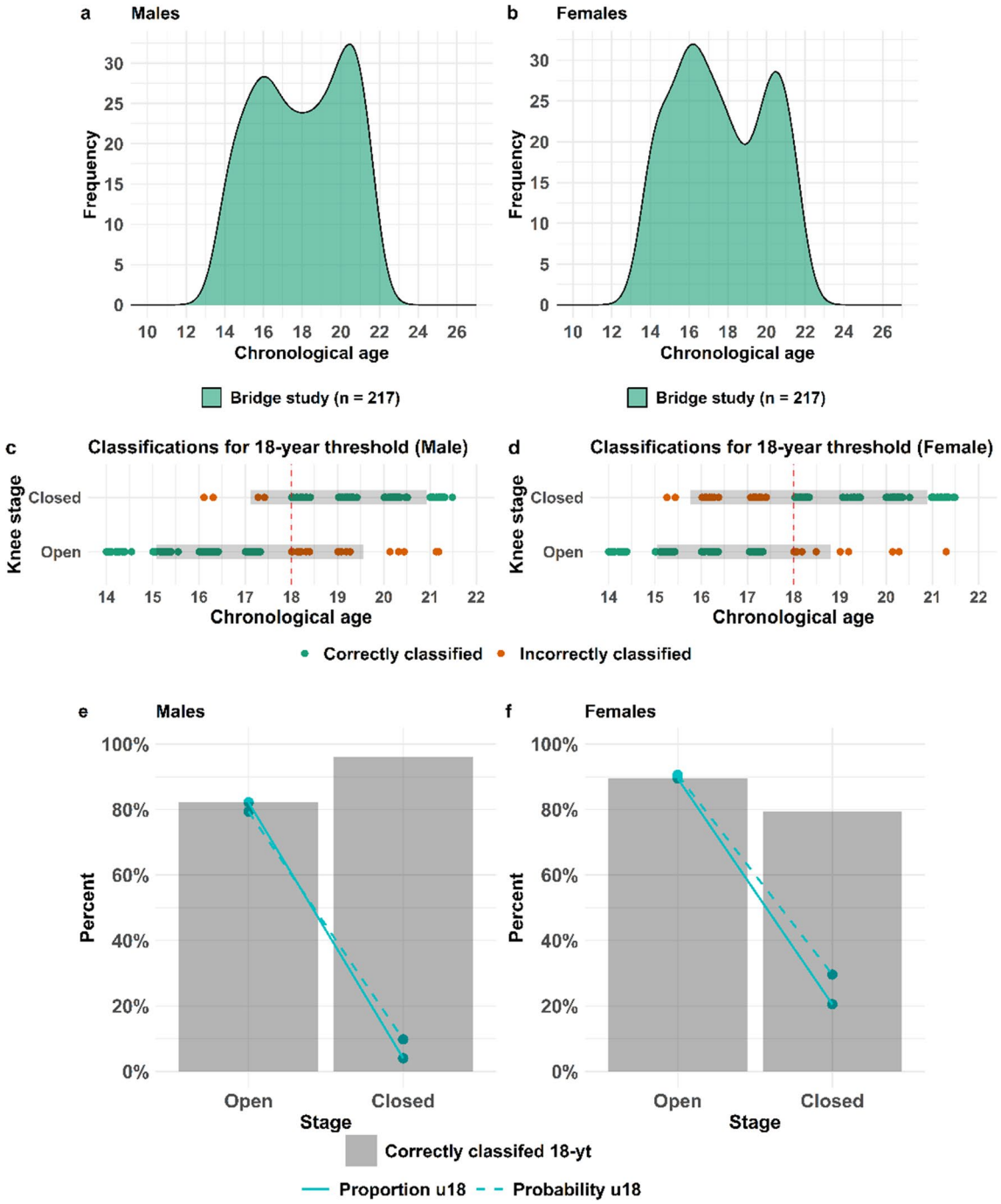



**Fig. 5** Validation of the distal femur model. Distribution of the full validation dataset for males (**a**) and females(**b**). Point plots displaying the chronological age and corresponding dichotomous development stage of the distal femur together with classification with regard to the 18-year threshold for males (**c**) and females (**d**). Grey bars in (**c-f**) represents the 95% PI for each development stage in the model. The proportion in the validation set (full line) being under 18 (blue) for the development stages together with the predicted probability according to the statistical model (dashed lines) for males (**e**) and females (**f**). The proportion of the validation set being correctly classified with regard to the 18-year threshold (dark grey bar) displayed for the two development stages for males (**e**) and females (**f**)

corresponds to how the model is recommended to be used in practice. This means that the result from the independent models underestimates the real precision when used in practice. Here, we test our model against one dataset where both third molars and hand/wrist development has been examined in the same individuals, along with CA. The validation data included an independent population of total 106 males and 116 females (Supplementary Fig. 12 (a-b) and Table 2,

spanning an age interval between 8–16 years (Supplementary Fig. 12). Classification with Demirjian's method of the lower left third molar together with the Greulich &Pyle grading of the hand skeleton were applied on individuals in this Lebanese population [28]. The validation of the combined model is limited in that the validation population mostly includes individuals younger than 15 years. However, it is a valuable dataset in that it confirms the higher specificity as demonstrated by a tighter PI compared to single indicators (Supplementary Fig. 9) and a high number of correctly classified under 15 represented by a high specificity for both males (Supplementary Fig. 12 (c) and Table 4 (e)) and females (Supplementary Fig. 12 (d) and Table 4 (e)). In total 96% of the independent male and 97% of the female populations were correctly classified with regard to the 15-year threshold representing the accuracy (Supplementary Fig. 12 (c-d) and Table 4 (e)).

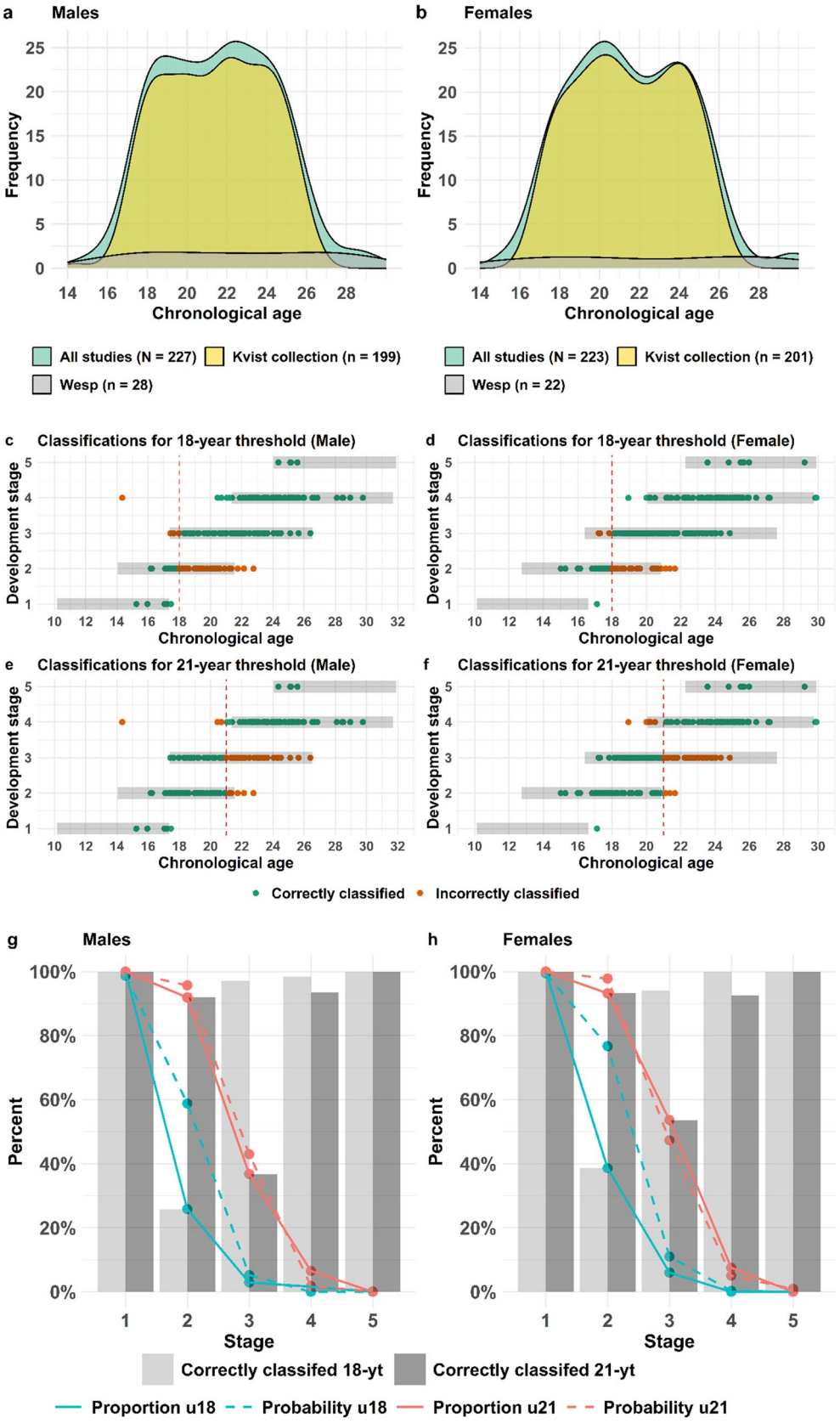
## Discussion

Reliable methods for age estimation in living individuals are of major importance in legal contexts when birth records or other official identification documents are missing. The main aim of this study is to generate and present a validated statistical model for estimating age in living individuals relative to the 15, 18 or 21-year old thresholds. To our knowledge, this is the first model to include several skeletal indicators combined with third molar development to provide assessments for several age thresholds that has been validated with independent datasets. It could be argued that our model addresses the knowledge gap concerning the objective utilization of multiple anatomical locations and statistical models to enhance the accuracy of estimating an individual's age. The spectrum of methods recommended by the Study Group on Forensic Age Diagnostics in Münster include radiography examination of the hand/wrist and third molars as well as CT clavicle, which may also be supplemented with MRI of distal femur in the future [29]. However, their recommended approach is to add CT clavicle if hand/wrist is fully developed and to use these examinations in a minimal age concept rather than a probability approach. Their recommended methods also include a physical examination and recording of sexual maturity [29], even though the latter is noticed to be against the EASO recommended guidelines [1]. In the statistical model investigated here, radiography of third molar is combined with either radiography hand/wrist, CT clavicle or MRI distal femur depending on the age threshold of interest. The estimation of age from dental radiographs is one of the most studied and widely used approaches, and the Demirjian staging technique is the most widely used staging method in studies focusing on age estimation [6, 30]. Demirjian's staging of the wisdom tooth is well suited to assess both the 15- and 18-year threshold (Fig. 1 (b and f). Due to a chosen upper age limit at 21 years for the third molar model, it is not suited to assess the 21-year threshold as a single indicator. However, in combination with the clavicle, a slightly older assumed age distribution has been included in the model that renders it suitable (Fig. 2). The higher age as a chosen upper age limit of the third molar in this combination is motivated by the fact that the PI in the combined model is tighter than the clavicle model alone (Supplementary Fig. 11). Radiography of the hand/wrist is internationally the most widely applied method to assess skeletal development [5, 16, 31]. The development stages of hand/wrist are suitable for assessing the 15-year threshold in males and females and possibly the 18-year threshold in males, based on the development stage distributions (Fig. 1 (a and e)). The dichotomous distal femur model is suitable for the 18-year threshold in males while an open development stage can be used in women to indicate minority status (Fig. 1 (c and g)). The medial clavicle epiphysis is considered useful for the 21-year threshold due to a continued development until around age 30 [32–35] (Fig. 1(d and h)).

To create reliable and detailed assessment models, a much larger data set than typically found in a single study is required. The underlying reference population needs to cover all relevant age cohorts that also allow a Bayesian approach to minimize the effect of age mimicry from the underlying studies [12]. Several probability methods have previously been presented in the literature [5, 7, 16]. All these methods have the advantage of relying on larger reference populations when providing age distributions, unlike other assessment approaches that compare with only one limited study population [36]. None of the models will provide a definite age for an individual but in the case of the probability methods, either an age span [5] or a probability of an age in relation to a threshold [7] will be provided, together with an error rate. These probabilities are the base to form the medical component for the overall assessment of an individual's age.

It has been argued that population-specific reference data is needed in age assessments. According to current scientific understanding, the ethnicity or genetic-geographic origin of an individual may not significantly impact the dental- or skeletal maturity [37–41]. It is noted that a study by Olze et al. [42] as well as a review on dental age estimation [43] cautions against possible differences in dental aging between populations and ethnicities. However, as pointed out before [7] and shown in Rolseth et al. [6], studies might be subject to age mimicry, meaning that the observed difference between populations is likely to reflect differences in the underlying age distributions of the study population rather than inherent differences in development.

Factors such as stress or living standard have been suggested to influence skeletal development [38, 44, 45]. Consequently, individuals from lower socioeconomic backgrounds


a Males
b Females
Frequency
Chronological age
All studies (N = 227)
Kvist collection (n = 199)
Wesp (n = 28)
All studies (N = 223)
Kvist collection (n = 201)
Wesp (n = 22)
c Classifications for 18-year threshold (Male)
d Classifications for 18-year threshold (Female)
e Classifications for 21-year threshold (Male)
f Classifications for 21-year threshold (Female)
Development stage
Correctly classified
Incorrectly classified
g Males
h Females
Percent
Stage
Correctly classifed 18-yt
Correctly classifed 21-yt
Proportion u18
Probability u18
Proportion u21
Probability u21

◂**Fig. 6** Validation of the clavicle model. Distribution of the full validation dataset for males (**a**) and females(**b**). Point plots displaying the chronological age and corresponding dichotomous development stage of the clavicle together with classification with regard to the 18- and 21- year threshold for males (**c**) and females (**d**). Grey bars in (**c-f**) represents the 95% PI for each development stage in the model. The proportion in the validation set (full line) being under 18 (blue) or 21 (orange) for the development stages together with the predicted probability according to the statistical model (dashed lines) for males (**e**) and females (**f**). The proportion of the validation set being correctly classified with regard to the 18-(grey bar) or 21-year threshold (dark grey bar) displayed for the five development stages for males (**e**) and females (**f**)

undergoing medical age assessments may face the risk of being estimated as younger than their CA. In line with the approach of the BioAlder tool [5], we have opted to incorporate a broad spectrum of individuals from chosen studies into the reference population. This decision aims to encompass the widest possible range of biological variations in age-dependent development, striving for thorough coverage. The single studies covering a single geographic region, socio-economic or other possible influencing factors are argued too small to provide reliable reference populations on their own. The total number of individuals included in the model is high (27,000), but is unequally distributed between the included indicators. The number of studies (34) is limited by covering 6 geographic regions and the main limitation factor is the availability of studies focusing on age in relation to development and fulfilling the pre-set criteria. Similar to the previous statistical models [5, 7], the results in this model are dependent on the assumptions for the underlying age distributions, conditional independence and simulations as well as study selection.

Given the inevitable diversity in underlying studies and limited ethnic representation, a key concern that arises when developing a prediction tool is: how accurately does the tool perform for the individuals we intend to predict? The availability of independent complete data sets is scarce, yet essential to perform a validation of the model compared to real world data. The validation of this model with collected independent populations indicates a high accuracy and precision for all indicators, particularly for the third molar model and the distal femur.

When combining dental and skeletal indicators, only a few individuals were wrongly classified with regard to the 15-year threshold in the validation of the combined third molar and hand/wrist model. Considering that the age span in this validation set is limited to a population almost exclusively under 15-years of age, it is possible to establish an adequate level of precision for these individuals, but not for individuals over 15. It has been concluded that a multifactorial age estimation is more accurate than one based on a single anatomical site [46, 47]. Multifactorial age estimation is also recommended by the Münster-based AGFAD study group [29]. An important consideration of multifactorial age estimation is the risk of increased ionizing radiation to a young individual which is against the EASO guidelines and ALARA (as low as reasonably achievable) principle. However, the availability of datasets containing concurrent grading of third molars with a skeletal indicator in the same individuals is limited, and efforts to simultaneously measure multiple developmental indicators would allow for more robust estimations of model accuracy.

The validation with the independent populations has pinpointed and confirmed the predicted development stages that are associated with the highest uncertainties. For instance, 30–40% of the individuals in third molar development stage D in both males and females are wrongly classified with regard to the 15-year threshold (Fig. 3 (g-h)), and this uncertainty agrees with the prediction provided by the model, that these individuals are below 15, with a margin of error of 30% and 35% for males and females, respectively. When applying the model on individuals with an unknown age, the degree of certainty in the statement needs to reflect the estimated age distribution and the probability of being below or above the age limit together with this margin of error that corresponds to the proportion of the reference population on the other side of the limit. The presented validation allows reliable assessments together with margin of errors to be provided.

To facilitate medical age assessments in routine practice using this complex statistical model, a user-friendly tool is advisable. Such a dashboard has been developed to streamline these assessments by forensic pathologists in Sweden. Dropdown menus allow the assessor to populate the model with the current combination of examinations performed together with gender and development stages. The corresponding distribution of the reference population is then displayed together with 95% PI, probability for the three age thresholds together with probabilities in one-year cohorts. This tool provides the probabilities and the measure of margin of error.

A promising tool for faster and more accurate radiological age assessments are artificial intelligence (AI) approaches [30, 35, 48–50]. Methods using AI necessitate a substantial volume of data for construction and are not exempt from the conventional questions inherent in age assessments, such as biologic variation, the socioeconomic dimension or other factors influencing development. An AI tool, based on third molar development in a Brazilian population, presents a binary assessment with high accuracy of being above or below a specific age threshold [49]. In addition, a high accuracy performing AI-model of age classification with regard to 18, 20, 21 and 22-year thresholds based on clavicle development was recently presented in a Chinese study [35]. Notably, a common feature of these methods is that they achieve a high level of accuracy. Even though

additional studies are required, deep learning approaches remain a promising vision for the future following validation on a broader scale.

## Limitations

The complex relationship between skeletal or dental development and CA presents an unavoidable barrier to achieving perfect accuracy in age assessment methods [6, 51]. Even though our approach has been to include a broad spectrum of studies performed in different countries and geographic regions in the reference population, the ethnic and socioeconomic variation is still limited. The retrospective nature of data collection and the fact that studies are conducted with slightly different protocols and/or data reporting, may introduce variations. The evaluation of the accuracy and precision of the probability model is limited by the access to independent validation populations where multiple indicators have been measured. Although one of the models is based on magnetic resonance imaging, this tool is not entirely devoid of potentially harmful ionizing radiation.

## Conclusion

In summary, our study presents a validated statistical model for estimating an age relative to key legal thresholds (15, 18, and 21 years) based on a skeleton (CT-clavicle, radiography-hand/wrist or MR-knee) and teeth (radiography-third molar) developmental stages allowing to provide reliable assessments with margin of errors. This probability model provides a most likely age distribution based on a large reference population rather than an indeterminable CA. The assessment based on the model generated probabilities form the medical component for the overall assessment of an individual's age.While statistical models are by nature complex, the creation of a dashboard may easier facilitate and streamline individual assessments in routine practice. Although AI approaches are in development, providing a validated probability method addresses a knowledge gap and is of high interest as currently, no available method can provide a reliable CA.

**Supplementary Information** The online version contains supplementary material available at https://doi.org/10.1007/s00414-024-03324-x.

**Acknowledgements** We thank Daria Medvedeva and SDS Life Science AB (Cytel Sweden) for critically reviewing the R-code and assumptions made in the statistical model. We would also like to thank Balthasar Maria Schachtner for data collection and Bastian Oliver Sabel for ossification stage reading in the Wesp 2024 data collection. We also thank Abdul Mueed Zafar, Ariane Maggio, Marianne Saade, Bernhard Knell and Jayakumar Jayaraman for providing data and Alexander Tyr for proofreading.

**Author contribution** Nina Heldring: Conceptualization, Methodology, Investigation, Writing—original draft, review & editing. Ali-Reza Rezaie: Methodology, Writing—review & editing. André Larsson: Methodology, Writing—review & editing**.** Rebecca Gahn: Writing—review & editing. Brita Zilg: Writing—review & editing. Simon Camilleri: Investigation, Writing-review & editing. Antoine Saade: Investigation, Writing—review & editing. Philipp Wesp: Investigation, Writing—review & editing. Elias Palm: Conceptualization, Writing—review & editing. Ola Kvist: Methodology, Investigation, Writing—review & editing.

**Funding** Open access funding provided by Swedish National Board of Forensic Medicine. This work was fully supported by the Swedish National Board of Forensic Medicine.

**Code availability** The source code for running all the modeling, simulations, and providing the results as well as the dashboard can be obtained by contacting the first author.

## Declarations

**Ethical approval** The retrospective collection and assessment of development stage of clavicle to generate the validation population was approved by the Swedish Ethics Review Authority (Approval number Dnr 2024–00531-01). The Ethics Committee, Medical Faculty, LMU Munich approved the sharing of a retrospectively collected and assessed clavicle dataset (20–324).

**Competing interest** The authors declare no conflict of interest.



## References

1. EASO practical guide on age assessment, 2nd edn (2018) European asylum support office. https://euaa.europa.eu/sites/default/files/easo-practical-guide-on-age-assesment-v3-2018.pdf
2. Report: Biological evaluation methods to assist in assessing the age of unaccompanied asylum-seeking children. Interim age estimation science advisory committee, home office UK. Published 0 January 2023. https://www.gov.uk/government/publications/methods-to-assess-the-age-of-unaccompanied-asylum-seeking-children
3. Anderson M (1971) Use of the Greulich-Pyle "Atlas of Skeletal Development of the Hand and Wrist" in a clinical context. Am J Phys Anthropol 35:347–352. https://doi.org/10.1002/ajpa.1330350309
4. Bachs L, Bleka Ø, Dahlberg PS, Rolseth V, Delaveris G-JM (2020) BioAlder: a tool for using biological tests to assess the age of unaccompanied minor asylum-seekers. BioAlder Manual Version 3b. Oslo Universitetssykehus

5. Bleka O, Rolseth V, Dahlberg PS, Saade A, Saade M, Bachs L (2019) BioAlder: a tool for assessing chronological age based on two radiological methods. Int J Legal Med 133:1177–1189. https://doi.org/10.1007/s00414-018-1959-5
6. Rolseth V, Mosdol A, Dahlberg PS et al (2019) Age assessment by Demirjian's development stages of the third molar: a systematic review. Eur Radiol 29:2311–2321. https://doi.org/10.1007/s00330-018-5761-z
7. Heldring N, Larsson A, Rezaie AR, Rasten-Almqvist P, Zilg B (2022) A probability model for assessing age relative to the 18-year old threshold based on magnetic resonance imaging of the knee combined with radiography of third molars in the lower jaw. Forensic Sci Int 330:111108. https://doi.org/10.1016/j.forsciint.2021.111108
8. Kramer JA, Schmidt S, Jurgens KU, Lentschig M, Schmeling A, Vieth V (2014) Forensic age estimation in living individuals using 3.0 T MRI of the distal femur. Int J Legal Med 128:509–514. https://doi.org/10.1007/s00414-014-0967-3
9. Schmeling A, Schulz R, Reisinger W, Muhler M, Wernecke KD, Geserick G (2004) Studies on the time frame for ossification of the medial clavicular epiphyseal cartilage in conventional radiography. Int J Legal Med 118:5–8. https://doi.org/10.1007/s00414-003-0404-5
10. Ankit R (n.d.) WebPlotDigitizer, ed 4.6. https://automeris.io
11. Mersmann O, Trautmann H, Steuer D, Bornkamp B (2023) Truncnorm: truncated normal distribution. R package version 1.0–9. https://CRAN.R-project.org/package=truncnorm
12. Boldsen JL, Milner GR, Konigsberg LW, Wood JW (2002) Transition analysis: a new method for estimating age from skeletons. Paleodemography. Cambridge University Press 2009:73–106. https://doi.org/10.1017/cbo9780511542428.005
13. Roberts GJ, McDonald F, Andiappan M, Lucas VS (2015) Dental Age Estimation (DAE): data management for tooth development stages including the third molar. Appropriate censoring of Stage H, the final stage of tooth development. J Forensic Leg Med 36:177–184. https://doi.org/10.1016/j.jflm.2015.08.013
14. Knell B, Ruhstaller P, Prieels F, Schmeling A (2009) Dental age diagnostics by means of radiographical evaluation of the growth stages of lower wisdom teeth. Int J Legal Med 123:465–469. https://doi.org/10.1007/s00414-009-0330-2
15. Olze A, Pynn BR, Kraul V et al (2010) Studies on the chronology of third molar mineralization in First Nations people of Canada. Int J Legal Med 124:433–437. https://doi.org/10.1007/s00414-010-0483-z
16. Bleka O, Wisloff T, Dahlberg PS, Rolseth V, Egeland T (2019) Advancing estimation of chronological age by utilizing available evidence based on two radiographical methods. Int J Legal Med 133:217–229. https://doi.org/10.1007/s00414-018-1848-y
17. Varkkola O, Ranta H, Metsaniitty M, Sajantila A (2011) 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 7:311–316. https://doi.org/10.1007/s12024-010-9173-x
18. Gelbrich B, Frerking C, Weiss S et al (2015) Combining wrist age and third molars in forensic age estimation: how to calculate the joint age estimate and its error rate in age diagnostics. Ann Hum Biol 42:389–396. https://doi.org/10.3109/03014460.2015.1046487
19. Kumari S, Sahu AK, Rajguru J, Bishnoi P, Garg AJ, Thakur R (2022) Age estimation by dental calcification stages and hand-wrist radiograph. Cureus 14:e29045. https://doi.org/10.7759/cureus.29045
20. Akaike H (1992) Information theory and an extension of the maximum likelihood principle. In: Kotz S, Johnson NL (eds) Breakthroughs in statistics: foundations and basic theory. Springer, New York New York, NY, pp 610–624
21. Team RC (2021) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria
22. Venables WN, Ripley BD (2002) Modern Applied Statistics with S, 4th edn. Springer, New York
23. Kellinghaus M, Schulz R, Vieth V, Schmidt S, Pfeiffer H, Schmeling A (2010) Enhanced possibilities to make statements on the ossification status of the medial clavicular epiphysis using an amplified staging scheme in evaluating thin-slice CT scans. Int J Legal Med 124:321–325. https://doi.org/10.1007/s00414-010-0448-2
24. Wickham H (2016) ggplot2: elegant graphics for data analysis. Springer-Verlag, New York. https://CRAN.R-project.org/package=ggplot2
25. Ensor J (2023) pmsampsize: sample size for development of a prediction model. R package version 1.1.3. https://CRAN.R-project.org/package=pmsampsize
26. Riley RD, Debray TPA, Collins GS et al (2021) Minimum sample size for external validation of a clinical prediction model with a binary outcome. Stat Med 40:4230–4251. https://doi.org/10.1002/sim.9025
27. Snell KIE, Archer L, Ensor J et al (2021) External validation of clinical prediction models: simulation-based sample size calculations were more reliable than rules-of-thumb. J Clin Epidemiol 135:79–89. https://doi.org/10.1016/j.jclinepi.2021.02.011
28. Saade A, Baron P, Noujeim Z, Azar D (2017) Dental and skeletal age estimations in Lebanese children: a retrospective cross-sectional study. J Int Soc Prev Community Dent 7:90–97. https://doi.org/10.4103/jispcd.JISPCD_139_17
29. Wittschieber D, Hahnemann ML, Mentzel HJ (2024) Forensic diagnostics of the skeletal age in the living - backgrounds and methodology. Rofo 196:254–261. https://doi.org/10.1055/a-2130-3162
30. Vila-Blanco N, Varas-Quintana P, Tomas I, Carreira MJ (2023) A systematic overview of dental methods for age assessment in living individuals: from traditional to artificial intelligence-based approaches. Int J Legal Med 137:1117–1146. https://doi.org/10.1007/s00414-023-02960-z
31. Manzoor Mughal A, Hassan N, Ahmed A (2014) Bone age assessment methods: a critical review. Pak J Med Sci 30:211–215. https://doi.org/10.12669/pjms.301.4295
32. Pattamapaspong N, Madla C, Mekjaidee K, Namwongprom S (2015) Age estimation of a Thai population based on maturation of the medial clavicular epiphysis using computed tomography. Forensic Sci Int 246(123):e1-5. https://doi.org/10.1016/j.forsciint.2014.10.044
33. Houpert T, Rerolle C, Savall F, Telmon N, Saint-Martin P (2016) Is a CT-scan of the medial clavicle epiphysis a good exam to attest to the 18-year threshold in forensic age estimation? Forensic Sci Int 260(103):e1–e3. https://doi.org/10.1016/j.forsciint.2015.12.007
34. Torimitsu S, Makino Y, Saitoh H et al (2019) Age estimation based on maturation of the medial clavicular epiphysis in a Japanese population using multidetector computed tomography. Leg Med (Tokyo) 37:28–32. https://doi.org/10.1016/j.legalmed.2018.12.003
35. Qiu L, Liu A, Dai X et al (2024) Machine learning and deep learning enabled age estimation on medial clavicle CT images. Int J Legal Med 138:487–498. https://doi.org/10.1007/s00414-023-03115-w
36. Schmeling A, Grundmann C, Fuhrmann A et al (2008) Criteria for age estimation in living individuals. Int J Legal Med 122:457–460. https://doi.org/10.1007/s00414-008-0254-2
37. Pechnikova M, Gibelli D, De Angelis D, de Santis F, Cattaneo C (2011) The "blind age assessment": applicability of Greulich and Pyle, Demirjian and Mincer aging methods to a population of

unknown ethnic origin. Radiol Med 116:1105–1114. https://doi.org/10.1007/s11547-011-0694-5
38. Schmeling A, Reisinger W, Loreck D, Vendura K, Markus W, Geserick G (2000) Effects of ethnicity on skeletal maturation: consequences for forensic age estimations. Int J Legal Med 113:253–258. https://doi.org/10.1007/s004149900102
39. Meijerman L, Maat GJ, Schulz R, Schmeling A (2007) Variables affecting the probability of complete fusion of the medial clavicular epiphysis. Int J Legal Med 121:463–468. https://doi.org/10.1007/s00414-007-0189-z
40. Cameriere R, De Luca S, Ferrante L (2021) Study of the ethnicity's influence on the third molar maturity index (I(3M)) for estimating age of majority in living juveniles and young adults. Int J Legal Med 135:1945–1952. https://doi.org/10.1007/s00414-021-02622-y
41. Thevissen PW, Alqerban A, Asaumi J et al (2010) Human dental age estimation using third molar developmental stages: accuracy of age predictions not using country specific information. Forensic Sci Int 201:106–111. https://doi.org/10.1016/j.forsciint.2010.04.040
42. Olze A, Schmeling A, Taniguchi M et al (2004) Forensic age estimation in living subjects: the ethnic factor in wisdom tooth mineralization. Int J Legal Med 118:170–173. https://doi.org/10.1007/s00414-004-0434-7
43. De Donno A, Angrisani C, Mele F, Introna F, Santoro V (2021) Dental age estimation: Demirjian's versus the other methods in different populations. A literature review. Med Sci Law 61:125–129. https://doi.org/10.1177/0025802420934253
44. Schmeling A, Schulz R, Danner B, Rosing FW (2006) The impact of economic progress and modernization in medicine on the ossification of hand and wrist. Int J Legal Med 120:121–126. https://doi.org/10.1007/s00414-005-0007-4
45. Cardoso HF (2007) Environmental effects on skeletal versus dental development: Using a documented subadult skeletal sample to test a basic assumption in human osteological research. Am J Phys Anthropol 132:223–233. https://doi.org/10.1002/ajpa.20482
46. De Tobel J, Bauwens J, Parmentier GIL et al (2020) Magnetic resonance imaging for forensic age estimation in living children and young adults: a systematic review. Pediatr Radiol 50:1691–1708. https://doi.org/10.1007/s00247-020-04709-x
47. De Tobel J, Fieuws S, Hillewig E et al (2020) Multi-factorial age estimation: a Bayesian approach combining dental and skeletal magnetic resonance imaging. Forensic Sci Int 306:110054. https://doi.org/10.1016/j.forsciint.2019.110054
48. Kim PH, Yoon HM, Kim JR et al (2023) Bone age assessment using artificial intelligence in Korean pediatric population: a comparison of deep-learning models trained with healthy chronological and Greulich-Pyle ages as labels. Korean J Radiol 24:1151–1163. https://doi.org/10.3348/kjr.2023.0092
49. Franco A, Murray J, Heng D et al (2024) Binary decisions of artificial intelligence to classify third molar development around the legal age thresholds of 14, 16 and 18 years. Sci Rep 14:4668. https://doi.org/10.1038/s41598-024-55497-5
50. Wesp P, Schachtner BM, Jeblick K et al (2024) Radiological age assessment based on clavicle ossification in CT: enhanced accuracy through deep learning. Int J Legal Med 138:1497–1507. https://doi.org/10.1007/s00414-024-03167-6
51. Dahlberg PS, Mosdol A, Ding Y et al (2019) A systematic review of the agreement between chronological age and skeletal age based on the Greulich and Pyle atlas. Eur Radiol 29:2936–2948. https://doi.org/10.1007/s00330-018-5718-2
52. Alcina M, Lucea A, Salicru M, Turbon D (2018) Reliability of the Greulich and Pyle method for chronological age estimation and age majority prediction in a Spanish sample. Int J Legal Med 132:1139–1149. https://doi.org/10.1007/s00414-017-1760-x
53. Bala M, Pathak A, Jain RL (2010) Assessment of skeletal age using MP3 and hand-wrist radiographs and its correlation with dental and chronological ages in children. J Indian Soc Pedod Prev Dent 28:95–99. https://doi.org/10.4103/0970-4388.66746
54. Büken B, Safak AA, Yazici B, Büken E, Mayda AS (2007) Is the assessment of bone age by the Greulich-Pyle method reliable at forensic age estimation for Turkish children? Forensic Sci Int 173:146–153. https://doi.org/10.1016/j.forsciint.2007.02.023
55. Cantekin K, Yilmaz Y, Demirci T, Celikoglu M (2012) Morphologic analysis of third-molar mineralization for eastern Turkish children and youth. J Forensic Sci 57:531–534. https://doi.org/10.1111/j.1556-4029.2011.02011.x
56. Chaumoitre K, Saliba-Serre B, Adalian P, Signoli M, Leonetti G, Panuel M (2017) Forensic use of the Greulich and Pyle atlas: prediction intervals and relevance. Eur Radiol 27:1032–1043. https://doi.org/10.1007/s00330-016-4466-4
57. Dembetembe KA, Morris AG (2012) Is Greulich-Pyle age estimation applicable for determining maturation in male Africans? S Afr J Sci 108. https://doi.org/10.4102/sajs.v108i9/10.1036
58. Elamin F, Abdelazeem N, Elamin A, Saif D, Liversidge HM (2017) Skeletal maturity of the hand in an East African group from Sudan. Am J Phys Anthropol 163:816–823. https://doi.org/10.1002/ajpa.23247
59. Hackman L, Black S (2013) The reliability of the Greulich and Pyle atlas when applied to a modern Scottish population. J Forensic Sci 58:114–119. https://doi.org/10.1111/j.1556-4029.2012.02294.x
60. Koc A, Karaoglanoglu M, Erdogan M, Kosecik M, Cesur Y (2001) Assessment of bone ages: is the Greulich-Pyle method sufficient for Turkish boys? Pediatr Int 43:662–665. https://doi.org/10.1046/j.1442-200x.2001.01470.x
61. Mora S, Boechat MI, Pietka E, Huang HK, Gilsanz V (2001) Skeletal age determinations in children of European and African descent: applicability of the Greulich and Pyle standards. Pediatr Res 50:624–628. https://doi.org/10.1203/00006450-200111000-00015
62. Paxton ML, Lamont AC, Stillwell AP (2013) The reliability of the Greulich-Pyle method in bone age determination among Australian children. J Med Imaging Radiat Oncol 57:21–24. https://doi.org/10.1111/j.1754-9485.2012.02462.x
63. Soudack M, Ben-Shlush A, Jacobson J, Raviv-Zilka L, Eshed I, Hamiel O (2012) Bone age in the 21st century: is Greulich and Pyle's atlas accurate for Israeli children? Pediatr Radiol 42:343–348. https://doi.org/10.1007/s00247-011-2302-1
64. Tise M, Mazzarini L, Fabrizzi G, Ferrante L, Giorgetti R, Tagliabracci A (2011) Applicability of Greulich and Pyle method for age assessment in forensic practice on an Italian sample. Int J Legal Med 125:411–416. https://doi.org/10.1007/s00414-010-0541-6
65. van Rijn RR, Lequin MH, Robben SG, Hop WC, van Kuijk C (2001) Is the Greulich and Pyle atlas still valid for Dutch Caucasian children today? Pediatr Radiol 31:748–752. https://doi.org/10.1007/s002470100531
66. Zabet D, Rerolle C, Pucheux J, Telmon N, Saint-Martin P (2015) Can the Greulich and Pyle method be used on French contemporary individuals? Int J Legal Med 129:171–177. https://doi.org/10.1007/s00414-014-1028-7
67. Ekizoglu O, Er A, Bozdag M et al (2021) Forensic age estimation via magnetic resonance imaging of knee in the Turkish population: use of T1-TSE sequence. Int J Legal Med 135:631–637. https://doi.org/10.1007/s00414-020-02402-0
68. Kramer JA, Schmidt S, Jurgens KU, Lentschig M, Schmeling A, Vieth V (2014) The use of magnetic resonance imaging to examine ossification of the proximal tibial epiphysis for forensic age estimation in living individuals. Forensic Sci Med Pathol 10:306–313. https://doi.org/10.1007/s12024-014-9559-2

69. Ottow C, Schulz R, Pfeiffer H, Heindel W, Schmeling A, Vieth V (2017) 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 27:5041–5048. https://doi.org/10.1007/s00330-017-4880-2
70. Saint-Martin P, Rerolle C, Pucheux J, Dedouit F, Telmon N (2015) Contribution of distal femur MRI to the determination of the 18-year limit in forensic age estimation. Int J Legal Med 129:619–620. https://doi.org/10.1007/s00414-014-1020-2
71. Ekizoglu O, Hocaoglu E, Can IO, Inci E, Aksoy S, Bilgili MG (2015) Magnetic resonance imaging of distal tibia and calcaneus for forensic age estimation in living individuals. Int J Legal Med 129:825–831. https://doi.org/10.1007/s00414-015-1187-1
72. Franklin D, Flavel A (2015) CT evaluation of timing for ossification of the medial clavicular epiphysis in a contemporary Western Australian population. Int J Legal Med 129:583–594. https://doi.org/10.1007/s00414-014-1116-8
73. Uysal Ramadan S, Gurses MS, Inanir NT, Hacifazlioglu C, Fedakar R, Hizli S (2017) Evaluation of the medial clavicular epiphysis according to the Schmeling and Kellinghaus method in living individuals: a retrospective CT study. Leg Med (Tokyo) 25:16–22. https://doi.org/10.1016/j.legalmed.2016.12.012
74. Zhang K, Chen XG, Zhao H, Dong XA, Deng ZH (2015) Forensic age estimation using thin-slice multidetector CT of the clavicular epiphyses among adolescent Western Chinese. J Forensic Sci 60:675–678. https://doi.org/10.1111/1556-4029.12739
75. Duangto P, Iamaroon A, Prasitwattanaseree S, Mahakkanukrauh P, Janhom A (2017) New models for age estimation and assessment of their accuracy using developing mandibular third molar teeth in a Thai population. Int J Legal Med 131:559–568. https://doi.org/10.1007/s00414-016-1467-4
76. Hassan FM, Moawad AM, Samir W, Helaly YR, Abu-Taleb NS (2021) Mandibular third molar maturation stage as indicator for the legal adult age in an Egyptian sample. Homo 72:87–97. https://doi.org/10.1127/homo/2021/1344
77. Hegde S, Patodia A, Dixit U (2016) Staging of third molar development in relation to chronological age of 5–16 year old Indian children. Forensic Sci Int 269:63–69. https://doi.org/10.1016/j.forsciint.2016.11.009
78. Johan NA, Khamis MF, Abdul Jamal NS, Ahmad B, Mahanani ES (2012) The variability of lower third molar development in Northeast Malaysian population with application to age estimation. J Forensic Odontostomatol 30:45–54
79. Kasper KA, Austin D, Kvanli AH, Rios TR, Senn DR (2009) Reliability of third molar development for age estimation in a Texas Hispanic population: a comparison study. J Forensic Sci 54:651–657. https://doi.org/10.1111/j.1556-4029.2009.01031.x
80. Lee SH, Lee JY, Park HK, Kim YK (2009) Development of third molars in Korean juveniles and adolescents. Forensic Sci Int 188:107–111. https://doi.org/10.1016/j.forsciint.2009.03.033
81. Li G, Ren J, Zhao S et al (2012) Dental age estimation from the developmental stage of the third molars in western Chinese population. Forensic Sci Int 219:158–164. https://doi.org/10.1016/j.forsciint.2011.12.015
82. Liu Y, Geng K, Chu Y, Xu M, Zha L (2018) Third molar mineralization in relation to chronologic age estimation of the Han in central southern China. Int J Legal Med 132:1427–1435. https://doi.org/10.1007/s00414-018-1804-x
83. Lopez TT, Arruda CP, Rocha M, Rosin AS, Michel-Crosato E, Biazevic MG (2013) Estimating ages by third molars: stages of development in Brazilian young adults. J Forensic Leg Med 20:412–418. https://doi.org/10.1016/j.jflm.2012.12.001
84. Quispe Lizarbe RJ, Solis Adrianzen C, Quezada-Marquez MM, Galic I, Cameriere R (2017) Demirjian's stages and Cameriere's third molar maturity index to estimate legal adult age in Peruvian population. Leg Med (Tokyo) 25:59–65. https://doi.org/10.1016/j.legalmed.2017.01.003
85. Maggio A, Flavel A, Hart R, Franklin D (2018) Assessment of the accuracy of the Greulich and Pyle hand-wrist atlas for age estimation in a contemporary Australian population. Aust J Forensic Sci 50:385–395. https://doi.org/10.1080/00450618.2016.1251970
86. Zafar AM, Nadeem N, Husen Y, Ahmad MN (2010) An appraisal of Greulich-Pyle Atlas for skeletal age assessment in Pakistan. J Pak Med Assoc 60:552–555
87. Socialstyrelsen (2018) Om magnetkamera vid bedömning av ålder: en studie av validiteten i radiologisk undersökning. Artikelnummer 2018-5-21. https://www.socialstyrelsen.se
88. Jayaraman J, Mendez MJC, Gakunga PT, Roberts G (2022) Age estimation of Hispanic children in the United States: development and validation of dental reference dataset based on two staging systems. Leg Med (Tokyo) 56:102033. https://doi.org/10.1016/j.legalmed.2022.102033