Forensic Science International 319 (2021) 110654

Contents lists available at ScienceDirect

Forensic Science International

journal homepage: www.elsevier.com/locate/forsciint

Bone age estimation with the Greulich-Pyle atlas using 3T MR images of hand and wrist

Thomas Widek[a,b,*], Pia Genet[c,d], Thomas Ehammer[a], Thorsten Schwark[a,e], Martin Urschler[f], Eva Scheurer[a,g]

[a] Ludwig Boltzmann Institute for Clinical Forensic Imaging, Universitätsplatz 4/II, 8010 Graz, Austria
[b] Gottfried Schatz Research Center, Division of Biophysics, Medical University of Graz, Neue Stiftingtalstraße 6/IV, 8010 Graz, Austria
[c] University Centre of Legal Medicine Lausanne, Lausanne University Hospital, Chemin de la Vulliette 4, 1000 Lausanne 25, Switzerland
[d] University Centre of Legal Medicine Geneva, Geneva University Hospital, Rue Michel-Servet 1, 1211 Geneva 4, Switzerland
[e] Department of Forensic Medicine, Laboratoire national de santé, Rue Louis Rech 1, 3555 Dudelange, Luxembourg
[f] School of Computer Science, The University of Auckland, 38 Princes Street, 1010 Auckland, New Zealand
[g] Institute of Forensic Medicine, University Basel - Health Department Basel-Stadt, Pestalozzistrasse 22, 4056 Basel, Switzerland

ARTICLE INFO

Article history:
Received 13 March 2020
Received in revised form 20 November 2020
Accepted 8 December 2020
Available online 9 December 2020

Keywords:
Age estimation by skeleton
Hand
Magnetic resonance imaging

ABSTRACT

The age estimation of the hand bones by means of X-ray examination is a pillar of the forensic age estimation. Since the associated radiation exposure is controversial, the search for ionizing radiation-free alternatives such as MRI is part of forensic research. The aim of the current study was to use the Greulich-Pyle (GP) atlas on MR images of the hand and wrist to provide reference values for assessing the age of the hand bones. 3T hand MR images of 238 male participants between the ages of 13 and 21 were acquired using 3D gradient echo sequences (VIBE, DESS). Two readers rated the images using the X-ray-based GP atlas method. A descriptive analysis and a transitional analysis were used for the statistical processing of the data. The agreement between and within the raters was assessed. In addition, a comparison was made with the chronological age and with X-ray studies. The descriptive analysis and the transition analysis showed similar results. Both evaluations showed good agreement with X-ray studies. The comparison with the chronological age showed a difference of 0.37 and 0.54 years for the two readers. The age estimate based on the cross-validated transition analysis showed a mean error of −0.28 years. Inter- and intra-rater agreement were good. In summary, it can be concluded that age estimation of hand bones with MR images is routinely applicable with the GP atlas as an alternative without ionizing radiation. However, in order to reduce the estimation error, a multi-factorial assessment based on examinations of several body regions is still recommended.

1. Introduction

Bone age assessment of the hand by means of X-ray examination represents an important pillar both for clinical purposes [1] and for forensic age estimation [2]. The recommendations of the Study Group on Forensic Age Diagnostics (AGFAD) for forensic age estimation suggest a holistic approach including a physical examination, an X-ray of the left hand, a dental examination including an X-ray examination of the teeth and, depending on the hand bone development, an X-ray or CT examination of the medial end of the clavicles [2]. This article focuses on the hand bone age assessment. There are several methods to perform hand bone age estimation. Among others [3–7] the atlas of Greulich and Pyle (GP) [8] is one of the most used methods in the clinical as well as in the forensic context [9,10]. However, the associated exposure to ionizing radiation is still subject to a controversial discussion. In forensics, the use of ionizing radiation is questionable as legal proceedings lack a medical indication. Therefore, the search for radiation-free alternatives such as MRI is a focus of many research groups. Recently, several MRI studies [11–23] have been published using different approaches for age estimation of the hand. However, only two pilot studies [11,20] have previously investigated if the X-ray based atlas method of Greulich and Pyle [8] is applicable to MR hand images. Both studies stated that the use of the GP atlas method is feasible, however, both lack a sufficient sample size. The aim of the current study was the use of the X-ray based GP staging method with 3T MR images in a male cohort to provide reference values for an ionizing radiation-free hand age assessment.

* Corresponding author at: Gottfried Schatz Research Center, Medical University of Graz, Neue Stiftingtalstraße 6/IV, 8010 Graz, Austria, Tel.: +43 (0)316/385-71462.
E-mail address: thomas.widek@medunigraz.at (T. Widek).

http://dx.doi.org/10.1016/j.forsciint.2020.110654

Descriptive statistics with mean and standard deviation values are, however, influenced by the size, age range and distribution of the sample examined. This could cause an effect known as age mimicry [24]. A possibility to get rid of these influences is transition analysis [25,26]. Therefore, reference values obtained through transition analysis are provided as well.

2. Materials and methods

2.1. Subjects

238 male volunteers from 13 to 20 years participated in the study (age range: 13.01 – 20.99 years, median: 17.39 years). All subjects were young Caucasian, healthy men with documented birth date. The study focused exclusively on males as the practice of age estimation shows that more than 90% of the examined persons are male. Their distribution of chronological age can be seen in Table 1. The upper age range limit was chosen to calculate accurate means and prediction intervals for the higher male GP standards [27]. Inclusion criteria were age from 13 to 20 years, informed consent, no medical history of developmental disorders (anamnestically raised) and no severe underweight. Exclusion criteria included MRI contraindications and non-compliance during the examination.

2.2. Imaging

All subjects underwent MRI of the left hand and wrist. The MRIs were performed with 3T MR scanners (Magnetom Trio, a TIM system & Magnetom Skyra, Siemens Healthcare, Erlangen, Germany). The volunteers were placed in prone position with the left arm outstretched, the hand was fixed and the standard 20-channel head and neck coil was used. The following sequences were applied in coronal orientation: T1 weighted 3D VIBE WE and T2 weighted 3D DESS WE (2). The Field-of-View of the VIBE sequence included the whole hand and wrist, whereas the DESS sequence was focused solely on the wrist to particularly support the evaluation of the higher GP standards. The total acquisition time of the two high-resolution sequences was 5:50 minutes. The selection of the VIBE was based on the studies of Urschler et al. [20] and Hojreh et al. [11]. The DESS sequence was selected as second 3D sequence which gives additional aspects to the wrist region due to the different weighting.

2.3. Image evaluation

The MR images of all subjects were evaluated independently by two blinded readers (T.E., T.W., one with more than 10 years MR reading experience and one with 20 years MR imaging experience) according to the defined standards of the GP atlas [8]. Both used open-source DICOM viewer software (OsiriX 4.1, https://www.osirix-viewer.com and Horos 3.3.5, https://horosproject.org). For the intra-rater agreement, the experienced reader (T.E.) re-evaluated 10% of the cases (randomly selected) six months after finishing the first evaluation.

Table 1
Chronological age distribution of subjects in groups with a bin size of one year, e.g., bin 14 contains subjects ≥14.00 and ≤14.99 years.

	Age (years)								Total (n)
	13	14	15	16	17	18	19	20	
n	27	19	27	35	38	26	30	36	238

2.4. Statistical analysis

All analyses were performed using the statistical software R v3.6.0 (https://www.r-project.org) with the R-packages “BlandAltmanLeh” [28], “irr” [29], “pscl” [30], “MASS” [31] and “VGAM” [32]) and with modified R-scripts provided by Lyle Konigsberg [33]. The age estimates of both readers were compared to chronological age using Bland-Altman plots [34]. Inter-observer (all cases) and intra-observer agreements were calculated using weighted Cohen's Kappa [35,36] and a contingency table is provided showing the agreement between the two observers. There was no consensus reading, therefore, the results from the experienced reader (T.E.) were used for descriptive statistics of the single data groups (according to the assigned GP standard including mean, standard deviation (SD), minimum (Min), maximum (Max), lower and upper quartile (LQ, UQ) and the median).

Additionally, transition analysis [25] using a cumulative probit model [37,38] with age on a log-scale was performed. Transition analysis, in contrast to mean value statistics, does not reflect the distribution of the reference sample in the results. [39] For the probit model, the GP standards (GPS) have been partly collapsed due to low counts (≤GPS24, GPS25, GPS26+27, GPS28, GPS29, GPS30 and GPS31). The “Goodness-of-fit” was tested using a Lagrange multiplier test and by calculating Cragg & Uhlers Pseudo-R^2 following a study of Konigsberg et al. [38]. Using “weight functions” (according to Love and Müller [40]) and normalization, normed maximum likelihood curves and 95% prediction intervals were calculated for estimation of the most likely age and age range per GP standard. [37]. A uniform prior distribution was assumed, thus the estimated age equals the age value which meets the normed maximum likelihood [26] (maximum likely age (MLAge)). The limits of the 95% prediction interval for the maximum likelihood estimation (MLEst) were calculated as the minimum and maximum age values, where the normed likelihood is greater than 0.1465001 [37].

To avoid overfitting and to validate the age prediction, a ten-fold cross-validation was used. Mean error (chronological age minus maximum likely age (MLAge)) and RMSE were calculated to show the accuracy of the age prediction. These values were obtained by averaging the results from the cross-validation.

A comparison of the averaged values gained by maximum likelihood estimation (MLEst) and the values of the descriptive statistics was performed. The prediction interval (95%) following the descriptive statistics was computed considering the number of subjects per group.

Normal distribution of the data was tested with a Shapiro-Wilk test, thus the following formula was used: $\bar{x} \pm t_{0.975,n-1} * s * \sqrt{1 + \frac{1}{n}}$. In the formula $\bar{x}$ equals the mean, s is the standard deviation, n is the group size and t was calculated as the 97.5 quantile of the Student's t-distribution with n-1 degrees of freedom [27]. For the comparison of MLAge and the mean values, a one-sample t-test using the MLAge as hypothesized mean was done. Statistical values of X-ray studies [27,41,42] were also compared to the current study and possible significant differences were investigated for [27] with one and two sample t-tests [43]. The limit of significance was set at $p < 0.05$.

2.5. Ethical considerations

The local ethics committee granted ethical clearance for the study. All participants gave written informed consent prior to study participation, with consent given by legal guardians for minors.

Table 2
MRI sequence parameters.

Sequence	FOV (mm)	Readout Matrix	Slice thickness (mm)	TR (ms)	TE (ms)	Flip angle (°)	Acquisition time (min:sec)
T1w 3D VIBE WE	230	512i	0.9	14	4.08	15	3:46
T2w 3D DESS WE	100	512i	0.8	14.28	5.18	30	2:04

FOV field-of-view, i interpolated, TR repetition time, TE echo time

3. Results

The study population consisted of 238 male subjects. The experienced reader judged the overall image quality as good. In some cases, less fat suppression was seen in the VIBE sequence. This occurred mainly in the fingers and made the GP standard allocation sometimes more difficult. Nevertheless, all hand MRIs were evaluable. Fig. 1 shows an example of MR images of the hand and wrist of a 15 years old male.

Descriptive statistics were calculated for the GP standards found (Table 3).

Fig. 2 displays the distribution of the GP-Standards per one-year age group. Here, the variability of GP standards is particularly evident in individuals under 17 years of age. No subject under 15 years has reached GP standard 30. At the age of 15 and 17 years, at least GP standard 24 and 28 is achieved, respectively.

For the comparison of the results of the two readers, a contingency table is given showing good agreement (Table 4). The

Fig. 1. MR images of the left hand and wrist of a 15 years old boy (GP standard 26). The T1w 3D VIBE WE sequence images showing the whole hand (a), the base of the first metacarpal bone (b) and the wrist (c). The T2w 3D DESS WE sequence image (d) shows the wrist on the same position as image (c).

Table 3
Descriptive statistics for the single GP standards in years.

GP standard	GP age	N	Mean	SD	Min	LQ	Median	UQ	Max
21	11.5	1	14.46		14.46	14.46	14.46	14.46	14.46
22	12.5	1	13.15		13.15	13.15	13.15	13.15	13.15
23	13	11	13.87	0.61	13.02	13.58	13.76	14.09	14.91
24	13.5	20	14.1	0.78	13.01	13.63	14.08	14.53	16.06
25	14	13	14.66	1.08	13.11	13.87	14.59	15.36	16.93
26	15	9	15.47	1.09	13.22	15	15.7	16.06	16.98
27	15.5	11	15.47	0.7	14.11	15.15	15.54	15.97	16.26
28	16	15	16.34	1.52	13.37	15.61	16.38	16.76	20.07
29	17	47	17.06	1.04	14.02	16.61	17.1	17.68	19.03
30	18	28	18.45	1.44	15.79	17.36	18.13	19.93	20.99
31	19	82	19.38	1.21	16.00	18.64	19.7	20.35	20.98

SD standard deviation, Min minimum, Max maximum, LQ lower quartile, UQ upper quartile

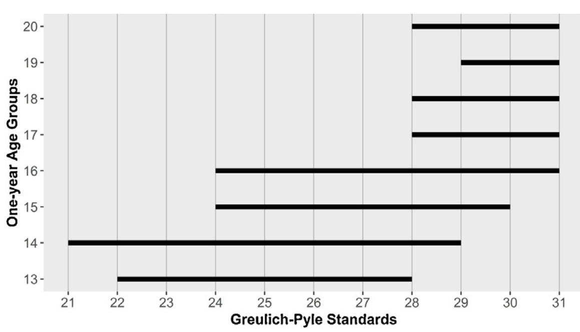

Fig. 2. Distribution of the GP Standards per one-year age group.

mean difference between reader 1 and 2 was almost zero ($d_{mean} = 0.16y$). The inter- and intra-rater agreement were very good with $\kappa_w = 0.85$ and $\kappa_w = 0.88$, respectively, according to the interpretation scheme by Altman [44].

The comparison to the chronological age was done separately for the two readers with Bland-Altman-like plots (Fig. 3). Both plots show no major deviation, the diagonal line pattern is explained by the assignment to the discrete GP standards. The mean difference between estimated and chronological age for the two readers was 0.37 and 0.54 years, respectively, which showed a slight underestimation. The limits of agreement were -1.90 and -1.58 years for the lower limit and for 2.65 and 2.66 years the upper limit, respectively.

Transition curves were calculated for the GP standards 24 - 31 with the help of a cumulative probit approach (Fig. 4). Normed maximum likelihood curves for the single GP standards were obtained within the age range 13.00 – 21.00 years (Fig. 5). The age values of maximum likelihood estimation and the corresponding prediction intervals as well as a comparison to the mean and 95% prediction interval of the descriptive statistics can be found in Table 5. For reasons of comparability, the GP standard groups have been equated and descriptive statistic values have been averaged.

The comparison between transition analysis and descriptive statistics showed mainly younger ages in the MLAge and similar prediction intervals. However, significant differences between MLAge values and mean values were found only for the first and the last GP standard group.

The ten-fold cross-validation showed a mean error (chronological age - MLA) of -0.28 years with a standard deviation of ± 0.53 years, which means a slight overestimation. The RMSE was 1.52 years showing the variability of the estimation. The average of the pseudo R^2 ($R^2_{Cragg\&Uhlers} = 0.75$) and the Lagrange Multiplier Test ($p = 0.26$) give evidence of a good fit of the cumulative probit model.

4. Discussion

The GP atlas is one of the most commonly used methods to assess bone age on hand radiographs [9,10]. Recently, two pilot studies [11,20] have shown the feasibility of applying the GP atlas to MR hand images. Based on these two studies the current study provides male reference values for assessing the bone age with hand MRI. Due to the good image quality, no subject had to be excluded. The acquisition time for the two MR sequences is rather

Table 4
Cumulative cross tabulation of frequencies of allocated GP ages by reader 1 compared with reader 2.

		Reader 2												
		11	11.5	12.5	13	13.5	14	15	15.5	16	17	18	19	Total
Reader 1	11	0	0	0	0	0	0	0	0	0	0	0	0	0
	11.5	0	1	0	0	0	0	0	0	0	0	0	0	1
	12.5	0	1	0	0	0	0	0	0	0	0	0	0	1
	13	1	0	2	5	3	0	0	0	0	0	0	0	11
	13.5	0	0	2	3	3	12	0	0	0	0	0	0	20
	14	0	0	1	1	1	7	3	0	0	0	0	0	13
	15	0	0	0	0	0	2	7	0	0	0	0	0	9
	15.5	0	0	0	0	0	0	5	5	1	0	0	0	11
	16	0	0	0	0	0	0	0	4	7	4	0	0	15
	17	0	0	0	0	0	0	0	0	13	33	1	0	47
	18	0	0	0	0	0	0	0	0	0	16	6	6	28
	19	0	0	0	0	0	0	0	0	1	2	8	71	82
	Total	1	2	5	9	7	21	15	9	22	55	15	77	238

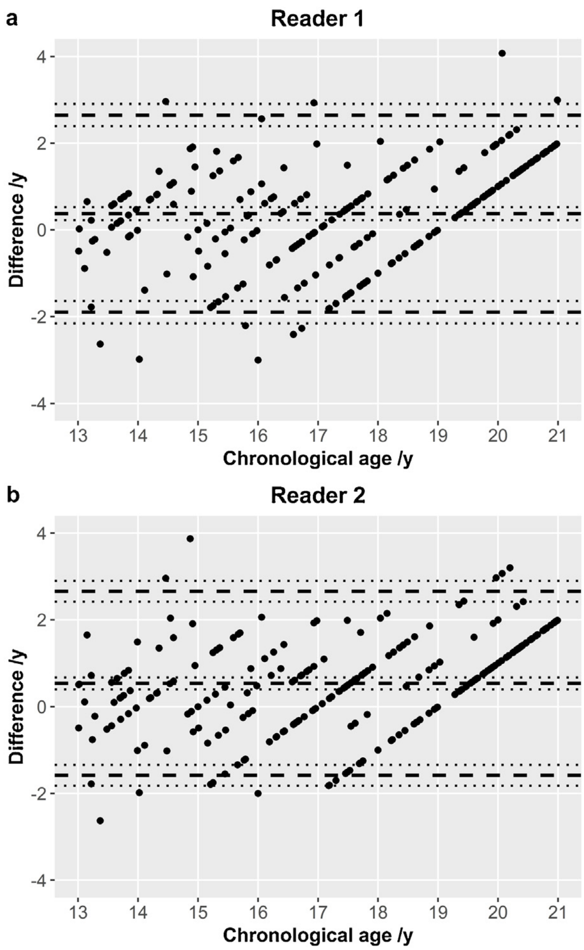

Fig. 3. Bland-Altman plots showing the comparison of the two readers with the chronological age. Difference means chronological age minus estimated age. The bold middle dashed line represents the mean difference; the two outer bold dashed lines are the limits of agreement (corresponding to two standard deviations). All lines are accompanied by their respective 95% confidence intervals (thin dotted lines).

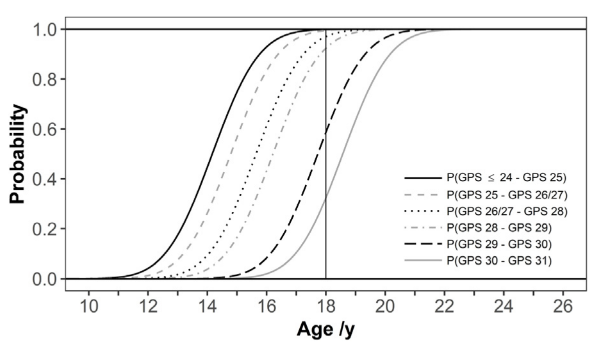

Fig. 4. Transition curves using the whole data set (GPS: Greulich-Pyle Standard). The black vertical line shows the 18-year threshold.

long compared to the time it takes to make a radiograph of the hand. However, recently a pilot study by Neumayer et al. [45] showed a possible reduction of the acquisition time of the VIBE sequence down to 15 seconds with promising results regarding age estimation. Shorter sequences would eliminate two major shortcomings of long MRI acquisition times: possible artifacts due to motion and the high costs of the exam. However, the results of Neumayer et al. have to be confirmed in a bigger cohort and as well on different MR scanners and field strengths.

The inter- and intra-rater agreement showed that GP bone age assessment from hand MRIs is reliable and reproducible. This is in concordance with the two before mentioned pilot studies. The good agreement of the two raters is also shown in the cross-tabulation (Table 3). The mean difference of the two readers was only 0.16 years. The mean relative disagreement (mean difference of the non-matches) was less than half a year, which also confirms the reliability of the method. However, it should be noted that the reading experience of the observers related to the specific sequences used may have contributed to the high inter-rater agreement.

The distribution of the GP standards shows a rather large variation in subjects under 17 years of age. There is no specific cut-off value for the 18-year threshold, although all subjects over 18 have at least a GP standard of 28. Due to the large deviation, age estimation by hand only is not recommended in the forensic context.

On average, a slight underestimation of about half a year was seen for both readers in comparison to the chronological age. This could be mainly explained by the last age category of 19 years and the upper age range limit of the sample (20.99 years). However, in the context of forensic age estimation, an underestimation is always in favor of a person and therefore far more acceptable in practice than an overestimation.

About twenty years ago, paleodemographers have introduced a method called transition analysis to circumvent the problem of age mimicry and small sample sizes in the age estimation of skeletal samples [25]. Recently, scientists have introduced this method also in the age estimation procedure of living adolescents [21,26,38,46,47]. In the current study, a cumulative probit model was used for fitting the transition curves. Tests showed that the model was appropriate and the cross-validation resulted in a mean error of about three months for the age estimation. The slight overestimation might be mainly explained by the estimated age of 21 for the final GP standard.

A comparison of the mean and predictive values of the transition analysis and the plain descriptive statistics showed no significant differences with the exception of the study's end stages. This is not surprising as the MLAge values of the first and last group are methodologically conditioned on minimal and maximal value of the age range of the transition analysis. Therefore, the transition analysis seems not to be superior over plain statistics, which could be expected due to the non-informative prior and the upper age range value. If the investigated age range would have included older individuals ($\geq$ 21y), the age mimicry effect would have been greater and the transition analysis would have shown its strength better. However, the prediction intervals of MLEst are mostly narrower and might correspond better to the real distribution, considering the sample size.

A comparison with three recent European GP X-ray studies [27,41,42] with at least similar or bigger cohorts showed the following differences. The X-ray studies showed a mean difference of the chronological age to the estimated age of 1.6, 2.16 and −1.45 months, respectively. In the current study, the mean difference (descriptive analysis) was about six month, which is - although higher - still in favor of an examined person. Urschler et al. [20] also saw this underestimation of the MRI assessments versus the X-ray assessments. Their explanation was the possibility of superimpositions in the area of the epiphyseal plate due to the projective nature of the X-ray images, which might lead to an earlier upstaging, compared to the non-superimposed MR images. In contrast, the mean error of our transition model showed an overestimation of about 3 months,

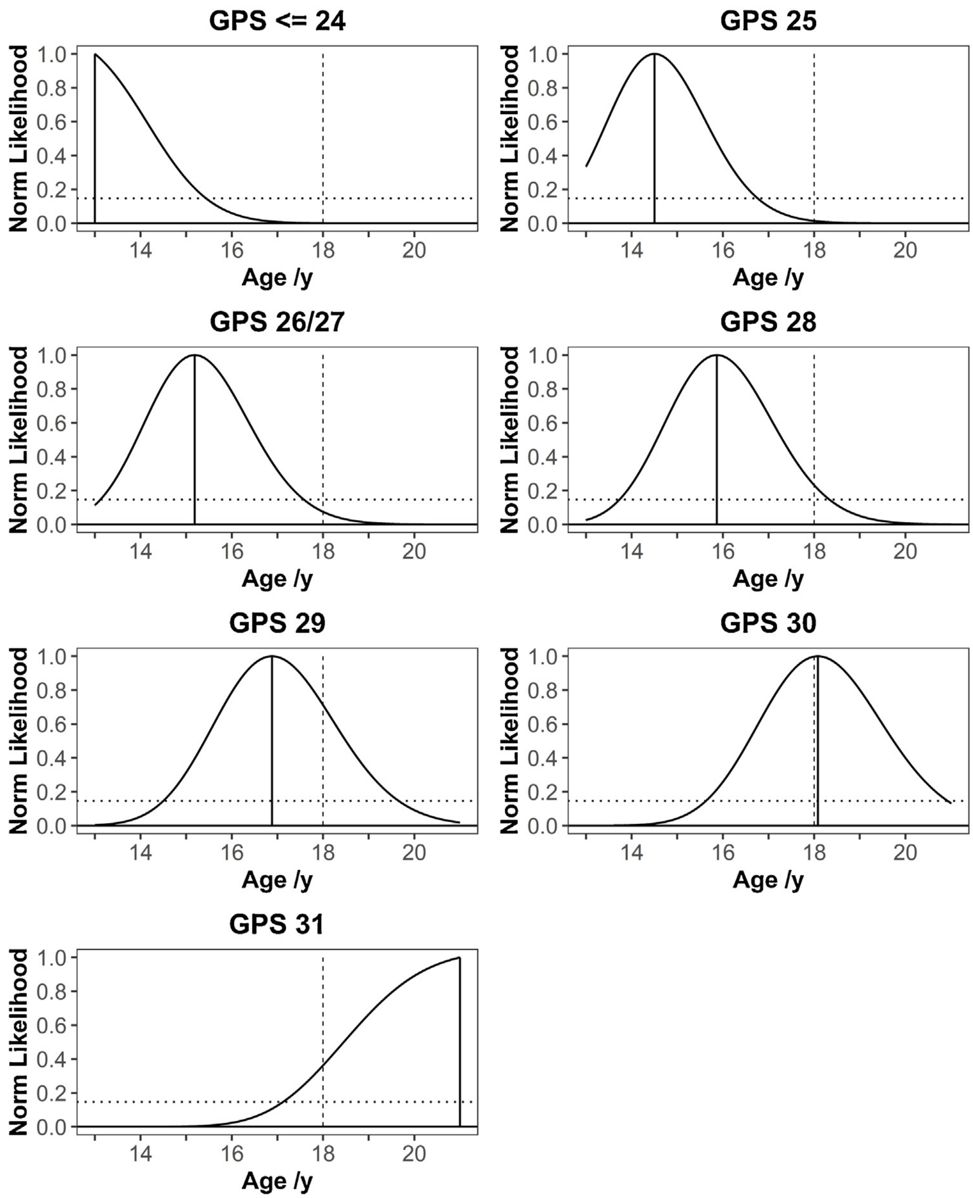

Fig. 5. Normed maximum likelihood curves (whole data set). The solid vertical line represents the maximum likely age (MLAge), the dashed vertical line displays the age of 18. The horizontal line shows the likelihood of 0.1465001 indicating the 95% prediction interval.

Table 5
Comparison of the MLAge and the mean, and the prediction intervals gained through maximum likelihood estimation and descriptive statistics (in years).

GP standard	N	MLAge	95% prediction interval MLEst	Mean age	95% prediction interval
≤ 24	33	13.00	13.00 – 15.44	13.89	12.38 – 15.40
25	13	14.50	13.00 – 16.76	14.66	12.22 – 17.10
26/27	20	15.19	13.13 – 17.57	15.47	13.56 – 17.38
28	15	15.87	13.74 – 18.33	16.34	12.97 – 19.71
29	47	16.88	14.50 – 19.64	17.06	14.94 – 19.18
30	28	18.08	15.63 – 20.89	18.45	15.44 – 21.46
31	82	21.00	17.12 – 21.00	19.38	16.96 – 21.80

MLAge maximum likely age, MLEst maximum likelihood estimation

which contradicts this explanation approach. However, the MLAge of the last stage, which might be mainly responsible for the overestimation, has to be used with caution, as it is dependent on the upper age range value. Therefore, the superimposition explanation might still be valid.

Comparing the descriptive statistic values and MLEst values with the very recent X-ray study of Chaumoitre et al. [27], which used a large French cohort, one can see that all descriptive mean values of the current study are higher (0.06–0.72 years). Significant differences could only be found for the GP standards

24, 29 and 30. However, the values of the current study fit well to the GP age categories. To be certain that there is no systematic deviation between MRI and X-ray a comparative study with a larger cohort would be beneficial. The MLAge values behaved differently with differences between −0.41 and 0.36 years except for GP 31 where the difference was 1.7 years. A significant difference was found for GP standards 24, 29 and 31, but regarding the study end stages, this must be interpreted with caution. For GP standard 29 the higher group size (Chaumoitre) and therefore lower standard deviation is one explanation for the significance, because the values are quite similar (16.88y vs. 16.52y). The range size of the prediction intervals of the two studies are on average similar, but there are differences in the interval position. For the lower limits of the current study (mean and std. deviation) the differences are between −0.38 and 0.94 years (median: 0.02y), but the upper limits are always higher than the one of the Chaumoitre study (0.03–1.5 years, median: 0.39 years). The same can be seen for the predicted lower limits of the current study's MLEst model (0.01–1.56 years, median: 0.25 years). For the upper limits, there are differences between −0.88 and 0.93 years (median: 0.21y).

The current findings propose the use of the X-ray based GP method on hand MR images. However, there are some limitations which need to be taken into account. First, the study focused on male subjects only. The main reason was the fact that more than 90% of the persons who undergo a forensic age assessment in practice are males, which leads to a higher urgency of optimising age estimation methods in males than in females. The socioeconomic aspect as a factor which could influence bone development and growth was not taken into account as the general economic level of the country of residence of the subjects was high. Also the investigated age range starting at 13 years was chosen in view of the practical forensic needs. However, even though no conclusions can be drawn on females, this study showed that the applied methodology, i.e., to use the GP method on MRI data, works in principle well and can as such be rolled out to investigate other cohorts, e.g., females, other ethnic groups and possibly also to younger subjects. The sample size was limited due to limited resources, and the sample distribution regarding age was not evenly balanced. The inclusion of subjects in the younger age groups turned out to be much more challenging than that of older individuals. However, by using a transition analysis model, this limitation was taken into account. The acquisition time of the MR sequences can be a limitation when assessing younger children and shorter sequences should be investigated. However, in the current study we did not see severe motion artifacts due to longer acquisition times. Additionally, it has to be mentioned that the sequence-specific experience of the observers can have an influence on the results. Since the applied sequences are not used as standard in radiological practice, not all radiologists will be used to them, which can affect the reading result.

In conclusion, the current study confirms that the X-ray based GP method can be used with hand images acquired with MR. Both, the plain descriptive statistics and transition model showed errors and prediction intervals very similar to X-ray studies. For clinical purposes, these values might be sufficient. However, in the forensic context narrower prediction intervals would be welcomed. The authors therefore recommend not to use the hand age estimation as the only source for an age assessment, especially for the forensically important 18-year threshold. A combination of different regions such as hand, clavicle and wisdom teeth, preferably using radiation-free modalities, should be sought [48]. In such a combined (multivariate) approach, transition analysis could further show its strengths. In addition, it should be noted that MRI has a number of different sequences that could lead to different results. Therefore, these study results should be used carefully and should only be applied when similar MRI parameters are used.

Credit author statement

Thomas Widek: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Data curation, Project administration, Resources, Visualization, Writing – original draft, Writing – review & editing.

Pia Genet: Investigation, Validation, Writing – review & editing.

Thomas Ehammer: Formal analysis.

Thorsten Schwark: Supervision, Writing – review & editing.

Martin Urschler: Supervision, Writing – review & editing.

Eva Scheurer: Conceptualization, Methodology, Validation, Project administration, Supervision, Writing – review & editing.

Funding

The authors state that this work has not received any funding.

Guarantor

The scientific guarantor of this publication is Eva Scheurer, University Basel

Conflict of interest

The authors of this manuscript declare no relationships with any companies, whose products or services may be related to the subject matter of the article.

Statistics and biometry

No complex statistical methods were necessary for this paper.

Informed consent

Written informed consent was obtained from all participants in this study. In case of minors, written informed consent was also obtained from the legal guardian.

Ethical approval

The study was approved by the Medical University of Graz Ethics Committee.

Methodology

- prospective
- cross sectional study / observational
- performed at one institution

References

[1] D.D. Martin, J.M. Wit, Z. Hochberg, L. Sävendahl, R.R. van Rijn, O. Fricke, N. Cameron, J. Caliebe, T. Hertel, D. Kiepe, K. Albertsson-Wikland, H.H. Thodberg, G. Binder, M.B. Ranke, The use of bone age in clinical practice – part 1, Horm. Res. Paediatr. 76 (1) (2011) 1–9, doi:http://dx.doi.org/10.1159/000329372.

[2] A. Schmeling, C. Grundmann, A. Fuhrmann, H.J. Kaatsch, B. Knell, F. Ramsthaler, W. Reisinger, T. Riepert, S. Ritz-Timme, F.W. Rosing, K. Rotzscher, G. Geserick, Criteria for age estimation in living individuals, Int. J. Legal Med. 122 (6) (2008) 457–460, doi:http://dx.doi.org/10.1007/s00414-008-0254-2.

[3] V. Gilsanz, O. Ratib, Hand Bone Age: A Digital Atlas Of Skeletal Maturity, Springer, Berlin, 2005.

[4] A. Roche, W. Chumlea, D. Thissen, Assessing the Skeletal Maturity of the Hand Wrist: Fels Method, Charles C. Thomas Publisher, Springfield, 1988.

[5] J.M. Tanner, R.H. Whitehouse, W.A. Marshall, M. Healy, H. Goldstein, Assessment of Skeletal Maturity and Prediction of Adult Height (TW2 Method), 2nd Edition, Academic Press, London, 1983.

[6] J. Tanner, M. Healy, N. Cameron, H. Goldstein, Assessment of Skeletal Maturity and Prediction of Adult Height (TW3 Method), W.B. Saunders, London, 2001.
[7] H.H. Thiemann, I. Nitz, A. Schmeling, Röntgenatlas der normalen Hand im Kindesalter, 3rd Edition, Thieme, Stuttgart, 2006.
[8] W. Greulich, S. Pyle, Radiographic Atlas of Skeletal Development of Hand and Wrist, 2nd Edition, Stanford University Press, Stanford, 1959.
[9] K. Alshamrani, A.C. Offiah, Applicability of two commonly used bone age assessment methods to twenty-first century uk children, Eur. Radiol. 30 (1) (2020) 504–513, doi:http://dx.doi.org/10.1007/s00330-019-06300-x.
[10] S. Schmidt, I. Nitz, S. Ribbecke, R. Schulz, H. Pfeiffer, A. Schmeling, Skeletal age determination of the hand: a comparison of methods, Int. J. Legal Med. 127 (3) (2013) 691–698, doi:http://dx.doi.org/10.1007/s00414-013-0845-4.
[11] A. Hojreh, J. Gamper, M.T. Schmook, M. Weber, D. Prayer, C.J. Herold, I.M. Noebauer-Huhmann, Hand mri and the greulich-pyle atlas in skeletal age estimation in adolescents, Skeletal Radiol. 47 (7) (2018) 963–971, doi:http://dx.doi.org/10.1007/s00256-017-2867-3.
[12] J. Dvorak, J. George, A. Junge, J. Hodler, Age determination by magnetic resonance imaging of the wrist in adolescent male football players, Br. J. Sports Med. 41 (1) (2007) 45–52, doi:http://dx.doi.org/10.1136/bjsm.2006.031021.
[13] J. Dvorak, J. George, A. Junge, J. Hodler, Application of mri of the wrist for age determination in international u-17 soccer competitions, Br. J. Sports Med. 41 (8) (2007) 497–500, doi:http://dx.doi.org/10.1136/bjsm.2006.033431.
[14] J. George, J. Nagendran, K. Azmi, Comparison study of growth plate fusion using mri versus plain radiographs as used in age determination for exclusion of overaged football players, Br. J. Sports Med. 46 (4) (2012) 273–278, doi: http://dx.doi.org/10.1136/bjsm.2010.074948.
[15] Y. Terada, S. Kono, D. Tamada, T. Uchiumi, K. Kose, R. Miyagi, E. Yamabe, H. Yoshioka, Skeletal age assessment in children using an open compact mri system, Magn. Reson. Med. 69 (6) (2013) 1697–1702, doi:http://dx.doi.org/10.1002/mrm.24439.
[16] E. Tomei, A. Sartori, D. Nissman, N. Al Ansari, S. Battisti, A. Rubini, A. Stagnitti, M. Martino, M. Marini, E. Barbato, R.C. Semelka, Value of mri of the hand and the wrist in evaluation of bone age: preliminary results, J. Magn. Reson. Imaging 39 (5) (2014) 1198–1205, doi:http://dx.doi.org/10.1002/jmri.24286.
[17] S. Serinelli, V. Panebianco, M. Martino, S. Battisti, K. Rodacki, E. Marinelli, F. Zaccagna, R.C. Semelka, E. Tomei, Accuracy of mri skeletal age estimation for subjects 12-19. potential use for subjects of unknown age, Int. J. Legal Med. 129 (3) (2015) 609–617, doi:http://dx.doi.org/10.1007/s00414-015-1161-y.
[18] S. Schmidt, V. Vieth, M. Timme, J. Dvorak, A. Schmeling, Examination of ossification of the distal radial epiphysis using magnetic resonance imaging. new insights for age estimation in young footballers in fifa tournaments, Sci. Justice 55 (2) (2015) 139–144, doi:http://dx.doi.org/10.1016/j.scijus.2014.12.003.
[19] J. Serin, C. Rerolle, J. Pucheux, F. Dedouit, N. Telmon, F. Savall, P. Saint-Martin, Contribution of magnetic resonance imaging of the wrist and hand to forensic age assessment, Int. J. Legal Med. 130 (4) (2016) 1121–1128, doi:http://dx.doi.org/10.1007/s00414-016-1362-z.
[20] M. Urschler, A. Krauskopf, T. Widek, E. Sorantin, T. Ehammer, M. Borkenstein, K. Yen, E. Scheurer, Applicability of greulich-pyle and tanner-whitehouse grading methods to mri when assessing hand bone age in forensic age estimation: A pilot study, Forensic Sci. Int. 266 (2016) 281–288, doi:http://dx.doi.org/10.1016/j.forsciint.2016.06.016.
[21] J. De Tobel, E. Hillewig, M.B. de Haas, B. Van Eeckhout, S. Fieuws, P.W. Thevissen, K.L. Verstraete, Forensic age estimation based on t1 se and vibe wrist mri: do a one-fits-all staging technique and age estimation model apply? Eur. Radiol. 29 (6) (2019) 2924–2935, doi:http://dx.doi.org/10.1007/s00330-018-5944-7.
[22] A. Er, M. Bozdag, C.D. Basa, I.E. Kacmaz, O. Ekizoglu, Estimating forensic age via magnetic resonance imaging of the distal radial epiphysis, Int. J. Legal Med. 134 (1) (2020) 375–380, doi:http://dx.doi.org/10.1007/s00414-019-02189-9.
[23] M. Timme, C. Ottow, R. Schulz, H. Pfeiffer, W. Heindel, V. Vieth, A. Schmeling, S. Schmidt, Magnetic resonance imaging of the distal radial epiphysis: a new criterion of maturity for determining whether the age of 18 has been completed? Int. J. Legal Med. 131 (2) (2017) 579–584, doi:http://dx.doi.org/10.1007/s00414-016-1502-5.
[24] J.-P. Bocquet-Appel, C. Masset, Farewell to paleodemography, J. Hum. Evol. 11 (4) (1982) 321–333, doi:http://dx.doi.org/10.1016/S0047-2484(82)80023-7.
[25] J.L. Boldsen, G.R. Milner, L.W. Konigsberg, J.M. Wood, Transition analysis: a new method for estimating age from skeletons, in: R.D. Hoppa, J.W. Vaupel (Eds.), Paleodemography: Age Distributions From Skeletal Samples, Cambridge University Press, New York, NY, 2002, pp. 73–106.
[26] S. Tangmose, P. Thevissen, N. Lynnerup, G. Willems, J. Boldsen, Age estimation in the living: transition analysis on developing third molars, Forensic Sci. Int. 257 (2015), doi:http://dx.doi.org/10.1016/j.forsciint.2015.07.049 512 e1-512 e7.
[27] K. Chaumoitre, B. Saliba-Serre, P. Adalian, M. Signoli, G. Leonetti, M. Panuel, Forensic use of the greulich and pyle atlas: prediction intervals and relevance, Eur. Radiol. 27 (3) (2017) 1032–1043, doi:http://dx.doi.org/10.1007/s00330-016-4466-4.
[28] B. Lehnert, BlandAltmanLeh: Plots (slightly extended) Bland-Altman Plots, R package version 0.3.1., 2015. https://CRAN.R-project.org/package=BlandAltmanLeh.
[29] M. Gamer, J. Lemon, I. Fellows, P. Singh, Irr: Various Coefficients Of Interrater Reliability And Agreement, R package version 0.84., 2012. https://CRAN.R-project.org/package=irr.
[30] S. Jackman, pscl: Classes and Methods for R. Developed in the Political Science Computational Laboratory, United States Studies Centre, University of Sydney, Sydney, New South Wales, Australia, 2017 R package version 1.5.2. https://github.com/atahk/pscl/.
[31] W.N. Venables, B.D. Ripley, Modern Applied Statistics with S, 4th Edition, Springer, New York, 2002.
[32] T.W. Yee, VGAM: Vector Generalized Linear and Additive Models, (2018) R package version 1.0-6. https://CRAN.R-project.org/package=VGAM.
[33] L.W. Konigsberg, Lyle W. Konigsberg's Webpage, (2019) [accessed 6 March 2020]. http://faculty.las.illinois.edu/lylek/.
[34] J.M. Bland, D.G. Altman, Statistical methods for assessing agreement between two methods of clinical measurement, Lancet 1 (8476) (1986) 307–310.
[35] J. Cohen, Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit, Psychol. Bull. 70 (4) (1968) 213–220.
[36] N.R. Shirley, P.A. Ramirez Montes, Age estimation in forensic anthropology: quantification of observer error in phase versus component-based methods, J. Forensic Sci. 60 (1) (2015) 107–111, doi:http://dx.doi.org/10.1111/1556-4029.12617.
[37] L.W. Konigsberg, N.P. Herrmann, D.J. Wescott, E.H. Kimmerle, Estimation and evidence in forensic anthropology: age-at-death, J. Forensic Sci. 53 (3) (2008) 541–557, doi:http://dx.doi.org/10.1111/j.1556-4029.2008.00710.x.
[38] L.W. Konigsberg, S.R. Frankenberg, H.M. Liversidge, Optimal trait scoring for age estimation, Am. J. Phys. Anthropol. 159 (4) (2016) 557–576, doi:http://dx.doi.org/10.1002/ajpa.22914.
[39] J. Buckberry, The (mis)use of adult age estimates in osteology, Ann. Hum. Biol. 42 (4) (2015) 323–331, doi:http://dx.doi.org/10.3109/03014460.2015.1046926 pMID: 26190373. arXiv:https://doi.org/10.3109/03014460.2015.1046926.
[40] B. Love, H.G. Müller, A solution to the problem of obtaining a mortality schedule for paleodemographic data, in: R.D. Hoppa, J.W. Vaupel (Eds.), Paleodemography: Age Distributions From Skeletal Samples, Cambridge University Press, New York, NY, 2002, pp. 181–192.
[41] L. Hackman, S. Black, The reliability of the greulich and pyle atlas when applied to a modern scottish population, J. Forensic Sci. 58 (1) (2013) 114–119, doi: http://dx.doi.org/10.1111/j.1556-4029.2012.02294.x.
[42] C. Santos, M. Ferreira, F.C. Alves, E. Cunha, Comparative study of greulich and pyle atlas and maturos 4.0 program for age estimation in a portuguese sample, Forensic Sci. Int. 212 (1-3) (2011), doi:http://dx.doi.org/10.1016/j.forsciint.2011.05.032 276 e1-7.
[43] E. Whitley, J. Ball, Statistics review 5: comparison of means, Crit. Care 6 (5) (2002) 424–428.
[44] D.G. Altman, Practical Statistics for Medical Research, Chapman & Hall/CRC, London, 1991.
[45] B. Neumayer, M. Schloegl, C. Payer, T. Widek, S. Tschauner, T. Ehammer, R. Stollberger, M. Urschler, Reducing acquisition time for mri-based forensic age estimation, Sci. Rep. 8 (1) (2018) 2063, doi:http://dx.doi.org/10.1038/s41598-018-20475-1.
[46] S. Fieuws, G. Willems, S. Larsen-Tangmose, N. Lynnerup, J. Boldsen, P. Thevissen, Obtaining appropriate interval estimates for age when multiple indicators are used: evaluation of an ad-hoc procedure, Int. J. Legal Med. 130 (2) (2016) 489–499, doi:http://dx.doi.org/10.1007/s00414-015-1200-8.
[47] M. Metsaniitty, J. Waltimo-Siren, H. Ranta, S. Fieuws, P. Thevissen, Dental age estimation in somali children and sub-adults combining permanent teeth and third molar development, Int. J. Legal Med. 133 (4) (2019) 1207–1215, doi: http://dx.doi.org/10.1007/s00414-019-02053-w.
[48] J. De Tobel, S. Fieuws, E. Hillewig, I. Phlypo, M. van Wijk, M.B. de Haas, C. Politis, K.L. Verstraete, P.W. Thevissen, Multi-factorial age estimation: a bayesian approach combining dental and skeletal magnetic resonance imaging, Forensic Sci. Int. 306 (2020) 110054, doi:http://dx.doi.org/10.1016/j.forsciint.2019.110054.