

Dissertation

Forensic Age Estimation in the Living using MRI

submitted by

Thomas WIDEK, MSc

for the Academic Degree of

Doctor of Medical Science

Dr.scient.med.

at the

Medical University of Graz

under the supervision of

PD Dr. Thorsten Schwark

2021

Dedicated to all my beloved ones who have supported me

"To the world you are anyone, but to someone you are the world."

Freely adapted after Erich Fried

Declaration

I hereby declare that this thesis is my own original work and that I have fully acknowledged by name all those individuals and organisations that have contributed to the research of this thesis. Acknowledgement has been made in the text to all other material used. Throughout this thesis and in all related publications I followed the guidelines of “Standards of Good Scientific Practice and Ombuds Committee at the Medical University of Graz“.

Graz, 2021-03-03

Thomas Widek

Disclosure

Partially, the introduction, materials & methods, results, discussion and limitations as well as individual tables and figures are derived with permission from the following two manuscripts published with the doctoral candidate as first author in Forensic Science International (Elsevier allows authors to include their articles in full or in part in a dissertation for non-commercial purposes, <https://www.elsevier.com/about/our-business/policies/copyright/permissions>, last accessed 2021-03-03).

Widek T, Genet P, Merkens H, Boldt J, Petrovic A, Vallis J, Scheurer E. Dental age estimation: The chronology of mineralization and eruption of male third molars with 3T MRI. Forensic Sci Int. 2019;297:228-35. DOI: 10.1016/j.forsciint.2019.02.019

Widek T, Genet P, Ehammer T, Schwark T, Urschler M, Scheurer E. Bone Age Estimation with the Greulich-Pyle Atlas using 3T MR Images of Hand and Wrist. Forensic Sci Int. 2021;319:110654 DOI: 10.1016/j.forsciint.2020.110654

All co-authors, who actively contributed to the research in this thesis agreed to the use of their data:

Julian Boldt, DD PhD

Department of Prosthodontics, University of Würzburg, Würzburg, Germany

Thomas Ehammer, MD, Graz, Austria

Pia Genet, MD

University Centre of Legal Medicine Lausanne, Lausanne University Hospital, Lausanne, Switzerland

University Centre of Legal Medicine Geneva, Geneva University Hospital, Geneva, Switzerland

Heiko Merkens, DD, Aachen, Germany

Andreas Petrovic, MSc

MSD Department of Radiology, University of Melbourne, Melbourne, Australia

Eva Scheurer, MD MSc

Institute of Forensic Medicine, University Basel, Basel, Switzerland

Thorsten Schwark, MD

Department of Forensic Medicine, Laboratoire national de santé, Dudelange, Luxembourg

Martin Urschler, PhD

School of Computer Science, University of Auckland, Auckland, New Zealand

Jacquie Vallis, MCSFS SFHEA

Borough of Darlington, UK Society and College of Radiographers, UK

Additional contribution and agreement to use his data was given by:

Jannick de Tobel, DD MD PhD

Department of Diagnostic Sciences - Radiology, Ghent University, Ghent, Belgium

Department of Imaging and Pathology - Forensic Odontology, KU Leuven, Leuven, Belgium.

Publications which are related to the thesis with the doctoral candidate as co-author:

Neumayer B, Schloegl M, Payer C, Widek T, Tschauner S, Ehammer T, Stollberger R, Urschler M. Reducing acquisition time for MRI-based forensic age estimation. *Sci Reports*. 2018;8(1):2063. DOI: 10.1038/s41598-018-20475-1

Neumayer B, Lesch A, Thaler F, Widek T, Tschauner S, De Tobel J, Ehammer T, Kirnbauer B, Boldt J, Van Wijk M, Stollberger R, Urschler M. The four-minute approach revisited: accelerating MRI-based multi-factorial age estimation. *Int J Legal Med*. 2020;134(4):1475-85. DOI: 10.1007/s00414-019-02231-w

Acknowledgments

This is it!

First of all, I would like to thank my family, my friends and everyone else who supported or just accompanied me. You all made this possible and made me who I am.

I am grateful to the members of my thesis committee Thorsten Schwark, Eva Scheurer and Franz Quehenberger for their help and contribution during this doctoral study. A special thanks goes to Eva Scheurer, who recruited me for the LBI-CFI, supported my path and my studies and awakened the spirit of research in me. Without you, I would not be where I am.

Another special thank you goes out to Pia Genet for her support, friendship and Berlin. I would like to give a big shout out to Martin Urschler for his help, support and accuracy; and must confess to being a little envious of the place you are currently working.

I would also like to thank all the other members of the former LBI-CFI. Unfortunately, this institute was closed in 2019. That is a great pity!

I would like to acknowledge the work of the people who evaluated data for this thesis: Julian Boldt, Jannick de Tobel, Thomas Ehammer and Heiko Merkens. Thank you!

I am grateful for all the MR physicists I have got to know over the past ten years. I have to put three of them in the foreground: Andreas Petrovic, Christian Langkammer and Peter Opriessnig. Thank you for your help and for many funny hours in Graz and around the world.

Not only for the sake of completeness, I would like to mention three persons who did not actively contributed to the thesis, but I am glad to have them with me. Kathrin Ogris and Isabella Klasinc, thank you for the hours of talking, discussing, lamenting and laughing. Simone Leski, for her friendship and for staying in the office with me through this crazy year of 2020.

Finally, I would like to thank the Medical University of Graz for providing a grant that allowed me to attend the ISMRM Congress in Montreal, Canada. I would also like to mention the LBG and the LBI-CFI for supporting me to present my research around the world and funding the last year through the "Age estimation 2020" project.

Last but not least I would like to thank all the musicians for making music which were very helpful during the days of evaluating, writing and correcting.

One very last point. Nothing would be as it is if I did not have three wonderful people around me: my wife and my two amazing sons. Eva, Raphael, Elias - Thank you!

Contents

Declaration	iii
Disclosure	iv
Acknowledgments	vii
Table of contents	viii
List of Abbrevations	xii
List of Figures	xv
List of Tables	xv
Zusammenfassung	xviii
Abstract	xx
1. Introduction	1
1.1. Reasons for age assessment of living individuals	2
1.1.1. Importance of age assessment	3
1.1.2. Significant age limits	3
1.2. Current state-of-the-art of the age assessment in living adolescents	4

1.3. The basics of medical age estimation in living adolescents	6
1.3.1. External physical examination	6
1.3.2. Plain X-ray of the left hand and wrist	7
1.3.3. Dental examination	8
1.3.4. Plain X-ray or computed tomography (CT) of the clavicles	10
1.4. Ethical considerations	11
1.5. Age estimation using Magnetic Resonance Imaging (MRI)	13
1.6. Multi-factorial age estimation (MFA)	13
1.7. Aim of the thesis	15
2. Materials and methods	16
2.1. Ethical issues	16
2.2. Study subjects	16
2.3. MR Imaging	18
2.4. Single region studies: Evaluation and analysis methods	20
2.4.1. Third Molars	20
2.4.2. Hand and Wrist	24
2.4.3. Clavicular	27
2.5. Multi-factorial age estimation (MFA)	31
2.5.1. MR Data	31
2.5.2. Statistical analysis	32
3. Results	34
3.1. Third Molars: Results	34
3.1.1. Mineralization results	36
3.1.2. Eruption results	41
3.2. Hand and Wrist: Results	45
3.3. Clavicular: Results	53
3.4. Multi-factorial age estimation (MFA): Results	59

Contents	xi
4. Discussion	66
4.1. Third Molars	66
4.1.1. Mineralization	67
4.1.2. Eruption	69
4.2. Hand and Wrist	70
4.3. Clavicular	73
4.4. Multi-factorial age estimation (MFA)	76
5. Limitations	79
6. Conclusion and Outlook	81
Bibliography	84
Appendix	XII
A. R-scripts	XIII
A.1. Normed likelihood curves	XIII
A.2. Markov Chain Monte Carlo (MCMC)	XV
A.3. Posterior density calculation	XVIII
B. Ethics vote	XIX

List of Abbreviations

2D	Two-dimensional
3D	Three-dimensional
AGFAD	Study group for forensic age diagnostics
BMI	Body Mass Index
CISS	Constructive Interference in Steady State sequence
CPC	Cloth Pin Coil
CT	Computed tomography
DESS	Dual Echo Steady State sequence
DIS-SLOPE	Discrimination slope
EASO	European Asylum Support Office
EU	European Union
FA	Flip angle
FAE	Forensic age estimation
FDI	World Dental Federation
FOV	Field-of-View
FS	Spectral fat saturation
GP	Greulich-Pyle
GPS	Greulich-Pyle Standard
LM	Lagrange multiplier test
LQ	Lower quartile

MAE	Mean absolute error
MAX	Maximum
MCMC	Markov Chain Monte Carlo
MFA	Multi-factorial age estimation
MIN	Minimum
MLAGE	Maximum likelihood of age
MLEST	Maximum likelihood estimation
MR	Magnetic Resonance
MRI	Magnetic Resonance Imaging
OPG	Orthopantomogram
PDw	Proton density weighted
PI	Prediction interval
RMSE	Root mean squared error
SAR	Specific absorption rate
SD	Standard deviation
SEM	Standard error of the mean
SLT	Slice thickness
T	Tesla
TA	Transition analysis
TE	Echo time
TR	Repetition time
TSE	Turbo Spin Echo sequence
UASC	Unaccompanied or separated children
UNHCR	United Nations High Commissioner for Refugees
UQ	Upper quartile
US	Ultrasound
VIBE	Volume Interpolated Breathhold Examination sequence
w	Weighted
WE	Water excitation

WHO World Health Organisation

List of Figures

1.1.	Examples of radiological images for conventional age assessment	11
2.1.	Age distribution of all study subjects in one-year age groups	18
2.2.	Stages and corresponding MR images for mineralization and eruption	22
2.3.	Comparison of an X-ray image of the hand with MR images	25
2.4.	Clavicular stages	29
3.1.	MR images of the third molars	35
3.2.	Box and whisker plot of the mineralization	37
3.3.	Box and whisker plot of the eruption	42
3.4.	Hand MRI example	45
3.5.	Range of the GP Standards per one-year age group	47
3.6.	Bland-Altman plots: Comparison of the two readers with the chronological age .	49
3.7.	Transition curves of the hand evaluation	50
3.8.	Normed maximum likelihood curves of the hand evaluation	51
3.9.	Example for a MR clavicle exam	53
3.10.	Box and whisker plot of the clavicle ossification	54
3.11.	Normed maximum likelihood curves of the max. stages clavicle evaluation . . .	58
3.12.	Box plot of the predicted probability to be a minor	61
3.13.	MFA example 1	62
3.14.	MFA example 2	63
3.15.	MFA example 3	64

List of Tables

1.1. Juridical age limits in Austria, UK and USA	4
1.2. Overview of methods used for assessing the age in European countries and in the USA	5
2.1. MRI sequence parameter	20
2.2. Written explanation for the mineralization stages	21
2.3. Written explanation for the eruption stages	21
2.4. Written explanation for the clavicle stages	28
2.5. Cross-validation and test-set constellations.	31
3.1. Number of in- and excluded teeth per third molar and evaluation	36
3.2. Descriptive statistical data of the mineralization of tooth 18, 28, 38, 48	38
3.3. Age at different probability values to reach stage H	39
3.4. Differing stages in subjects with all four third molars.	40
3.5. Probability calculation for mineralization stage E, F, G and H	40
3.6. Descriptive statistical data of the eruption of tooth 18, 28, 38, 48	41
3.7. Age at different probability values to reach stage D	43
3.8. Pairwise comparison using Conover post-hoc test	43
3.9. Differing stages in subjects with all four third molars.	44
3.10. Probability calculation for eruption stage C and D	44
3.11. Descriptive statistics for the single GP standards	46
3.12. Cumulative cross tabulation of frequencies of allocated GP ages	48

3.13. Comparison of values gained through maximum likelihood estimation and descriptive statistics	52
3.14. Descriptive statistical data for the individual ossification stages of the medial end of the clavicle	55
3.15. Inter-rater agreement between the individual raters (clavicles)	55
3.16. Cross-validated performance of different model approaches	56
3.17. Extended results from the evaluation of the model using the max. stages of the clavicles	57
3.18. Performance of the different models using different age indicators' combinations	59
3.19. Comparison of the absolute errors of the H-Cmax-UJmax-LJmax model with the other models	60
3.20. Evaluation of all study participants with the corresponding model	65

Zusammenfassung

Die forensische Altersschätzung bei Lebenden ist in jüngster Zeit durch die stetige Zunahme der grenzüberschreitenden Migration in den Vordergrund der forensischen Forschung gerückt. Derzeit basiert das in vielen Ländern etablierte Standardverfahren für die forensische Altersschätzung auf radiologischen Untersuchungen unter Verwendung ionisierender Strahlung. Die hauptsächlich untersuchten Körperregionen sind die Hand, die Schlüsselbeine und die dritten Molaren. Die Bewertung stützt sich auf die Entwicklungsstadien der Knochen und Zähne. Die Verwendung ionisierender Strahlung zum Zwecke der forensischen Altersschätzung ist jedoch aufgrund fehlender medizinischer Indikation Gegenstand kontroverser Diskussionen. Daher ist u.a. die Suche nach radiologischen Untersuchungen ohne ionisierende Strahlung ein wichtiger Schwerpunkt der forensischen Forschung. Ziel dieser Arbeit war es, die 3T Magnetresonanztomographie (MRT) als alternative Methode für das Altersschätzungsverfahren zu untersuchen. Für die Auswertung radiologischer Daten sollten etablierte Klassifizierungssysteme verwendet und ein Datenmodell für die Kombination der drei Regionen gefunden werden.

Die vorgestellte Arbeit kann in vier Teile unterteilt werden: Auswertungen der drei genannten Einzelregionen und eine Kombination der Regionen im Sinne einer multifaktoriellen (Alters) Analyse. Für das gesamte Projekt wurden 338 männliche Kaukasier im Alter von 13 – 25 Jahren (Altersbereich: 13,01 – 24,98 Jahre, Median: 18,90 Jahre, Mittelwert: 18,93 Jahre, Standardabweichung: 3,28 Jahre) prospektiv mit einem 3T MR-Scanner untersucht. Für die linke Hand und das Handgelenk, die beiden Schlüsselbeine und die dritten Molaren wurden unterschiedlichen MR-Sequenzen verwendet. Die Analyse der MR-Datensätze wurde von verschiedenen Auswert-

ern mit etablierten Klassifizierungssystemen durchgeführt. Der Greulich-Pyle Atlas wurde zur Hand- und Handgelenksbewertung verwendet, das mediale Ende der Schlüsselbeine wurde unter Verwendung der Schmeling- und Kellinghaus-Stadien analysiert, und schließlich wurden die dritten Molaren unter Verwendung der von Demirjian bzw. Olze eingeführten Klassifikationen bewertet. Die statistische Analyse der einzelnen Regionen erfolgte mit einer Übergangsanalyse und/oder deskriptiver Statistik. Für die multifaktorielle Kombination wurde eine Übergangsanalyse mit einem „Latent trait“-Ansatz unter Berücksichtigung des Bayes-Theorems angewendet. Die Ergebnisse zeigten die Anwendbarkeit der 3T MRT für die einzelnen Regionen unter Verwendung der derzeit verwendeten Klassifizierungssysteme. Referenzwerte für Männer für die drei Regionen wurden präsentiert. Das verwendete Modell für die Kombination der drei Regionen zeigte gute Ergebnisse und war der Einzelbetrachtung der jeweiligen Regionen überlegen. Einschränkungen des Projekts sind hauptsächlich die männliche Kohorte kaukasischer Probanden und die langen MR-Akquisitionszeiten. Daher sind weitere Studien erforderlich, um auch Frauen sowie andere ethnische Gruppen zu untersuchen. Darüber hinaus sind Studien zur Verkürzung der MRT-Sequenzen notwendig, um die Implementierung der MRT in den Altersbewertungsprozess zu vereinfachen.

Abstract

Forensic age estimation in the living has recently been brought to the forefront of forensic research due to the steady increase in cross-border migration. Currently, the standard procedure for forensic age estimation established in many countries is based on radiological examinations using ionizing radiation. The main body regions examined are the hand, the clavicles and the third molars. The assessment is based on the developmental stages of the bones and teeth. The use of ionizing radiation for the purpose of forensic age estimation is, however, the subject of controversial discussions due to the lack of a medical indication. Therefore, among others, the search for radiological examinations without ionizing radiation is an important focus of forensic research. The aim of this thesis was to investigate 3T magnetic resonance imaging (MRI) as an alternative method for the age estimation procedure. Established classification systems should be used for the evaluation of radiological data and a data model for the combination of the three regions should be found.

The presented work can be divided into four parts: Evaluations of the three mentioned single regions and a combination of the regions in terms of a multi-factorial (age) analysis. For the entire project, 338 male Caucasians aged 13 – 25 years (range: 13.01 – 24.98y, median: 18.90y, mean: 18.93y, standard deviation: 3.28y) were prospectively examined with a 3T MR scanner. Different MR sequences were used for the left hand and wrist, the two clavicles, and the third molars. The analysis of the MR data sets was carried out by different evaluators using established classification systems. The Greulich-Pyle atlas was used for hand and wrist evaluation, the medial end of the clavicles was analyzed using the Schmeling and Kellinghaus stages, and finally the

third molars were scored using the classifications introduced by Demirjian and Olze, respectively. Statistical analysis of the individual regions was done with a transition analysis and/or descriptive statistics. For the combination, transition analysis in a latent trait approach was applied, taking Bayes' theorem into account.

The results showed the applicability of 3T MRI for the individual regions using the classification systems currently in use. Reference values for males for the three regions were presented. The model used for the combination of the three regions showed good results and was superior to single region evaluation.

Limitations of the project are mainly the male only cohort of Caucasian subjects and the long MR acquisition times. Therefore, further studies are needed to examine females as well as other ethnic groups. In addition, studies are needed to shorten the MRI sequences to facilitate the implementation of MRI in the age assessment process.

1. Introduction

Parts of the introduction have been used in previously published manuscripts as a result of this thesis (1, 2).

In the last few decades, cross-border migration has steadily increased not only in Austria but also in many other countries. Due to this development age assessment of living individuals, claiming to be minors, has become an inherent part of forensic work (3). However, the age estimation procedure is needed not only in criminal and civil proceedings, but also in youth sports competition and for clinical purposes.

Regarding the court proceedings it is crucial for the court to know the chronological age of the individuals, because law application changes if they are classified as adults or as adolescents. Often, these individuals have no reliable documents regarding their age. However, for the well-being of minors in particular, knowing their age is necessary as treatment and accommodation are different from adults. Additionally, many countries offer education for minors and they cannot be deported. Since it can be advantageous to be a minor, there is a tendency to pretend to be a minor. Therefore, it is imperative to make a reasonable assessment of a person's age.

In 2000 an interdisciplinary study group for forensic age diagnostics (AGFAD) was formed to develop and publish guidelines for age estimation procedure to ensure quality standards. The recommended model consists of a physical examination, an X-ray of the left hand and an inspection of the dental status by a dentist including the evaluation of an orthopantomogram (OPG) . In the case that the development of the hand has already been completed, an examination of the medial end of the clavicles should be additionally done (3).

The following sections will provide information about the necessity of the age assessment procedure, the legal issues, the current state-of-the-art, the bone and dental development and the aim of the research.

1.1. Reasons for age assessment of living individuals

According to the UNHCR (4) about 3.5 million people worldwide asked for asylum by the end of 2018. About 138,000 refugees were minors, so-called unaccompanied or separated children (UASC), although the dark figure is considered to be much higher. In 2020, there were around 14,000 asylum seekers in Austria, about 40% claimed to be less than 18 years of age (5). Asylum-seeking minors or those claiming to be minors are one group of individuals who require an age estimation when their stated age is in doubt or they cannot prove their stated age. However, there are other legal reasons for age estimations in the living. These are among others the following (6, p. 33):

- Age of criminal responsibility
- Stateless individual
- Human trafficking
- Human smuggling
- Sex workers
- Competency and legal capacity
- Consent for marriage
- Employment recruitment
- Retirement benefits
- Consent for medical treatment or procedures
- Pornographic images
- Complainant of sexual offence
- Compulsory school attendance or impending exclusion from school

Recently another reason for age estimation has risen: fraud in youth sport events, where older “sportsmen” claim to be younger to get an advantage over the competitors in their chronological subgroup. There are also younger athletes which claim to be older to be accepted at a tournament. This happens mainly in gymnastic sports (7, 8).

Beside the forensic age estimation (FAE) and the sports related assessments, there are of course

bone age estimations in clinical practice. These assessments are mainly used for the diagnosis and therapy monitoring of growth disorders in pediatrics (9). But they are also used for planning orthopedic and orthodontic treatments (10, 11).

1.1.1. Importance of age assessment

Chronological age has an important role in dealing with people, especially in First World countries (12).

In legal proceedings for example age can only be proven by providing a birth certificate. However, in many countries, birth registration is not consistently performed. An estimated 237 million children at age 5 were not in possession of a birth certificate by the end of 2019. Most of them, over 85%, were born in South-Asia or in sub-Saharan Africa (13).

In addition, refugees, although they have documents, often leave them behind or lose them along the way due to the circumstances of their flight (12). As minors receive special protection, education, health care and have the right to get asylum independent of their origin, it is important that minority can be proven or be excluded based on an objectively performed age estimation procedure.

1.1.2. Significant age limits

In the UN Convention of the Rights of the Child, children are defined as all “human beings below the age of eighteen years” (14). There is no law definition for other terms such as ‘minor’, ‘youth’ and ‘adolescents’, although these terms are used for a specific age group. In the EU the age of majority is 18 years in all member states (15). Additional to the age of majority, there are different relevant age limits in countries all over the world. They depend on the jurisdiction of the country concerned and the nature of the legal proceedings (criminal or civil) and range from 6 to 21 years. Table 1.1 shows as an example the comparison of the age limits in Austria, in the UK and in the USA (16–21).

Table 1.1.: Juridical age limits in Austria, UK and USA (in years)

Limit	Austria	UK ^a	USA
Criminal responsibility	14	10	6-10 ^b
Competence to act	16	16	–
Age of Majority	18	18	18 ^c
Applicability of adult law only	21	–	20 ^d
State pension age	60/65	66 to 68	66 to 67

^aUK w/o Scotland. ^bnot specified in 32 States and D.C.

^cnot in Alabama, Nebraska and Mississippi. ^din 35 states and D.C.

1.2. Current state-of-the-art of the age assessment in living adolescents

There is no common concept for (forensic) age assessments in the European countries. This is surprising as the AGFAD was founded over 20 years ago. The interdisciplinary group (forensic pathologists, dentists, radiologist, anthropologists, etc.) with members from all over Europe was established to define and to publish guidelines for forensic age assessment, to harmonize the different approaches and to increase quality of the methods and procedure (3). Although guidelines have been in place for about 20 years, there are still different approaches - medical and non-medical - to age assessment procedures in Europe, and there are no legally binding guidelines, e.g., from the EU. Table 1.2 shows the age assessment concepts of thirty European countries according to a survey conducted by European Asylum Support Office (EASO) in 2017 (22) updated through personal communications (23–25) and including the USA (26, 27). The asylum proceedings in Austria in cases of a doubtful minority are governed by section 12 of the Aliens' Police Act 2005 and by section 2 of the Asylum Act 2005 (28, 29). Together the two acts permit medical methods for age determination. It should be done with a multi-factorial examination methodology that explicitly includes the use of ionizing radiation. However, the asylum-seeking individual cannot be forced to take part in the medical examination.

Table 1.2.: Overview of methods used for assessing the age in European countries and in the USA (adapted from (22), (23–25)). Note: not all methods are used for every case.

Country	Non-medical methods			Medical methods								
	Documents submitted	Estimations based on physical appearance	Age assessment interview	Social service assessment	Psychological interviews	Dental observation	Physical development	Sexual maturity observation	Hand/Wrist X-ray	Clavicle X-ray/CT	Dental X-ray	Other
Austria	x	x	x			x	x	x	x	x ^a	x	
Belgium	x		x	x		x			x	x	x	
Bulgaria	x	x				x	x		x			
Croatia	x	x		x	x	x	x	x	x		x	
Cyprus	x		x	x		x			x		x	
Czech Republic	x								x			
Denmark	x		x				x		x		x	
Estonia	x	x	x	x	x		x	x	x	x	x	
Finland	x					x			x		x	
France	x		x		x				x	x	x	
Germany	x	x	x	x	x	x	x	x	x ^b	x ^{a,b}	x ^b	
Greece	x	x	x	x	x	x	x				x	
Hungary	x	x	x			x		x	x			Pelvic bone X-ray
Ireland	x	x	x	x								
Italy	x	x	x	x	x ^c	x	x	x	x		x	
Latvia						x	x		x	x	x	
Lithuania	x	x							x	x		
Luxembourg	x						x	x	x	x ^a	x	
Malta	x		x						x			
Netherlands	x	x								x		
Norway	x	x	x			x			x		x	
Poland	x	x				x	x		x		x	
Portugal	x	x				x			x		x	Fourth rib X-ray
Romania		x				x		x	x	x	x	
Slovakia	x	x ^d		x ^e		x			x	x	x	
Slovenia	x	x										
Spain	x		x						x	x		
Sweden	x		x	x							x	MRI Knee
Switzerland	x		x			x	x	x	x	x ^a	x	
United Kingdom	x	x	x	x								
USA	x		x						x		x	

^aIf the hand/wrist bones are fully mature. ^bRadiological examinations are only allowed in criminal proceedings ^cFor victims of human trafficking or vulnerable individuals. ^dThe initial age estimation is only a directive if the refugee is thought to be a minor. If he/she claims to be a minor and is suspected to be an adult, a medical examination is always undertaken. So the principle of benefit of the doubt is used in this initial estimation. ^eOnly some consultations are with social workers.

1.3. The basics of medical age estimation in living adolescents

The principles for age estimation are based on biological ontogeny. The ontogeny describes the process of aging, which is different for every species and genetically controlled (30). Additionally there are individual, population specific, environmental and sociology-controlled factors which influence the aging process as well (30). The progress until adulthood is detectable and measurable based on morphological changes and allows for an estimation of the biological age, which can then be transferred/translated to an estimated chronological age (30). Of course, this estimate cannot give an exact date of birth. Age estimation is always specified as an age range, which is defined as a most likely age and a range of uncertainty. The real age is most likely within this range, with a defined probability (30).

The age estimation process itself should be based on four pillars: skeletal age, dental age, sexual maturity and body constitution (30). According to Schmeling et al. (3) (AGFAD recommendation for criminal proceedings) the combination of the following medical examinations is best suited to fulfill the four-pillar model:

1. External physical examination
2. Plain X-ray of the left hand and wrist
3. Dental examination including an OPG
4. Plain X-ray or CT of the medial end of the clavicles
(if the bone development of the hand is completed)

1.3.1. External physical examination

The examination should include the measurement of anthropometric dimensions like body weight, body height and body type, as well as the evaluation of the sexual maturity. There are different publications for staging sexual maturity, one of the most used are those from Marshall et al. (31,

32) and Tanner (33) who introduced a staging for genital development, breast development and pubic hair. In addition, the individual should be examined for developmental disorders (3) due to malnutrition, infectious diseases and genetic diseases, as these could disrupt bone and/or tooth development.

1.3.2. Plain X-ray of the left hand and wrist

The reason for the examination of the skeleton is that the development sequence of human bones is always the same. There are two types of ossification: Intra-membranous and endochondral ossification. In endochondral ossification, each "bone" is initially composed mainly of hyaline cartilage, which is replaced by bone tissue as the human skeleton develops (34). The appearance of primary and secondary ossification centers and the replacement of cartilage are good indicators for the age estimation of small children up to the age of ten. Around this age the approximate final bone shape is reached, secondary ossification centers have been formed and the bones then mainly grow in length (35, 36). This growth takes place in the so-called epiphyseal or growth plates. The attainment of the final shape and the ossification of the gaps between the metaphysis and epiphysis are good indicators for age estimation in adolescents (34, 36).

The advantage of the hand and wrist for the purpose of age estimation is the high number of individual bones (29 pieces), which account for about 10% of the entire skeleton. In addition, the X-ray examination is a routine procedure which can be carried out under standardized conditions (35, 36). Among others (37–39) the atlas-based method by Greulich and Pyle (40) and the Tanner-Whitehouse single-bone method (41, 42) are the most frequently used classification studies (43, 44).

The atlas-based method, which will be used in this dissertation, was introduced by Greulich and Pyle (40) in 1959 using a North American Caucasian cohort, recruited from a population with good socio-economic background. Using this method the whole hand/wrist radiograph is compared to standard radiographs within the atlas and is assigned to a Greulich-Pyle Standard (GPS) that corresponds to a specific bone age. Since the introduction of the Greulich-Pyle (GP) atlas, many validation studies with different populations have been carried out. Although the

study (40) is now over 50 years old many more recent studies show that it is still applicable, at least in populations with good socio-economic background (45). The development of the bones seems not to be influenced by the ethnic origin, but rather by the socio-economic status (46). A less good status can ultimately lead to a delay in the growth of the bones. In the age estimation process, this could lead to an underestimation of age, but in the legal context this is in favor of the person and, thus, judged as acceptable (46, 47). However, in a very recent MR study by Timme et al. (48) who investigated a German cohort of diverse socio-economic status no influence was found on the development of the bones and teeth. Nevertheless, these results could be biased, as even the low socio-economic standard in a German cohort is higher than most standards in economically weak countries.

1.3.3. Dental examination

This examination should consist of an orthopantomogram (OPG) and an oral dental examination to determine the status of dental development. The OPG serves for the evaluation of mineralization and eruption of the teeth. The oral examination evaluates only the erupted teeth and other dental characteristics, such as dental health and tooth wear (30).

The tooth development is a complex process of organ evolution and consists of the tooth mineralization and the tooth eruption. Contrary to the skeletal development the teeth emerge from a different germ layer, the ectoderm (49). This fact makes the teeth very interesting for the age estimation as an additional complementary factor. The development of the tooth is equal both for deciduous and permanent teeth and is a constant process with less influence from the outside as long as the teeth are inside the jaws (49, 50). While the tooth mineralization refers to the development of the crown and root, tooth eruption describes the path of the tooth from the jaw to the chewing level in the oral cavity. Mineralization can only be evaluated radiographically (30) and there are numerous publications with different stages and staging techniques including, but not limited to (51–56). There are two main evaluation methods: an atlas style technique and a measurement employing technique. The most used publication for the atlas method is the one from Schour et al. (55) and updated versions of it (57). The measurement techniques are based on

the measurement of the developmental stage of the teeth which is then used for two calculation methods. One method converts the measurements into a score that leads to an estimated age using tables. The other possibility is to use the measurements in a regression equation to calculate the age. There are two basic approaches for the measurement: a) relative measurements compared against a scheme and b) percentage of the final root. The system introduced by Demirjian et al. (52) is representative for the first approach, the one from Moorrees et al. (54) for the second. For mid-teen and older individuals the only remaining radiographically relevant teeth are the third molars (57). However, the third molar is known as the most variable tooth in dentition, which makes its use in the age assessment procedure more difficult (58, 59). Mincer et al. (59) showed the applicability of the Demirjian approach, which was originally developed for teeth 41-47 (World Dental Federation (FDI) two digit classification), to the third molars. In a comparison of five different staging systems for mineralization of third molars Olze et al. (60) showed that the Demirjian classification was the most accurate, maybe because the stages are not based on a speculative estimation of the final root length. Therefore, it was used in this dissertation.

Tooth eruption can be evaluated radiographically, but partly also through oral examination (when the teeth have broken into the oral cavity). Publications on radiologically investigated eruption of third molars are scarce. The first who evaluated the eruption on orthopantomograms were Levesque et al. (61). However, Olze et al. (62) who investigated the course of eruption of third molars in three ethnic cohorts introduced the first classification system. The few studies regarding eruption showed that this feature is very variable, with high standard deviations.

According to Olze et al. (62, 63) for both mineralization and eruption the ethnic origin is a relevant factor in dental age estimation, and population specific reference tables should be used to prevent systematic errors and to improve the estimation results. This was confirmed by studies done by Orhan et al. (64), Rai et al. (65), and Zeng et al. (66). However, this topic is also controversial. Liversidge (67), Liversidge et al. (68) and Thevissen et al. (69, 70) compared up to nine different populations and did not find relevant statistically significant differences in the development between the cohorts. Also the estimation with non-population specific references showed only minor differences. They concluded that a big, evenly distributed non-specific reference cohort is as good or even better than population-specific references. It should be noted

that these studies have used other classification systems than the comparison study of Olze. Another topic that is controversially discussed is the influence of tooth impaction on the tooth development. The eruption is obviously not assessable. For the tooth mineralization there are again two opinions. Friedrich et al. (71) stated that there is no influence respectively that the overall variation is bigger than the influence of the impaction. On the other hand Olze et al. (72, 73) found that impacted teeth are developing slower than non-impacted teeth. In a recent study done by Guo et al. (74), a slower development of impacted teeth was also shown in a Chinese cohort, but the results were significant only in early stages.

In contrast to skeletal development, tooth development seems not to be influenced by the socio-economic status of the individual (75). The impact of nutrition on the development of the teeth was investigated in a study conducted by Cameriere et al. (76) and as well as in a study conducted by Elamin et al. (77). Both studies compared the tooth development of well-nourished and undernourished children and concluded that nutrition seemed to have no impact on the dental development.

1.3.4. Plain X-ray or CT of the clavicles

Whether an examination of the clavicles is performed depends on the developmental status of the hand. If all epiphyseal gaps of the hand and wrist are closed and therefore its bone development is finished, an additional radiological examination of the clavicles is recommended. The development of the hand skeleton is typically finished at about 18 - 19 years. The epiphyseal plate of the medial end of the clavicle is normally still open at this age and can therefore be a useful feature for the age estimation. The examination of the clavicles should be performed either with a plain X-ray or a CT. Schulz et al. (78) did a comparison between the two modalities and found no significant difference, but some X-ray images were not assessable due to the superposition of anatomical structures. These superpositions could possibly be avoided by additional anterior-oblique images. However, there are no reference data on combined assessments (78). Therefore, plain X-ray exams of the clavicle bones are not recommended anymore and CT is the modality

of choice in the age estimation procedure (79, 80).

First CT studies on the medial clavicle ends were performed by Kreitner et al. (81, 82), who used a four-stage classification. Schmeling et al. (83) expanded this classification to a five-stage system when investigating X-ray images of the clavicles. Schulz et al. (84) used the same five stages for the first time on CT exams and, finally, Kellinghaus et al. (85) adapted the reference values to the nowadays used thin-slice multi-detector CT-systems. In the same year, Kellinghaus et al. (86) also introduced a modified five-stage classification with three substages each for stages 2 and 3.

Fig. 1.1 shows exemplary images regarding the three radiological examinations.

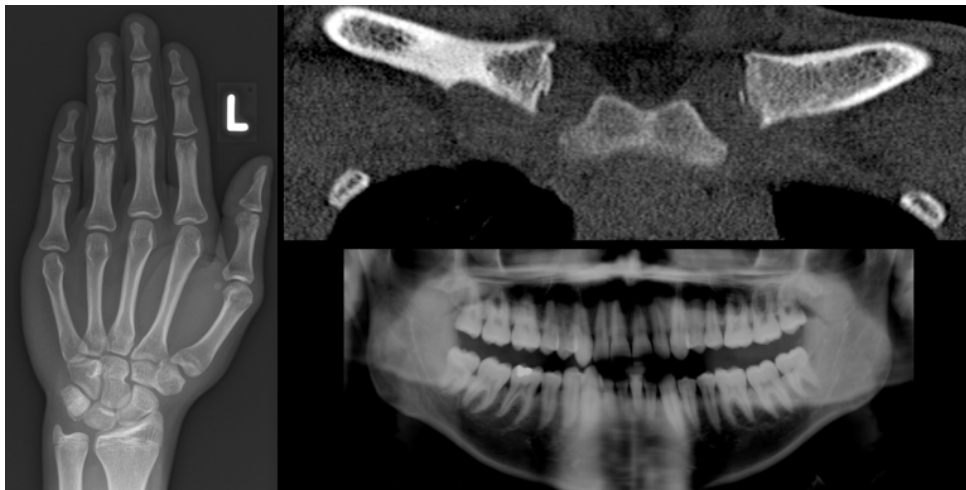


Figure 1.1.: Examples of radiological images for conventional age assessment. X-ray of the left hand, CT of the clavicles and OPG (Images are not from the same person).

1.4. Ethical considerations

There are several points in the current age estimation procedure in the living which raise ethical questions. Regarding this thesis two points should be stressed. On one hand there is the use of ionizing radiation and on the other hand there is the question of accuracy of the current procedure. Both points are subject of controversial discussions (87–89).

One major problem with the current age estimation procedure using X-ray examinations is the application of ionizing radiation. Although the risk of the radiation used is not considered to be significantly high, the biological risk of low-dose radiation is not sufficiently known (90). The use of X-ray examinations without medical indication for the purpose of age estimation is, therefore, at the center of many public debates. In the UK for example the debate rose again in 2012 as the UK border agency wanted to start a trial on age assessment by dental X-rays. It was put on hold after a public outcry by UK child commissioners and immigration lawyers. They criticized the use of ionizing radiation without medical indication and the possible inability of refugees to give informed consent (91–93). Shortly prior to the great wave of refugees in 2015, the discussion about age determination with X-ray examinations was also resumed in Germany on the basis of an article in the Journal of the German Medical Association. The authors of the article were against the current practice of age estimation, in particular, the use of ionizing radiation is criticized (94). The article led to a heated debate within the German Medical Society. Forensic pathologists in particular were outraged, and in a statement (95), the chairman of the AGFAD accused the authors of the article of having provided false information on this sensitive topic. All of these debates point to the urgent need for alternative, non-X-ray based methods for age assessment.

The accuracy of the age estimation is as well a hot topic in the discussions. Indeed, as the name suggests, the age estimation process leads to an estimate and not to an exact age determination. This is due to differences in individual development and may also be influenced by other circumstances such as nutrition, socio-economic influences and stress effects (46, 47, 67, 77, 87). However, age estimation should be as accurate as possible and should have the least possible range of uncertainty. It should be performed by experts according to the current state of science (96, 97). Age estimation carried out in this way is important both for the person examined and for society.

1.5. Age estimation using MRI

In the last decade, research has increased in non-ionizing radiation imaging techniques for age estimation. Two modalities were in the focus of research: ultrasound (US) and MRI. Schulz et al. (98) and Schmidt et al. (99, 100) showed that US might be a good alternative at least for skeletal regions. Others pointed out that US-procedures are highly user dependent and that there are technical and morphological limitations for age estimations (101, 102).

In contrast, MRI is a user-independent alternative based on non-ionizing radiation. As a cross-sectional imaging technique, it generates superposition-free images and the images can also be reconstructed in other planes with the appropriate imaging acquisition technique. Several studies using MRI for age estimation have been conducted in the last years. They were mainly performed on skeletal regions but a few were also used for dental age estimation.

Regarding hand and wrist different approaches were used (103–116). Only four did a comparison to X-ray images (107, 108, 114, 116) and three of them were pilot studies (108, 114, 116) using the Greulich-Pyle atlas. Studies of the clavicles have been published by groups of Germany (117–120), Belgium (121, 122) and Denmark (123). There are also studies regarding other regions of the skeleton such as the distal femur, the proximal and distal tibia, the proximal humerus and the iliac crest (e.g. (124–128)). Concerning dental age estimation of the wisdom teeth, again three groups (from Germany, Belgium and Austria) have published results (129–134). Baumann et al. (129) were the first and only to use 3D MR sequences and to study both mineralization and eruption. They also performed the first comparison of MRI with conventional OPG.

1.6. Multi-factorial age estimation (MFA)

Although a combination of the single regions investigated is a major point in the recommendations of the AGFAD (3), there are not many publications on this issue yet. This may be due to ethical problems (ionizing radiation) for prospective studies, too few cases in corpse studies and too little clinical data from people in whom all regions were examined within a certain, short period of

time (135). However, a good combination is very important because it can partially compensate for the inter-individual biological variation of each region (136).

There are different methodological/statistical approaches for the combination in MFA such as regression formulas, weighted average, transition analysis in combination with Bayes' rule and linear discrimination analysis. Another quite pragmatic approach, which is used in some European countries, is the so called "minimum age concept" that is based on the extrema of the descriptive statistics within the stages of the individual body regions reference studies. The highest minimum age of the stages allocated to the regions is taken as the minimum age of the examined subject (137). However, this approach is highly dependent on the sample size and distribution of the applied reference studies (138) and might, thus, be prone to outliers.

Currently, there is no scientific consensus on the best approach.

To circumvent the ethical issue on prospective studies, three groups are working on an MFA using MRI. The Belgian group used transition analysis in combination with Bayes' rule combining anthropological results with MRIs of the wrist, the clavicles and the third molars (135). The Austrian group published so far results on an automatic, deep learning based approach combining MRIs of the wrist, the clavicles and the third molars (139). Finally, the German group published only the execution of a MFA and may stick to their kind of combination (minimum age concept) currently used with X-ray based modalities. They performed MRIs of the wrist, the knee, the clavicles and one third molar (140–143).

1.7. Aim of the thesis

The aim of this thesis was to investigate the applicability of MRI for forensic age estimation using X-ray based staging systems. The so far in age-estimation practice applied X-ray exams should be replaced by MRI exams. Furthermore, reference values for male adolescent Central Europeans should be provided for MRI examination of the three current used regions (hand, third molars, medial end of the clavicles). In addition, on the basis of the data collected, a prediction model should be applied to get an approach to distinguish between minors and adults.

The following hypotheses formed the basis of the thesis and were tested in studies that form this dissertation:

1. MRI enables detailed image and data acquisition of bones and teeth of developing male adolescents and young adults, allowing optimal assessment of current stages of bone and tooth development with MRI without the use of ionizing radiation.
2. The three investigated regions (hand and wrist, medial end of the clavicles, third molars) can be evaluated with X-ray based classification systems and reference values can be provided.
3. A combination of the evaluation results of the three regions leads to a good prediction model for age estimation purposes.

2. Materials and methods

Parts of the described materials and methods are taken from previously published manuscripts resulting from this thesis (1, 2).

The project is divided into three sections:

- Recruiting of individuals and acquisition of the MRI data
- Evaluation of the single sites and providing results for the single regions
- Multi-factorial combination of the data

2.1. Ethical issues

Ethical clearance for the study was granted by the ethics committee of the Medical University of Graz (No. 21-399 ex 09/10, Appendix B). All participants gave written informed consent prior to study participation, with consent given by legal guardians for minors.

2.2. Study subjects

Study participants were 338 male volunteers between 13 and 25 years (range: 13.01 – 24.98y, median: 18.90y, mean: 18.93y, standard deviation: 3.28y). All participants were young Caucasian

healthy men with documented birth date. All participants had to fulfill the following inclusion criteria:

- Age between 13 and 25 years
- Written consent of the participant or the legal guardian after having been informed about the study
- Documented birth date
- Caucasian ethnicity
- Normal Body Mass Index (BMI)
- No extraction of all third molars

The age range was chosen because it includes all in Austria relevant legal age limits (14, 16, 18 and 21 years). Only Caucasians were included to avoid a possible influence due to ethnicity. The BMI was calculated using the weight and height of the participant. The age distribution of the study subjects can be seen in figure 2.1.

Individuals have been excluded from the study participation in cases of:

- Contraindication to stay inside a 3T magnetic field (e.g. implanted medical devices, claustrophobia, etc.) based on general clinical specifications
- Ferromagnetic implants
- Persons, which could not lie still during the MR examination
- Persons with a relevant trauma in the examined regions
- Persons with physical development disorders
- Severe underweighted persons (≥ 19 years: $\text{BMI} \leq 16.5$ (144);
 < 19 years: $\text{BMI} \leq 3$ standard deviations of the WHO mean values for male children aged 5-19 years (145))

Contraindications were monitored with high attention, as sloppiness can bring serious harm to the persons involved. Depending on the material (non-ferromagnetic), metallic implants are primarily no contraindication for an MRI exam. However, if they are near or inside the investigated area,

they cause massive artefacts, which influence the evaluation. Movement is also a big issue in MR imaging, because it generates blurry images, not suitable for the evaluation. Trauma or developmental disorders may cause a retarded growth, therefore subjects with a corresponding history were excluded. Severe underweight may also cause a retarded growth and therefore, subjects with the above explained BMI values were excluded.

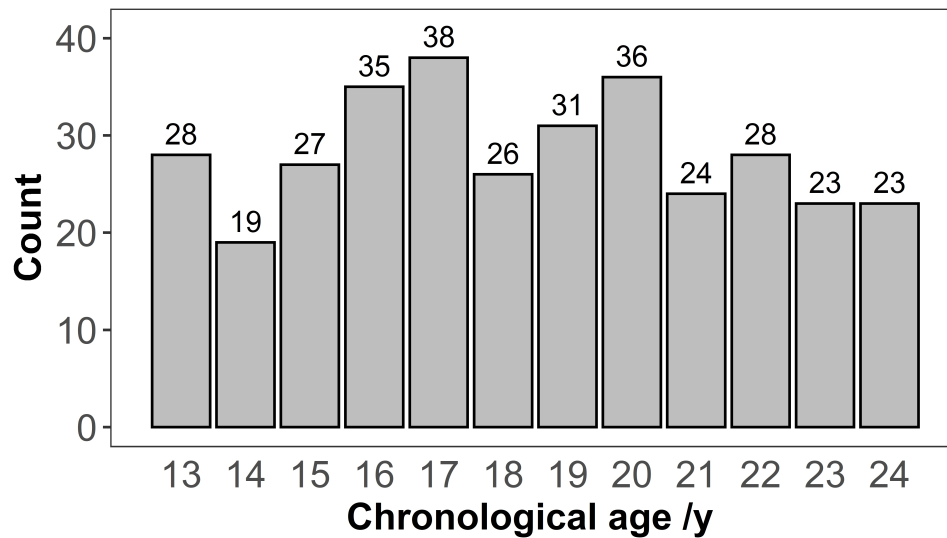


Figure 2.1.: Age distribution of all study subjects in one-year age groups

2.3. MR Imaging

All subjects underwent an MRI of the left hand and wrist, the third molars and the sternoclavicular joints. The MRIs were performed at three sites (University Hospital Graz, MRI-Lab Graz University of Technology, MRI Center Privat Hospital of the Sisters of Mercy of the Holy Cross Graz) with two types of 3T MR scanners (MAGNETOM Trio, a TIM system & MAGNETOM Skyra, SIEMENS Healthcare, Erlangen, Germany).

The sternoclavicular joints were examined in supine position using the neck coil and a flex/body matrix coil. Two coronary sequences were planned that were inclined parallel to the sternal bone. The first sequence, a T2w 2D TSE, had a bigger Field-of-View (FOV) to image the

whole clavicles. For the second sequence, a T1w 3D VIBE FS the flex/body matrix coil was switched off and the FOV was focused on the sternoclavicular joints only. The VIBE was the main sequence used for evaluation, as a similar sequence had also been used by other groups (118, 146) who investigated the clavicles. The acquisition of the two sequences took about 11 minutes.

For the hand imaging the participants were placed in prone position with the left arm outstretched. The hand was fixed with a sand bag and the standard head and neck coil was used. Two sequences were applied in coronal orientation: T1w 3D VIBE WE and T2w 3D DESS WE. The FOV of the VIBE sequence included the entire hand and wrist, while the DESS sequence focused solely on the wrist. The total time for the acquisition of the two high-resolution sequences was 5:50 minutes. The VIBE sequence selection followed studies done by Hojreh et al. (108) and Urschler et al. (114). The DESS sequence was selected to acquire a second 3D sequence with a different weighting in order to obtain additional image data for evaluation of the wrist region.

Dental MRI was performed in supine position using a Cloth Pin Coil (CPC) (Noras MRI products GmbH, Hoechberg, Germany). The two coil parts were placed left and right along the jaw. Two sequences, a PDw 3D TSE FS and a T2w 3D CISS, were positioned identically, sagittally oblique, to image the distal portions of both jaws. Both 3D sequences were nearly isotropic and allowed multi-planar reformations and were used beforehand in the pilot study of Baumann et al. (129). The examination of the third molars was the longest part of the examination and took approx. 20 minutes. Both sequences were newly adapted from existing SIEMENS standard protocols to get an optimal contrast between tooth and surrounding tissue.

The whole MRI examination (three parts) lasted approx. 50 minutes including positioning and planning time. All sequences are modified standard sequences, which were implemented by an MR-physicist, a radiologist and the doctoral candidate. Table 2.1 shows the parameters of the six MR sequences. Minor parameter changes were necessary due to the different weight of the study subjects and the different SAR monitoring of the two scanner types. However, this did not have any influence on the image quality and evaluability. The MR examinations were performed exclusively by one radiographer (the doctoral candidate).

Table 2.1.: MRI sequence parameter

Region	Sequence	FOV (mm)	Readout Matrix (px)	SLT (mm)	TR (ms)	TE (ms)	FA (°)	Acquisition Time (min)
Clavicles	T2w 2D	250	256	2.0	2910	65	150 ^a	5:57
	TSE							
	T1w 3D VIBE FS	170	192	0.9	9.41- 9.77	3.69- 3.72	12	≈5:00
Hand & Wrist	T1w 3D	230	512 ^b	0.9	14.00	4.01- 4.08	15	3:46
	VIBE WE							
	T2w 3D DESS WE	100	256 ^b	0.8	14.28	5.18	30	2:04
Third Molars	PDw 3D	150	512 ^b	1.0	172-	10-12	150 ^a	≈10:00
	TSE FS				254			
	T2w 3D CISS	150	256	1.0	5.41- 5.44	2.33- 2.37	30	≈9:00

FOV Field-of-View SLT slice thickness TR repetition time TE echo time FA flip angle

^a refocussing ^b interpolated

2.4. Single region studies: Evaluation and analysis methods

In this section the methods for the single region evaluation and data analysis will be presented.

2.4.1. Third Molars

Dental MR evaluation

For the dental evaluation data of 316 subjects (range: 13.01–24.98 y, median: 19.32 y) were used. The evaluation of the MR images was done independently by Heiko Merkens and Julian Boldt,

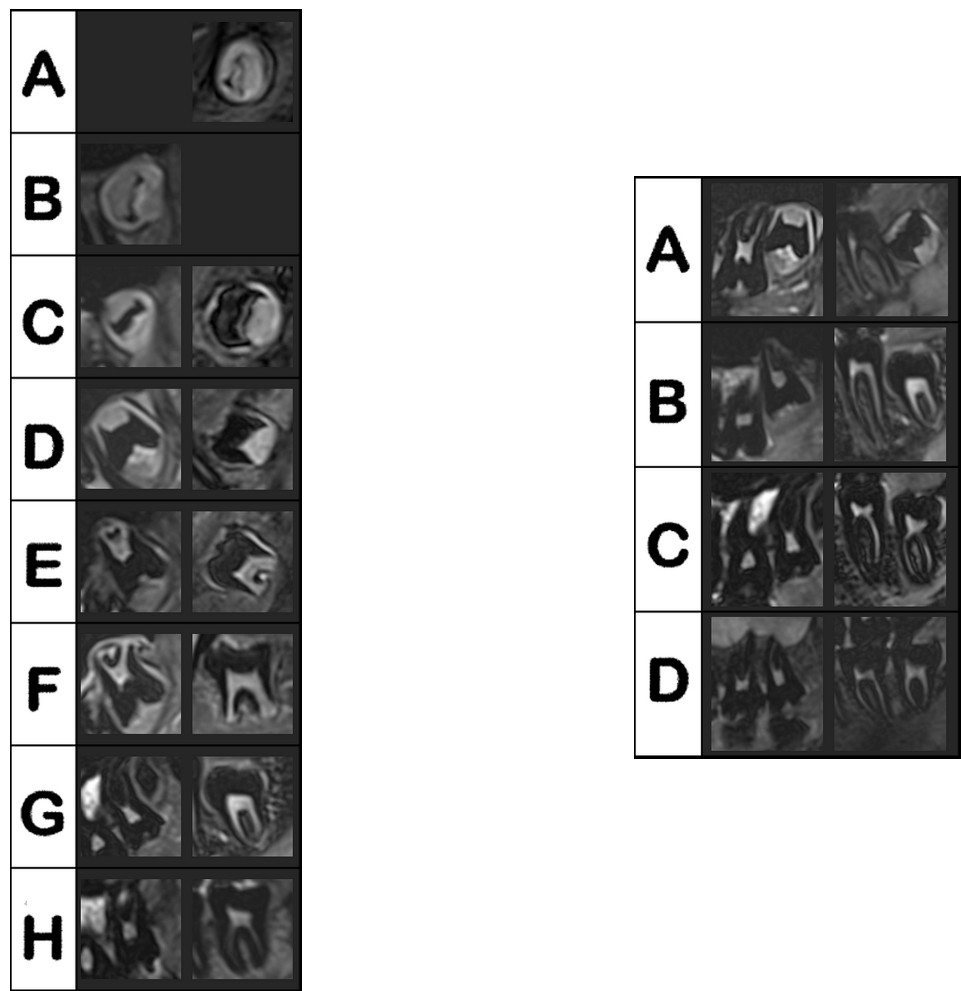
two board certified dentists. They used an open-source DICOM viewer software (OsiriX 4.1, <http://www.osirix-viewer.com>). The MR images were anonymized and both readers were blinded to the birth date of the participants and to the scores of each other. Both raters evaluated the teeth 18, 28, 38 and 48 (according to the FDI two-digit classification) regarding mineralization and eruption stages. The scoring of the mineralization followed the eight grade staging system established by Demirjian et al. (52) and modified by Mincer et al. (59) for the third molars. Eruption was assessed using the stages introduced by Olze et al. (62). Figs. 2.2a and 2.2b show the stages and corresponding MR images for mineralization and eruption, respectively. Tables 2.2 and 2.3 give the written explanation for the mineralization and eruption stages.

Table 2.2.: Written explanation for the mineralization stages as described by Mincer et al. (59).

Stage	Description
A	Cusp tips are mineralized but have not yet coalesced
B	Mineralized cusps are united so the mature coronal morphology is well-defined
C	The crown is about half formed; the pulp chamber is evident and dentinal deposition is occurring
D	Crown formation is complete to the dentoenamel junction. The pulp chamber has a trapezoidal form
E	Formation of the inter-radicular bifurcation has begun. Root length is less than the crown length
F	Root length is at least as great as crown length. Roots have funnel-shaped endings
G	Root walls are parallel, but apices remain open
H	Apical ends of the roots are completely closed, and the periodontal membrane has a uniform width around the root

Table 2.3.: Written explanation for the eruption stages as described by Olze et al. (62).

Stage	Description
A	Occlusal plane covered with alveolar bone
B	Alveolar emergence; complete resorption of alveolar bone over occlusal plane
C	Gingival emergence; penetration of gingiva by at least one dental cusp
D	Complete emergence in occlusal plane



(a) Mineralization stages and corresponding MR images of the upper (left hand side) and lower (right hand side) jaw according to the Demirjian et al. (52) classification. In the current study there was seen no stage A in the upper jaw and no stage B in the lower jaw.

(b) Eruption stages and corresponding MR images of the upper (left hand side) and lower (right hand side) jaw according to Olze et al. (147)

Figure 2.2.: Stages and corresponding MR images for mineralization and eruption (Figure adapted from (1) with permission from Elsevier).

Statistical analysis

All statistical analyses were performed using the software R v3.6.3 (148) and Microsoft Office Excel 2016 (Microsoft Corp., Redmond, WA, USA).

There was no consensus reading, therefore, the results from the more experienced reader (H.M.) were used for descriptive statistics of the single data groups according to the assigned stages. The statistics was done separately for each third molar (18, 28, 38, 48) for mineralization and eruption, respectively. It includes mean, standard deviation (SD), standard error of the mean (SEM), minimum (MIN), maximum (MAX), lower quartile (LQ), upper quartile (UQ) and the median. For the description of the correlation between the stages and the chronological age box plots of the data for each third molar were used. In addition, the 10, 50 and 90% probabilities for reaching the end stages H and D, respectively, were calculated by means of logistic regression (149).

Chronological age at the time of the MRI was calculated by subtracting the examination date from the date of birth and recorded as years with two decimals. In subjects in whom all four third molars were assessable, possible significant differences in development between teeth were tested with a Friedman test. If further differentiation was needed between the groups a Conover post-hoc test with a Holm p-value correction was used.

For the calculation of the ratios to be 14, 16 and 18 years of age, respectively, when reaching the mineralization stages E, F, G or H or the eruption stages B, C or D an empirical probability approach was used for each third molar separately (150).

Linear weighted Cohen's Kappa was taken to compute inter- and intra-rater agreement. The weighting takes into account the degree of disagreement (151, 152). For the intra-rater reliability one dentist (H.M.) re-evaluated 30% of the data sets .

2.4.2. Hand and Wrist

Hand MR evaluation

For the hand and wrist evaluation only subjects up to 21 years were included due to growth of hand and wrist finishing in general in the range between 16 and 20 years (34). Therefore, 238 male subjects between the age of 13 and 21 years were evaluated (age range: 13.01 – 20.99 years, median: 17.39 years). The upper age range limit was chosen for an accurate calculation of the means and the prediction intervals for the higher male GP standards (45).

All hand MR images were evaluated independently by two blinded raters (Thomas Ehammer, board certified radiologist with more than 10 years MR reading experience and the doctoral candidate with 20 years MR imaging experience) according to the defined standards of the GP atlas (40). Fig. 2.3 provides an exemplary comparison of an X-ray image of the hand with MR hand images for GPS 25. For the evaluation open-source DICOM viewer applications (OsiriX 4.1, <https://www.osirix-viewer.com> and Horos 3.3.5, <https://horosproject.org>) were used. Six months after completing the first assessment, the radiologist (T.E.) re-evaluated 10% randomly selected cases for the intra-rater agreement.

Statistical analysis

Statistical analyses were performed using the software R v3.6.3 (148) with the R-packages “BlandAltmanLeh” (153), “irr” (154), “pscl” (155), “MASS” (156) and “VGAM” (157)) and with modified R-scripts provided by Lyle Konigsberg (158).

As no consensus reading was done the results from the radiologist were used for descriptive statistics of the single Greulich-Pyle Standard (GPS) groups (mean, SD, MIN, MAX, LQ, UQ and the median). Again, weighted Cohen’s Kappa (151, 152) was used for intra-observer and inter-observer agreements. Additionally, a contingency table was calculated to show the agreement in detail. A comparison of the age estimates of both readers with chronological age is shown using Bland-Altman plots (159).



Figure 2.3.: Comparison of an X-ray image of the hand with MR hand images of the same individual (MR sequence: T1w VIBE / GPS 25).

As an additional statistical evaluation transition analysis (160) was performed. Transition analysis (TA) is an approach which fits a parametric regression model to describe the transition from one stage to the next higher developmental stage (161). The probability to be in one stage as opposed to another can be described with a general linear model: $Pr(y_j \geq i | a_j) = \Lambda(\alpha + \beta a_j)$, where y_j is the stage of the j-th subject, a_j is the age of the j-th subject, α and β are the parameters to be estimated from the cohort and the symbol Λ is the link function (160). The intercept (α) and the slope (β) of the above equation mark the time for the transition from one stage to the next higher stage and can be transferred to the mean and standard deviation of a distribution (160). For the current evaluation a cumulative probit model (161, 162) with the age on a log-scale was taken as link function. As some of the GPS groups showed low counts they were collapsed according to the following scheme: $\leq$ GPS24, GPS25, GPS26+27, GPS28, GPS29, GPS30 and GPS31. The "goodness-of-fit" of the cumulative probit model was tested. For this purpose a Lagrange multiplier test and the calculation of Cragg & Uhlers Pseudo- R^2 was performed (162). The maximum likelihood estimation of α and β were done with a vectorized generalized linear model (157) with the probit as link function. The estimated intercepts and the slope were used to calculate the probability to be in a stage, which were then normalized and normed maximum likelihood curves with 95% prediction intervals were drawn per GPS (R-script in the Appendix A.1). Normed likelihoods are the probabilities to be in one stage, scaled such that the maximum likelihood equals 1.0 (161). Following Tangmose et al. (163) a uniform prior distribution was assumed, thus the estimated age (point prediction) equals the age value which meets the normed maximum likelihood (maximum likelihood of age (MLAGE)). The calculation of the limits of the 95% prediction interval for the likelihood curves was carried out according to (161), which means that all values greater than a normed likelihood of 0.1465 lie within the 95% prediction interval (161).

For validation and to avoid overfitting a ten-fold cross-validation was applied. The accuracy of the age prediction was shown through mean error (chronological age minus MLAGE) and RMSE. These results were obtained by averaging the resulting values from the cross-validation. The values of the TA and the values of the descriptive statistics were then compared. For the calculation of the 95% prediction interval (PI) of the descriptive statistics the data was tested

with a Shapiro-Wilk test for normal distribution. With confirmed normality the interval was calculated using the following equation: $95\% PI = \bar{x} \pm t_{0.975, n-1} * s * \sqrt{1 + \frac{1}{n}}$

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 (45). Additional both the descriptive and the TA values were compared to X-ray studies (45, 164, 165). A more detailed comparison including significance investigation was done with the big french study of Chaumoitre et al. (45).

For all comparisons one and two sample t-tests were used (166, 167). The limit of significance was set at $p < 0.05$.

2.4.3. Clavicular

Sternoclavicular joint MR evaluation

338 MRI clavicle datasets have been evaluated for this study part (range: 13.01y – 24.98y, median: 18.90y). For the evaluation open-source DICOM viewers (OsiriX 4.1, <https://www.osirix-viewer.com> and Horos 3.3.5, <https://horosproject.org>) were used. The data was first evaluated by two blinded raters (Thomas Ehammer MD, radiologist with more than 10 years MR reading experience and Jannick De Tobel DD MD PhD, forensic age estimation specialist with more than 7 years experience). After applying a Cohen's Kappa inter-rater evaluation only a fair agreement was seen ($\kappa \approx 0.3$) (168). Therefore, additionally a third blinded rater (the doctoral candidate, 20 years MR imaging experience) evaluated the datasets. All three raters used the staging scheme of Schmeling et al. (83) and the sub-classification introduced by Kellinghaus et al. (86) (Table 2.4). Fig. 2.4 shows representative MR images for the stages. For the evaluation the acquired coronal slices and multi-planar reformations of the VIBE sequence were used. The final stage per side was determined by a majority vote. In cases ($\approx 13\%$) were all three assigned a different stage, the third rater made the decision by re-evaluating the cases knowing the assignments of all three raters.

Table 2.4.: Written explanation for the clavicle stages as described by Kellinghaus et al. (85, 86).

Stage	Description
1	Ossification centre not ossified
2	Ossification centre ossified, the epiphyseal cartilage not ossified
2a	The lengthwise epiphyseal measurement is one third or less compared to the widthwise measurement of the metaphyseal ending
2b	The lengthwise epiphyseal measurement is over one third until two thirds compared to the widthwise measurement of the metaphyseal ending
2c	The lengthwise epiphyseal measurement is over two thirds compared to the widthwise measurement of the metaphyseal ending
3	Epiphyseal cartilage is partly ossified
3a	The epiphyseal-metaphyseal fusion completes one third or less of the former gap between epiphysis and metaphysis
3b	The epiphyseal-metaphyseal fusion completes over one third until two thirds of the former gap between epiphysis and metaphysis
3c	The epiphyseal-metaphyseal fusion completes over two thirds of the former gap between epiphysis and metaphysis
4	Epiphyseal cartilage fully ossified, epiphyseal scar visible
5	Epiphyseal cartilage fully ossified, epiphyseal scar no longer visible

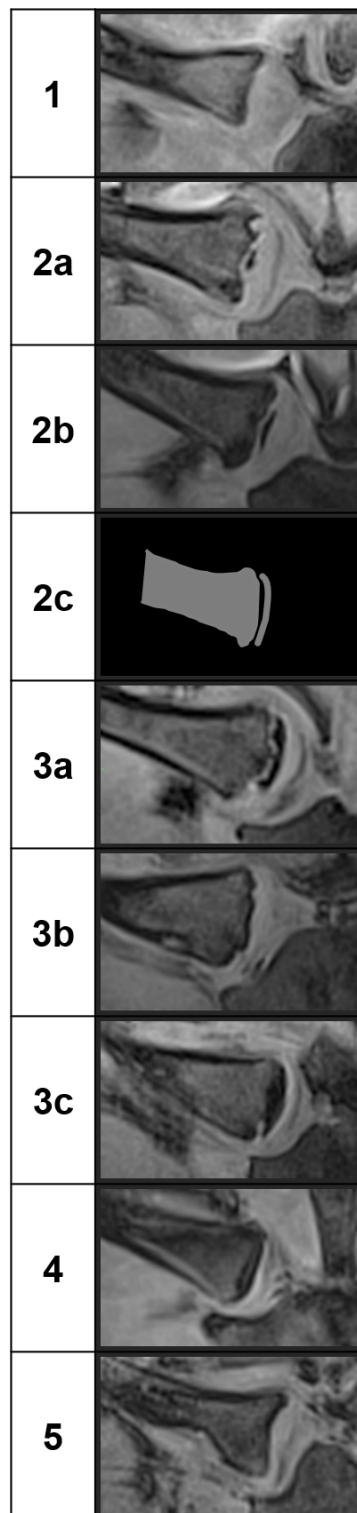


Figure 2.4.: Representative MR images of the clavicle stages. Stage 2c was not found in the current study, therefore, a sketch is presented.

Statistical analysis

All analyses were done using the software R v3.6.3 (148) with the R-packages “irr” (154), “pscl” (155), “MASS” (156) and “VGAM” (157)) and with modified R-scripts provided by Lyle Konigsberg (158).

First, descriptive statistics of the individual stages (mean, SD, MIN, MAX, LQ, UQ and the median) were calculated, both separately for the two sides and jointly for both sides. Possible statistical differences between the right and left side were evaluated using a paired Wilcoxon test. For the inter-rater agreement Fleiss’ Kappa and Cohen’s Kappa calculations were applied.

Similar to the hand evaluation transition analysis using a cumulative probit model was performed. Using the model normed likelihoods prediction intervals were calculated individually per stage. Cross-validation and test-set validation were performed, and the accuracy of the age prediction was shown by calculating the absolute mean error (chronological age minus MLAGe) and the RMSE. The different test constellations can be seen in Table 2.5); only cases with both sides staged were used. Due to low counts some stages were collapsed resulting in the following staging groups: 1, 2(abc), 3a, 3b, 3c and 4/5. The best model was chosen to present more detailed results and plots. For these calculations no cross-validation and all assessed clavicles were used. The fit of the model was tested with a Lagrange multiplier test and also Cragg & Uhlers Pseudo-R² was calculated (162). Additionally, to judge the models ability to differentiate between adults and minors, the accuracy (percentage of correctly classified individuals), the specificity (percentage of correctly classified minors) and the sensitivity (percentage of correctly classified adults) were determined with respect to the point prediction (122). Finally, similar to the hand and wrist evaluation normed likelihood curves were plotted following the description by Konigsberg et al. (161) and Tangmose et al. (163).

Table 2.5.: Cross-validation and test-set constellations.

Model trainings data set	Count	Crossvalidation fold	Test set
Clavicula right	288	10	Clavicula right
Clavicula left	288	10	Clavicula left
Whole set ^a	576	10	Whole set
Clavicula right	288	—	Clavicula left
Clavicula left	288	—	Clavicula right
Clavicula right/left ^b	288	—	Half set max ^c
Clavicula right/left	288	—	Half set min ^d
Half set max	288	10	Half set max
Half set min	288	10	Half set min

^a Whole set: merged dataset of right and left clavicula stages

^b Clavicula right/left: Two models one for the right side, one for the left side (separate evaluation)

^{c,d} Half set max/min: Maximal/Minimal stage of both sides have been taken

2.5. Multi-factorial age estimation (MFA)

2.5.1. MR Data

337 MRI datasets were used for the MFA evaluation. The datasets included an MRI of the hand and wrist, an MRI of both clavicles and an MRI of the third molars. All three MRIs were acquired in one session per individual. The exact parameters can be seen in the MR imaging section (see page 18). The evaluation was done as described in the above sections of the single region MR evaluation. From the third molar evaluation only the mineralization stages were included.

2.5.2. Statistical analysis

For the combination of the individual anatomical regions a latent trait approach according to a study by Konigsberg (169) was used. The aim was to get a point estimate, a prediction interval and a probability to be over the age of 18 years for an individual. In latent trait evaluation, the ordinal categorical trait is assumed to be generated by an unobserved underlying continuous trait that is discretized by thresholds (169). The latent trait is normally distributed and the cumulative probit regression on age/log(age) determines the mean of the distribution with a variance of 1. The intercepts of the probit regression forms the thresholds of the latent trait (169).

The current approach applies a multivariate cumulative probit model fit with age on the log-scale assuming a log-normal distribution. The "Goodness-of-the-fit" was explored with a Lagrange multiplier test (162, 170, 171) for the single traits. The single anatomical regions are not conditionally independent of each other; where conditional independence in relation to age estimation means that all in-between correlations are due to correlations with age. Consequently the residual correlations between the single regions must be zero, but this is not the case (169). Therefore, a residual correlation matrix has to be calculated after accounting for log(age). The correction due to the non-independence is very important to get the right point-estimate (maximum likelihood of age (MLAGE)) and the correct prediction interval. The uncorrected prediction interval will be too narrow and, therefore, will underestimate the uncertainty (169, 172).

In a multivariate case the calculation of the parameters for the individual traits (thresholds and slopes) and the calculation of the correlation matrix is complex, therefore a Markov Chain Monte Carlo (MCMC) method (169) was used for the estimation (R-script in the Appendix A.2). For the MCMC 200000 iterations were taken, by retaining every 40th iteration a total of 5000 iterations remained. To avoid a long "burn-in", starting values for the traits were calculated by univariate cumulative probit regression. The first 500 iterations were nevertheless discarded and the average of the remaining 4500 values per variable was used for the individual results. The results of the MCMC calculation are a table of the thresholds and slopes for the individual traits and the correlation matrix.

Posterior densities were then calculated by combining the traits by numerical integration under a uniform prior taking into account the thresholds for the individual stages of the subjects. During the calculation the results were corrected using the correlation matrix (R-script in the Appendix A.3). The point estimate equaled then the MLAG. The prediction interval and the probability were computed using highest density region computation (173) and linear interpolation (using R-function `auc` with composite trapezoid rule) (174), respectively.

Different combinations of the hand, clavicle and third molars were used and evaluated to identify the best model. As a result of the evaluation of the individual regions, only minimal or maximal stages of the clavicles and third molars were included. For comparison purposes, the above described approach was also applied to the individual regions (hand (univariate), combined clavicles (2C), combined third molars (4T)). As target values mean absolute error (MAE), RMSE, accuracy, sensitivity, specificity, discrimination slope and the mean prediction interval were analyzed. Accuracy (percentage of correctly classified individuals), specificity (percentage of correctly classified minors), sensitivity (percentage of correctly classified adults) and discrimination slope (the difference between adults and minors in mean predicted probability to be a minor) judge the models ability to differentiate adults and minors (122). These calculations refer to the point estimation. Additionally, the absolute errors (absolute difference between chronological age and MLAG) of the different models were pairwise compared with a Wilcoxon signed-rank test, to further evaluate the performance of the models.

All statistical calculations were done using the software R v3.6.3 (148) with the R-packages "coda" (175), "fitR" (176), "hdrcde" (177), "MESS" (174), "mvtnorm" (178), "PredictABEL" (179) and "VGAM" (157)). Additionally, modified R-scripts provided by Konigsberg (158) and Lynch (180) were applied.

3. Results

Parts of the results chapter are taken from previously published manuscripts that resulted from the thesis (1, 2). Figures and Tables are (slightly adapted) reused with a right granted from Elsevier where indicated.

In this chapter the results of the single region studies and the multi-factorial age estimation (MFA) approach are presented.

3.1. Third Molars: Results

316 male subjects participated in this study. About three quarters (n=242) of them had all four third molars, 31 subjects possessed three third molars, 26 had two third molars and in 5 cases three third molars were missing resulting in only one third molar. The rest (n=12) had no third molars at all.

Fig. 3.1 shows an exemplary case of a 17.69 years old study participant. Both used sequences were attested a good image quality by the two readers. The image quality assessment was based on the dentists' ability to recognize the various criteria they needed to assess the corresponding mineralization or eruption stage. There was no scale applied. Although image quality was generally good, there was an exclusion of single teeth due to three reasons: Technical causes (e.g. motion, metal artifacts), anatomical causes (e.g. peg-shaped tooth) and stage definition causes

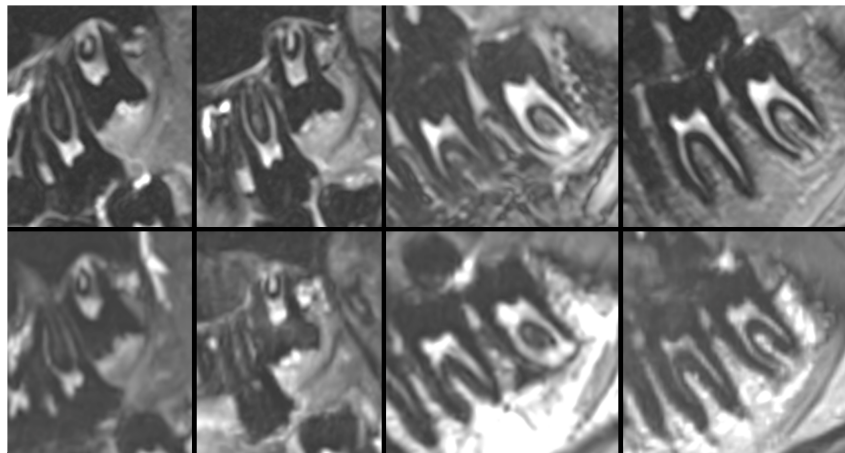


Figure 3.1.: MR images of the third molars of a 17.69y old male. From left to right tooth 18, 28, 38 and 48 are shown (Mineralization - Demirjian stages: F F F G; Eruption - Olze stages: B B D D). The first row displays the PDw 3D TSE FS, second row the T2w 3D CISS.

(e.g. mesial tilt of the tooth resulting in an undefined eruption stage). The number of included and excluded teeth per individual third molar group (18, 28, 38, 48), separated by mineralization and eruption, is shown in Table 3.1.

Missing teeth among the four third molars were almost equally distributed. Mineralization classification was approximately twice as much affected by artifact exclusion than eruption evaluation, but non quadrant specific. Stage definition problems occurred almost exclusively in the analysis of eruption and predominately in the lower jaw. An anatomically determined exclusion occurred in only one tooth. For the mineralization assessment, an average of 15% of the teeth per third molar had to be excluded, approximately one quarter were not evaluable in the classification of the eruption stage.

Table 3.1.: Number of in- and excluded teeth per third molar and evaluation.

Table adapted from Widek et al. (1).

		18		28		38		48	
		Miner	Erup	Miner	Erup	Miner	Erup	Miner	Erup
Included		262	256	277	276	255	224	259	211
Excluded	Missing	35	35	26	26	42	42	43	43
	Artifact	18	8	12	5	19	9	14	8
	Undefined stage	1	17	0	8	0	41	0	54
	Anatomical cause	0	0	1	1	0	0	0	0
Sum		316	316	316	316	316	316	316	316

Miner = Mineralization Erup = Eruption

3.1.1. Mineralization results

Mean, SD, SEM, MIN, MAX, LQ, UQ and median were computed for each third molar group separately for the single Demirjian stages (A – H) (Table 3.2 see page 38).

Due to a low count single stages were excluded from further statistical analysis (A, B, C). Stage A was only seen in the lower jaw, whereas stage B appeared only in the upper jaw. The minimum age found for stages D and E was equal to the cohorts minimum age (13.01 years). For terminal stage H, the minimum age on the left side of the jaw (28, 38) was 17.3 years, but 18.7 years on the right side of the jaw (18, 48). The SEM was in the range of 0.1 – 0.3 years and the SD varied between 0.6 and 2.3 years for stages D – H. Overall, the SEM and the SD, were very stable for all third molars and all stages.

The interrelation between the mineralization stages and chronological age is shown in Figure 3.2. The boxes cover the lower and upper quartile, and the thick transverse line displays the median. The maximum and minimum values are shown by the whiskers, and outliers are indicated by a

black dot. Outliers are defined as values which lie outside 1.5 times the interquartile range (IQR) of the lower or upper quartile, respectively.

The values of the first stage and the final stage (A and H) are of course only guiding values as they depend on the age range of the sample. In particular, stage H can be seen from achievement of the stage till death or loss of the tooth. Therefore, the 10, 50 and 90% probability to reach stage H was calculated for each wisdom tooth. Table 3.3 shows the correspondent values gained by means of logistic regression.

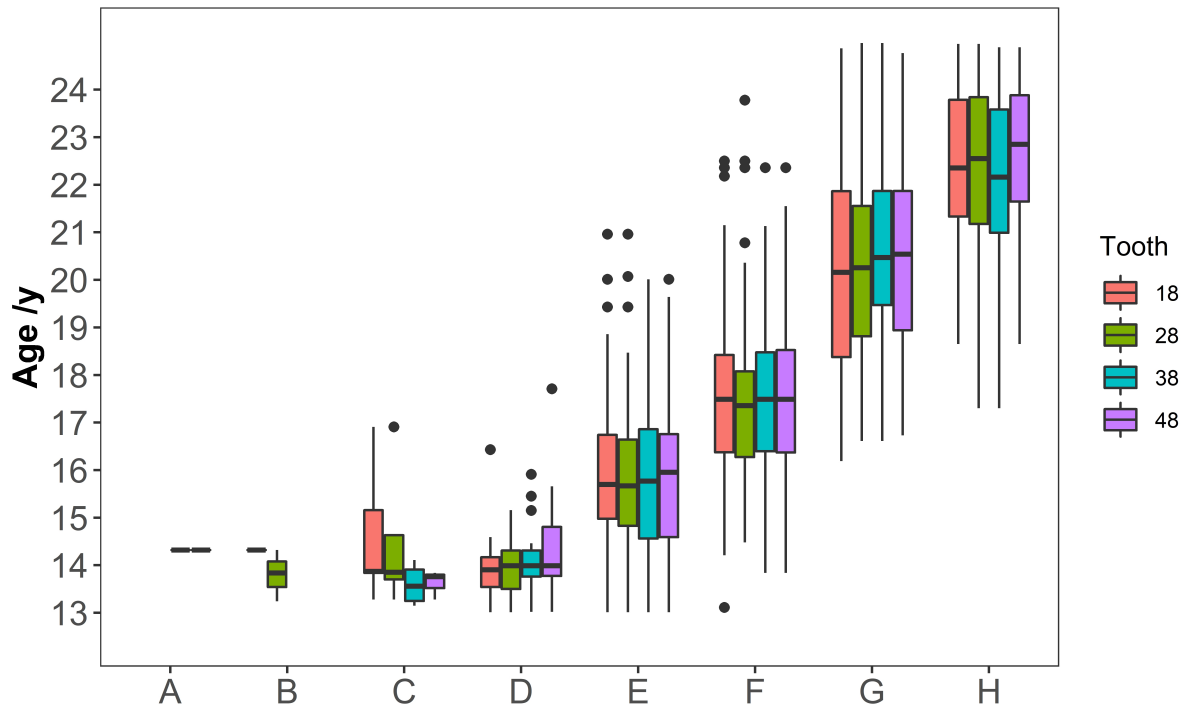


Figure 3.2.: Box and whisker plot of the mineralization. On the x-axis are the Demirjian stages, the y-axis displays chronological age. Figure adapted from Widek et al. (1).

Table 3.2.: Descriptive statistical data of the mineralization of tooth 18, 28, 38, 48 (in years) concerning the different Demirjian stages (A-H). Table adapted from Widek et al. (1).

	Tooth	18	28	38	48		Tooth	18	28	38	48
stage A	N	0	0	1	1	stage E	N	39	41	43	46
	Mean			14.32	14.32		Mean	16.03	15.81	15.8	15.89
	SEM						SEM	0.28	0.28	0.26	0.25
	SD						SD	1.74	1.81	1.69	1.72
	MIN						MIN	13.01	13.01	13.01	13.01
	LQ						LQ	14.97	14.83	14.56	14.59
	Median						Median	15.7	15.67	15.77	15.96
	UQ						UQ	16.74	16.64	16.86	16.76
	MAX						MAX	20.96	20.96	20.01	20.01
stage B	N	1	3	0	0	stage F	N	51	60	64	55
	Mean	14.32	13.8				Mean	17.56	17.47	17.53	17.59
	SEM		0.31				SEM	0.29	0.24	0.21	0.25
	SD		0.54				SD	2.04	1.89	1.7	1.85
	MIN		13.24				MIN	13.11	14.48	13.84	13.84
	LQ		13.54				LQ	16.38	16.27	16.40	16.37
	Median		13.84				Median	17.49	17.36	17.49	17.49
	UQ		14.08				UQ	18.42	18.07	18.47	18.52
	MAX		14.32				MAX	22.5	23.78	22.36	22.36
stage C	N	5	4	4	3	stage G	N	94	96	73	79
	Mean	14.61	14.47	13.59	13.63		Mean	20.22	20.28	20.62	20.48
	SEM	0.65	0.82	0.23	0.17		SEM	0.23	0.21	0.24	0.23
	SD	1.46	1.65	0.46	0.3		SD	2.26	2.02	2.08	2.05
	MIN	13.28	13.28	13.15	13.28		MIN	16.19	16.61	16.61	16.73
	LQ	13.84	13.7	13.25	13.52		LQ	18.38	18.81	19.47	18.94
	Median	13.87	13.86	13.56	13.76		Median	20.16	20.26	20.47	20.54
	UQ	15.16	14.63	13.91	13.8		UQ	21.87	21.55	21.87	21.87
	MAX	16.91	16.91	14.11	13.84		MAX	24.87	24.98	24.98	24.77
stage D	N	20	18	17	19	stage H	N	52	55	53	56
	Mean	13.95	13.92	14.13	14.35		Mean	22.39	22.36	22.15	22.61
	SEM	0.17	0.14	0.18	0.25		SEM	0.23	0.25	0.25	0.21
	SD	0.75	0.59	0.76	1.1		SD	1.66	1.89	1.84	1.57
	MIN	13.01	13.01	13.02	13.02		MIN	18.65	17.3	17.3	18.65
	LQ	13.54	13.5	13.76	13.78		LQ	21.33	21.17	20.99	21.65
	Median	13.91	13.99	13.99	13.99		Median	22.35	22.55	22.16	22.85
	UQ	14.17	14.31	14.31	14.8		UQ	23.79	23.84	23.58	23.88
	MAX	16.43	15.16	15.91	17.71		MAX	24.96	24.96	24.89	24.89

SD standard deviation SEM standard error of the mean MIN minimum MAX maximum

LQ lower quartile UQ upper quartile

Table 3.3.: Age at different probability values to reach stage H (in years). Table adapted from Widek et al. (1).

	Tooth 18	Tooth 28	Tooth 38	Tooth 48
10%	19.26	19.17	18.83	19.46
50%	22.69	22.61	22.43	22.36
90%	26.13	26.06	26.04	25.27

The evaluation of subjects in whom all four third molars have been staged (n=210) showed that less than half (n=95) of them had the equal stage for all four teeth. The entire distribution can be seen in Table 3.4 The different mineralization stage patterns showed no tendency for the upper or lower jaw and for the left or right side, respectively. An applied Friedman test showed no significant developmental differences between the four third molars.

Further, the probability to be over 14, 16 or 18 years, when a tooth is in a dedicated root stage, was calculated separately for all four third molars. Results are shown in Table 3.5.

The weighted Kappa for the inter-rater agreement was good (168) with $\kappa_w = 0.76$. The intra-rater agreement was also good with a weighted Kappa of $\kappa_w = 0.77$.

Table 3.4.: Differing stages in subjects with all four third molars.

	Count	%
Four teeth with same stage ^a	95	45.2
One tooth with one stage difference ^b	51	24.3
Upper jaw (18, 28) same stage – Lower jaw (38, 48) one stage difference ^c	37	17.6
Two teeth with one stage ^d	8	3.8
Three different stages within the four teeth ^e	19	9.1
Sum	210	100

^a e.g. Tooth 18 28 38 48 Stage D D D D^b e.g. Tooth 18 28 38 48 Stage D D D E^c e.g. Tooth 18 28 38 48 Stage D D E E^d e.g. Tooth 18 28 38 48 Stage D E D E^e e.g. Tooth 18 28 38 48 Stage D E F E**Table 3.5.:** Empirical probability calculation for mineralization stage E, F, G and H. Table adapted from Widek et al. (1).

	Tooth 18	Tooth 28	Tooth 38	Tooth 48
stage E $\geq 14y$	89.7%	82.9%	83.7%	87.0%
stage F $\geq 14y$	98.0%	100%	98.4%	98.2%
stage F $\geq 16y$	78.4%	78.3%	81.3%	80.0%
stage G $\geq 16y$	100%	100%	100%	100%
stage G $\geq 18y$	77.7%	83.3%	86.3%	86.1%
stage H $\geq 18y$	100%	98.2%	98.1%	100%

3.1.2. Eruption results

Table 3.6 shows the results of the descriptive statistical evaluation (Mean, SD, SEM, MIN, MAX, LQ, UQ and median) separated for the eruption stages A – D and for the four third molars.

Table 3.6.: Descriptive statistical data of the eruption of tooth 18, 28, 38, 48 (in years), concerning the different Olze stages (A – D). Table adapted from Widek et al. (1).

	Tooth	18	28	38	48		Tooth	18	28	38	48
stage A	N	125	138	62	58	stage C	N	25	27	41	34
	Mean	16.65	16.92	15.49	15.50		Mean	20.94	20.28	20.56	19.82
	SEM	0.23	0.23	0.28	0.30		SEM	0.41	0.41	0.36	0.37
	SD	2.52	2.74	2.17	2.32		SD	2.05	2.15	2.28	2.15
	MIN	13.01	13.01	13.01	13.01		MIN	16.73	16.73	16.73	14.92
	LQ	14.59	14.88	13.98	13.85		LQ	20.25	18.34	19.37	17.77
	Median	16.43	16.61	14.96	14.71		Median	21.15	20.57	20.54	20.34
	UQ	18.04	18.61	16.66	16.63		UQ	22.33	21.66	22.39	21.17
	MAX	24.67	24.77	24.40	24.40		MAX	24.56	24.29	24.63	23.55
stage B	N	41	45	59	48	stage D	N	65	66	62	71
	Mean	20.01	19.82	18.94	18.61		Mean	21.67	21.87	21.54	21.65
	SEM	0.40	0.40	0.36	0.40		SEM	0.26	0.25	0.27	0.25
	SD	2.56	2.65	2.80	2.74		SD	2.10	2.02	2.16	2.13
	MIN	15.54	15.54	13.84	15.00		MIN	16.59	17.36	17.30	16.85
	LQ	17.53	17.49	16.61	16.57		LQ	19.92	20.25	19.96	20.10
	Median	19.83	19.60	18.64	17.82		Median	22.03	22.10	21.70	21.80
	UQ	22.10	21.57	20.85	20.53		UQ	23.47	23.73	23.32	23.45
	MAX	24.58	24.98	24.15	24.15		MAX	24.96	24.96	24.89	24.89

SD standard deviation SEM standard error of the mean MIN minimum MAX maximum

LQ lower quartile UQ upper quartile

The correlation between the eruption stages of the four third molars and the chronological age is displayed in Figure 3.3. Again, the maximum and minimum values are shown by the whiskers, and outliers are indicated by a black dot. Outliers are defined as values which lie outside 1.5 times the interquartile range (IQR) of the lower or upper quartile, respectively.

Similar to the mineralization, the terminal stages values are conditional on the sample's age range. This is particularly true for the upper final stage D. Therefore, 10, 50 and 90% probability values to reach stage D were calculated for each third molar (Table 3.7, see page 43).

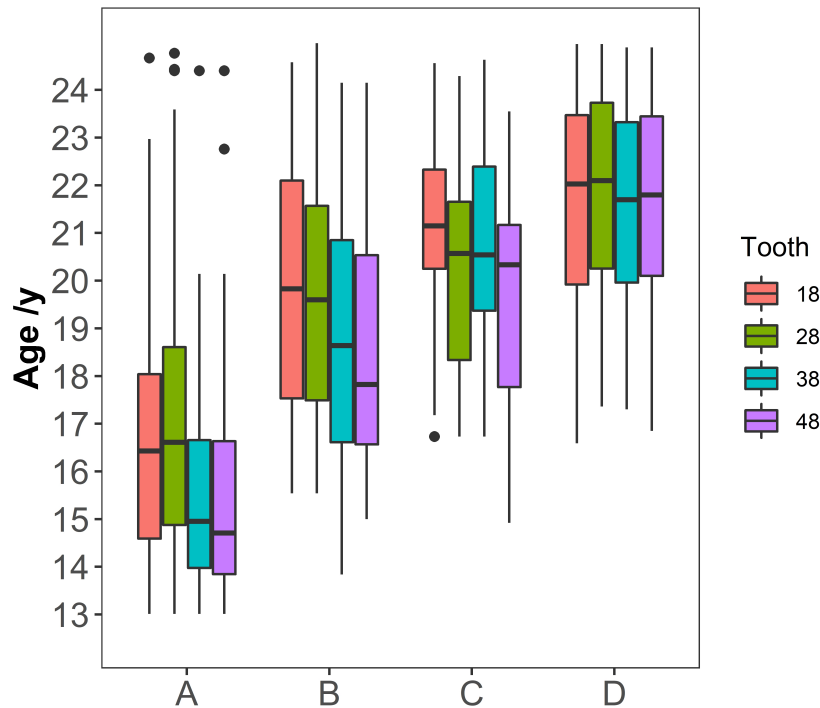


Figure 3.3.: Plot of the distribution of the eruption stages vs. the chronological age of tooth 18, 28, 38 and 48. On the x-axis are the Olze stages and on the y-axis the chronological age is displayed. Figure adapted from Widek et al. (1).

The comparison of the four third molars regarding different eruption stages showed significant differences. The further differentiation with a post-hoc test showed significant differences between upper and lower jaw ($p < 0.05$) (Table 3.8). This can also be seen in the boxplot (Fig. 3.3) where the median values of stages A and B for the lower jaw is clearly lower, that means an earlier transition to the next stage.

The individual tooth comparison of study participants with all four molars ($n = 171$) demonstrates that about half of them ($n = 86$) had the same stages for all four teeth. The other subjects showed different eruption patterns. A weak tendency of higher stages for the lower jaw was found in cases, where the same stage for the upper jaw (18, 28) and a higher or lower stage for the lower jaw (38, 48) was seen. The distribution can be seen in Table 3.9.

Table 3.7.: Age at different probability values to reach stage D (in years). Table adapted from Widek et al. (1).

	Tooth 18	Tooth 28	Tooth 38	Tooth 48
10%	17.54	17.95	16.79	16.74
50%	22.25	22.41	22.31	21.14
90%	26.96	26.87	27.83	25.54

Table 3.8.: Pairwise comparison using Conover post-hoc test with Holm p-value adjustment.

	Z18	Z28	Z38
Z28	0.282	-	-
Z38	0.002*	<0.001*	-
Z48	<0.001*	<0.001*	0.114

* significant

Similar to mineralization the probability to be over 14, 16 or 18 years for the four individual third molars when the tooth is in a dedicated eruption stage (B, C or D) is shown in Table 3.10.

The inter-rater and intra-rater agreement for the eruption evaluation were good to very good (168) with $\kappa_w = 0.80$ and $\kappa_w = 0.89$, respectively.

Table 3.9.: Differing stages in subjects with all four third molars.

	Count	%
Four teeth with same stage ^a	86	50.3
One tooth with one stage difference ^b	33	19.3
Two teeth with one stage ^c	28	16.4
Two teeth with more than one stage difference or three different stages within the four teeth ^d	24	14.0
Sum	171	100

^a e.g. Tooth 18 28 38 48 Stage D D D D^b e.g. Tooth 18 28 38 48 Stage D D D C^c e.g. Tooth 18 28 38 48 Stage D E D E or C C D D^d e.g. Tooth 18 28 38 48 Stage B B D D or B C D C**Table 3.10.:** Empirical probability calculation for eruption stage B, C and D. Table adapted from Widek et al. (1).

	Tooth 18	Tooth 28	Tooth 38	Tooth 48
stage B $\geq 14y$	100%	100%	98.3%	100%
stage B $\geq 16y$	97.6%	95.6%	86.4%	79.2%
stage C $\geq 16y$	100%	100%	100%	97.1%
stage C $\geq 18y$	84.0%	74.1%	82.9%	73.5%
stage D $\geq 16y$	100%	100%	100%	100%
stage D $\geq 18y$	93.9%	95.5%	90.3%	93.0%

3.2. Hand and Wrist: Results

This study included 238 male subjects. The experienced reader rated the overall image quality as good, based on a 1 – 5 scale. Less fat suppression was sometimes seen in the VIBE sequence. The insufficient fat suppression was mainly observed in the finger,s and made the GP standard assessment more difficult. However, all hand MRIs were evaluable. Fig. 3.4 shows an example of MR images of the hand and wrist of a 15 years old male.



Figure 3.4.: 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). Figure adapted from Widek et al. (2).

Table 3.11 displays the descriptive statistics that were calculated for the GP standards found. The statistics include number of subjects, mean, SD, MIN, LQ, median, UQ and MAX for the single GP standards and corresponding GP ages.

Table 3.11.: Descriptive statistics for the single GP standards in years. Table adapted from Widek et al. (2).

GP standard	GP age	N	Mean	SD	MIN	LQ	Median	UQ	MAX
21	11.5	1	14.46		14.46	NA	14.46	NA	14.46
22	12.5	1	13.15		13.15	NA	13.15	NA	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

NA not applicable

Fig. 3.5 shows the range of the GP-Standards per one-year age group. The variability of the GP standards is particularly evident here in people under the age of 17. No subject under 15 years has achieved the GP standard 30. At the age of 15 and 17 years, at least GP standard 24 respectively 28 is reached.

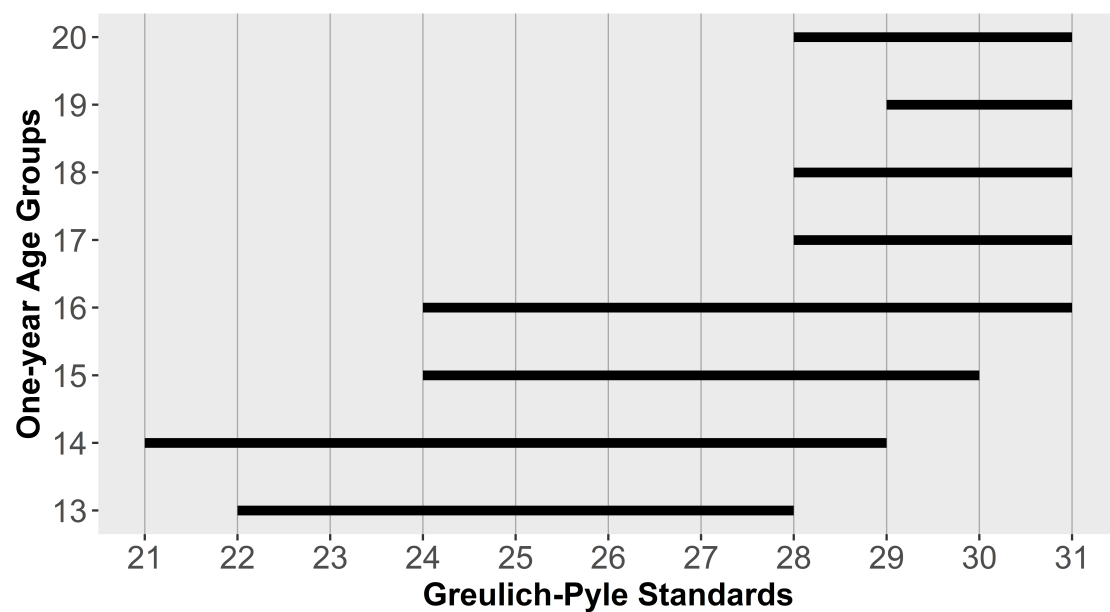


Figure 3.5.: Range of the GP Standards per one-year age group. Figure adapted from Widek et al. (2).

The comparison of the results of the two readers showed a good agreement. The mean difference between reader 1 and 2 was below two months ($d_{mean} = 0.16y$). The weighted Kappa regarding inter- and intra-rater agreement was very good (168) with $\kappa_w = 0.85$ and $\kappa_w = 0.88$, respectively. Table 3.12 displays the agreement in a contingency table in detail.

Table 3.12.: Cumulative cross tabulation of frequencies of allocated GP ages by reader 1 compared with reader 2. Table adapted from Widek et al. (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

Bland-Altman-like plots show the comparison of the estimated to the chronological age separately for the two readers (Fig.3.6). No major deviation was seen for both plots (the digital line pattern can be explained by the assignment to the discrete GP standards). Both readers showed a mean difference of about half a year (R1: 0.37y, R2: 0.54y), which means a slight underestimation. The limits of agreement for the lower limit lay at -1.90 and -1.58 years and for the upper limit at 2.65 and 2.66 years, respectively.

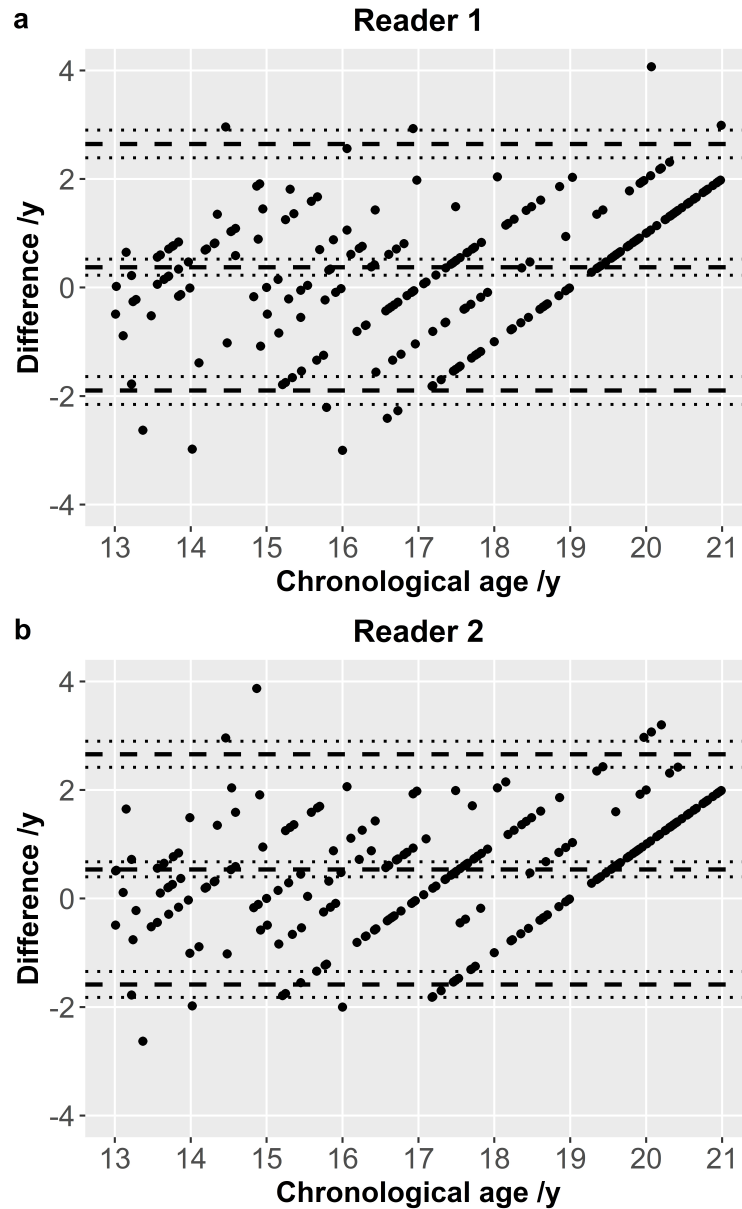


Figure 3.6.: Bland-Altman-like plots showing the comparison of the estimated ages of the two readers with the chronological age. Difference means chronological age minus estimated age. The bold middle dashed line represents the mean difference; the two outer bold dashed lines are the limits of agreement (corresponding to two standard deviations). All lines are accompanied by their respective 95% confidence intervals (thin dotted lines). Figure adapted from Widek et al. (2).

For the GPS $\leq 24 - 31$ transition curves were calculated based on a cumulative probit approach (Fig. 3.7). Within the range of 13 to 21 years normed maximum likelihood curves were computed and plotted for each (combined) GPS (Fig. 3.8).

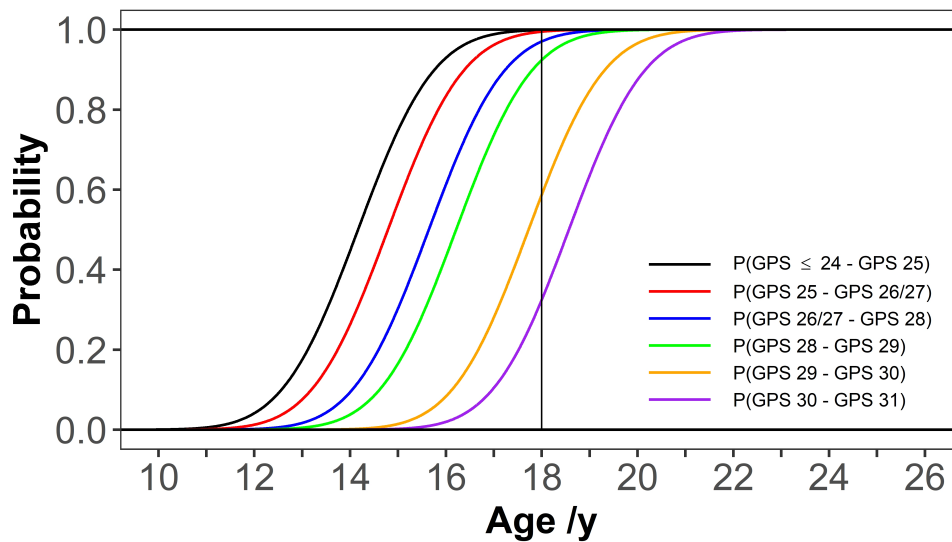


Figure 3.7.: Transition curves using the whole data set (GPS Greulich-Pyle Standard). Figure adapted from Widek et al. (2).

Table 3.13 (see page 52) presents the age values of maximum likelihood estimations and the corresponding prediction intervals as well as a comparison to the mean age and its 95% prediction interval. For reasons of comparability, the GP standard groups have been equated, and descriptive statistic values have been averaged. In the transition analysis, final stages have no maximum likelihood for age because the probit curves are open. For reasons of comparability, the values for the comparison were set to the minimum and maximum value of the sample, respectively. Comparing the descriptive statistics data and the transition analysis data mainly younger ages were seen for the MLAG but similar prediction intervals. Significant differences were not found except for the first and the last GPS group.

The ten-fold cross-validation showed a slight overestimation with a mean error (chronological age - MLAG) of -0.28 years ($SD = \pm 0.53$ years). The RMSE was 1.52 years showing the variability of the estimation. The averaged results of the Lagrange multiplier test ($p = 0.26$) and the pseudo R^2 ($R^2_{Cragg\&Uhlers} = 0.75$) indicate the good fit of the cumulative probit model.

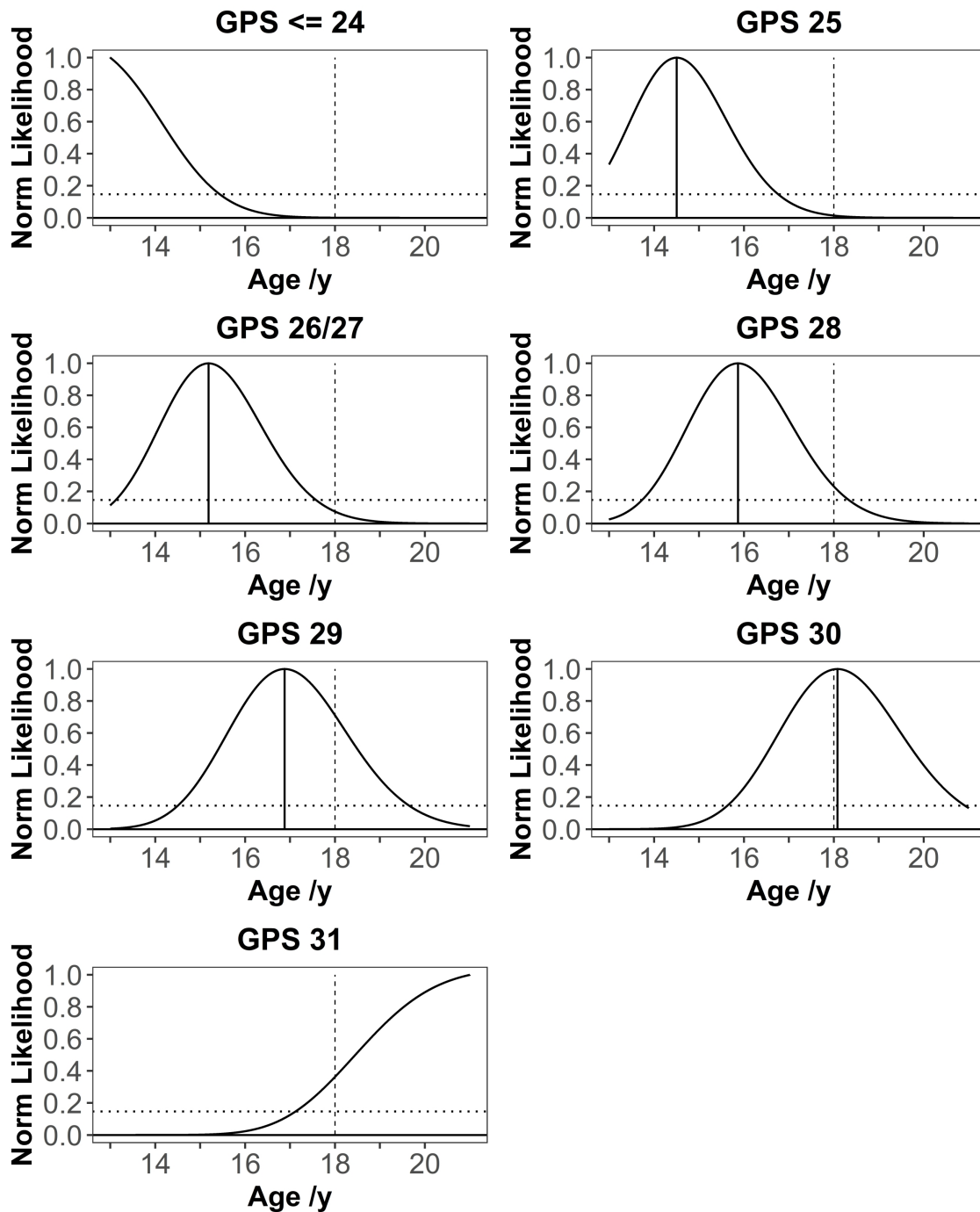


Figure 3.8.: Normed maximum likelihood curves (whole data set). The solid vertical line represents the maximum likelihood of age (MLAGE), the dashed vertical line displays the age of 18. The horizontal line shows the likelihood of 0.1465 indicating the 95% prediction interval. Figure adapted from Widek et al. (2).

Table 3.13.: Comparison of the MLAGe and the mean, and the prediction intervals gained through maximum likelihood estimation and descriptive statistics (in years). Table adapted from Widek et al. (2). Note that the lower bound of the prediction interval in the lowest GPSs reflects the minimum age in the study sample. Similarly, the upper bound in the highest GPS corresponds to the maximum age in the study sample.

GP standard	N	MLAGE	95% prediction interval TA	Mean age	95% prediction interval
≤ 24	33	–	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	–	17.12 – 21.00	19.38	16.96 – 21.80

MLAGE maximum likelihood of age TA transition analysis

3.3. Clavicular: Results

Data of 338 male subjects were analyzed for this part. Some cases respectively single clavicles had to be excluded due to motion artefacts or anatomical shape variants (e.g. “fish-mouth-like” depression). Finally, of 314 subjects, 602 (R: 304, L: 298) clavicles could be included in the descriptive statistics. Fig. 3.9 shows one example for the used sequences in a 17.74 years old male individual.

In Table 3.14 mean, SD, MIN, MAX, LQ, UQ and median of the individual stages found are displayed, separately for the right side, the left side and for both sides together. In Figure 3.10 boxplots of the left and right clavicle staging results are shown. The boxes cover the lower and upper quartile and the thick transverse line displays the median. The maximum and minimum values are shown by the whisker ends and outliers are indicated by a black dot. Outliers are defined as values which lie outside 1.5 times the interquartile range (IQR) of the lower or upper quartile, respectively.

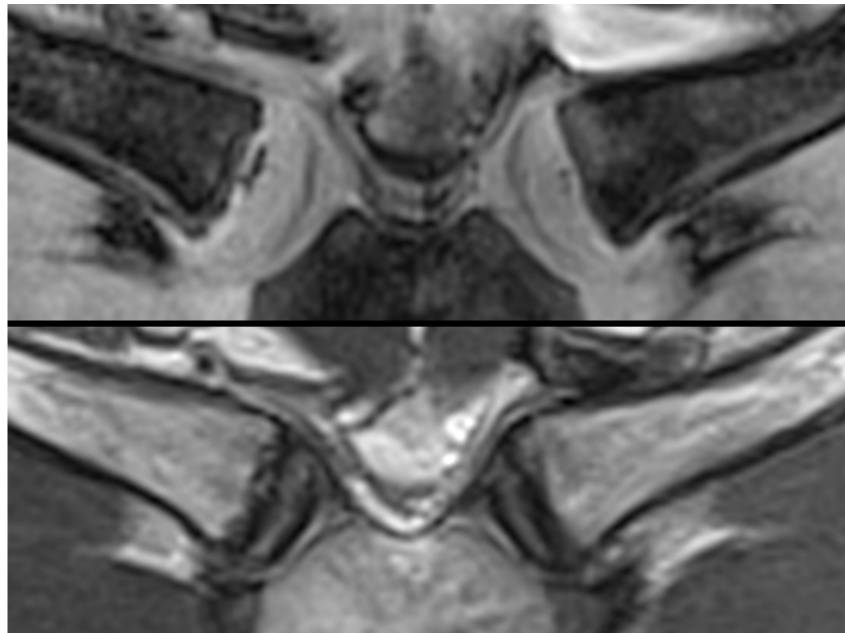


Figure 3.9.: MR images of a 17.74y old male with stage 3a (right) and 2a (left) according to the enhanced Kellinghaus classification. Top: T1w 3D VIBE. Bottom: T2w TSE.

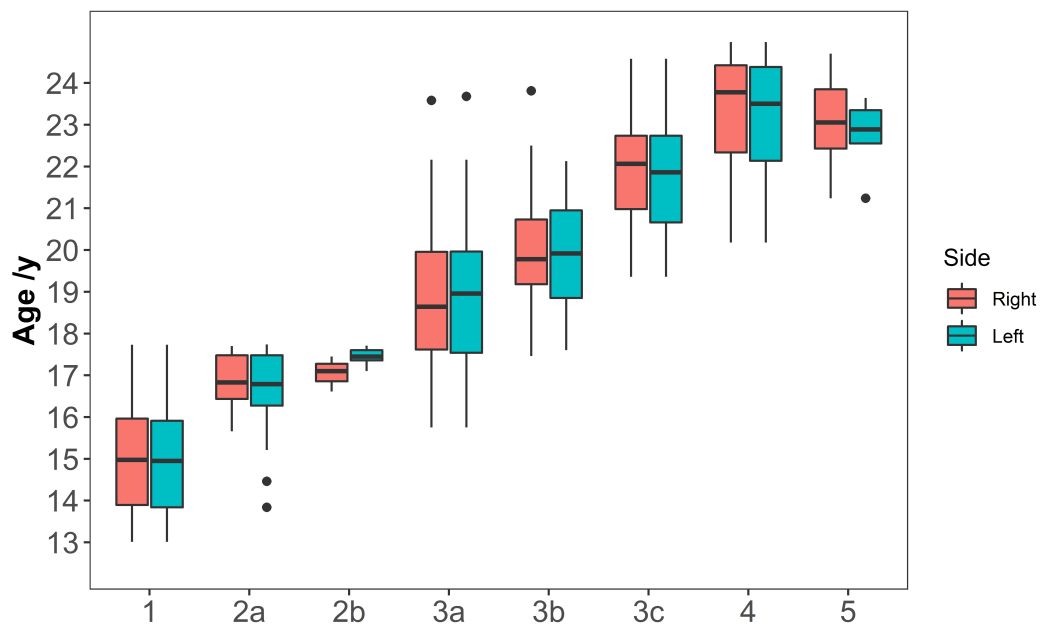


Figure 3.10.: Box and whisker plot of the clavicle ossification. The x-axis shows the stages, the y-axis displays the chronological age. Note, that stage 2c was not found in the current sample

In 288 cases both sides have been evaluated. Of these cases, 230 ($\approx 80\%$) individuals showed the same stage for the right and left side. Of the remaining 58 cases, 47 subjects showed a difference of one stage and 11 cases had a difference of two or more stages. There was no tendency for the left or right side for an accelerated or retarded development. An applied paired Wilcoxon test showed no significant results (significance level $p < 0.05$).

The Fleiss' Kappa regarding the inter-rater reliability of the three raters was moderate (168) with $\kappa = 0.43$ ($p < 0.05$) including also non-staged cases. The individual agreement (only staged cases) between two raters was calculated with weighted Cohen's Kappa and varied between moderate and good (see Table 3.15).

For the transition analysis models different approaches (Table 2.5 see page 31) were used, and the results of the cross-validation can be seen in Table 3.16. In the TA, the final stages have no MLAG, so these MLAG values were set to the minimum and maximum value of the sample, respectively, for the calculation.

Table 3.14.: Descriptive statistical data (in years) for the individual ossification stages of the medial end of the clavicle (separated for the right side, the left side and for both sides together).

Stage		Right	Left	Right+Left	Stage		Right	Left	Right+Left	Stage		Right	Left	Right+Left
1	N	86	81	167	3a	N	88	78	166	4	N	38	46	84
	Mean	15,04	14,98	15,01		Mean	18,82	18,84	18,83		Mean	23,33	23,16	23,24
	SD	1,29	1,31	1,3		SD	1,57	1,54	1,55		SD	1,41	1,37	1,38
	MIN	13,01	13,01	13,01		MIN	15,75	15,75	15,75		MIN	20,18	20,18	20,18
	LQ	13,89	13,84	13,86		LQ	17,62	17,54	17,61		LQ	22,34	22,13	22,18
	Median	14,97	14,95	14,95		Median	18,64	18,96	18,77		Median	23,78	23,51	23,71
	UQ	15,96	15,91	15,95		UQ	19,95	19,96	19,96		UQ	24,42	24,38	24,41
	MAX	17,73	17,73	17,73		MAX	23,58	23,68	23,68		MAX	24,98	24,98	24,98
2a	N	12	16	28	3b	N	23	17	40	5	N	10	5	15
	Mean	16,92	16,58	16,72		Mean	20,14	19,89	20,03		Mean	23,01	22,73	22,92
	SD	0,63	1,18	0,98		SD	1,47	1,36	1,41		SD	1,11	0,93	1,03
	MIN	15,66	13,84	13,84		MIN	17,46	17,6	17,46		MIN	21,24	21,24	21,24
	LQ	16,44	16,28	16,42		LQ	19,18	18,85	18,95		LQ	22,43	22,55	22,47
	Median	16,83	16,79	16,81		Median	19,78	19,92	19,84		Median	23,05	22,89	22,89
	UQ	17,48	17,48	17,48		UQ	20,73	20,95	20,84		UQ	23,85	23,35	23,64
	MAX	17,7	17,74	17,74		MAX	23,81	22,13	23,81		MAX	24,7	23,64	24,7
2b	N	3	7	10	3c	N	44	48	92					
	Mean	17,05	17,45	17,33		Mean	21,96	21,82	21,89					
	SD	0,42	0,21	0,32		SD	1,27	1,42	1,34					
	MIN	16,61	17,1	16,61		MIN	19,36	19,36	19,36					
	LQ	16,86	17,36	17,16		LQ	20,98	20,66	20,89					
	Median	17,1	17,45	17,41		Median	22,06	21,86	21,99					
	UQ	17,27	17,6	17,53		UQ	22,73	22,73	22,73					
	MAX	17,45	17,71	17,71		MAX	24,58	24,58	24,58					
2c	N	0	0	0										
SD standard deviation		MIN minimum	MAX maximum	LO lower quartile	UQ upper quartile									

SD standard deviation MIN minimum MAX maximum LQ lower quartile UQ upper quartile

Table 3.15.: Inter-rater agreement between the individual raters. Only staged cases are included.

	N	κ_w	p-value
R1R2	573	0.54	<0.05
R1R3	582	0.75	<0.05
R2R3	561	0.64	<0.05
R rater			
N count			
κ_w weighted Kappa			

Table 3.16.: Cross-validated performance of different model approaches (in years).

Model	Test set	MAE	RMSE
Clavicula right	Clavicula right	1.50	1.86
Clavicula left	Clavicula left	1.53	1.90
Whole set ^a	Whole set	1.51	1.88
Clavicula right	Clavicula left	1.53	1.91
Clavicula left	Clavicula right	1.50	1.88
Clavicula right	Half set max ^b	1.49	1.86
Clavicula left	Half set max	1.49	1.85
Clavicula right	Half set min ^c	1.53	1.92
Clavicula left	Half set min	1.54	1.93
Half set max	Half set max	1.49	1.84
Half set min	Half set min	1.52	1.91

^a Whole set: merged dataset of right and left clavicula stages

^{b,c} Half set max/min: Maximal/Minimal stage of both sides have been taken

MAE mean absolute error

RMSE root mean squared error

Although no model seems to outperform the other, the nominal best model after cross-validation ("Half set max") was chosen for a more detailed investigation. For this purpose all data where at least one side was staged were included. The highest reached stage (max. stage) was chosen resulting in 314 cases. The results can be seen in Table 3.17 a, b and in Figure 3.11.

Table 3.17.: Extended results from the evaluation of the model using the max. stages of the clavicles. Note that the lower bound of the prediction interval in the lowest stage reflects the minimum age in the study sample. Similarly, the upper bound in the highest stage corresponds to the maximum age in the study sample.

(a)		(b)		
N	314	Stage	MLAGE	95% prediction interval TA
MAE	1.482 y	1	–	13.00 - 17.43
RMSE	1.852 y	2 (abc)	16.59	14.50 - 18.99
LM test	$p = 0.948$	3a	18.35	15.62 - 21.55
$R^2_{Cragg\&Uhlers}$	0.839	3b	20.22	17.68 - 23.11
Accuracy	88.5%	3c	21.56	18.70 - 24.86
Specificity	73.7%	4/5	–	20.97 - 25.00
Sensitivity	100%	in years		
MAE mean absolute error		MLAGE maximum likelihood of age		
RMSE root mean squared error		TA transition analysis		
LM test Lagrange multiplier test				

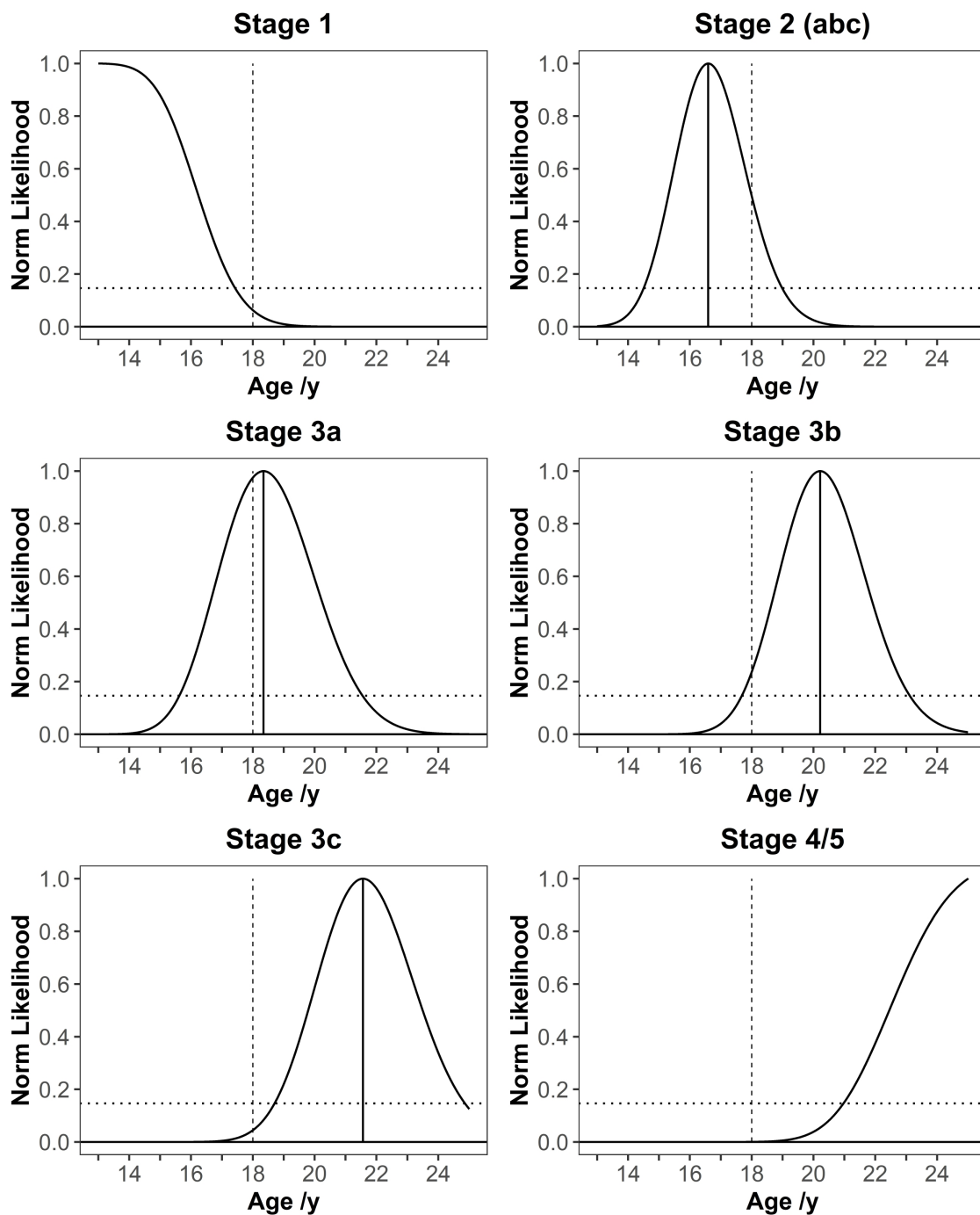


Figure 3.11.: Normed maximum likelihood curves (max. stages model). The solid vertical line represents the maximum likelihood of age (MLAGe), the dashed vertical line displays the age of 18. The horizontal line shows the likelihood of 0.1465 indicating the 95% prediction interval.

3.4. Multi-factorial age estimation (MFA): Results

Different models (combinations) were tested. The included subjects per model varied due to the exclusion of data for different reasons such as agenesis, MR-related artifacts or anatomical norm variants. Table 3.18 shows the results of the calculations including MAE, RMSE, accuracy, sensitivity, specificity, discrimination slope and the mean prediction interval. In the calculations, in cases where all levels of each characteristic were in the first or last level, the MLAGe was set to the minimum and maximum ages of the sample, respectively. Also, the smallest and largest limits of the PI were set to the minimum and maximum ages of the sample, respectively.

Table 3.18.: Performance of the different models using different combinations of age indicators.

Combination	N	MAE ^a	RMSE ^a	PI mean ^a	PI SD ^a	PI MIN ^a	PI MAX ^a	Accuracy ^b	Sensitivity ^b	Specificity ^b	DIS-SLOPE	LM test ^c
H-Cmax-UJmax-LJmax	261	0.924	1.180	3.96	0.82	1.81	5.01	92	97	85	0.749	ok
H-Cmin-UJmin-LJmin	261	0.878	1.122	3.87	0.79	1.87	4.92	93	99	86	0.767	signif.
H-Cmax-Tmax	295	1.001	1.275	4.11	0.83	1.98	5.14	91	98	81	0.733	ok
H-Cmin-Tmin	295	0.934	1.207	4.09	0.82	2.16	5.09	93	96	88	0.748	signif.
H-Cmax	313	1.138	1.451	4.42	0.89	2.57	5.41	91	95	85	0.728	ok
H-Cmin	313	1.089	1.408	4.36	0.86	2.65	5.38	92	95	88	0.747	ok
H-UJmax-LJmax	283	1.318	1.825	5.01	1.25	1.88	6.68	89	92	87	0.697	ok
H-UJmin-LJmin	283	1.251	1.647	5.01	1.29	1.92	6.68	91	96	86	0.696	ok
H-Tmax	319	1.58	2.196	5.43	1.30	2.1	7.09	90	94	84	0.663	ok
H-Tmin	319	1.355	1.744	5.29	1.30	2.23	6.91	88	96	79	0.67	ok
Cmax-UJmax-LJmax	261	0.994	1.256	4.20	0.78	2.32	5.18	91	94	86	0.729	ok
Cmin-UJmin-LJmin	261	0.981	1.221	4.15	0.75	2.56	5.10	90	94	85	0.747	signif.
Cmax-Tmax	295	1.051	1.348	4.31	0.80	2.47	5.24	90	95	84	0.717	ok
Cmin-Tmin	295	1.066	1.348	4.38	0.77	2.91	5.34	90	93	86	0.728	signif.
H	337	2.467	3.271	6.06	1.31	2.83	6.89	89	95	81	0.633	ok
2C	287	1.418	1.795	4.61	0.77	3.52	5.52	91	100	80	0.721	ok
4T	225	1.346	1.733	5.61	1.38	2.45	7.31	85	86	83	0.608	ok

^a in years ^b Percentage ^c signif. = not all traits have a good fit ($p < 0.05$)

MAE mean absolute error RMSE root mean squared error PI prediction interval DIS-SLOPE Discrimination slope LM Lagrange multiplier test

SD standard deviation, MIN minimum, MAX maximum

H hand/wrist C clavicle UJ upper jaw third molars LJ lower jaw third molars T third molars

In general, it can be said that the performance of the model increases when more age indicators are included. In addition to the visual comparison of the results, the absolute errors, absolute difference between chronological age and MLAGe, of the individual models were statistically compared using the Wilcoxon signed-rank test. Table 3.19 shows a part of these tests (comparison of the H-Cmax-UJmax-LJmax model with the other models). Summarizing all results including

the "Goodness-of-the-fit" test, the H-Cmax-UJmax-LJmax model was seen as the best choice and, therefore, the following illustrations are based on this model.

Fig. 3.12 presents a box plot of the predicted probability to be a minor if you are a minor or adult with the individual age indicators using the H-Cmax-UJmax-LJmax model. The difference of the box plot mean values is the discrimination slope. The higher the difference, the better the discrimination is between minors and adults. The coverage of the PI of this model was also evaluated; and showed that the percentage of the chronological age of the subjects that is within the predicted intervals was approximately 95% (0.946%), which means a very good coverage. In addition, the applicability of the model was also assessed by a ten-fold cross-validation. The results with $MAE_{CV} = 0.942y$ and $RMSE_{CV} = 1.202y$ were quite similar to the overall result and the applied LM test showed a good fit for all folds' traits.

Table 3.19.: Comparison of the absolute errors (absolute difference between chronological age and MLAG) of the H-Cmax-UJmax-LJmax model with the other models.

Combination 1	MAE 1 ^a	SD _{MAE} 1 ^a	Combination 2	MAE 2 ^a	SD _{MAE} 2 ^a	p.value ^{bc}	N ^d
H-Cmax-UJmax-LJmax	0,924	0.74	H-Cmin-UJmin-LJmin	0,878	0.70	0,188	261
H-Cmax-UJmax-LJmax	0,924	0.74	H-Cmax-Tmax	0,981	0.78	0,028*	261
H-Cmax-UJmax-LJmax	0,924	0.74	H-Cmin-Tmin	0,906	0.74	0,859	261
H-Cmax-UJmax-LJmax	0,924	0.74	Cmin-UJmin-LJmin	0,981	0.73	0,137	261
H-Cmax-UJmax-LJmax	0,924	0.74	Cmax-UJmax-LJmax	0,994	0.77	0,007*	261
H-Cmax-UJmax-LJmax	0,924	0.74	Cmax-Tmax	1,030	0.84	0,006*	261
H-Cmax-UJmax-LJmax	0,924	0.74	Cmin-Tmin	1,049	0.81	0,003*	261
H-Cmax-UJmax-LJmax	0,924	0.74	H-Cmin	1,050	0.85	0,002*	261
H-Cmax-UJmax-LJmax	0,924	0.74	H-Cmax	1,116	0.85	<0,001*	261
H-Cmax-UJmax-LJmax	0,924	0.74	H-UJmin-LJmin	1,234	1.06	<0,001*	261
H-Cmax-UJmax-LJmax	0,924	0.74	H-UJmax-LJmax	1,308	1.27	<0,001*	261
H-Cmax-UJmax-LJmax	0,863	0.74	4T	1,318	1.06	<0,001*	207
H-Cmax-UJmax-LJmax	0,924	0.74	H-Tmin	1,307	1.04	<0,001*	261
H-Cmax-UJmax-LJmax	0,924	0.74	H-Tmax	1,556	1.52	<0,001*	261
H-Cmax-UJmax-LJmax	0,915	0.74	2C	1,391	1.08	<0,001*	240
H-Cmax-UJmax-LJmax	0,924	0.74	H	2,497	2.20	<0,001*	261

^a in years ^b p.value from the Wilcoxon signed-rank test

^c * significant (p.value <0.05) ^d number of included cases

N count MAE mean absolute error SD standard deviation

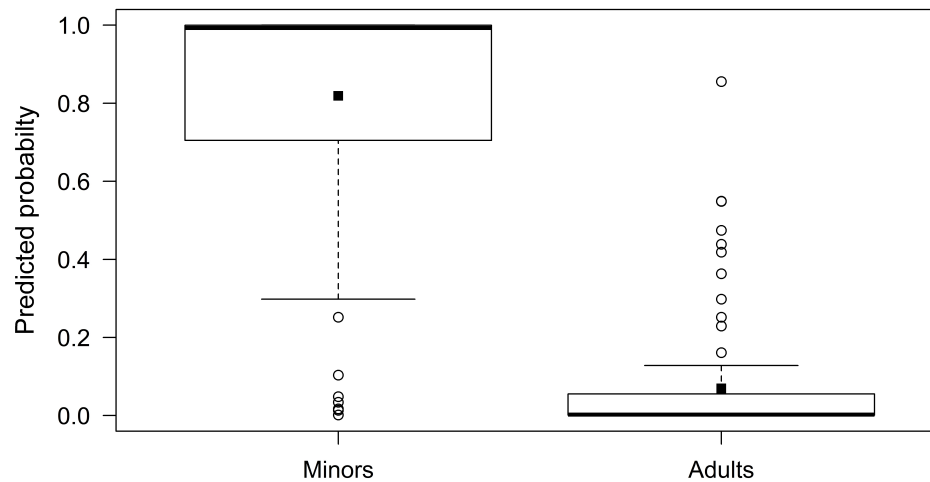
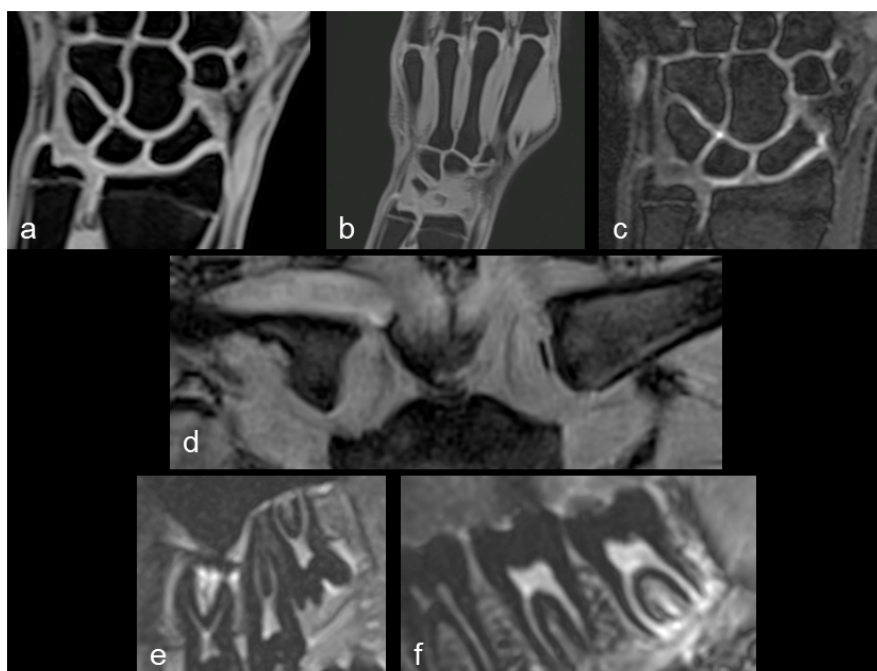
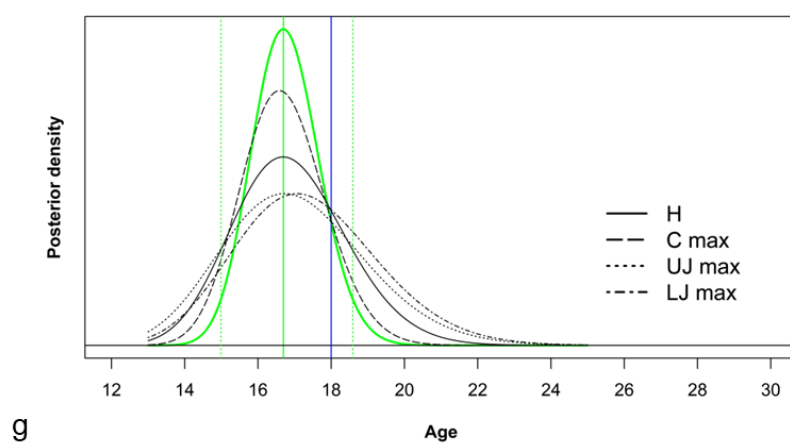


Figure 3.12.: Box plot of the predicted probability to be a minor (used age indicators: Hand, max. clavicle stage, max. upper and max. lower jaw stage). The black square shows the mean predicted probability. The difference between the means is the discrimination slope (= 0.75).

Figs. 3.13-3.15 show three examples of the MFA approach using the H-Cmax-UJmax-LJmax model. In the figures the staged anatomical regions and in the diagrams the MFA (green curve) and the individual traits (black curves) can be seen. Additionally, MLAG (green line), PI (green dotted lines) and 18 years threshold (blue line) are displayed in the plots. At the end of each figure there is a table with the results of the single region evaluation and the MFA evaluation. These include the point prediction, the prediction interval and the probability of being over 18 years old. Fig. 3.15 shows a good example why the application of a multi-factorial approach is superior over a single site approach. The hand evaluation showed a non complete development of the radius. It is common practice (at least in Austria) to stop the age estimation procedure if the hand development is not finished. In the present case, this would have lead to the person being classified as a minor, since the uncertainty (prediction interval) has to be taken into account. By using the MFA approach the individual was classified as an adult with a very high probability (>99.8%). Among the reasons to stop the age estimation process when hand development has not yet been completed is high exposure to ionizing radiation from OPG and CT. Therefore, an MFA approach with MRI is a very good alternative.



H = 29 , C max = 2a , UJ max = F , LJ max = F



g

	Hand	C max	UJ max	LJ max	MFA
MLAge	16.69	16.59	16.72	17.07	16.70
LL PI	13.95	14.50	13.46	13.76	15.00
UL PI	19.96	18.98	20.78	21.17	18.59
P ≥ 18a	0.23	0.13	0.30	0.36	0.094

h

Figure 3.13.: Example for an MFA age estimation approach of a 16.73y old male individual.

a,b,c MRI of Hand and Wrist (VIBE, DESS): GPS 29.

d MRI of the sterno-clavicular region (VIBE): stage 2a.

e,f MRI of the third molars (TSE): tooth 28: stage F, tooth 38: stage F.

g Posterior density plot of the single region (black curves) and the MFA (green curve) age estimation.

h Table of the results of the age estimation calculation.

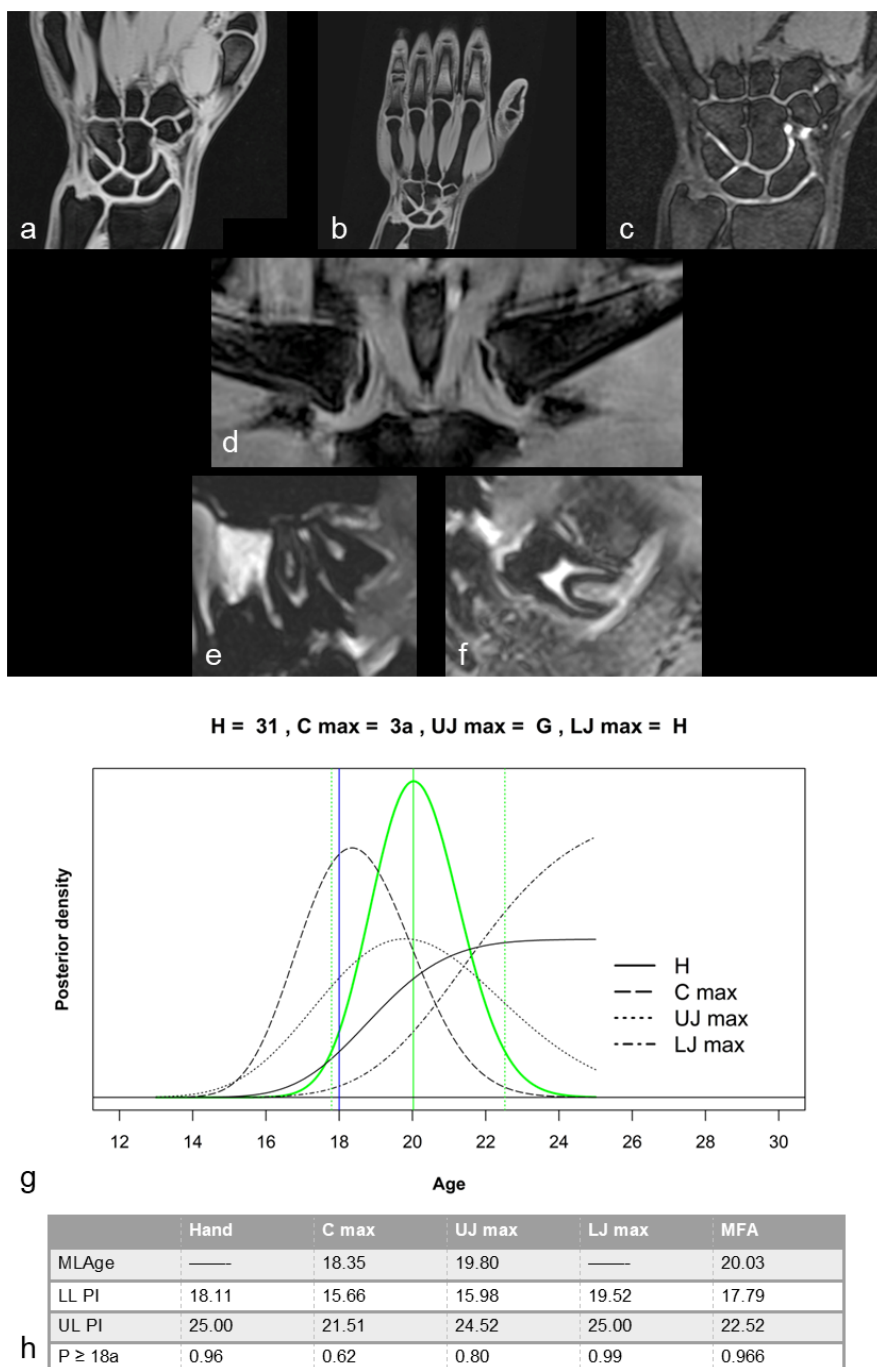


Figure 3.14.: Example for an MFA age estimation approach of a 19.95y old male individual.

a,b,c MRI of Hand and Wrist (VIBE, DESS): GPS 31.

d MRI of the sterno-clavicular region (VIBE): stage 3a.

e,f MRI of the third molars (TSE): tooth 28: stage G, tooth 38: stage H.

g Posterior density plot of the single region (black curves) and the MFA (green curve) age estimation.

h Table of the results of the age estimation calculation.

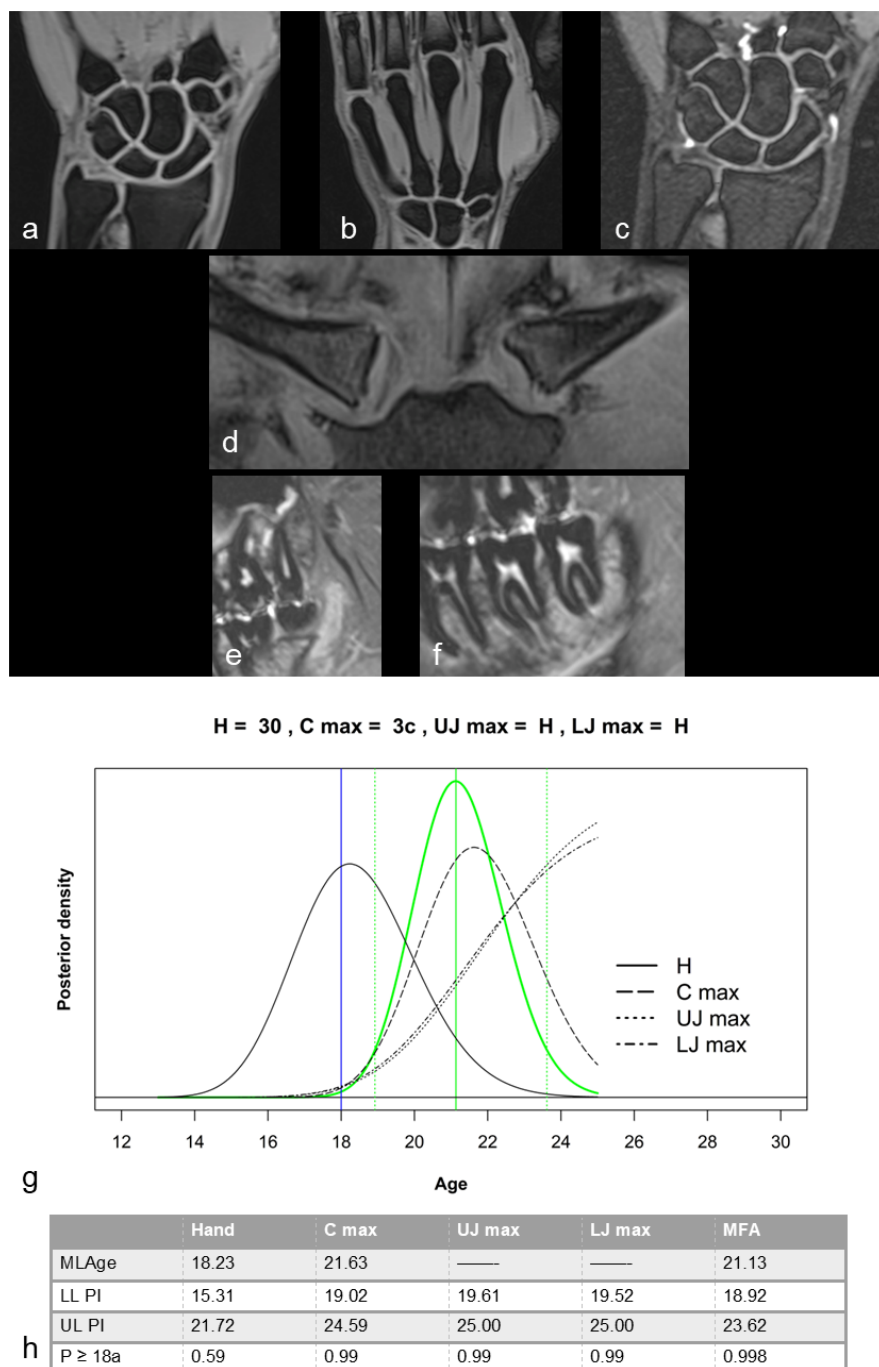


Figure 3.15.: Example for an MFA age estimation approach of a 22.06y old male individual.

a,b,c MRI of Hand and Wrist (VIBE, DESS): GPS 30.

d MRI of the sterno-clavicular region (VIBE): stage 3c.

e,f MRI of the third molars (TSE): tooth 28: stage H, tooth 38: stage H.

g Posterior density plot of the single region (black curves) and the MFA (green curve) age estimation.

h Table of the results of the age estimation calculation.

An evaluation of all subjects included in this study ($n = 337$) with different models, depending on available age indicators, showed also good results and are presented in Table 3.20 a and b. 126 out of 146 minors were classified correctly with respect to the point prediction. Taking into account the lower limit of the prediction interval, only 3 minors remained incorrectly classified as adults. Of 191 adults only 10 were classified as minors. However, when further evaluating the 181 correctly classified adults, 47 had a lower prediction interval limit under the age of 18 years.

Table 3.20.: Evaluation of all study participants with the corresponding model.

(a) Individuals per model

Model	N
H-Cmax-UJmax-LJmax	261
H-Cmax-Tmax	34
H-Cmax	18
H-UJmax-LJmax	22
H-Tmax	2
Sum	337
N count	

(b) Results including all study participants

N	337
MAE	0.995y
SD _{MAE}	0.82y
Accuracy	91%
Sensitivity	95%
Specificity	86%
N count	
MAE mean absolute error	
SD standard deviation	

4. Discussion

The discussion chapter uses parts from previously published manuscripts resulting from this thesis (1, 2).

In this chapter the individual sections of the results chapter are discussed individually.

4.1. Third Molars

Radiological evaluation of dental mineralization and eruption is done until now by using radiological methods based on ionizing radiation. The aim of this part of the project was to investigate the chronological course of the dental development by means of MRI and the comparison to currently used reference values, which are based on X-ray studies.

Image quality was judged in general as good, nevertheless, some teeth had to be excluded. MR-specific were all exclusions caused by metal artifacts. These exclusions affected predominantly the mineralization evaluation as the analysis of the teeth's roots was more affected by these artifacts. In contrast, exclusions due to stage definition problems occurred primarily in the eruption evaluation. This was caused by tilting of the teeth and appeared predominately in the lower jaw. However, this is not specific for MRI.

4.1.1. Mineralization

About half of the participants showed different mineralization stages if all four third molars were present. Nevertheless, the statistics did not show any significant differences between all four third molars. This confirms the findings of X-ray based studies (59, 61, 181–183). However, this also means that there is no evidence that any particular wisdom tooth is systematically accelerated or decelerated in terms of its development. It is therefore initially suggested to include all third molars in the assessment of dental age, in order to overcome the variations in the development of the individual teeth.

Since the age of 18 years is an important legal limit in many countries, it would be of great benefit to have a cut-off stage for this age. The empirical probability to be over 18 years when reaching the final stage H resulted in a probability between 98 – 100% for the four third molars. Therefore, it can be concluded that a male person of Central European origin is most likely to be at least 18 years old when one of the four third molars reached the last degree of mineralization (stage H), which is in accordance with Meinel et al. (181), who stated a value of 99.5%. For other legally important age limits, it could be shown that there is a high probability of being over 14 or 16 years old if stage F or G, respectively, have been reached.

There is a strong dependence of the terminal stage's mean and median values on the upper limit of the age sample. Therefore, it might be better to use a different indicator, the age at the 50% probability for this stage (184). In the current study, these age values lay over 22 years of age for all four third molars, which is in comparison to studies with a Swiss (149) and a German cohort (63) up to three years higher. The difficulty in differentiating between stages G and H described by Mincer et al. (59) for conventional OPG could explain this large difference, as this was also determined by the evaluating dentists in this MR study. De Tobel et al. (133) did a comparison study of their dental MR protocol and the one used in this study regarding the assessment of apical closure (stage G/H). Although they could not exactly reproduce this study's protocol due to differing MR coils they concluded that age assessment with MRI needs sequence specific reference data.

So far only one other European MR study on dental age estimation (134) provides reference values. Nevertheless, a comparison to the current study is only conditionally possible as the study by Guo et al. (134) used only teeth from the lower jaw without differing between right and left. However, similar results with an average difference of four months were seen when comparing Guo's results to the averaged mean values of the two lower third molars of this study. The standard deviations were also similar. In the current study, the first appearance of the stages C–E (minimal values) showed slightly higher values, that could be explained by the higher minimum age of the cohort. For stages F–H only minor differences were observed.

Numerous X-ray studies investigated the course of mineralization of the third molars. However, they differ in methodology, sample size and ethnicity. Therefore, only two studies were included in the comparison with the current study. Both, the study conducted by Meinl et al. (181) (only mandibular third molars) and the study performed by Olze et al. (182) used the same classification system and investigated a comparable Central European sample.

Good correlation between the studies was found for the maxillary values in stages F and G, and for all teeth in stage H. For stage D a mere advanced mineralization (mean values) of up to two years was seen in this study. With regard to the German sample, the explanation can be found in the small number of subjects with stage D in the current study, but this is not a valid explanation for the study carried out by Meinl et al. (181). Gelbrich et al. (185), however, criticized in particular the high average age of stage D in various reference studies and explained it mainly with the used lower limit of the age range. However, the lower limit used in the comparative studies roughly corresponds to the current study. Stage E showed a good agreement with Meinl's study but an advance in mineralization of up to eleven month in comparison to the German cohort.

The minimum age of the cohort in the current study is the main explanation for the low frequency of stages A–C. The mean values of stages A–D are approximately the same and can be attributed to the low number of subjects in these stages.

4.1.2. Eruption

Similar to the mineralization, partly different eruption stages were seen in participants with all four third molars, but no tendency for accelerated or decelerated dentition for a particular third molar was seen. In the statistics of these subjects, significant differences were found between the maxillary and mandibular teeth. A possible explanation is the different morphology in the upper and lower jaw.

There was found a high empirical probability (100%) to be over 14 or 16 years of age, when reaching stage B or C, respectively. However, the probability to be over 18 years when reaching stage D was only 90% for all four third molars, which is not high enough for court proceedings. In comparison with the study conducted by Caldas et al. (186), the values are slightly lower.

As mentioned in the previous section, the final stage's values are dependent on the sample age range. The age values calculated for this reason with a probability of 50% for stage D were over 22 years, with the exception of tooth 48, where it was approximately 21 years. Compared to a Portuguese cohort (186), the values of teeth 18, 28 and 38 were about one year higher in the current study.

Up to date no other MR studies were conducted on third molars eruption, except the pilot study of Baumann et al. (129). X-ray studies regarding dental eruption are equally rare as well. These studies were mainly performed by Olze and co-workers (62, 147, 187–189) who investigated samples from different regions and ethnicities.

In comparison to X-ray studies with a Central European cohort (Olze et al. (187) (German sample) and Brkic et al. (190) (Croatian sample)) there was not a good agreement. The reasons for this can be mainly explained by the different age range and distribution of the investigated cohorts. However, dental eruption has a higher variance compared to mineralization, and there might be a need for population-specific data as stated by some authors (186, 190).

4.2. Hand and Wrist

Alshamrani et al. (43) and Schmidt et al. (44) determined the GP atlas method as one of the most used approaches to assess bone age on hand X-ray images. In two pilot studies, Hojreh et al. (108) and Urschler et al. (114) recently showed that it is possible to use the GP atlas on MR hand and wrist images. These two studies were the basis for this study part to provide male reference values for the age estimation of hand bones using MRI. No participant had to be excluded due to the good image quality.

Compared to an X-ray of the hand and wrist, the acquisition time for the MR scan is quite long. However, recent MRI research provided techniques to shorten MR sequences dramatically. This was shown in a pilot study by Neumayer et al. (191), where the VIBE Sequence used in the current study was reduced down to 15 seconds with promising results regarding age estimation. This approach will eliminate two major deficits of MR scans with a long acquisition time: artifacts due to patient movement and the high cost of long MR examinations. However, the study conducted by Neumayer et al. (191) used only a small sample and has to be validated in a larger cohort and with other MR scanners and field strengths.

The estimate of the bone age using the GP atlas from hand MRIs is reliable and reproducible, which is shown by the good agreement between and within the raters. This was also observed in the two previously mentioned pilot studies. The cross-tabulation (Table 3.12) also confirms the good agreement, showing no major discrepancies between the raters. The mean difference between reader one and reader two was less than two months. The calculation of the mean relative disagreement (mean difference of the non-matches) resulted in less than half a year. Both values underline the good agreement between the two evaluators and the reliability of the method.

Large variation could be seen in the distribution of the GP standards (Fig. 3.5) particularly in subjects under 17 years of age. All subjects older than 18 years had at least a GPS of 28. However, this cannot be used as a cut-off value as there are also younger subjects who already reached this development stage. Conversely, no one with a GPS less than or equal to 27 was

shown to be older than 16 years, which appears to be a strong indicator of minority. Nevertheless, due to the large variation, it cannot be recommended to use hand bones as the only source for the forensic age assessment procedure to distinguish minors from adults, at least not from a GPS of 28.

Both readers underestimated the subjects by about half a year on average compared to their chronological age. The main explanation is the upper age limit (20.99 years), which far exceeds the last age category of 19 years, and thus promotes an underestimation. However, an underestimation is more acceptable in the context of forensic age estimation in the living than an overestimation, since this is in favor of the person being evaluated.

Age mimicry and small sample sizes have always been an issue with age estimation. To circumvent these problems, paleodemographers have introduced a method called transition analysis (160), which was recently also transferred into the age estimation procedure of living adolescents (103, 162, 163, 172, 192). One approach for this method is to use a cumulative probit model for the fitting of the transition curves. In the current study, the applicability of the model was demonstrated by means of a Lagrange multiplier tests. For the age estimation, the cross-validation resulted in a mean error of about plus three months, which is mainly explained by the estimated (set) age of 21 years for the final GPS 31.

No significant differences with the exception of the study's terminal stages were found in the comparison of the mean and predictive values of the plain descriptive statistics and the transition analysis. These differences are, however, due to the methodologically conditioned MLAG values of the first and last group of the transition analysis, which correspond to the minimum and maximum values of the sample's age range. As a result, neither of the two methods, transition analysis and descriptive statistics, respectively, seems to be superior to the other. This was expected due to the non-informative prior and the upper limit of the age range. However, if the sample would include older individuals (≥ 21 y), there would be a stronger age mimicry effect and the strength of the transition analysis would be more prominent. Nevertheless, the maximum likelihood estimation (MLEST) of the transition analysis generally showed somewhat narrower

prediction intervals which, taking the sample size into account, could correspond more to the actual distribution.

In a comparison of the current study with three recent European GP X-ray studies (45, 164, 165) with at least similar or larger cohorts the following differences were found. The X-ray studies showed a mean difference between the chronological age and the estimated age of 1.6, 2.16 and -1.45 months, respectively. In the MR study the descriptive analysis showed a mean difference of about six months, which is slightly higher but is in favor of the examined person. This underestimation of the MR examinations compared to the X-ray examinations was also seen by Urschler et al. (114). One possible explanation is that due to the projective nature of X-ray images, there are superimpositions that can give the impression that a higher stage has already been reached. In comparison, MR images are free of superimpositions and a clearer view of the epiphyseal plate is possible (114).

The mean error of the transition model of the MR study, however, showed an overestimation of about 3 months. This is contradictory but mainly due to high MLAG value set for the last stage. Therefore, the before mentioned superimposition explanation might be still valid.

Both, the descriptive statistic values and the TA values were compared in more detail with a very recent big X-ray study of Chaumoitre et al. (45). The descriptive mean values of the current study are all higher (0.06–0.72 years). However, significant differences were only found for GPS 24, 29 and 30, and all values of the current study agree well with the GP age categories. When comparing the Chaumoitre study and the values obtained by the transition model, the result was different, with differences ranging from -0.41 to 0.36 years, except for GPS 31, where the difference was 1.7 years. Significant differences were again found for GPS 24 and 29 but not for GPS30, instead GPS 31 showed a significant difference. However, with regard to the final stages (GPS 24, 31) of the MR study, the interpretation must be made with caution, because here the values were set artificially. Regarding GPS 29, an explanation could be the larger group size of the Chaumoitre study and the resulting smaller standard deviation, since the values differ only slightly from each other (MLAGE: 16.88y vs. mean: 16.52y). The comparison of the prediction intervals showed on average a similar range but with differences in the interval position. The

differences for the lower limits lay for the current descriptive analysis between -0.38 and 0.94 years (median: $0.02y$) and for the study's TA model between $0.01 - 1.56$ years (median: $0.25y$). The upper limits of the descriptive analysis were always higher than the Chaumoitre study values ($0.03 - 1.5$ years, median: 0.39 years), in contrast, the differences in the upper limits of the TA model were between -0.88 and 0.93 years (median: 0.21 years).

The results show that MRI can be used to investigate age in adolescent males with similar variability to X-ray exams, but with the advantage of not using ionizing radiation.

4.3. Clavicular

The examination of the medial end of the clavicles for the purpose of age estimation in the living is currently carried out using ionizing radiation with either CT or plain X-ray. Recently, a few groups have started to use MRI as an alternative modality (118, 120–123, 146, 193). The current study also used MRI and a 5-stage classification with subdivisions of stage 2 and 3 according to (83, 86). About 8% of the clavicles had to be excluded due to anatomical shape variants and about 3% due to motion artefacts. The total exclusions are similar to the study of Schmidt et al. (118). The number of exclusions due to shape variants seem to be highly dependent on the observers. There was a high variance between different studies, where the numbers vary between 2.1% (122) and 21.3% (120), but there was also a high variance in the results of the individual readers in the current study. Motion related exclusions were similar to the study of De Tobel et al. (122) and Hillewig et al. (193) but only half compared to Vieth et al. (120). Since these three studies positioned the subjects in prone position and the current study used the supine position, there appears to be no advantage to using either position.

Clavicula staging seems to be more challenging than the other two regions investigated in this project. This could be seen in the moderate Fleiss' Kappa as well as in the individual weighted Kappa statistics. The results are similar to the study by De Tobel et al. (122) and their proposal to evaluate the clavicles in consensus of at least two observers is also one of the conclusions of this study.

Similar to other MR studies (120, 122, 193) stage 2a and 2b were found rarely and stage 2c was not detected at all. Also in the CT study by Wittschieber et al. (194) the number of individuals with stage 2 was low. One reason might be that stage 2 and its subdivisions are quick transition stages which are seldom found, and single bone bridges build very early in the ossification process. There might be also a chance that current MR sequences and equipment are able to detect bone bridges earlier and therefore it comes to an upstaging. However, this could not be confirmed in smaller postmortem comparative studies (195, 196) between MR and CT. Nevertheless the decision if a small bone bridge exists and is completed seems to be more difficult in MR images than in CT images according to the current study's observers. This might be due to lower resolution, possible blurring and other artifacts in the MR data. This might also be one explanation for the lower agreement of the readers.

Similar to De Tobel et al. (122) the substages 2a, b and c were merged to stage 2 in the transition analysis. De Tobel et al. (122) also indicate that the subdivision of stage 2 is not necessary for MR images, which seems plausible following this study's results. On the other hand the suggestion of the same authors to subdivide stage 3a according to the width of the epiphyseal ossification seems to be promising.

Stage 5 was found in total only 15 times for the left and right clavicle. This is in concordance with other studies (117, 118) which found similar numbers or no stage 5 at all. One reason could be the upper age limit of the cohort that was also mentioned by Schmidt et al. (118). Another reason might be the long visibility of the epiphyseal scar in the used MR sequences. Hillewig et al. (193) reported difficulties in distinguishing between stage 4 and 5, and therefore decided to use a 4 stage system. However, as also stated by Schmidt et al. (118) stage 5 might be helpful in certain legal circumstances as the minimum age of stage 5 in the current study was over 21 years of age. Notable is the fact that in the current study the mean age of stage 5 was lower than for stage 4. However, as stage 5 is a terminal stage this should not be overrated. For the transition analysis stage 4 and 5 were merged due to the low numbers of stage 5.

Both, De Tobel et al. (122) and Hillewig et al. (193) reported difficulties in distinguishing stage 1 from stage 4/5. Schmidt et al. (118) contradicted this by pointing out that the shape of the medial

clavicular ends, as described in (117), clearly allows a distinction. In the current study there were also individual cases where a clear differentiation was difficult. In these cases, it is helpful to consider the thickness of the clavicular cartilage and the shape and structure of the surrounding bones, such as the manubrium or the first rib, since there are also age-related changes, as noted by Martínez Vera et al. (197) for the manubrium.

Different constellations were used for the transition analysis. A combination of the right and left clavicle after transition analysis was not done, as over 80% showed the same stage and no significant difference was found in the applied paired Wilcoxon test. The different transition models did not show much difference in their performance. However, the model using the highest stage found in both medial clavicle ends was marginally better than the rest and more detailed results were calculated only for this constellation. The selection of the highest stage if the two clavicles showed different stages was also applied in other studies (85, 118). A Lagrange multiplier test showed the good fitting of the probit model to the data. Accuracy for minor and adult determination was over 88% in terms of point prediction, and no adult was classified as a minor. However, about one quarter of the minors were identified as adults when looking only at the MLAG. Compared to the results of a forward continuation transition analysis approach (122) the results are slightly better. Nevertheless, miss-classification of minors should always be avoided. Therefore, relying on the MLAG alone is inappropriate, and a prediction interval should always be used. When using the lower prediction interval value 100% of the minors are classified as minors. On contrary, the accuracy and the classification of the adults is consequently decreased to about 80% and to 61%, respectively.

In conclusion, MRI and a transition analysis approach were shown to be applicable for age estimation of the medial end of the clavicle. However, the staging is more demanding than in other regions and a consensus reading of at least two raters is recommended. Transition analysis seems to be a good tool to reliably classifying minors when prediction intervals are used, but the high number of misclassified adults indicates that the prediction intervals are too wide. Therefore, multi-factorial age assessment is recommended as it will narrow the prediction intervals. In addition, older subjects should be studied to obtain a better differentiation between stage 4/5.

4.4. Multi-factorial age estimation (MFA)

The evaluation of different models has shown that a multi-factorial approach including all investigated regions is superior to single region models and combinations of fewer regions. The single region models showed significant higher MAEs, higher RMSEs and wider PIs. This conclusion is in concordance with the study done by De Tobel et al. (135) Regarding the distinction of minors and adults the MFA models were better as well, particularly in comparison to the third molars-only model. However, the clavicular-only model showed a similar accuracy and superior sensitivity as the best MFA models, but had a lower specificity. The high sensitivity, which means the correct classification of an adult as adult, is mainly explained by the fact that adults in this study have at least reached stage 3a and the point prediction of this stage is over the age of 18 years. However, the point prediction (MLAGE) should not be decisive. In contrast, the prediction interval and the probability of being over 18 years are much more important.

There were no significant differences between the MAEs of models using maximum or minimum stages of the clavicles and jaws. This could be due to the similarity of the data used, as there were many individuals who had the same stages for both sides of the regions studied. This was especially the case for the clavicles. However, all of the minimum stage models suffer from an insufficient fit for single traits (regions) as was seen in the Lagrange multiplier test. Taken all factors into account the H-Cmax-UJmax-LJmax model was the best.

There are currently only two other published MFA approaches using MR images, despite the "Minimum-age-concept" of a German group (143). De Tobel et al. (135) used a comparable approach as the current study. They used as well a Bayesian approach but with a continuation ratio model and an ad-hoc correction procedure to take the violation of the conditional independence into account. There were also differences in the classification systems of the hand, clavicles and third molars. In addition, they also discarded the first and the two last clavicle stages. Nevertheless, a comparison of the male results were made showing that in general the current study had lower MAE and RMSE values. For the models including all age indicators the MAEs were 0.924y (current study) vs. 1.357y. The accuracy of the point prediction to distinguish minors

from adults was similar for both studies ($\approx 92\%$), but the correctly classified minors (specificity) was a bit lower for the current study (85% vs 90%). However, the more important prediction interval was on average about 1.5y narrower with only half the standard deviation. Possible explanations for the lower/narrower values are the larger sample size, the inclusion of all clavicle stages and maybe a better fit of the model to the data, but this needs to be validated with an independent sample.

The second MFA study was done by Stern et al. (139), who used 322 individuals from this study's cohort. Their approach was a fully automated evaluation process for the assessment of the MR images, the fusion of the regions and the calculation of a point prediction using a deep convolutional neural network. Their aim was to avoid subjectivity in the staging process and the use of discrete stages. Additionally, they tried different fusion models to get a good overall age prediction. Currently, they only provide point predictions and no PI which is of high importance in the forensic age estimation. The comparison of the MAE of the two studies showed rather similar values ($0.92y \pm 0.74$ vs. $1.01y \pm 0.74$) with a slightly better value for the current study when using all age indicators (H-Cmax-UJmax-LJmax model). However, the automated approach did also include individuals, in which not all regions could be assessed in the conventional evaluation. The MAE of this study using individual models for 337 subjects was with $1.00y \pm 0.82$ nearly equal to Stern's study. The comparison of correctly classified minors at a threshold of 18 years showed a slightly better value (86% vs. 82%) for the current study, but for the correctly classified adults it was the other way round (95% vs. 97%).

In the evaluation of all individuals it was shown that including the PI in the minor/adult classification the results became better for the minors but worse for the adults. This reflects the biological variation in age indicators, especially in the range between the ages of 17 and 20, where most of the misclassified people were. However, introducing a threshold for the probability to be over 18 years would maybe improve the results.

In conclusion, it could be shown that a MFA with the current approach is superior over single region assessments. The approach showed similar or slightly better results than other approaches

and despite some limitations , which will be shown in the next section, it is applicable in routine.

5. Limitations

Parts of the limitations chapter are taken from previously published manuscripts (1, 2).

The limitations of this project are the same for all parts and will therefore be discussed together. A major limitation is the focus on male subjects only. However, there were reasons for this decision. In practice, a very large proportion of individuals undergoing forensic age assessment are male, which makes it more urgent to optimize age determination methods for males first. Another reason to focus on male subjects was the limited resources, both in terms of time and finances.

Other limitations are the chosen age range. The lower limit (13y) was taken for two reasons: the practical forensic needs and the MRI practicability. The forensic needs are the ability to differentiate people regarding specific age limits. This is mainly the age of 18, but 14, 16 and 21 years can also be relevant, at least in some countries. MRI practicability primarily relates to the higher likelihood of seeing motion artifacts in younger children. This is mostly expected in MR sequences with longer acquisition times. However, in the current study we did not see severe motion artifacts due to longer acquisition times. Nevertheless, it must be noted that the stages seen (number and type) would have changed with both younger and older study participants, but hardly the overall result.

As always, a larger sample would have been beneficial, but the current size is due to limited resources. The distribution of the sample regarding age could have been more balanced. However, especially the recruitment of younger individuals was very challenging and led to the current distribution. To circumvent these limitations to a certain amount (except the male only cohort),

transition analysis was implemented in most parts of the project.

Two more aspects should be mentioned. First the socio-economic status of the study participants was not documented; only weight and height were reported. There are studies that suggest that the socio-economic status has an impact on bone development. However, the economic level of the country of residence of the study participants is in general high, so no influence that would be higher than the normal variation was expected.

Another remark should be made to the sequences that have been used in the project. With the exception of the 3D PDW TSE sequence used for the third molars, all sequences are frequently used sequences, even if they are not all applied by default in the regions examined. Therefore, the sequence-specific experience of the readers can influence the results of the evaluation. Additionally, it should be mentioned that MR is still a new modality for most dentists. However, according to the AGFAD guidelines (3) the evaluation of a forensic age assessment should be done by specialists who are trained for the specific exams.

Further, it should be noted that different MR sequences could lead to differing evaluation results. Therefore, care should be taken when implementing the reference values and evaluations of the current project, if one is not using similar MR parameters for the image acquisition.

6. Conclusion and Outlook

The work documented in this thesis describes a forensic age estimation project in the living using MRI, which can be divided into four parts. The project investigated three regions (hand and wrist, clavicles and third molars) using a 3T MR scanner. The first three parts evaluated the single regions and the last part combined the single regions into a multi-factorial age estimation (MFA) procedure.

Analysis of the individual regions showed that data acquisition of bones and teeth by MRI can be used to assess the developmental status of male adolescents and young adults. All three regions can be assessed with at least the same variability and accuracy as radiographic examinations using established X-ray based classification systems without the use of ionizing radiation. In addition, it should be noted that, in contrast to radiographic examinations, all three regions can be examined in one session with one device. Therefore, hypothesis one and two are confirmed.

However, regarding the evaluation of the clavicles the reliability was not as high as expected and therefore a consensus reading of at least two observers is recommended. In the evaluation of the clavicle stages it seems also to be helpful to include the surrounding tissue in the staging process. Although the reliability was high for the other two regions, it seems as well beneficial to do a consensus reading, to circumvent possible errors.

The third hypothesis was also confirmed as a good prediction model was found for the combination of the three regions. Studies regarding MFA using MRI are scarce and only two other projects have currently published their approaches. In our study we used a simple but testable

model for the fitting of the single regions data. Primarily, a decision was made to only include maximal and minimal stages based on the results of the single regions evaluation. Additionally, some stages were collapsed due to a low count of some stages. A comparison to the other studies showed similar or slightly better results, although this was limited due to different approaches.

There are limitations in this work as discussed in the previous chapter. A major concern is the restriction to Caucasian male individuals only. However, even though no conclusion can be drawn on other groups, the project showed that the applied methodology works in principle well and can be used to investigate other groups such as other ethnic groups or females.

One point should be emphasized here as the AGFAD recommends the use of modality-specific reference values. When using MRI this should even be expanded, i.e. not only MR reference values, but also the same or almost similar sequence parameters of the reference publication should be used, since the image impression can change dramatically with different parameters and incorrect classifications can occur.

Before concluding this work with an outlook some remarks regarding MR exams must be made. MRI is generally considered slow and expensive compared to other modalities. There are also contraindications for performing an MR examination. Without going into more detail here, there are of course numerous advantages of the MRI such as e.g. radiation-free imaging or cartilage imaging. However, the availability of an MRI scanner is not always given. Taken together, this makes it difficult to implement MRI into the routine operation of the age estimation process, and additional research and studies must be carried out to advance the development.

Further studies and developments have already started in particular with the data of and in combination with this thesis. Neumayer, the doctoral candidate and other co-workers already investigated the shortening of acquisition time of the applied MR sequences in two feasibility studies. (191, 198) Both showed very promising results and the approach should be applied to a larger cohort in the future. Another development deals with the subjectivity of staging and result combination. Urschler, Stern and co-workers investigated on parts of this thesis cohort a fully automated age estimation approach using a machine learning procedure. Currently published is

only a MFA approach with point prediction results (139). However, they are currently working also on an uncertainty estimation (199), which is absolutely necessary for the implementation of this objective approach. Further future work needs also to be done regarding the study of a female cohort as well as other ethnic groups.

Forensic age estimation is not an exact science because of biological variations. However, good approaches can limit the interval of prediction, and a probability definition in relation to the 18 year threshold can be helpful in correctly classifying individuals regarding age.

For a comparison of the different approaches, such as Bayesian models or the “minimum age concept”, new independent cohorts are needed and a comparison study would have to be conducted.

In conclusion, MFA using MRI is a viable approach, and with further research, it will be possible to implement an objective, rapid procedure without the use of ionizing radiation for multi-factorial age estimation.

Bibliography

1. Widek T, Genet P, Merkens H, Boldt J, Petrovic A, Vallis J, Scheurer E. Dental age estimation: The chronology of mineralization and eruption of male third molars with 3T MRI. *Forensic Sci Int.* 2019;297:228–35. DOI: 10.1016/j.forsciint.2019.02.019.
2. Widek T, Genet P, Ehammer T, Schwark T, Urschler M, Scheurer E. Bone Age Estimation with the Greulich-Pyle Atlas using 3T MR Images of Hand and Wrist. *Forensic Sci Int.* 2021;319:110654. DOI: 10.1016/j.forsciint.2020.110654.
3. Schmeling A, Grundmann C, Fuhrmann A, Kaatsch HJ, Knell B, Ramsthaler F, Reisinger W, Riepert T, Ritz-Timme S, Rosing FW, Rotzsch K, Geserick G. Criteria for age estimation in living individuals. *Int J Legal Med.* 2008;122(6):457–60. DOI: 10.1007/s00414-008-0254-2.
4. UNHCR. Global Trends 2018 - Forced Displacement in 2018 [Internet]. Geneva: UNHCR; c2020 [cited 2021 Mar. 3]. Available from: <https://www.unhcr.org/globaltrends2018/>.
5. Eurostat. Asylbewerber und erstmalige Asylbewerber nach Staatsangehörigkeit, Alter und Geschlecht [Internet]. Brussels: Eurostat; c2021 [updated 2021 Mar 2; cited 2021 Mar. 3]. Available from: http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=migr_asyappctza&lang=de.

6. Beh P, Payne-James J. Clinical and legal requirements for age determination in the living. In: Black S, Aggrawal A, Payne-James J, editors. *Age Estimation in the Living: The Practitioner's Guide*. Oxford: Wiley-Blackwell; 2010. p. 30–42.
7. Malina RM. Skeletal age and age verification in youth sport. *Sports Medicine*. 2011;41(11):925–47. DOI: 10.2165/11590300-000000000-00000.
8. Malina RM, Peña Reyes ME, Figueiredo AJ, Coelho E Silva MJ, Horta L, Miller R, Chamorro M, Serratos L, Morate F. Skeletal age in youth soccer players: implication for age verification. *Clin J Sport Med*. 2010;20(6):469–74. DOI: 10.1097/JSM.0b013e3181f827ea.
9. Martin DD, Wit JM, Hochberg Z, Säwendahl L, Rijn RR van, Fricke O, Cameron N, Caliebe J, Hertel T, Kiepe D, Albertsson-Wikland K, Thodberg HH, Binder G, Ranke MB. The Use of Bone Age in Clinical Practice – Part 1. Hormone Research in Paediatrics. 2011;76(1):1–9. DOI: 10.1159/000329372.
10. Baccetti T, Franchi L, McNamara James A. J. The Cervical Vertebral Maturation (CVM) Method for the Assessment of Optimal Treatment Timing in Dentofacial Orthopedics. *Seminars in Orthodontics*. 2005;11(3):119–29. DOI: 10.1053/j.sodo.2005.04.005.
11. Horter MJ, Friesen S, Wacker S, Vogt B, Leidiger B, Roedl R, Schiedel F. [Determination of skeletal age : comparison of the methods of Greulich and Pyle and Tanner and Whitehouse]. *Der Orthopade*. 2012;41(12):966–76. DOI: 10.1007/s00132-012-1983-y. German.
12. Aynsley-Green A. The assessment of age in undocumented migrants. A report for the Office of the Defensor del Pueblo, Madrid, Spain. London: Aynsley-Green Consulting; 2011.
13. United Nations Children's Fund. Birth Registration for Every Child by 2030: Are we on track? [Internet]. New York: UNICEF; c2019 [cited 2021 Mar. 3]. Available from: <https://www.unicef.org/media/62981/file/Birth-registration-for-every-child-by-2030.pdf>.

14. Office of the High Commissioner for Human Rights. Convention on the Rights of the Child. Part I Article 1 [Internet]. Geneva: OHCHR; c1996-2021 [cited 2021 Mar. 3]. Available from: <https://www.ohchr.org/en/professionalinterest/pages/crc.aspx>.
15. European Union Agency for Fundamental Rights. Age of majority [Internet]. Vienna: European Union Agency for Fundamental Rights; c2007-2021 [cited 2021 Mar. 3]. Available from: <https://fra.europa.eu/en/publication/2017/mapping-minimum-age-requirements/age-majority>.
16. Department for Work and Pension. State Pension age timetable [Internet]. London: Government Digital Service; c2014 [updated 2014 May 15 cited 2021 Mar. 3]. Available from: <https://www.gov.uk/government/publications/state-pension-age-timetable/state-pension-age-timetable>.
17. Fernie G, Payne-James J. Legal Implications of Age Determination: Consent and Other Issues. In: Black S, Aggrawal A, Payne-James J, editors. Age Estimation in the Living: The Practitioner's Guide. Oxford: Wiley-Blackwell; 2010. p. 43–54.
18. Kainz S, Fischer F, Scheurer E. Forensische Altersdiagnostik bei Lebenden in Österreich. In: Grassberger M, Türk E, Yen K, editors. Klinisch-forensische Medizin: Interdisziplinärer Praxisleitfaden für Ärzte, Pflegekräfte, Juristen und Betreuer von Gewaltopfern. Vienna: Springer; 2013. p. 483–6. German.
19. National Center for Juvenile Justice. Juvenile Justice, Geography, Policy, Practice & Statistics. Jurisdictional boundaries [Internet]. MacArthur Foundation; c2017 [cited 2021 Mar. 3]. Available from: <http://www.jjgps.org/jurisdictional-boundaries>.
20. Social Security Administration. Retirement Benefits [Internet]. Social Security Administration; 2021. Publication No. 05-10035 [cited 2021 Mar. 3]. Available from: <https://www.ssa.gov/pubs/EN-05-10035.pdf>.
21. USLegal.com. Age of Majority [Internet]. Brookline (MA): airSlate Legal Forms, Inc.; c1997-2021 [cited 2021 Mar. 3]. Available from: <https://minors.uslegal.com/age-of-majority/>.

22. European Asylum Support Office. Practical Guide on age assessment [Internet]. 2nd ed. La Valetta (Malta): EASO; 2018. [cited 2021 Mar. 3]. Available from: <https://www.easo.europa.eu/sites/default/files/easo-practical-guide-on-age-assessment-v3-2018.pdf>.
23. Scheurer, E. (Institute of Forensic Medicine, University Basel, Basel, Switzerland). Personal communication with Thomas Widek (Medical University of Graz). 2021 Feb 5.
24. Genet, P. (University Centre of Legal Medicine Lausanne, Lausanne University Hospital, Lausanne, Switzerland). Personal communication with Thomas Widek (Medical University of Graz). 2021 Jan 12.
25. Schwark, T. (Department of Forensic Medicine, Laboratoire national de santé, Dudelange, Luxembourg). Personal communication with Thomas Widek (Medical University of Graz). 2021 Feb 5.
26. Office of Refugee Resettlement. Children Entering the United States Unaccompanied: Section 1. [Internet]. US Department of Health and Human Services; c2021 [cited 2021 Feb. 26]. Available from: <https://www.acf.hhs.gov/orr/report/children-entering-united-states-unaccompanied-section-1>.
27. Outten-Mills D, Schmidt A, Bobman H, Brown S. Age Determination Practices for Unaccompanied Alien Children in ICE Custody. Washington, DC: Department of Homeland Security - Office of Inspector General; 2009. OIG-10-12. Available from: https://www.oig.dhs.gov/assets/Mgmt/OIG_10-12_Nov09.pdf.
28. Bundesgesetz über die Ausübung der Fremdenpolizei, die Ausstellung von Dokumenten für Fremde und die Erteilung von Einreiseteil . Fremdenpolizeigesetz 2005 – FPG (Austria) [Internet]. [cited 2021 Mar. 3]. Available from: <https://www.ris.bka.gv.at/GeltendeFassung.wxe?Abfrage=Bundesnormen&Gesetzesnummer=20004241>. German.
29. Bundesgesetz über die Gewährung von Asyl. Asylgesetz 2005 - AsylG 2005. (Austria) [Internet]. [cited 2021 Mar. 3]. Available from: <https://www.ris.bka.gv.at>.

- at/GeltendeFassung.wxe?Abfrage=Bundesnormen&Gesetzesnummer=20004240.
German.
30. Ramsthaler F, Zedler B. Forensische Altersdiagnostik Lebender - Ein Methodenvergleich aus Sicht der Forensischen Medizin. In: Parzeller M, Bratzke H, Ramsthaler F, editors. Praxishandbuch Forensische Altersdiagnostik bei Lebenden: Medizinische und Rechtliche Grundlagen. Stuttgart: Boorberg; 2008. p. 21–126. German.
 31. Marshall WA, Tanner JM. Variations in the Pattern of Pubertal Changes in Boys. Arch Dis Child. 1970;45(239):13–23. DOI: 10.1136/adc.45.239.13.
 32. Marshall WA, Tanner JM. Variations in the Pattern of Pubertal Changes in Girls. Arch Dis Child. 1969;44(235):291–303. DOI: 10.1136/adc.44.235.291.
 33. Tanner JM. Growth at adolescence. 2nd ed. Oxford: Blackwell; 1962.
 34. Hackman S, Buck A, Black S. Age Evaluation from the Skeleton. In: Black S, Aggrawal A, Payne-James J, editors. Age Estimation in the Living: The Practitioner's Guide. Wiley-Blackwell; 2010. p. 202–35.
 35. Bilanz D. [Vergleichende Analyse verschiedener Stadieneinteilungen zur Beurteilung der Weisheitszahnmineralisation]. [dissertation]. Berlin: Freie Universität Berlin; 2010. Available from: http://www.diss.fu-berlin.de/diss/receive/FUDISS_thesis_000000013885. German.
 36. Schmidt S, Fracasso T, Pfeiffer H, Schmeling A. [Skelettaltersbestimmung der Hand]. Rechtsmedizin. 2010;20(6):475–82. DOI: 10.1007/s00194-010-0708-y. German.
 37. Gilsanz V, Ratib O. Hand Bone Age: A Digital Atlas Of Skeletal Maturity. Berlin: Springer; 2005.
 38. Roche A, Chumlea W, Thissen D. Assessing the Skeletal Maturity of the Hand Wrist: Fels Method. Springfield: Charles C. Thomas Publisher; 1988.
 39. Thiemann HH, Nitz I, Schmeling A. Röntgenatlas der normalen Hand im Kindesalter. 3rd edn. Stuttgart: Thieme; 2006. German.

40. Greulich W, Pyle S. Radiographic Atlas of Skeletal Development of Hand and Wrist. 2nd ed. Stanford: Stanford University Press; 1959.
41. Tanner JM, Whitehouse RH, Marshall WA, Healy M, Goldstein H. Assessment of Skeletal Maturity and Prediction of Adult Height (TW2 Method). 2nd ed. London: Academic Press; 1983.
42. Tanner J, Healy M, Cameron N, Goldstein H. Assessment of Skeletal Maturity and Prediction of Adult Height (TW3 Method). London: W.B. Saunders; 2001.
43. Alshamrani K, Offiah AC. Applicability of two commonly used bone age assessment methods to twenty-first century UK children. *Eur Radiol.* 2020;30(1):504–13. DOI: 10.1007/s00330-019-06300-x.
44. Schmidt S, Nitz I, Ribbecke S, Schulz R, Pfeiffer H, Schmeling A. Skeletal age determination of the hand: a comparison of methods. *Int J Legal Med.* 2013;127(3):691–8. DOI: 10.1007/s00414-013-0845-4.
45. Chaumoitre K, Saliba-Serre B, Adalian P, Signoli M, Leonetti G, Panuel M. Forensic use of the Greulich and Pyle atlas: prediction intervals and relevance. *Eur Radiol.* 2017;27(3):1032–43. DOI: 10.1007/s00330-016-4466-4.
46. Schmeling A, Reisinger W, Loreck D, Vendura K, Markus W, Geserick G. Effects of ethnicity on skeletal maturation: consequences for forensic age estimations. *International journal of legal medicine.* 2000;113(5):253–8. DOI: 10.1007/s004149900102.
47. Schmeling A, Schulz R, Danner B, Rosing FW. The impact of economic progress and modernization in medicine on the ossification of hand and wrist. *Int J Legal Med.* 2006;120(2):121–6. DOI: 10.1007/s00414-005-0007-4.
48. Timme M, Karch A, Shay D, Ottow C, Schmeling A. [Zur Altersdiagnostik lebender Personen: der Einfluss des sozioökonomischen Status auf die Skelett- und Zahnentwicklung in einer deutschen Studienkohorte]. *Rechtsmedizin.* 2020. DOI: 10.1007/s00194-020-00444-7. German.

49. Radlanski R. Orale Struktur- und Entwicklungsbiologie. Quintessenz-Verlag; 2011. German.
50. Röttscher K. Forensische Zahnmedizin. Berlin: Springer; 2003. German.
51. Cameriere R, Ferrante L, Cingolani M. Age estimation in children by measurement of open apices in teeth. *Int J Legal Med.* 2006;120(1):49–52. DOI: 10.1007/s00414-005-0047-9.
52. Demirjian A, Goldstein H, Tanner JM. A new system of dental age assessment. *Hum Biol.* 1973;45(2):211–27.
53. Gleiser I, Hunt Jr. EE. The permanent mandibular first molar: Its calcification, eruption and decay. *Am J Phys Anthropol.* 1955;13(2):253–83. DOI: 10.1002/ajpa.1330130206.
54. Moorrees C, Fanning E, Hunt E. Age Variation of Formation Stages for Ten Permanent Teeth. *J Dent Res.* 1963;42:1490–502. DOI: 10.1177/00220345630420062701.
55. Schour I, Massler M. The development of the human dentition. *J Am Dent Assoc.* 1941;28(7):p. 1153–60.
56. Willems G, Van Olmen A, Spiessens B, Carels C. Dental Age Estimation in Belgian Children: Demirjian's Technique Revisited. *J Forensic Sci.* 2001;46(4):893–5.
57. Taylor J, Blenkin M. Age Estimation in the Living: The Practitioner's Guide. In: Black S Aggrawal A PJJ, editor. Oxford: Wiley-Blackwell; 2010. p. 176–201.
58. Lopez TT, Arruda CP, Rocha M, Oliveira Rosin ASA de, Michel-Crosato E, Biazzevic MGH. Estimating ages by third molars: Stages of development in Brazilian young adults. *J Forensic Legal Med.* 2013;20(5):412–8. DOI: 10.1016/j.jflm.2012.12.001.
59. Mincer HH, Harris EF, Berryman HE. The A.B.F.O. study of third molar development and its use as an estimator of chronological age. *J Forensic Sci.* 1993;38(2):379–90.
60. Olze A, Bilang D, Schmidt S, Wernecke K, Geserick G, Schmeling A. Validation of common classification systems for assessing the mineralization of third molars. *Int J Legal Med.* 2005;119(1):22–8. DOI: 10.1007/s00414-004-0489-5.

61. Levesque GY, Demirjian A, Tanguay R. Sexual dimorphism in the development, emergence, and agenesis of the mandibular third molar. *J Dent Res.* 1981;60(10):1735–41. DOI: 10.1177/00220345810600100201.
62. Olze A, Niekerk P van, Ishikawa T, Zhu BL, Schulz R, Maeda H, Schmeling A. Comparative study on the effect of ethnicity on wisdom tooth eruption. *Int J Legal Med.* 2007;121(6):445–8. DOI: 10.1007/s00414-007-0171-9.
63. Olze A, Schmeling A, Taniguchi M, Maeda H, Niekerk P, Wernecke KD, Geserick G. Forensic age estimation in living subjects: the ethnic factor in wisdom tooth mineralization. *Int J Legal Med.* 2004;118(3):170–3. DOI: 10.1007/s00414-004-0434-7.
64. Orhan K, Ozer L, Orhan AI, Dogan S, Paksoy CS. Radiographic evaluation of third molar development in relation to chronological age among Turkish children and youth. *Forensic Sci Int.* 2007;165(1):46–51. DOI: 10.1016/j.forsciint.2006.02.046.
65. Rai B, Kaur J, Jafarzadeh H. Dental age estimation from the developmental stage of the third molars in Iranian population. *J Forensic Legal Med.* 2010;17(6):309–11. DOI: 10.1016/j.jflm.2010.04.010.
66. Zeng D, Wu Z, Cui M. Chronological age estimation of third molar mineralization of Han in southern China. *Int J Legal Med.* 2010;124(2):119–23. DOI: 10.1007/s00414-009-0379-y.
67. Liversidge HM. Similarity in dental maturation in two ethnic groups of London children. *Ann Hum Biol.* 2011;38(6):702–15. DOI: 10.3109/03014460.2011.609565.
68. Liversidge HM, Peariasamy K, Folayan MO, Adeniyi AO, Ngom PI, Mikami Y, Shimada Y, Kuroe K, Tsvete IF, Kvaal SI. A radiographic study of the mandibular third molar root development in different ethnic groups. *J Forensic Odontostomatol.* 2017;35(2):97–108.
69. Thevissen PW, Fieuws S, Willems G. Human third molars development: Comparison of 9 country specific populations. *Forensic Sci Int.* 2010;201(1-3):102–5. DOI: 10.1016/j.forsciint.2010.04.054.

70. Thevissen PW, Algerban A, Asaumi J, Kahveci F, Kaur J, Kim YK, Pittayapat P, Van Vlierberghe M, Zhang Y, Fieuws S, Willems G. Human dental age estimation using third molar developmental stages: Accuracy of age predictions not using country specific information. *Forensic Sci Int.* 2010;201(1-3):106–11. DOI: 10.1016/j.forsciint.2010.04.040.
71. Friedrich RE, Ulbricht C, Ljuba ABvM. The influence of wisdom tooth impaction on root formation. *Ann Anat.* 2003;185(5):481–92. DOI: 10.1016/s0940-9602(03)80112-7.
72. Olze A, Otto A, Tsokos M. [Influence of retention on the rate of mineralization of third molars]. *Rechtsmedizin.* 2012;22(2):110–4. DOI: 10.1007/s00194-012-0808-y. German.
73. Olze A, Van Niekerk P, Schulz R, Ribbecke S, Schmeling A. The influence of impaction on the rate of third molar mineralisation in male black Africans. *Int J Legal Med.* 2012;126(6):869–74. DOI: 10.1007/s00414-012-0753-z.
74. Guo Y, Yan C, Lin X, Zhang W, Zhou H, Pan F, Wei L, Tang Z, Liang F, Chen T. The influence of impaction to the third molar mineralization in northwestern Chinese population. *Int J Legal Med.* 2014;128(4):659–65. DOI: 10.1007/s00414-014-0979-z.
75. Conceicao EL, Cardoso HF. Environmental effects on skeletal versus dental development II: further testing of a basic assumption in human osteological research. *Am J Phys Anthropol.* 2011;144(3):463–70. DOI: 10.1002/ajpa.21433.
76. Cameriere R, Flores-Mir C, Mauricio F, Ferrante L. Effects of nutrition on timing of mineralization in teeth in a Peruvian sample by the Cameriere and Demirjian methods. *Ann Hum Biol.* 2007;34(5):547–56. DOI: 10.1080/03014460701556296.
77. Elamin F, Liversidge HM. Malnutrition has no effect on the timing of human tooth formation. *PLoS One.* 2013;8(8):e72274. DOI: 10.1371/journal.pone.0072274.
78. Schulz R, Mühler M, Reisinger W, Schmidt S, Schmeling A. Radiographic staging of ossification of the medial clavicular epiphysis. *Int J Legal Med.* 2008;122(1):55–8. DOI: 10.1007/s00414-007-0210-6.

79. Wittschieber D, Ottow C, Vieth V. Projection radiography of the clavicle: still recommendable for forensic age diagnostics in living individuals? *Int J Legal Med.* 2015;129(1):187–93. DOI: 10.1007/s00414-014-1067-0.
80. Wittschieber D, Ottow C, Schulz R. Forensic age diagnostics using projection radiography of the clavicle: a prospective multi-center validation study. *Int J Legal Med.* 2016;130(1):213–9. DOI: 10.1007/s00414-015-1285-0.
81. Kreitner KF, Schweden F, Schild HH, Riepert T, Nafe B. [Computerized tomography of the epiphyseal union of the medial clavicle: an auxiliary method of age determination during adolescence and the 3d decade of life?]. *RoFo.* 1997;166(6):481–6. DOI: 10.1055/s-2007-1015463. German.
82. Kreitner KF, Schweden FJ, Riepert T, Nafe B, Thelen M. Bone age determination based on the study of the medial extremity of the clavicle. *Eur Radiol.* 1998;8(7):1116–22. DOI: 10.1007/s003300050518.
83. Schmeling A, Schulz R, Reisinger W, Muhler M, Wernecke KD, Geserick G. Studies on the time frame for ossification of the medial clavicular epiphyseal cartilage in conventional radiography. *Int J Legal Med.* 2004;118(1):5–8. Available from: PM: 14534796.
84. Schulz R, Mühler M, Mutze S, Schmidt S, Reisinger W, Schmeling A. Studies on the time frame for ossification of the medial epiphysis of the clavicle as revealed by CT scans. *Int J Legal Med.* 2005;119(3):142–7. DOI: 10.1007/s00414-005-0529-9.
85. Kellinghaus M, Schulz R, Vieth V, Schmidt S, Schmeling A. Forensic age estimation in living subjects based on the ossification status of the medial clavicular epiphysis as revealed by thin-slice multidetector computed tomography. *Int J Legal Med.* 2010;124(2):149–54. DOI: 10.1007/s00414-009-0398-8.
86. Kellinghaus M, Schulz R, Vieth V, Schmidt S, Pfeiffer H, Schmeling A. 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.* 2010;124(4):321–5. DOI: 10.1007/s00414-010-0448-2.

87. Axelsson I. Bone maturation cannot be used to estimate chronological age in asylum-seeking adolescents. *Acta Paediatr.* 2019;108(4):590–2. DOI: 10.1111/apa.14645.
88. Aynsley-Green A, Cole T, Crawley H, Lessof N, Boag L, Wallace R. Medical, statistical, ethical and human rights considerations in the assessment of age in children and young people subject to immigration control. *Brit Med Bull.* 2012;102(1):17–42. DOI: 10.1093/bmb/lds014.
89. Ording Müller L, Offiah A, Adamsbaum C, Barber I, Di Paolo PL, Humphries P, Shelmerdine S, Tanturri De Horatio L, Toma P, Treguier C, Rosendahl K. Bone age for chronological age determination - statement of the European Society of Paediatric Radiology musculoskeletal task force group. *Pediatr Radiol.* 2019;49(7):979–82. DOI: 10.1007/s00247-019-04379-4.
90. Schmeling A, Reisinger W, Wormanns D, Geserick G. [Strahlenexposition bei Röntgenuntersuchungen zur forensischen Altersschätzung Lebender]. *Rechtsmedizin.* 2000;10(4):135–7. DOI: 10.1007/s001940000060. German.
91. BBC. Dental X-rays for asylum children criticised. *BBC News.* 2012. [cited 2021 Mar. 3]. Available from: <http://www.bbc.com/news/uk-17567082>.
92. Meikle J. UK Border Agency to trial x-rays to determine age of asylum seekers [Internet]. *Guardian.* 2012. [cited 2021 Mar. 3]. Available from: <http://www.theguardian.com/uk/2012/mar/30/uk-border-agency-x-rays-asylum-seekers>.
93. Meikle J. Border Agency halts x-ray programme for child asylum seekers [Internet]. *Guardian.* 2012. [cited 2021 Mar. 3]. Available from: <http://www.theguardian.com/uk/2012/apr/27/border-agency-xray-asylum-seekers>.
94. Nowotny T, Eisenberg W, Mohnike K. [Unbegleitete minderjährige Flüchtlinge: Strittiges Alter – strittige Altersdiagnostik]. *Deutsches Aerzteblatt.* 2014;111(18):786–8. German.
95. Schmeling A, Geserick G, Tsokos M, Dettmeyer R, Rudolf E, Püschel K. [Aktuelle Diskussionen zur Altersdiagnostik bei unbegleiteten minderjährigen Flüchtlingen]. *Rechtsmedizin.* 2014;24(6):475–9. DOI: 10.1007/s00194-014-0986-x. German.

96. Schmeling A. [Die aktuelle medizinethische Debatte über forensische Altersdiagnostik bei unbegleiteten minderjährigen Flüchtlingen]. [master thesis]. Münster (Germany): University of Münster; 2018. Available from: https://www.dgrm.de/fileadmin/PDF/AG_FAD/masterarbeit_schmeling.pdf. German.
97. Thevissen P, Kvaal S, Diericks K, Willems G. Ethics in age estimation of unaccompanied minors. *J Forensic Odontostomatol*. 2012;30(Suppl 1):84–102.
98. Schulz R, Zwiesigk P, Schiborr M, Schmidt S, Schmeling A. Ultrasound studies on the time course of clavicular ossification. *Int J Legal Med*. 2008;122(2):163–7. DOI: 10.1007/s00414-007-0220-4.
99. Schmidt S, Schiborr M, Pfeiffer H, Schmeling A, Schulz R. Sonographic examination of the apophysis of the iliac crest for forensic age estimation in living persons. *Sci Justice*. 2013;53(4):395–401. DOI: 10.1016/j.scijus.2013.05.004.
100. Schmidt S, Schiborr M, Pfeiffer H, Schmeling A, Schulz R. Age dependence of epiphyseal ossification of the distal radius in ultrasound diagnostics. *Int J Legal Med*. 2013;127(4):831–8. DOI: 10.1007/s00414-013-0871-2.
101. Tolsgaard M, Todsén T, Sørensen J. International multispecialty consensus on how to evaluate ultrasound competence: a Delphi consensus survey. *PLoS One*. 2013;8(2):57687. DOI: 10.1371/journal.pone.0057687.
102. Gonsior M, Ramsthaler F, Gehl A, Verhoff M. Morphology as a cause for different classification of the ossification stage of the medial clavicular epiphysis by ultrasound, computed tomography, and macroscopy. *Int J Legal Med*. 2013;127(5):1013–21. DOI: 10.1007/s00414-013-0889-5.
103. De Tobel J, Hillewig E, Haas MB de, Van Eeckhout B, Fieuws S, Thevissen PW, Verstraete KL. 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*. 2019;29(6):2924–35. DOI: 10.1007/s00330-018-5944-7.

104. Dvorak J, George J, Junge A, Hodler J. Age determination by magnetic resonance imaging of the wrist in adolescent male football players. *Br J Sports Med.* 2007;41(1):45–52. DOI: 10.1136/bjsm.2006.031021.
105. Dvorak J, George J, Junge A, Hodler J. Application of MRI of the wrist for age determination in international U-17 soccer competitions. *Br J Sports Med.* 2007;41(8):497–500. DOI: 10.1136/bjsm.2006.033431.
106. Er A, Bozdog M, Basa CD, Kacmaz IE, Ekizoglu O. Estimating forensic age via magnetic resonance imaging of the distal radial epiphysis. *Int J Legal Med.* 2020;134(1):375–80. DOI: 10.1007/s00414-019-02189-9.
107. George J, Nagendran J, Azmi K. 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.* 2012;46(4):273–8. DOI: 10.1136/bjsm.2010.074948.
108. Hojreh A, Gamper J, Schmook MT, Weber M, Prayer D, Herold CJ, Noebauer-Huhmann IM. Hand MRI and the Greulich-Pyle atlas in skeletal age estimation in adolescents. *Skeletal Radiol.* 2018;47(7):963–71. DOI: 10.1007/s00256-017-2867-3.
109. Schmidt S, Vieth V, Timme M, Dvorak J, Schmeling A. 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.* 2015;55(2):139–44. DOI: 10.1016/j.scijus.2014.12.003.
110. Serinelli S, Panebianco V, Martino M, Battisti S, Rodacki K, Marinelli E, Zaccagna F, Semelka RC, Tomei E. Accuracy of MRI skeletal age estimation for subjects 12-19. Potential use for subjects of unknown age. *Int J Legal Med.* 2015;129(3):609–17. DOI: 10.1007/s00414-015-1161-y.
111. Serin J, Rerolle C, Pucheux J, Dedouit F, Telmon N, Savall F, Saint-Martin P. Contribution of magnetic resonance imaging of the wrist and hand to forensic age assessment. *Int J Legal Med.* 2016;130(4):1121–8. DOI: 10.1007/s00414-016-1362-z.

112. Terada Y, Kono S, Tamada D, Uchiumi T, Kose K, Miyagi R, Yamabe E, Yoshioka H. Skeletal age assessment in children using an open compact MRI system. *Magn Reson Med*. 2013;69(6):1697–702. DOI: 10.1002/mrm.24439.
113. Tomei E, Sartori A, Nissman D, Al Ansari N, Battisti S, Rubini A, Stagnitti A, Martino M, Marini M, Barbato E, Semelka RC. Value of MRI of the hand and the wrist in evaluation of bone age: preliminary results. *J Magn Reson Imaging*. 2014;39(5):1198–205. DOI: 10.1002/jmri.24286.
114. Urschler M, Krauskopf A, Widek T, Sorantin E, Ehammer T, Borkenstein M, Yen K, Scheurer E. 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*. 2016;266:281–8. DOI: 10.1016/j.forsciint.2016.06.016.
115. Timme M, Ottow C, Schulz R. 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*. 2017;131(2):579–84. DOI: 10.1007/s00414-016-1502-5.
116. Diete V, Wabitsch M, Denzer C, Jäger H, Hauth E, Beer M, Vogele D. Applicability of Magnetic Resonance Imaging for Bone Age Estimation in the Context of Medical Issues. *Rofo*. 2020. DOI: 10.1055/a-1313-7664.
117. Schmidt S, Henke C, Wittschieber D. Optimising magnetic resonance imaging-based evaluation of the ossification of the medial clavicular epiphysis: a multi-centre study. *Int J Legal Med*. 2016;130(6):1615–21. DOI: 10.1007/s00414-016-1442-0.
118. Schmidt S, Ottow C, Pfeiffer H. Magnetic resonance imaging-based evaluation of ossification of the medial clavicular epiphysis in forensic age assessment. *Int J Legal Med*. 2017;131(6):1665–73. DOI: 10.1007/s00414-017-1676-5.
119. Schmidt S, Mühler M, Schmeling A, Reisinger W, Schulz R. Magnetic resonance imaging of the clavicular ossification. *Int J Legal Med*. 2007;121(4):321–4. DOI: 10.1007/s00414-007-0160-z.

120. Vieth V, Schulz R, Brinkmeier P, Dvorak J, Schmeling A. Age estimation in U-20 football players using 3.0 tesla MRI of the clavicle. *Forensic Sci Int.* 2014;241:118–22. DOI: 10.1016/j.forsciint.2014.05.008.
121. De Tobel J, Van Wijk M, Alberink I. The influence of motion artefacts on magnetic resonance imaging of the clavicles for age estimation. *Int J Legal Med.* 2020;134(2):753–68. DOI: 10.1007/s00414-019-02230-x.
122. De Tobel J, Hillewig E, Van Wijk M. Staging Clavicular Development on MRI: Pitfalls and Suggestions for Age Estimation. *J Magn Reson Imaging.* 2020;51(2):377–88. DOI: 10.1002/jmri.26889.
123. Tangmose S, Jensen K, Villa C, Lynnerup N. Forensic age estimation from the clavicle using 1.0T MRI—preliminary results. *Forensic Sci Int.* 2014;234:7–12. DOI: 10.1016/j.forsciint.2013.10.027.
124. Dedouit F, Auriol J, Rousseau H, Rougé D, Crubézy E, Telmon N. Age assessment by magnetic resonance imaging of the knee: A preliminary study. *Forensic Sci Int.* 2012;217(1–3):232.e1–232.e7. DOI: 10.1016/j.forsciint.2011.11.013.
125. Ekizoglu O, Inci E, Ors S, Hocaoglu E, Can IO, Basa CD, Kacmaz IE, Kranioti EF. Forensic age diagnostics by magnetic resonance imaging of the proximal humeral epiphysis. *Int J Legal Med.* 2019;133(1):249–56. DOI: 10.1007/s00414-018-1952-z. Available from: <https://doi.org/10.1007/s00414-018-1952-z>.
126. Krämer JA, Schmidt S, Jürgens KU, Lentschig M, Schmeling A, Vieth V. Forensic age estimation in living individuals using 3.0T MRI of the distal femur. *Int J Legal Med.* 2014;128(3):509–14. DOI: 10.1007/s00414-014-0967-3. Available from: <https://doi.org/10.1007/s00414-014-0967-3>.
127. Krämer JA, Schmidt S, Jürgens KU, Lentschig M, Schmeling A, Vieth V. 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.* 2014;10(3):306–13. DOI: 10.1007/s12024-014-9559-2.

128. Wittschieber D, Vieth V, Timme M, Dvorak J, Schmeling A. Magnetic resonance imaging of the iliac crest: age estimation in under-20 soccer players. *Forensic Sci Med Pathol.* 2014;10(2):198–202. DOI: 10.1007/s12024-014-9548-5.
129. Baumann P, Widek T, Merkens H, Boldt J, Petrovic A, Urschler M, Kirnbauer B, Jakse N, Scheurer E. Dental age estimation of living persons: Comparison of MRI with OPG. *Forensic Sci Int.* 2015;253:76–80. DOI: 10.1016/j.forsciint.2015.06.001.
130. De Tobel J, Hillewig E, Bogaert S, Deblaere K, Verstraete K. Magnetic resonance imaging of third molars: developing a protocol suitable for forensic age estimation. *Ann Hum Biol.* 2017;44(2):130–9. DOI: 10.1080/03014460.2016.1202321.
131. De Tobel J, Hillewig E, Verstraete K. Forensic age estimation based on magnetic resonance imaging of third molars: converting 2D staging into 3D staging. *Ann Hum Biol.* 2017;44(2):121–9. DOI: 10.1080/03014460.2016.1223884.
132. De Tobel J, Phlypo I, Fieuws S, Politis C, Verstraete K, Thevissen P. Forensic age estimation based on development of third molars: a staging technique for magnetic resonance imaging. *J Forensic Odontostomatol.* 2017;35(2):117–40.
133. De Tobel J, Parmentier G, Phlypo I. Magnetic resonance imaging of third molars in forensic age estimation: comparison of the Ghent and Graz protocols focusing on apical closure. *Int J Legal Med.* 2019;133(2):583–92. DOI: 10.1007/s00414-018-1905-6.
134. Guo Y, Olze A, Ottow C, Schmidt S, Schulz R, Heindel W, Pfeiffer H, Vieth V, Schmeling A. Dental age estimation in living individuals using 3.0T MRI of lower third molars. *Int J Legal Med.* 2015;129(6):1265–70. DOI: 10.1007/s00414-015-1238-7.
135. De Tobel J, Fieuws S, Hillewig E, Phlypo I, Wijk M van, Haas MB de, Politis C, Verstraete KL, Thevissen PW. Multi-factorial age estimation: A Bayesian approach combining dental and skeletal magnetic resonance imaging. *Forensic Sci Int.* 2020;306:110054. DOI: 10.1016/j.forsciint.2019.110054.
136. Gelbrich B, Frerking C, Weiss S. 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.* 2015;42(4):389–96. DOI: 10.3109/03014460.2015.1046487.

137. Schmeling A, Dettmeyer R, Vieth R, Geserick V, null G. Forensic Age Estimation. *Dtsch Arztebl Int.* 2016;113(4):44–50. DOI: 10.3238/arztebl.2016.0044.
138. Nowotny T. Minimum Age Concept: Full of Pitfalls [Letter]. *Dtsch Arztebl Int.* 2016;113(27-28):486–7. DOI: 10.3238/arztebl.2016.0486b.
139. Stern D, Payer C, Giuliani N, Urschler M. Automatic Age Estimation and Majority Age Classification From Multi-Factorial MRI Data. *IEEE J Biomed Health Inform.* 2019;23(4):1392–403. DOI: 10.1109/JBHI.2018.2869606.
140. Ottow C, Krämer JA, Olze A, Schmidt S, Schulz R, Wittschieber D, Heindel W, Pfeiffer H, Ribbecke S, Vieth V, Schmeling A. [Magnetresonanztomographiestudie zur Altersschätzung von unbegleiteten minderjährigen Flüchtlingen]. *Rechtsmedizin.* 2015;25(1):12–20. DOI: 10.1007/s00194-014-0991-0. German.
141. Schmeling A, Kreitner KF, Heindel W, Vieth V. [Bildgebung zur forensischen Altersdiagnostik bei Jugendlichen und jungen Erwachsenen]. *Radiologie up2date.* 2019;19(01):63–75. DOI: 10.1055/a-0795-1838. German.
142. Schumacher G, Schmeling A, Rudolf E. Medical Age Assessment of Juvenile Migrants [Internet]. Publications Office of the European Union; 2018. JRC109278 [cited 2021 Mar. 3]. Available from: <http://publications.jrc.ec.europa.eu/repository/handle/JRC109278>.
143. De Tobel J, Ottow C, Widek T, Klasinc I, Mörnstad H, Thevissen PW, Verstraete KL. Dental and Skeletal Imaging in Forensic Age Estimation: Disparities in Current Approaches and the Continuing Search for Optimization. *Seminars in musculoskeletal radiology.* 2020;24(5):510–22. DOI: 10.1055/s-0040-1701495.
144. Weir C, Jan A. BMI Classification Percentile And Cut Off Points. [Internet]. StatPearls Publishing; c2020 [updated 2020 Jul 10; cited 2021 Mar. 3]. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK541070/>.
145. World Health Organisation. Growth reference 5-19 years. Geneva: World Health Organisation; c2020 [cited 2021 Mar. 3]. Available from: http://www.who.int/growthref/who2007_bmi_for_age/en/.

146. Hillewig E, De Tobel J, Cuche O, Vandemaele P, Piette M, Verstraete K. Magnetic resonance imaging of the medial extremity of the clavicle in forensic bone age determination: a new four-minute approach. *Eur Radiol.* 2011;21(4):757–67. DOI: 10.1007/s00330-010-1978-1.
147. Olze A, Ishikawa T, Zhu BL, Schulz R, Heinecke A, Maeda H, Schmeling A. Studies of the chronological course of wisdom tooth eruption in a Japanese population. *Forensic Sci Int.* 2008;174(2-3):203–6. DOI: 10.1016/j.forsciint.2007.04.218.
148. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; c2020 [cited 2021 Mar. 3]. Available from: <https://www.R-project.org/>.
149. Knell B, Ruhstaller P, Prieels F, Schmeling A. Dental age diagnostics by means of radiographical evaluation of the growth stages of lower wisdom teeth. *Int J Legal Med.* 2009;123(6):465–9. DOI: 10.1007/s00414-009-0330-2.
150. Mood A, Franklin A, Graybill D, Boes. *Introduction to the Theory of Statistics*. 3rd ed. New York: McGraw-Hill; 1974.
151. Cohen J. Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. *Psychol Bull.* 1968;70(4):213–20.
152. Shirley NR, Ramirez Montes PA. Age estimation in forensic anthropology: quantification of observer error in phase versus component-based methods. *J Forensic Sci.* 2015;60(1):107–11. DOI: 10.1111/1556-4029.12617.
153. Lehnert B. BlandAltmanLeh: Plots (Slightly Extended) Bland-Altman Plots. R package version 0.3.1. c2015 [cited 2021 Mar. 3]. Available from: <https://CRAN.R-project.org/package=BlandAltmanLeh>.
154. Gamer M, Lemon J, Fellows I, Singh P. irr: Various Coefficients of Interrater Reliability and Agreement. R package version 0.84.1. c2019 [cited 2021 Mar. 3]. Available from: <https://CRAN.R-project.org/package=irr>.

155. Jackman S. *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. R package version 1.5.5. c2020 [cited 2021 Mar. 3]. Available from: <https://github.com/atahk/pscl/>.
156. Venables WN, Ripley BD. *Modern Applied Statistics with S*. 4th ed. New York: Springer; 2002.
157. Yee TW. *VGAM: Vector Generalized Linear and Additive Models*. R package version 1.1-2. c2019 [cited 2021 Mar. 3]. Available from: <https://CRAN.R-project.org/package=VGAM>.
158. Konigsberg LW, Lyle W. Konigsberg's webpage. c2020 [updated 2020 Nov 27; cited 2021 Mar. 3]. Available from: <http://faculty.las.illinois.edu/lylek/>.
159. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;1(8476):307–10.
160. Boldsen JL, Milner GR, Konigsberg LW, Wood JM. Transition analysis: a new method for estimating age from skeletons. In: Hoppa RD, Vaupel JW, editors. *Paleodemography: age distributions from skeletal samples*. New York, NY: Cambridge University Press; 2002. p. 73–106.
161. Konigsberg LW, Herrmann NP, Wescott DJ, Kimmerle EH. Estimation and evidence in forensic anthropology: age-at-death. *J Forensic Sci*. 2008;53(3):541–57. DOI: 10.1111/j.1556-4029.2008.00710.x.
162. Konigsberg LW, Frankenberg SR, Liversidge HM. Optimal trait scoring for age estimation. *Am J Phys Anthropol*. 2016;159(4):557–76. DOI: 10.1002/ajpa.22914.
163. Tangmose S, Thevissen P, Lynnerup N, Willems G, Boldsen J. Age estimation in the living: Transition analysis on developing third molars. *Forensic Sci Int*. 2015;257:512 e1–512 e7. DOI: 10.1016/j.forsciint.2015.07.049.

164. Hackman L, Black S. The reliability of the Greulich and Pyle atlas when applied to a modern Scottish population. *J Forensic Sci.* 2013;58(1):114–9. DOI: 10.1111/j.1556-4029.2012.02294.x.
165. Santos C, Ferreira M, Alves FC, Cunha E. Comparative study of Greulich and Pyle Atlas and Matusos 4.0 program for age estimation in a Portuguese sample. *Forensic Sci Int.* 2011;212(1-3):276 e1–7. DOI: 10.1016/j.forsciint.2011.05.032.
166. Student. The probable error of a mean. *Biometrika.* 1908:1–25.
167. Whitley E, Ball J. Statistics review 5: Comparison of means. *Critical Care.* 2002;6(5):424–8. DOI: 10.1186/cc1548.
168. Altman DG. *Practical Statistics for Medical Research.* London: Chapman & Hall/CRC; 1991.
169. Konigsberg LW. Multivariate cumulative probit for age estimation using ordinal categorical data. *Annals of human biology.* 2015;42(4):368–78. DOI: 10.3109/03014460.2015.1045430.
170. Bera A, Jarque CM, Lee LF. Testing the Normality Assumption in Limited Dependent Variable Models. *Int Eco Rev.* 1984;25(3):563–78. DOI: 10.2307/2526219.
171. Johnson P. A Test of the Normality Assumption in the Ordered Probit Model. *Metron.* 1996;54:213–21.
172. Fieuws S, Willems G, Larsen-Tangmose S, Lynnerup N, Boldsen J, Thevissen P. Obtaining appropriate interval estimates for age when multiple indicators are used: evaluation of an ad-hoc procedure. *Int J Legal Med.* 2016;130(2):489–99. DOI: 10.1007/s00414-015-1200-8.
173. Hyndman RJ. Computing and Graphing Highest Density Regions. *The American Statistician.* 1996;50(2):120–6. DOI: 10.2307/2684423.
174. Ekstrøm CT. MESS: Miscellaneous Esoteric Statistical Scripts. R package version 0.5.6. c2019 [cited 2021 Mar. 3]. Available from: <https://CRAN.R-project.org/package=MESS>.

175. Plummer M, Best N, Cowles K, Vines K. CODA: Convergence Diagnosis and Output Analysis for MCMC. *R News*. 2006;6(1):7–11.
176. Camacho A, Funk S. fitR: Tool box for fitting dynamic infectious disease models to time series. R package version 0.1. c2019 [cited 2021 Mar. 3]. Available from: <https://github.com/sbfnk/fitR>.
177. Hyndman RJ. hrdcde: Highest Density Regions and Conditional Density Estimation. R package version 3.3. c2018 [cited 2021 Mar. 3]. Available from: <http://pkg.robjhyndman.com/hrdcde>.
178. Genz A, Bretz F, Miwa T, Mi X, Leisch F, Scheipl F, Hothorn T. mvtnorm: Multivariate Normal and t Distributions. R package version 1.1-0. c2020 [cited 2021 Mar. 3]. Available from: <https://CRAN.R-project.org/package=mvtnorm>.
179. Kundu S, Aulchenko YS, Janssens ACJ. PredictABEL: Assessment of Risk Prediction Models. R package version 1.2-4. c2020 [cited 2021 Mar. 3]. Available from: <https://CRAN.R-project.org/package=PredictABEL>.
180. Lynch S. Introduction to Applied Bayesian Statistics and Estimation for Social Scientists. *Statistics for Social and Behavioral Sciences*. New York: Springer; 2007.
181. Meinel A, Tangl S, Huber C, Maurer B, Watzek G. The chronology of third molar mineralization in the Austrian population—a contribution to forensic age estimation. *Forensic Sci Int*. 2007;169(2-3):161–7. DOI: 10.1016/j.forsciint.2006.08.014.
182. Olze A, Schmeling A, Rieger K, Kalb G, Geserick G. [Studies on the Chronology of Third Molar Mineralization in a German Population]. *Rechtsmedizin*. 2003;13(1):5–10. DOI: 10.1007/s00194-002-0182-2. German.
183. Prieto JL, Barberia E, Ortega R, Magana C. Evaluation of chronological age based on third molar development in the Spanish population. *Int J Legal Med*. 2005;119(6):349–54. DOI: 10.1007/s00414-005-0530-3.

184. Olze A, Pynn BR, Kraul V, Schulz R, Heinecke A, Pfeiffer H, Schmeling A. Studies on the chronology of third molar mineralization in First Nations people of Canada. *Int J Legal Med.* 2010;124(5):433–7. DOI: 10.1007/s00414-010-0483-z.
185. Gelbrich B, Lessig R, Lehmann M, Dannhauer KH, Gelbrich G. Altersselektion in Referenzstichproben. *Rechtsmedizin.* 2010;20(6):459–63. DOI: 10.1007/s00194-010-0703-3.
186. Caldas IM, Carneiro JL, Teixeira A, Matos E, Afonso A, Magalhaes T. Chronological course of third molar eruption in a Portuguese population. *Int J Legal Med.* 2012;126(1):107–12. DOI: 10.1007/s00414-011-0600-7.
187. Olze A, Peschke C, Schulz R, Schmeling A. Studies of the chronological course of wisdom tooth eruption in a German population. *J Forensic Leg Med.* 2008;15(7):426–9. DOI: 10.1016/j.jflm.2008.02.008.
188. Olze A, Pynn BR, Kraul V, Schulz R, Heinecke A, Pfeiffer H, Schmeling A. Dental age estimation based on third molar eruption in First Nation people of Canada. *J Forensic Odontostomatol.* 2010;28(1):32–8.
189. Olze A, Van Niekerk P, Schulz R, Schmeling A. Studies of the Chronological Course of Wisdom Tooth Eruption in a Black African Population. *J Forensic Sci.* 2007;52(5):1161–3. DOI: 10.1111/j.1556-4029.2007.00534.x.
190. Brkic H, Vodanovic M, Dumancic J, Lovric Z, Cukovic-Bagic I, Petroveckii M. The Chronology of Third Molar Eruption in the Croatian Population. *Collegium Antropol.* 2011;35(2).
191. Neumayer B, Schloegl M, Payer C, Widek T, Tschauner S, Ehammer T, Stollberger R, Urschler M. Reducing acquisition time for MRI-based forensic age estimation. *Sci Rep.* 2018;8(1):2063. DOI: 10.1038/s41598-018-20475-1.
192. Metsaniitty M, Waltimo-Siren J, Ranta H, Fieuws S, Thevissen P. Dental age estimation in Somali children and sub-adults combining permanent teeth and third molar development. *Int J Legal Med.* 2019;133(4):1207–15. DOI: 10.1007/s00414-019-02053-w.

193. Hillewig E, Degroote J, Van der Paelt T, Visscher A, Vandemaele P, Lutin B, D’Hooghe L, Vandriessche V, Piette M, Verstraete K. Magnetic resonance imaging of the sternal extremity of the clavicle in forensic age estimation: towards more sound age estimates. *Int J Legal Med.* 2013;127(3):677–89. DOI: 10.1007/s00414-012-0798-z.
194. Wittschieber D, Schulz R, Vieth V, Küppers M, Bajanowski T, Ramsthaler F, Püschel K, Pfeiffer H, Schmidt S, Schmeling A. The value of sub-stages and thin slices for the assessment of the medial clavicular epiphysis: a prospective multi-center CT study. *Forensic science, medicine, and pathology.* 2014;10(2):163–9. DOI: 10.1007/s12024-013-9511-x.
195. Tangmose S, Jensen KE, Lynnerup N. Comparative study on developmental stages of the clavicle by postmortem MRI and CT imaging. *J Forensic Radiol Imaging.* 2013;1(3):102–6. DOI: 10.1016/j.jofri.2013.05.008.
196. Vieth V, Kellinghaus M, Schulz R, Pfeiffer H, Schmeling A. [Ossification stage of the medial clavicular epiphysis. Comparison of projectional radiography, computed tomography and magnetic resonance imaging]. *Rechtsmedizin.* 2010;20(6):483–8. DOI: 10.1007/s00194-010-0709-x. German.
197. Martínez Vera NP, Höller J, Widek T, Neumayer B, Ehammer T, Urschler M. Forensic age estimation by morphometric analysis of the manubrium from 3D MR images. *Forensic Sci Int.* 2017;277:21–9. DOI: 10.1016/j.forsciint.2017.05.005.
198. Neumayer B, Lesch A, Thaler F, Widek T, Tschauner S, De Tobel J, Ehammer T, Kirnbauer B, Boldt J, Wijk M van, Stollberger R, Urschler M. The four-minute approach revisited: accelerating MRI-based multi-factorial age estimation. *Int J Legal Med.* 2020;134(4):1475–85. DOI: 10.1007/s00414-019-02231-w.
199. Eggenreich S, Payer C, Urschler M, Štern D. Variational Inference and Bayesian CNNs for Uncertainty Estimation in Multi-Factorial Bone Age Prediction. Paper presented at: Medical Imaging meets NeurIPS. 33rd Conference on Neural Information Processing Systems (NeurIPS); 2019 Dec. 14; Vancouver, Canada. Available from: <https://arxiv.org/abs/2002.10819>.

Appendix

A. R-scripts

All scripts are modified from (158).

A.1. Normed likelihood curves

The following shows the R-script for the normed likelihood plot of the hand and wrist evaluation (Fig. 3.8), the plot for the clavicle normed likelihood (Fig. 3.11) used a similar script.

```
### maximum likelihood of age determination ###
max.like.hand =function (stage.conv=stage.conv, truage=truage){
  library(VGAM)
  #####
  theta.hand.prob.log = coef(vglm(stage.conv ~log(truage),cumulative(link='probit',parallel = TRUE,
    reverse = TRUE)))
  sd=1/unnname(theta.hand.prob.log[7])
  mu=unnname(-theta.hand.prob.log[1:6])*sd

  age1.seq=0.0001
  a<-seq(13,21,age1.seq)
  #####
  ptran<-function(t,ii){
    if(ii==1){p<-1-plnorm((t),mu[1],sd)
      return(p)}
    if(ii==7){p<-plnorm((t),mu[6],sd)
      return(p)}
    p<-plnorm((t),mu[ii-1],sd)-plnorm((t),mu[ii],sd)
    return(p)
  }
  #####
  x<-a
  max.like=0
  for(i in 1:7){
    y<-ptran(x,i)
    norm<-optimize(ptran,a,maximum=T,ii=i)$obj
    max.like[i]<-optimize(ptran,a,maximum=T,ii=i)$max
  }
  return(invisible(max.like))
}

max.like.GP<-max.like.hand(stage.conv=Hand_GP_stage, truage=Trueage)
max.like.GP
max.like.graph<-c(0,max.like.GP[2:6],99) #get rid of false max.like at endstages
```

```
##### normed likelihood plot ####

norm.like.hand.gg = function(stage.conv=stage.conv, truage=truage, max.like=max.like){
  library(VGAM)
  library(gridExtra)
  library(ggplot2)
  #=====
  theta.hand.prob.log = coef(vglm(stage.conv ~log(truage),cumulative(link='probit',parallel = TRUE,
    reverse = TRUE)))
  sd=1/unnname(theta.hand.prob.log[7])
  mu=unnname(-theta.hand.prob.log[1:6])*sd

  age1.seq=0.0001
  a<-seq(13,21,age1.seq)
  #=====
  ptran<-function(t,ii){
    if(ii==1){p<-1-plnorm((t),mu[1],sd)
      return(p)}
    if(ii==7){p<-plnorm((t),mu[6],sd)
      return(p)}
    p<-plnorm((t),mu[ii-1],sd)-plnorm((t),mu[ii],sd)
    return(p)
  }
  #=====
  maxlike<-max.like
  tit=c('GPS<=24','GPS<25','GPS<26/27','GPS<28','GPS<29','GPS<30','GPS<31')
  x<-a
  p.norm<-list()

  for(i in 1:7){
    y<-ptran(x,i)
    norml<-optimize(ptran,a,maximum=T,ii=i)$obj
    y<-y/norml
    s.tit<-tit[i]
    #=====
    df<-data.frame(cbind(x,y))

    p.norm[[i]]<-(ggplot(df, aes(x=x, y=y))+geom_line()+
      theme_bw()+ ylab ("NormLikelihood")+ xlab("Age<math>y</math>")+
      theme(axis.title.y = element_text(face="bold", size=14, vjust=1),
        axis.text.x = element_text( size=14),
        axis.text.y = element_text( size=14),
        axis.title.x = element_text(face="bold", size=14, vjust=0.9))+
      ggtitle(paste(s.tit))+
      theme(plot.title = element_text(vjust=1,face="bold",size=15, hjust = 0.5))+
      theme(panel.grid.major = element_blank(),panel.grid.minor = element_blank()+
        geom_segment(aes(x=18,y=0,xend=18,yend=1), color="black", linetype=2, size=0.3)+
        geom_hline(yintercept = 0.1465, color="black", linetype=3)+
        geom_hline(yintercept = 0, color="black", linetype=1)+
        geom_segment(x=maxlike[i],y=0,xend=maxlike[i],yend=1)+
        scale_y_continuous(breaks=seq(0, 1, 0.2))+
        scale_x_continuous(breaks=seq(13, 20.99, 1),labels = c("",14,"",16,"",18,"",20))
      )
    }
    return(invisible(p.norm))
  }
  p.norm.list<-norm.like.hand.gg(stage.conv=Hand_GP_stage, truage=Trueage, max.like = max.like.graph)

  do.call(grid.arrange,c(p.norm.list, ncol = 2, nrow = 4))
}
```

A.2. Markov Chain Monte Carlo (MCMC)

The following shows the R-script for the MCMC parameter estimation for the multi-factorial age estimation (section 2.5.2).

```
### startvalues ###

pre.MCMC = function (sto=sto)
{
  options(warn = -1)
  library(VGAM)
  NS=unname(apply(sto[,2:length(sto)],2,max))
  beta=matrix(NA,nr=2,nc=ncol(sto)-1)
  alpha=matrix(Inf,nr=ncol(sto)-1,nc=max(NS)+1)
  alpha[,1]=-Inf
  alpha[,2]=0
  sto = sto
  N = NROW(sto)
  age <- log(sto[, 1])
  for(i in 1:(ncol(sto)-1)){
    t1 = sto[, i+1]

    probit = coef(vglm(factor(t1)-age,cumulative(link='probit',parallel=T,reverse = F)))

    beta[1,i]=unname(-probit[1])
    beta[2,i]=unname(-probit[NS[i]])
    alpha_subtrahend<-probit[1]
    jj=1
    for(j in 3:NS[i]){
      alpha[i,j] = unname(probit[-1])[jj]-unname(alpha_subtrahend)
      jj=jj+1}
  }
  beta=matrix(beta,nc=1)
  return(list(beta,alpha))
}

pars = pre.MCMC(sto=c("Trueage", "stages"))

#####

### auxiliary function for corr matrix ###

to.symm = function(x)
{
  # A utility function to take the vector of correlation coefficients and place into a
  # symmetric matrix
  k=length(x)
  n.dim=(sqrt(8*k+1)+1)/2
  sym.mat=diag(n.dim)
  ip=0
  for(i in 1:(n.dim-1)){
    for(j in (i+1):n.dim){
      ip=ip+1
      sym.mat[i,j]=sym.mat[j,i]=x[ip]
    }
  }
  return(sym.mat)
}
```

```

### markov chain monte carlo ###
m.MCMC = function (N,saved.iters=5000,jump.sd=.1,x=a_g_e,
z=data.region.stage,thin=5,start.b=start.b,start.tz=start.tz)
{
  set.seed(123456)
  #R program for multivariate probit model
  N = length(x)
  x = cbind(rep(1,N),x)

  # d is the dimensionality of the probit (from # of columns in "z"
  # k is the number of explanatory variables (plus 1 for the intercept). Ordinarily k=2, for age and intercept.
  d=NCOL(z)
  k=NCOL(x)
  n.stages = apply(z,2,max)
  max.stages=max(n.stages)

  #create variables and starting values
  zstar=matrix(0,nrow(z),d)
  b=matrix(0,(d*k))
  # starting value 1. Intercept + slope
  b=start.b
  s=cs=diag(d)
  tz=matrix(0,d,max.stages+1)
  tz=start.tz
  ctz=tz
  acc1=acc2=acc3=acc4=acctot=0
  n.corr=d*(d-1)/2
  nc.results=sum(n.stages)+n.corr
  temp = rep(NA,sum(n.stages))
  results=matrix(0,nc=nc.results,nr=N.saved.iters)

  #begin Gibbs sampling
  ip=0
  acc=0
  iter=0
  for(i in 1:(thin*N.saved.iters)){
    #draw latent data: one-iteration gibbs sampler for truncated mvn simulation
    bb=matrix(b,2,d)
    m=x%*%bb
    for(j in 1:d){
      mm = m[,j] + (zstar[,j]-m[,j])%*%solve(s[-j,-j])%*%s[j,-j]
      ss = s[j,j]-t(s[j,-j])%*(solve(s[-j,-j]))%*%s[j,-j]
      zstar[,j]=qnorm(runif(nrow(z),
min=pnorm(tz[j,z[,j]],mm,sqrt(ss)),
max=pnorm(tz[j,(z[,j]+1)],mm,sqrt(ss))),mean=mm,sd=sqrt(ss))
    }

    # draw new thresholds from uniforms defined by max(zstar) for stage below
    # and min(zstar) for stage above threshold. Note that Cowles MH sampler would be
    # needed for large reference samples (typically not the case that they are)

    for(ii in 1:d){
      if(n.stages[ii]>2){
        for(jj in 3:(n.stages[ii])){
          tz[ii,jj]= runif(1,max(zstar[z[,ii]==(jj-1),ii]),min(zstar[z[,ii]==jj,ii)))
        }
      }
    }

    #draw b from mvn
    vb=solve(solve(s)%x%(t(x)%*%x))
    mn=vb%*(as.vector(t(x)%*%zstar)%*%t(solve(s)))
    b=mn+t(rnorm((d*k),0,1)%*%chol(vb))

    #use metropolis-hastings sampling to draw sigma
    e=matrix((as.vector(zstar)-(diag(d)%x%x%*%b)),nrow(z),d)
    v=t(e)%*%e
  }
}

```

```

    for(j in 1:(d-1)){
      for(jj in (j+1):d){
        cs=s
        iter=iter+1
        like=-.5*(nrow(z)+n.corr+1)*log(det(s))-.5*sum(diag(v%*%solve(s)))
        cs[j,j]=cs[jj,j]=cs[j,j]+rnorm(1,mean=0,sd=jump.sd)
        if(abs(cs[j,j])<1){
          det.cs = det(cs)
          if(det.cs<=0) {clike=-Inf}
          else {clike=-.5*(nrow(z)+n.corr+1)*log(det.cs)-.5*sum(diag(v%*%solve(cs)))}
          if((clike-likelike)>log(runif(1,0,1))) {
            acc=acc+1
            s=cs
          }
        }
      }
    }

    if(i%%thin==0){
      cat(i/thin, ' ', round(acc/iter*100,2), '\r')
      flush.console()
      ip=ip+1
      b.vec=as.vector(b)
      ip2 = 0
      for(i.trait in 1:d){
        ip2=ip2+1
        temp[ip2] = -b.vec[2*i.trait-1]
        if(n.stages[i.trait]>2){
          for(j.stage in 3:(n.stages[i.trait])){
            ip2=ip2+1
            temp[ip2] = tz[i.trait,j.stage]-b[2*i.trait-1]
          }
          ip2=ip2+1
          temp[ip2] = b.vec[i.trait*2]
        }
        results[ip,]=c(temp,s[lower.tri(s)])
      }
    }
    cat('\n\n')
    find.it = matrix(NA,nr=d,nc=3)
    ip2 = 0
    for(i.trait in 1:d){
      find.it[i.trait,1]=i.trait
      ip2=ip2+1
      find.it[i.trait,2]=ip2
      if(n.stages[i.trait]>2){
        for(j.stage in 3:(n.stages[i.trait])){
          ip2=ip2+1
        }
      }
      ip2=ip2+1
      find.it[i.trait,3]=ip2
    }
    ip2=ip2+1
    colnames(find.it)=c('Trait', 'start', 'stop')
    print(find.it)
    cat(paste('Corr matrix starts at: ', ip2, ' and ends at: ', ip2+n.corr-1, '\n\n'))
    return(results)
  }

theta<-m.MCMC(N.saved.iters=5000, jump.sd=.1, x=log(Trueage), z=stages, thin=40, start.b = pars[[1]], start.tz = pars[[2]])

library(coda)
library(fitR)
theta.mcmc<-mcmc(theta, thin = 40)
theta.mcmc.burn<-burnAndThin(theta.mcmc, burn = 500)
theta.mcmc.burn.sum<-summary(theta.mcmc.burn)
theta.burn.matrix<-matrix(NA,nrow = , ncol = ) # Fill matrix according to the MCMC parameter results
theta.burn.corr<-to.symm(theta.mcmc.burn.sum$statistics[]) # Fill with the correlation results

```

A.3. Posterior density calculation

The following shows the R-script for the posterior density calculations for the multi-factorial age estimation (section 2.5.2).

```
### function calculation MLAge + predicted values for one subject ###

age1.seq=0.01
age1<-seq(13,25,age1.seq)

Post.dens_tt = function (x,stages=stages, right=right){
  set.seed(123456)
  stages=as.matrix(stages)

  threshes1=c(-Inf,theta.burn.matrix[1,1:6],Inf)
  threshes2=c(-Inf,theta.burn.matrix[2,1:5],Inf)
  threshes3=c(-Inf,theta.burn.matrix[3,1:4],Inf)
  threshes4=c(-Inf,theta.burn.matrix[4,1:4],Inf)

  slopes=theta.burn.matrix[,7]
  r=theta.burn.corr

  bot = c(threshes1[stages[1]],threshes2[stages[2]],threshes3[stages[3]],threshes4[stages[4]])
  top = c(threshes1[stages[1]+1],threshes2[stages[2]+1],threshes3[stages[3]+1],threshes4[stages[4]+1])

  library(mvtnorm)

  get.densr = function(x){
    return(pmvnorm(mean=slopes*log(x),lower=bot,upper=top,corr=r)[1]*dunif(x,13,right))
  }
  denomr = integrate(Vectorize(get.densr),13,right)$val
  get.postr = function(x){
    return(pmvnorm(mean=slopes*log(x),lower=bot,upper=top,corr=r)[1]*dunif(x,13,right)/denomr)
  }
  age0c=optimize(get.postr,int=c(13,right),maximum=T)$max

  p=0
  j=x
  pred.values<-matrix(NA,nr=length(j),nc=2)
  for(i in 1:length(j)){
    p[i] = get.postr(j[i])
    test1[i,1]<-p[i]
  }
  test1[,2]<-j

  return(list(age0c,pred.values))
  # age0C... point prediction
  # pred. values... predicted values for the posterior density curve
}

ttt<-Post.dens_tt(age1,stages=stages,right=25)
```

B. Ethics vote

Ethikkommission



Medizinische Universität Graz

Auenbruggerplatz 2, A-8036 Graz
ethikkommission@medunigraz.at
Tel.: +43 / 316 / 385-13928, Fax: -14348

FOLGEVOTUM gültig bis 18.06.2020

EK-Nummer: 21-399 ex 09/10
Studientitel: Validierung von MR-basierten Methoden zur forensischen Altersschätzung
Prüfer: Prof.Dr. Johann Pfeifer
LBI für klinisch-forensische Bildgebung
Sponsor: Ludwig Boltzmann Gesellschaft GmbH
Ansprechpartner: Dr. Peter Mayerhofer, 1090 Wien, Nussdorferstr. 64/4
CRO: -
Antragsteller: LBI für klinisch-forensische Bildgebung
Ansprechpartner: Prof.Dr. Johann Pfeifer, 8010 Graz, Universitätsplatz 4/2

Die o.a. Studie wurde von der Ethikkommission erstmals im 'expedited Review' am 18.06.2010 behandelt. Die Ethikkommission ist zu folgendem Schluss gekommen:

Es besteht kein Einwand gegen die Durchführung der Studie in der vorliegenden Form.

Kommissionsmitglieder, die für diesen Tagesordnungspunkt als befähigt anzusehen waren und daher gemäß Geschäftsordnung an der Entscheidungsfindung und Abstimmung nicht teilgenommen haben: keine

Zur Beurteilung vorliegende Dokumente:

Dokumente eingegangen am 26.05.2010, begutachtet im 'expedited Review' am 18.06.2010	
✓ Antragsformular	25.05.2010
✓ Originalprotokoll 2	25.05.2010
✓ Informed Consent Form Probanden 2	25.05.2010
✓ Informed Consent Form Eltern 2	25.05.2010
✓ Informed Consent Form Jugendliche 2	25.05.2010
✓ Informed Consent Form Kinder 1	25.05.2010
✓ Fragebögen Informations- und Fragebogen zur Magnetresonanztomographie	
Sonstiges: Informations- und Fragebogen zur Magnetresonanztomographie, undatiert	
✓ Sonstiges: Antrag auf Kostenbefreiung	25.05.2010
✓ Sonstiges: Gesucht! Probanden für Studie	25.05.2010
Dokumente eingegangen am 18.07.2011, begutachtet im 'expedited Review' am 21.07.2011	
✓ Zwischenbericht	18.07.2011
Dokumente eingegangen am 22.05.2012, begutachtet im 'expedited Review' am 25.05.2012	
✓ Zwischenbericht	22.05.2012
Dokumente eingegangen am 08.05.2013, begutachtet im 'expedited Review' am 17.05.2013	
✓ Zwischenbericht	08.05.2013
Dokumente eingegangen am 16.06.2014, begutachtet im 'expedited Review' am 26.06.2014	
✓ Zwischenbericht	10.06.2014
Dokumente eingegangen am 06.02.2015 (in der nächsten Begutachtung mitbegutachtet)	
Sonstiges: EK-Meldeformular - Prüferwechsel	04.02.2015

EK-Nummer: 21-399 ex 09/10 Votum (21.06.2019) Seite 1 von 3

Medizinische Universität Graz, Auenbruggerplatz 2, A-8036 Graz, www.medunigraz.at

Rechtsform: Juristische Person öffentlichen Rechts gem. Universitätsgesetz 2002. Information: Mitteilungsblatt der Universität und www.medunigraz.at, DVR-Nr. 210 9494.
UID: ATU 576 111 79, Bankverbindung: Bank Austria Creditanstalt BLZ 12000 Konto-Nr. 500 949 400 04, Raiffeisen Landesbank Steiermark BLZ 38000 Konto-Nr. 49510

Dokumente eingegangen am 16.03.2015 (in der nächsten Begutachtung mitbegutachtet)	
✓ Protokoll Amendment 1 zu Protokollversion 2	16.03.2015
✓ Informed Consent Form Jugendliche 3	16.03.2015
✓ Informed Consent Form Eltern 3	16.03.2015
✓ Informed Consent Form Probanden 3	16.03.2015
✓ Sonstiges: EK-Meldeformular	
Dokumente eingegangen am 25.03.2015, begutachtet im 'expedited Review' am 01.04.2015	
✓ Werbematerial Aushang für Schule 1	25.03.2015
✓ Werbematerial Aushang für Universität, Sportvereine 1	25.03.2015
Dokumente eingegangen am 19.05.2015, begutachtet im 'expedited Review' am 29.05.2015	
✓ Zwischenbericht	19.05.2015
Dokumente eingegangen am 02.03.2016, begutachtet im 'expedited Review' am 16.03.2016	
✓ Protokoll Amendment 2	02.03.2016
✓ Informed Consent Form Jugendliche 4	02.03.2016
✓ Informed Consent Form Kinder 2	02.03.2016
✓ Informed Consent Form Eltern 4	02.03.2016
✓ Informed Consent Form Probanden 4	02.03.2016
✓ Werbematerial Aushang 2	29.02.2016
✓ Sonstiges: EK-Meldeformular	02.03.2016
Dokumente eingegangen am 24.05.2017, begutachtet im 'expedited Review' am 02.06.2017	
✓ Zwischenbericht	24.05.2017
Dokumente eingegangen am 03.11.2017 (in der nächsten Begutachtung mitbegutachtet)	
✓ Originalprotokoll 3	30.10.2017
✓ CV Prüfer	02.11.2017
✓ Sonstiges: EK-Meldeformular	03.11.2017
Dokumente eingegangen am 12.12.2017, begutachtet im 'expedited Review' am 04.01.2018	
✓ Informed Consent Form Eltern 5	12.12.2017
✓ Informed Consent Form Probanden 5	12.12.2017
Dokumente eingegangen am 28.05.2018, begutachtet im 'expedited Review' am 30.05.2018	
✓ Zwischenbericht	28.05.2018
Dokumente eingegangen am 03.06.2019 (in der nächsten Begutachtung mitbegutachtet)	
✓ Zwischenbericht	03.06.2019
Dokumente eingegangen am 12.06.2019, begutachtet im 'expedited Review' am 21.06.2019	
✓ Informed Consent Form Jugendliche 5	12.06.2019
✓ Informed Consent Form Eltern 6	12.06.2019
✓ Informed Consent Form Probanden 6	12.06.2019

Datum Erstvotum: 18.06.2010

Die Ethikkommission geht - rechtlich unverbindlich - davon aus, dass es sich um keine klinische Prüfung nach AMG bzw. MPG handelt.

Das Votum der Ethikkommission berührt in keiner Weise die alleinige Verantwortung der Prüferin / des Prüfers / der Prüfer für die ordnungsgemäße Durchführung der Studie unter Einhaltung aller einschlägiger gesetzlicher Bestimmungen und Richtlinien.

Weiters machen wir darauf aufmerksam, dass der Kommission unverzüglich zu melden sind:

- Abweichungen vom Protokoll aus Sicherheitsgründen oder Protokolländerungen
- Änderungen, die das Risiko der Teilnehmer/-innen erhöhen oder die Durchführung der Studie wesentlich beeinflussen
- Mutmaßliche unerwartete schwerwiegende Nebenwirkungen - SUSARs (AMG-Studien ab 1.5.2004) oder schwerwiegende unerwünschte Ereignisse - SAEs (andere Studien)
- Jegliche Information über sonstige Umstände, die die Sicherheit der Teilnehmer/-innen oder die Durchführung der Studie beeinträchtigen können

EK-Nummer: 21-399 ex 09/10

Votum (21.06.2019)

Seite 2 von 3

Graz, 21. Juni 2019



Univ.Prof.Dr. Josef Haas
Vorsitzender



Univ.Prof.Dr. Hans Dimai
Stv. Vorsitzender

Achtung: Bitte bei allen das Projekt betreffende Schreiben oder telefonischen Anfragen die EK-Nummer angeben!